



JOURNAL OF
APPLIED RESEARCH
ON SCIENCE
AND TECHNOLOGY



INSTITUTE OF RESEARCH AND DEVELOPMENT
RAJAMANGALA UNIVERSITY OF TECHNOLOGY THANYABURI

ISSN Online
2773-9473

Volume 25 Issue 2
May – August 2026

JARST
RMUTT

Journal of Applied Research on Science and Technology (JARST)

E-ISSN : 2773-9473

Vol. 25 No. 2 (May – August 2026)

The Journal of Applied Research on Science and Technology (JARST) is an international open-access journal dedicated to disseminating high-quality scientific and technological research. It provides a platform for academics, professionals, and industrial practitioners to exchange knowledge, promote innovation, and strengthen academic and research collaborations at both national and international levels. All submissions must be original, unpublished, and not under consideration elsewhere. The authors bear full responsibility for the accuracy, originality, and integrity of their work.

Aims and Scope

The Journal of Applied Research on Science and Technology (JARST) aims to disseminate high-quality research and innovative ideas in science and technology, fostering academic and industrial collaboration at both national and international levels.

The scope of the journal covers General Engineering, Materials Science, Agricultural and Biological Sciences, Computer Science, and Mathematics, with particular emphasis on research that integrates fundamental understanding with practical applications.

JARST publishes original research and review articles that contribute to scientific advancement and sustainable development. All accepted articles are published online immediately and freely accessible worldwide without restriction.

Review Process

1. Evaluation by the Editor-in-Chief

The Editor-in-Chief will see whether the topic and theme of the article are appropriate and congruent with the stipulated objectives and format of the Journal. Plagiarism and benefits relating to theory and business contribution will also be investigated. The submitted paper may be returned to the author for preliminary revising or, if the aforesaid criteria are not met, rejected.

2. Evaluation by Review

All submitted manuscripts must be reviewed by at least two expert reviewers in the related fields. Reviewers will evaluate the quality of submitted article for publication via the double-blinded review system.

3. Evaluates the Reviews

The Editor-in-Chief make decision for article publication based on the external readers' evaluation. The said decision is either accepting the article for publication, rejecting it, or resending it back to the author for further elaborating revision.

Period of Issued Journal

The Journal of Applied Research on Science and Technology (JARST) will be 3 issues/year, as follows:

1st issue: January - April

2nd issue: May - August

3rd issue: September - December

Advisory Board

Sommai Pivsa-Art
Krischonme Bhumkittipich
Kiattisak Sangpradit
Sorapong Pavasupree
Boonyang Plangklang
Syuji Fujii

Rajamangala University of Technology Thanyaburi, Thailand
Rajamangala University of Technology Thanyaburi, Thailand
Rajamangala University of Technology Thanyaburi, Thailand
Rajamangala University of Technology Thanyaburi, Thailand
Rajamangala University of Technology Thanyaburi, Thailand
Osaka Institute of Technology, Japan

Editor-in-Chief

Amorn Chaiyasat Rajamangala University of Technology Thanyaburi, Thailand

Assistant Editors

Chatchai Ponchio Rajamangala University of Technology Thanyaburi, Thailand
Chatchai Veranitisagul Rajamangala University of Technology Thanyaburi, Thailand
Thammasak Rojviroon Rajamangala University of Technology Thanyaburi, Thailand

Editorial Board

Arunachala Mada Kannan	Arizona State University, USA
Chatthai Kaewtong	Maharakham University, Thailand
Clastinruselraj I. Sathish	The University of Newcastle, Australia
Chaudry Masood Khaliq	Shandong University of Science and Technology, South Africa
Fernando Alberício	University of KwaZulu-Nata, South Africa
Fiazuddin Zaman	University of the Witwatersrand, South Africa
Hideto Minami	Kobe University, Japan
Katsuyuki Takahashi	Iwate University, Japan
Manoj Khandelwal	Federation University Australia, Australia
Md. Shahidul Islam	University of KwaZulu-Natal, South Africa
Pakorn Opaprakasit	Sirindhorn International Institute of Technology, Thailand
Kelvin Huang-Chou Chen	National Pingtung University, Taiwan
Peeyush Soni	Indian Institute of Technology Kharagpur, India
Ryo Honda	Kanazawa University, Japan
Siti Afiqah Binti Zainuddin	Universiti Malaysia Kelantan, Indonesia
Suhardono, Sapta	Universitas Sebelas Maret, Indonesia
Sukono, null	Universitas Padjadjaran, Indonesia
Tran Hung Tra	Nha Trang University, Viet Nam
Warayuth Sajomsang	Thailand National Nanotechnology Center, Thailand
Yukiya Kitayama	Osaka Metropolitan University, Japan

Managing Department

Jittima Singto	Rajamangala University of Technology Thanyaburi, Thailand
Monticha Ruttanapan	Rajamangala University of Technology Thanyaburi, Thailand
Nuthawan Thamawatchakorn	Rajamangala University of Technology Thanyaburi, Thailand
Praphatson Thongchai	Rajamangala University of Technology Thanyaburi, Thailand
Saranya Suwinai	Rajamangala University of Technology Thanyaburi, Thailand
Yuvada Sangjan	Rajamangala University of Technology Thanyaburi, Thailand

Contact

Institute of Research and Development, Rajamangala University of Technology Thanyaburi (RMUTT)
39 Moo 1, Klong 6, Khlong Luang Pathum Thani 12110 Thailand
Website: <https://ph01.tci-thaijo.org/index.php/rmutt-journal/index>
Phone: +66 2 5494492, +66 2 5494681
Fax: +66 2 5494680
Email: jarst@rmutt.ac.th

Editorial Note

The Journal of Applied Research on Science and Technology (JARST) is an academic journal published by the Institute of Research and Development, Rajamangala University of Technology Thanyaburi (RMUTT). The journal aims to disseminate high-quality research and innovative ideas in the fields of General Engineering, Materials Science, Agricultural and Biological Sciences, Computer Science, and Mathematics, serving as a platform for researchers, academics, faculty members, and students at both national and international levels.

This issue presents nine research articles encompassing a diverse range of topics, such as Impact of abiotic stress and micronutrient supplementation on the 2-acetyl-1-pyrroline content in KDML105 rice (*Oryza sativa* L.), Fake news detection through multi-level contextual and sequential modeling, Gamma shielding efficiency of sedimentary rock-based concrete blocks under Co-60 irradiation, A reinforcement learning model for route allocation optimization: A case study of C-130H transport aircraft in the Royal Thai Air Force, Product development of mangosteen (*Garcinia mangostana* L) with karonda (*Carissa carandas* L.) and rice protein sorbet, Effective detection of Thai fake news using machine learning method, Physical property of 925 sterling silver metal injection molding and morphology of brown parts under varying solvent-based debinding conditions, Isolation and selection of *Lactiplantibacillus plantarum* CLT-L01 (TISTR P088) from traditional fermented vegetables for solid-state fermentation of soybean meal (SBM) and Enhancing lung cancer screening accuracy through a dual-pipeline deep learning and machine learning framework for integrated risk assessment. These articles collectively contribute to the dissemination of knowledge across the physical and life sciences, providing valuable insights for future applications and developments.

Lastly, the editorial team extends its sincere appreciation to all contributors, reviewers, and supporters whose efforts have made this publication possible. We look forward to your continued cooperation and support in advancing the mission of JARST in the years to come.

Editorial Team

Contents

Research Articles	Page
Impact of abiotic stress and micronutrient supplementation on the 2-acetyl-1-pyrroline content in KDML105 rice (<i>Oryza sativa</i> L.) <i>Paweena Saleethong, Supatra Khabuanchalad, Laddawan Kammaphana and Noppawan Nounjan</i>	264595
Fake news detection through multi-level contextual and sequential modeling <i>Pimpa Cheewaparakobkit and Rindra Razafinjatovo</i>	264248
Gamma shielding efficiency of sedimentary rock-based concrete blocks under Co-60 irradiation <i>Pitchpilai Khoonphunnarai, Sutthisa Konruang and Phayao Yongsiriwit</i>	263233
A reinforcement learning model for route allocation optimization: A case study of C-130H transport aircraft in the Royal Thai Air Force <i>Patikorn Anchuen, Phummipat Daungklang, Nuntipat Phisutthangkoon, Nattawat Tanomchad and Hatsadin Jantaboon</i>	265635
Product development of mangosteen (<i>Garcinia mangostana</i> L.) with karonda (<i>Carissa carandas</i> L.) and rice protein sorbet <i>Jinussiga Monchaising, Sunisa Dorndi, Kamonchat Wanthongchan, Thiwaporn Jitdee, Amornrat Suwanposri, Warunya Nonmuang and Kannikar Charoensuk</i>	265905
Effective detection of Thai fake news using machine learning method <i>Nopphadol Sumleeroung, Komsan Kanjanasit and Nithizethe Mhuadthongon</i>	266903
Physical property of 925 sterling silver metal injection molding and morphology of brown parts under varying solvent-based debinding conditions <i>Montri Kawsuk, Witthaya Daodon and Surat Wannasri</i>	266270
Isolation and selection of <i>Lactiplantibacillus plantarum</i> CLT-L01 (TISTR P088) from traditional fermented vegetables for solid-state fermentation of soybean meal (SBM) <i>Phongsupha Chanthachaiyaphum, Wang Yue and Wang Wenbo</i>	268321
Enhancing lung cancer screening accuracy through a dual-pipeline deep learning and machine learning framework for integrated risk assessment <i>Patthanan Uthaitipitak</i>	267600



Impact of abiotic stress and micronutrient supplementation on the 2-acetyl-1-pyrroline content in KDML105 rice (*Oryza sativa* L.)

Paweena Saleethong^{1*}, Supatra Khabuanchalad¹, Laddawan Kammaphana¹ and Noppawan Nounjan²

¹Faculty of Agriculture and Technology, Rajamangala University of Technology Isan Surin Campus, Surin 32000, THAILAND

²International College, Khon Kaen University, Khon Kaen 40002, THAILAND

Corresponding author: paweena.sl@rmuti.ac.th

ABSTRACT

This study investigated the effects of abiotic stress and mineral supplementation on 2-acetyl-1-pyrroline (2-AP) content in Thai jasmine rice (*Oryza sativa* L. cv. 'Khao Dawk Mali 105'; KDML105). Rice aroma quality fundamentally influences the market value of aromatic rice. Therefore, enhancing the aromatic compound, 2-acetyl-1-pyrroline (2-AP) could improve rice aroma quality. In this study, KDML105 rice plants at the reproductive stage were used to investigate the effects of abiotic stress (drought, salinity and combined salt-drought stress) and mineral supplementation (Ca, Cu, Mg, Fe, and Zn) on 2-AP content compared to the control plants which were not treated with abiotic stress and exogenous minerals supplementation. The results found that proline levels in plants which were treated with salt stress or combined salt-drought stress accumulated higher leaf proline content (7.7-fold and 9.7-fold increases) compared to those of control plants (30.15 µg/g FW). Although the proline content increased, there were no significant differences in 2-AP levels of the plants. In contrast, the proline content of plants subjected to minerals did not increase while the 2-AP content increased compared to the control plants (3.37 ppm), particularly in plants supplemented with Ca and Cu (6.10 and 5.26 ppm, respectively). These findings indicated that mineral supplements did not stimulate proline production. However, it could induce or suppress other compounds in 2-AP synthesis pathway, resulting in high 2-AP accumulation. This was probably due to the downregulation of the betaine aldehyde dehydrogenase 2 gene (*BADH2*). Not only gene regulations in the pathway, but also external factors, such as the drying process after harvesting, affect 2-AP content. The results revealed that the solar-dried grain method caused higher 2-AP content compared to hot-air-dried and undried grains. Thus, targeted mineral supplementation and appropriate post-harvest processes are involved in enhancing aromatic quality in jasmine rice. In addition, supplemented plants with Ca significantly enhanced FRAP activity accompanied by remarkable increases in both total phenolic and total flavonoid contents.

Keywords: 2-acetyl-1-pyrroline, Aromatic rice, Abiotic stress, Antioxidant capacity, Nutrient supplementation

INTRODUCTION

Khao Dawk Mali 105 (KDML105) rice, internationally recognized as "jasmine rice", holds a distinguished position in Thailand cuisine and the global rice market due to its exceptional physical appearance, cooking quality, and distinctive aroma [1, 2]. The 12th World Rice Conference 2020 reported that jasmine rice dominated the global market [3]. However, Jasmine Rice still faces the challenge of high competition with other rice varieties such as Malys Angkor from Cambodia and ST24 from Vietnam [4]. The significant price premium commanded by KDML105, approximately double that of other rice cultivars in international markets [5], which raises the critical importance of maintaining and enhancing its quality characteristics.

Rice aroma quality fundamentally influences market value [6]. KDML105's distinctive fragrance stems from the accumulation of 2-acetyl-1-pyrroline (2-AP), a volatile compound whose synthesis correlates with proline content [7]. Multiple environmental and physiological factors modulate 2-AP synthesis and degradation, resulting in substantial variations in grain aromatic content [5, 8].

The 2-AP biosynthesis in aromatic rice is composed of complex networks. It is activated by the reformation of glutamate, proline, and ornithine to Δ^1 -pyrroline-5-carboxylate (P5C) by proline dehydrogenase (PDH), Δ^1 -pyrroline-5-carboxylate synthetase (P5CS), and ornithine aminotransferase (OAT), respectively. The pathway culminates in a non-enzymatic reaction between P5C/ Δ^1 -pyrroline and methylglyoxal to form 2-AP [9, 10]. Previous studies examining 2-AP accumulation in various

above-ground tissues of rice plants have demonstrated that the flag leaf exhibits the highest 2-AP content. Consequently, utilizing this specific tissue for analyzing the relationship and gene expression patterns associated with aroma compound biosynthesis could enhance our understanding of the underlying mechanisms of 2-AP formation [11].

It is reported that when aromatic rice grows under unappreciated growth conditions such as rain-fed, drought, or dry soil, it results in an increase in 2-AP content in its grains [7]. Moreover, high 2-AP levels in rice grains may depend on farming system. [5] presented that aromatic rice plants grown under organic farming systems had high 2-AP levels which were caused by high levels of nitrogen, humic acid and microbial populations found in the system. Not only are growth conditions and farming systems involved in 2-AP accumulation in grains, but mineral elements that plants absorb also influence 2-AP content. Each mineral element contributed to variation in 2-AP accumulation in rice plants depending on the growth stage of rice plants. For example, boron could induce high accumulated 2-AP during the heading stage [12]. Previous studies have showed that 2-AP increases under various stress conditions, which are the same conditions that activate the antioxidant system. Furthermore, the biosynthesis of 2-AP is associated with the metabolic pathways of proline and other precursors, which are directly linked to the plant's antioxidant system [13, 14].

The objective of this work was to induce high 2-AP accumulation in rice grains by optimizing growth conditions and evaluating the effects of individual micronutrient supplementation on 2-AP accumulation in rice grains. This included studying how individual micronutrient supplementation affects 2-AP accumulation in rice grains. Post-harvesting processes were also determined to compare the best way to preserve 2-AP in rice grains. The effects of micronutrient supplementation on antioxidant compounds were also investigated.

MATERIALS AND METHODS

Plant material, growth conditions and treatments

Seeds of rice (*Oryza sativa* L. cv. 'KDML105') used in this study were provided by the Surin Rice Research Center, Thailand. Preliminary experiments tested salt concentrations of 25, 50, 75, 100, 125 and 150 mM, and water deficit periods of 3, 5, 7, 9 and 11 days. The concentration of 100 mM NaCl and 7-day water deficit were selected as they enabled rice growth to the reproductive stage with maximum proline accumulation. The rice seeds were soaked in water for 2 days following by germinating in water-soaked tissues for 7 days. Then, four seedlings were transferred to a plastic pot containing 5 kg of soil and grown under natural sunlight in a greenhouse. There were five replications

per group. When the growth of rice plants was in the booting stage (approximately 60 days after transplantation), plants were divided into 9 groups as follows: control, no abiotic stress and no minerals supplementation, drought stress (water withheld for 7 days), then rewatering until the heading stage, salt stress (plants were watered with 100 mM NaCl for 10 days), then watered with water without NaCl solution until they reached the heading stage. For the combined stress treatment, rice plants were watered with 100 mM NaCl solution for 7 days followed by 7 days of water withholding. For the exogenous mineral applications that have been modified from Yoshida's rice nutrient solution formula [15], calcium 0.8 M ($\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$), copper 1.24 mM ($\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$), magnesium 1.3 M ($\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$), iron 28.5 mM ($\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$) and zinc 1.22 mM ($\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$). Each pot was supplied with 5 L per day for 10 days until the plants reached the heading stage. The plants in these groups were then allowed to grow until grain formation stage. There were five replications per group.

2-Acetyl-1-pyrroline (2-AP) content analysis

To analyze 2-acetyl-1-pyrroline (2-AP), first configure the Gas Chromatography (GC) system by setting the Nitrogen-Phosphorus Detector (NPD) to 250 °C, the column oven to 200 °C, and the Helium carrier gas flow to 3 ml/min, while maintaining detector gas flows of 80 ml/min for Zero Air and 4 ml/min for Hydrogen. Next, set the Headspace (HS) parameters with an oven temperature of 120 °C, an injector at 130 °C, and a transfer line at 140 °C. Once both systems reach their setpoint temperatures and the NPD baseline signal is stable, load the prepared sample vials into the HS autosampler tray and initiate the sequence by clicking "Start Run" to analyze all samples sequentially. Finally, quantify the 2-AP concentration by comparing the sample peak area against a 2-AP standard calibration curve [16].

RNA extraction and gene expression analysis

Total RNA was extracted from rice leaf tissues (0.1 g) using TRI reagent (Thermo Fisher Scientific, Waltham, MA, USA). DNase treatment was performed to eliminate genomic DNA contamination by RQ1 RNase-Free DNase kit, Promega. RNA quantity was measured with a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Fisher Scientific, Inc.) following protocols described by [17]. Then, the expressions levels of genes associated with 2-AP biosynthesis, ornithine aminotransferase (OAT), proline dehydrogenase (PDH), Δ 1-pyrroline-5-carboxylate synthetase (P5CS), and betaine aldehyde dehydrogenase 2 (BADH2), were determined. The actin gene was used for calculating relative expression level.

Grain drying process

To evaluate the effect of drying methods on 2-AP content, rice grains were treated under three drying conditions: (1) solar drying for 3 days, (2) hot air drying at 60°C for 3 days, and (3) no drying (control). Following the drying treatments, grain samples were analyzed for 2-AP content using the previously described analytical method to determine the influence of post-harvest processing on aromatic compound retention.

Total phenolics content, total flavonoid content and antioxidant capacity in rice grains

To determine bioactive compounds and antioxidant properties, two grams of dehulled rice seeds were used to quantify total phenolic compounds (TP) and total flavonoid compounds (TF). In addition, ferric reducing antioxidant power (FRAP) was measured as an indicator of antioxidant capacity using the method described by [18, 19].

Proline content determination

Fresh rice leaves (0.1 g) and roots (0.2 g) were cut into small pieces, placed in a mortar, and ground thoroughly. Five milliliters of 3% sulphosalicylic acid were added and grinding continued until the tissue turned brown. The mixture was filtered through filter paper, and the filtrate volume was measured. Two milliliters of filtrate were transferred to a screw-cap test tube, mixed with 2 mL glacial acetic acid and 2 mL ninhydrin using a vortex mixer, then boiled in water for 1 hour. The reaction was stopped by cooling in ice water. Four milliliters of toluene were added and mixed using a vortex mixer. After allowing air bubbles to dissipate, the pink-colored toluene layer was collected and absorbance was measured at 520 nm. Proline content was determined using a standard curve and calculated using the formula (Eq.1):

$$\text{Proline content (mg/g tissue)} = \frac{[V \times \text{proline from standard curve (mg)}]}{(2 \times W)} \quad (1)$$

where V = total volume of extract (mL)
and W = leaf weight (g) [20].

Statistical analysis

The experiment employed a randomized complete block design (RCBD) with five replications for each treatment. Statistical significance among treatment means was determined using analysis of variance (ANOVA), with significance determined at $P < 0.05$. Multiple comparisons were performed using Duncan's multiple range test (DMRT) to identify specific differences between treatment groups.

RESULTS AND DISCUSSION

The effects of abiotic stress (drought, salt, and combined drought and salt stress) and mineral supplementation (Ca, Cu, Mg, Fe, and Zn) on proline

and 2-AP contents were systematically investigated in this study. It was found that abiotic stress treatments caused a marked increase in proline accumulation in rice leaves at the booting stage (Table 1). The highest proline content was observed in combined salt and drought stress (291.54 µg/g FW), followed by those exposed to salt stress alone (231.43 µg/g FW). Notably, there was no significant difference between those groups. In contrast, mineral supplementation treatments resulted in proline concentrations ranging from 25.81 to 32.37 µg/g FW, with no significant differences observed among treatments.

These findings indicated that exposure of rice plants to abiotic stress, particularly salt stress or combination between salt and drought stresses, induced significant accumulation of proline, while mineral supplementation did not significantly alter proline levels compared to the control.

Conversely, 2-AP content in rice grains harvested from plants grown under abiotic stress treatment did not show any significant difference. These findings are inconsistent with previous studies that presented that high 2-AP content was found in rice grains collected from plants subjected to salt or drought stress [21, 22]. This may be caused by differences between NaCl concentrations and stress duration employed in this experiment.

Nevertheless, the highest 2-AP content (3.97 ppm) was found in grains collected from salt-stressed plants compared to other abiotic stress treatments which were range approximately (3.22 – 3.37 ppm). Under the mineral supplementation, exogenous Ca was able to enhance the highest 2-AP content (6.10 ppm) followed by supplementation with copper (5.26 ppm) (Table 1).

It is generally known that a marked increase in proline level is found when plants grow under inappropriate conditions, such as abiotic stress, to protect plant cells from abiotic stress [23]. In this case, it acts as an osmoprotectant that is why the proline level observed in this experiment was higher in plants treated with drought, NaCl, and combined drought and salt stress. According to the result, it suggested that high proline accumulation in the booting stage did not associate with high 2-AP accumulation in rice grains, though proline is the molecule involved in 2-AP synthesis. [24] suggested that forming 2-AP in grains is associated with the balance of carbon and nitrogen, grain filling conditions, and the expression level of the *BADH2* gene, which occurred later.

In contrast, supplementing plants with micronutrients did not induce stress and resulted in low proline accumulation. As mentioned earlier, 2-AP formation in grains depends on several factors. Exogenous Ca may stimulate supporting factors such as carbon and nitrogen assimilation and sugar metabolism and enhance enzyme activities leading

to high 2-AP accumulation [25, 26]. These results were aligned with previous research reporting that appropriate concentrations of mineral elements such as nitrogen, phosphorus, calcium, manganese, zinc, boron, copper, and selenium can promote 2-AP content in

aromatic rice grains [27]. The research by [28] pointed out that nano fertilizers containing calcium (and phosphorus; hydroxyapatite nanoparticles) improved 2-AP content in rice grains, suggesting a beneficial role for calcium in stimulating 2-AP production.

Table 1 Proline content in leaves at the booting stage and 2-AP content in rice grains following drought treatment, salt treatment, combined salt and drought stresses and various mineral supplementation treatments.

Treatment	Leaf of booting stage rice Proline ($\mu\text{g} / \text{gFW}$)	Rice grains 2-AP (ppm)
Control	30.15 $\pm$ 1.65 ^b	3.37 $\pm$ 0.48 ^c
Drought (7days)	50.85 $\pm$ 6.45 ^b	3.22 $\pm$ 0.28 ^c
NaCl (100 mM)	231.43 $\pm$ 61.15 ^a	3.97 $\pm$ 0.10 ^{bc}
Drought + NaCl	291.54 $\pm$ 53.63 ^a	3.37 $\pm$ 0.54 ^c
Ca	25.81 $\pm$ 3.66 ^b	6.10 $\pm$ 0.40 ^a
Cu	26.91 $\pm$ 3.74 ^b	5.26 $\pm$ 1.46 ^{ab}
Mg	32.37 $\pm$ 4.85 ^b	3.94 $\pm$ 0.05 ^{bc}
Fe	27.16 $\pm$ 2.05 ^b	3.65 $\pm$ 0.82 ^{bc}
Zn	26.87 $\pm$ 5.17 ^b	4.28 $\pm$ 0.71 ^{bc}

Different letters in the same column indicate significant differences at the 95% confidence level ($p < 0.05$) using Duncan's Multiple Range Test (DMRT).

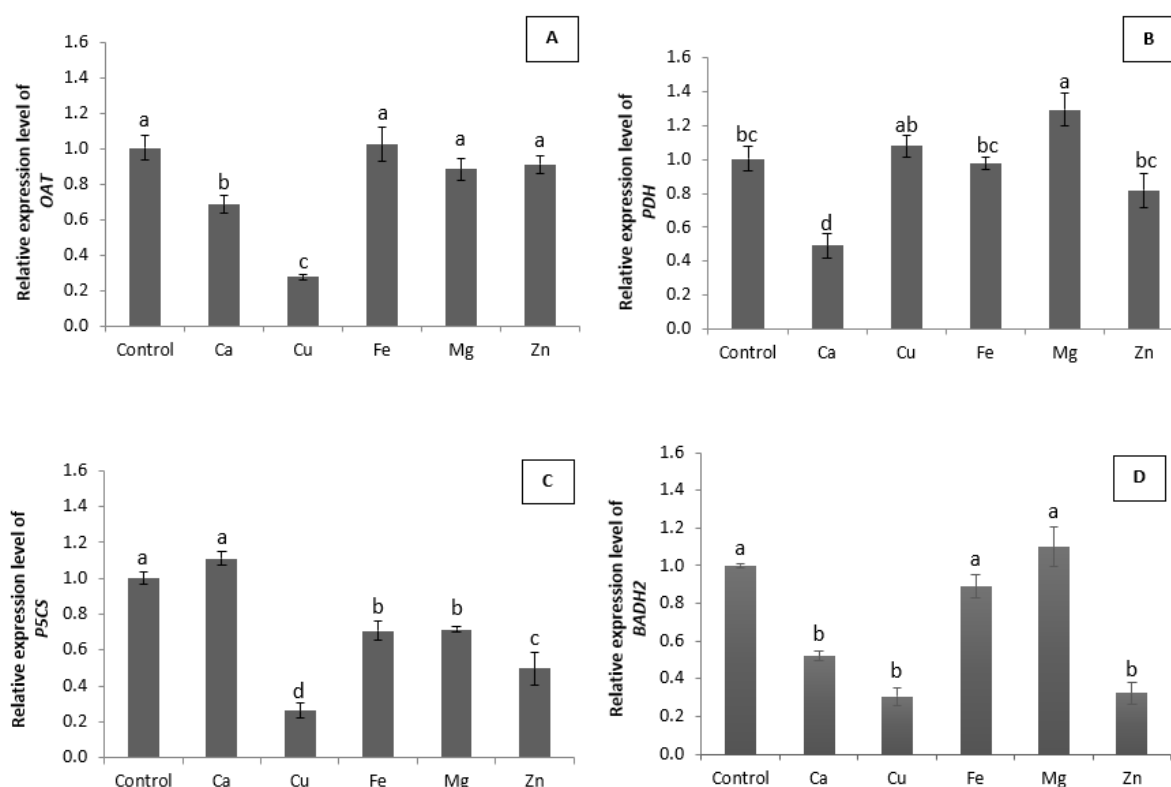


Figure 1 Relative gene expression levels of genes involved in the 2-AP biosynthesis pathway (a) *OAT*, (b) *PDH*, (c) *P5CS*, and (d) *BADH2* in the leaf of the booting stage of rice plants supplemented with no micronutrient supplementation (control), Ca, Cu, Fe, Mg, and Zn. The histogram shows the relative abundance of mRNA for each gene and each treatment after normalization with *OsActin*. Different letters in the bar were significantly different at the level of 95% ($p < 0.05$) using Duncan's New Multiple Range Test (DMRT).

Since mineral supplementation led to higher 2-AP accumulation than abiotic stress treatments, we examined the expression of genes associated with the 2-AP biosynthesis pathway. There are several

amino acid precursors (proline, ornithine, and glutamate) and enzymes (proline dehydrogenase (PDH), Δ 1-pyrroline-5-carboxylic acid synthase (P5CS), and ornithine aminotransferase (OAT)) involved in this

pathway. Among these, inactivation of betaine aldehyde dehydrogenase (*BADH2*) gene expression is considered to play an important role in contributing to 2-AP production [29, 30].

In this study the transcription levels of ornithine aminotransferase (*OAT*) were significantly decreased when plants were supplied with exogenous Cu and Ca by 3.6- and 1.5-fold, respectively, compared to the control (Figure 1a). For proline dehydrogenase (*PDH*), the highest expression was found in plants applied with exogenous Mg (1.3-fold higher than the control) whereas its expression was suppressed approximately 2.0-fold when plants were treated with exogenous Ca (Figure 1b). The expression of Δ^1 -pyrroline-5-carboxylate synthetase (*P5CS*) was noticeably down regulated in response to Mg, Zn, Cu, and Fe treatments compared to the control group. The Cu treatment resulted in the most significant decrease in *P5CS* transcript levels, which were 3.8-fold lower than the control (Figure 1c). Betaine aldehyde dehydrogenase 2 (*BADH2*) expression was markedly reduced in plants supplied with Ca, Cu, and Zn compared to the control, with *BADH2* transcript levels 3.1-fold lower in Cu and Zn treatments. No significant differences in *BADH2* expressions were observed.

Although supplementation with Ca, Cu, and Zn resulted in varying expression levels of *OAT*, *PDH*, and *P5CS*, *BADH2* expression was significantly downregulated in these treatment groups compared to the control (Figure 1). Previous studies have established that *BADH2* expression is negatively associated with 2-AP accumulation in scented rice cultivars [29, 31]. Therefore, the observed reduction in *BADH2* gene expression may contribute to the elevated 2-AP content in these treatment groups (Table 1).

Furthermore, [32] reported that the highest 2-AP content occurs in mature rice grains, which typically exhibit low proline levels and reduced *P5CS* expression. This suggests that proline produced in leaves might be directly translocated to rice grains and subsequently converted into 2-AP. Two mechanisms have been proposed for 2-AP accumulation in mature grains: (1) 2-AP synthesized in leaves and stem sheaths is transported to grains, and (2) proline is translocated from leaves to grains, where 2-AP synthesis subsequently occurs [33, 34]. The lack of significant upregulation of the three genes involved in the initial steps of 2-AP synthesis in leaf tissue supports these translocation hypotheses.

Table 2 Total phenolic content, total flavonoid content and FRAP value in grains of rice plants under various minerals treatments.

Treatment	TP (mg GAE/100 g DW)	TF (mg Cat E/100 g DW)	FRAP (mmol Fe ²⁺ /100 g DW)
Control	22.10 $\pm$ 2.34 ^{bc}	18.27 $\pm$ 6.64 ^{ab}	9.27 $\pm$ 0.9 ^b
Ca	27.60 $\pm$ 2.28 ^a	20.80 $\pm$ 0.00 ^a	14.93 $\pm$ 2.2 ^a
Cu	25.35 $\pm$ 1.61 ^{ab}	17.82 $\pm$ 2.84 ^{ab}	9.23 $\pm$ 0.9 ^b
Mg	25.00 $\pm$ 0.62 ^{abc}	14.50 $\pm$ 3.93 ^{ab}	8.20 $\pm$ 1.2 ^b
Fe	24.85 $\pm$ 0.36 ^{abc}	12.60 $\pm$ 2.90 ^b	8.10 $\pm$ 1.2 ^b
Zn	22.58 $\pm$ 1.62 ^{bc}	19.53 $\pm$ 3.93 ^{ab}	8.16 $\pm$ 1.1 ^b

TP: Total Phenolic content; TF: Total Flavonoid content; FRAP: Ferric Reducing Antioxidant Power. Difference letters in the same column were significantly different at the level of 95% ($p < 0.05$) using Duncan's New Multiple Range Test (DMRT).

The effects of different mineral treatments on bioactive compounds and antioxidant properties in rice grains are presented in Table 2. The antioxidant capacity and bioactive compounds in rice grains are primarily attributed to polyphenols [35, 36], which have been reported to confer health benefits against various diseases [37]. Calcium supplementation enhanced total phenolic content, total flavonoid content, and FRAP activity in rice grains (Table 2). Rice grains from plants treated with calcium exhibited the highest concentration of phenolic compounds (27.60 mg GAE/100 g DW), although no significant differences in phenolic content were observed among treatments. Similarly, calcium supplementation resulted in the highest flavonoid content in rice grains, though no statistically significant difference was detected between the Ca treatment and control. For FRAP, its

activity was also highest in grains from plants treated with calcium (14.93 mmol Fe²⁺/100 g DW), significantly exceeding values observed in other mineral treatments and the control group. These findings were consistent with previous studies demonstrating that calcium pretreatment positively influences polyphenol content and antioxidant capacity [38, 39]. These results indicate that calcium supplementation enhanced total phenolic content, total flavonoid content, and FRAP activity in rice grains (Table 2), consistent with previous studies demonstrating that calcium pretreatment positively influences polyphenol content and antioxidant capacity [38, 39]. Therefore, calcium supplementation during rice cultivation not only increases 2-AP content but also enhances the accumulation of bioactive compounds and improves antioxidant capacity.

The association between drying methods and 2-AP content was shown in Table 3, solar-dried rice grains contained significantly higher 2-AP content compared to both oven-dried grains at 60°C and the non-dried control group. These findings indicate that the drying method significantly influences the retention and possibly the formation of 2-AP in rice grains during post-harvest processing.

It is reported that temperature significantly influences 2-AP level in rice grain by impacting enzyme activity, carbon availability, and the metabolic capacity of the grain [40]. This suggests that temperature is a critical factor affecting 2-AP content, with moderate temperatures being optimal. [41] reported that low temperatures (22/16°C) induced high accumulation of 2-AP in rice grains, which explains why hot air drying at 60°C might negatively affect 2-AP synthesis or retention.

The average temperature during solar drying in our study was 35-38°C, indicating that this moderate temperature range may be optimal for enhancing or preserving 2-AP content. However, it is important to note that the response of 2-AP content to temperature variations depends on rice varieties, geographical factors, and exposure duration.

Table 3 2-AP content in rice grains with different grains drying methods.

Drying methods (day)	2-AP (ppm)
No drying	3.22 ± 0.28 ^b
Solar drying for 3 days	5.40 ± 0.76 ^a
Heated air drying at 60 °C for 3 days	3.51 ± 0.79 ^b

Difference letters in the same column were significantly different at the level of 95% ($p < 0.05$) using Duncan's New Multiple Range Test (DMRT).

CONCLUSIONS

The aromatic quality of rice grains, 2-AP level, after harvesting depends on many factors, for example abiotic stress induction, exogenous mineral application and post-harvest processes. The results of this study revealed that treating plants with abiotic stress did not provide any beneficial effects on 2-AP content. Conversely, supplementing plants with minerals especially, Ca had positive effects on 2-AP production since it could suppress the expression of BADH2, the key gene in the 2-AP synthesis pathway. Furthermore, exogenous Ca induced high accumulation of antioxidant capacity in rice grains. The process of drying grains after harvesting is also associated with 2-AP content. Among the methods used in this study, the solar drying method was the most effective for preserving aroma quality in rice grains.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

The authors used Grammarly Premium for English grammar and used Claude.ai to create the graphical abstract. All scientific content, data analysis, and interpretation were performed solely by the authors without AI assistance.

ACKNOWLEDGEMENT

This research was funded National Research Council of Thailand, grant number SRN2562REV007.

REFERENCES

1. Sarkarung S, Somrith B, Chitrakorn S. Chapter Ten-Aromatic rice of Other Countries. In: Singh RK, Singh US, Khush GS, editors. Aromatic Rices. Oxford & IBH Publishing Co. Pvt. Ltd.; 2000. p. 179-3.
2. Vanavichit A, Kamolsukyeunyong W, Siangliw M, Siangliw JL, Traprab S, Ruengphayak S. Thai Hom Mali rice: Origin and breeding for subsistence rainfed lowland rice system. Rice. 2018;11(20): 1-12.
3. Arunmas P. Thai Hom Mali nets rice prize [Internet]. Bangkok Post. 2020 [cited 2021 Sep 17]. Available from: <https://www.bangkokpost.com/thailand/general/2030339/thai-hom-mali-nets-rice-prize>.
4. Arunmas P. Farmers urge action after top rice fails to win prize [Internet]. Bangkok Post. 2019 [cited 2021 Sep 17]. Available from: <https://www.bangkokpost.com/business/1796249/farmers-urge-action-after-top-rice-fails-to-win-prize>.
5. Chinachanta K, Herrmann L, Lesueur D, Jongkaewwattana S, Santasup C, Shutsrirung A. Influences of farming practices on soil properties and the 2-acetyl-1-pyrroline content of Khao Dawk Mali 105 rice grains. Appl Environ Soil Sci. 2020;2020:8818922.
6. Verma DK, Srivastav PP. Science and technology of aroma, flavour and fragrance in rice. Boca Raton, FL: Apple Academic Press; 2018.
7. Yoshihashi T, Huong NT, Inatomi H. Precursors of 2-acetyl-1-pyrroline, a potent flavor compound of an aromatic rice variety. J Agric Food Chem. 2002;50(7):2001-4.
8. Yoshihashi T, Nguyen TT, Kabaki N. Area dependency of 2-acetyl-1-pyrroline content in an aromatic rice variety, Khao Dawk Mali 105. Food Technol. 2004;38(2):105-9.
9. Chen M, Wei X, Shao G, Tang S, Luo J, Hu P. Fragrance of the rice grain achieved via artificial

- microRNA-induced down-regulation of *OsBADH2*. Plant Breed. 2012;131:584-90.
10. Wakte K, Zanan R, Hinge V, Khandagale K, Nadaf A, Henry R. Thirty-three years of 2-acetyl-1-pyrroline, a principal basmati 2-AP in scented rice (*Oryza sativa* L.): A status review. J Sci Food Agric. 2017; 97(2):384-95.
 11. Poonlaphdecha J, Gantet P, Maraval I, Notton C, Lebrun M, Cazevieille C, et al. Fertilizer deep placement significantly affected yield, rice quality, 2-AP biosynthesis and physiological characteristics of the fragrant rice cultivars. Agronomy. 2022; 12(1):162.
 12. Nawara W, Bennett C, Norkaew O, Sookwong P, Moolkam S, Naruebal S, et al. Variation of the impact 2-AP, 2-acetyl-1-pyrroline, content in Thai fragrant rice plants and its enhanced accumulation by soil nutritional elements. J Agric Sci. 2020;12 (6):88-99.
 13. Deng S, Xing P, Peng L, Lu J, Wu Y, Zhang Y, et al. Exogenous application of abscisic acid enhanced 2-acetyl-1 pyrroline biosynthesis, accumulation, and antioxidant activities in aromatic rice. BMC Plant Biology. 2025;25(1):302.
 14. Renuka N, Barvkar VT, Ansari Z, Zhao C, Wang C, Zhang Y, et al. Co-functioning of 2AP precursor amino acids enhances 2-acetyl-1-pyrroline under salt stress in aromatic rice (*Oryza sativa* L.) cultivars. Sci Rep. 2022;12(1):3911.
 15. Yoshida S, Forno DA, Cock JH, Gomez KA. Laboratory manual for physiological studies of rice. 3rd ed. Manila: International Rice Research Institute; 1976.
 16. PerkinElmer. Analysis of 2-acetyl-1-pyrroline (2AP) in rice using headspace-solid phase microextraction (HS-SPME) and gas chromatography-mass spectrometry (GC-MS) [Internet]. Shelton (CT): PerkinElmer; [cited 2024 May 22]. Available from: www.perkinelmer.com.
 17. Nounjan N, Mahakham W, Siangliw JL, Toojinda T, Theerakulpisut P. Chlorophyll retention and high photosynthetic performance contribute to salinity tolerance in rice carrying drought tolerance quantitative trait loci (QTLs). Agriculture. 2020; 10:620.
 18. Atanassova M, Georgieva S, Ivancheva K. Total phenolic and total flavonoid contents, antioxidant capacity and biological contaminants in medicinal herbs. J Univ Chem Technol Metall. 2011;46(1): 81-8.
 19. Benzie IF, Strain JJ. The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": The FRAP assay. Anal Biochem. 1996;239(1):70-6.
 20. Bates LS, Waldren RP, Teare ID. Rapid determination of free proline for water-stress studies. Plant Soil. 1973;39(1):205-7.
 21. Dangthaisong P, Jaidee R, Kongruang S, Couvillion SP, Pongprayoon P, Chitikineni A. Molecular mechanisms of 2-acetyl-1-pyrroline (2-AP) accumulation in fragrant rice grains under combined drought and salinity stresses. Plants. 2023;12(8): 1618.
 22. Renuka N, Mathure SV, Zanan RL, Thengane RJ, Nadaf AB. Accumulation of key metabolites under salt stress to determine 2-acetyl-1-pyrroline content and physiological responses in aromatic rice (*Oryza sativa* L.). Metabolomics. 2022;18 (1):1-19.
 23. Ghosh UK, Islam MN, Siddiqui MN, Cao X, Khan MAR. Proline, a multifaceted signalling molecule in plant responses to abiotic stress: understanding the physiological mechanisms. Plant Biol (Stuttg). 2022;24(2):227-39.
 24. Imran M, Shafiq S, Ashraf U, Qi J, Mo Z, Tang X. Biosynthesis of 2-Acetyl-1-pyrroline in Fragrant Rice: Recent Insights into Agro-management, Environmental Factors, and Functional Genomics. J Agric Food Chem. 2023;71(10):4201-15.
 25. Knez M, Stangoulis JCR. Calcium biofortification of crops—challenges and projected benefits. Front Plant Sci. 2021;12:669053.
 26. Wang Q, Long W, Wu X, Wu C, Gong Y, Wang Z, et al. Study on Differences in 2-AP Synthesis and Metabolism Among Fragrant Rice Varieties. Int J Mol Sci. 2025;26(20):10102.
 27. Monggoot S, Popluechai S, Gentekaki E, Pripdeevech P. Relationships between soil microorganisms and nutrient elements in rice seeds by chemometric tools. Int J Agric Biol Eng. 2014;16:550-4.
 28. Pongpeera S, Maensiri D. Effects of hydroxyapatite nanoparticles on growth, physiological responses, aroma concentration, and yield of Thai jasmine rice. J Plant Nutr. 2023;46(4):693-709.
 29. Chen S, Yang Y, Shi W, Ji Q, He F, Zhang Z, et al. Badh2, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major component in rice fragrance. Plant Cell Online. 2008;20:1850-61.
 30. Tong W, Yoshihashi T, Gao F, Li Z, Chen G, Yang W, et al. Molecular mechanisms for the synthesis of 2-acetyl-1-pyrroline, the key 2-AP in fragrant rice. Mol Breed. 2023;43(7):1-14.
 31. Chen M, Wei X, Shao G, Tang S, Luo J, Hu P. Fragrance of the rice grain achieved via artificial

- microRNA-induced down-regulation of *OsBADH2*. Plant Breed. 2012;131:584-90.
32. Wei X, Handoko DD, Pather L, Methven L, Elmore JS. Evaluation of 2-acetyl-1-pyrroline in foods, with an emphasis on rice flavour. Food Chem. 2017;232:531-44.
 33. Hinge VR, Patil HB, Nadaf AB. Aroma volatile analyses and 2-AP characterization at various developmental stages in Basmati and non-Basmati scented rice (*Oryza sativa* L.) cultivars. Rice. 2016;9:38.
 34. Mo Z, Huang J, Xiao D, Ashraf U, Duan M, Pan S, et al. Supplementation of 2-Ap, Zn and La improves 2-acetyl-1-pyrroline concentrations in detached aromatic rice panicles in vitro. PLoS One. 2016;11(2):e0149523.
 35. Ma Z, Wang Y, Jia J, Li Y, Zhang J, Fan H, et al. Polyphenols in whole grain rice: Identification, characterization, bioactivities, and health benefits. Crit Rev Food Sci Nutr. 2022;62(15): 4164-84.
 36. Ragaei S, Seetharaman K, Abdel-Aal ESM. The impact of milling and thermal processing on phenolic compounds in cereal grains. Crit Rev Food Sci Nutr. 2014;54(7):837-49.
 37. Moisa C, Copolovici DM, Bungau S, Lupitu A, Copolovici L, Cioca G. Rice (*Oryza sativa* L.) bran as a source of bioactive compounds and their health-promoting benefits. Front Nutr. 2024; 11:1-15.
 38. Okpala LC, Okoli EC, Olojede AO. Temperature cycling enhanced 2-acetyl-1-pyrroline accumulation in low-temperature stored fragrant rice. Food Chem. 2020;315:126285.
 39. Yang X, Jiao W, Hao R, Guo H, Cao Y, Xu W. Effects of calcium ion on the synthesis of phenolic compounds and anthocyanins in grape cell cultures. Food Chem. 2024;430:137131.
 40. Luo H, Zhang Q, Lai R, Zhang S, Yi W, Tang X. Regulation of 2-Acetyl-1-pyrroline Content in Fragrant Rice under Different Temperatures at the Grain-Filling Stage. J Agric Food Chem. 2024;72(18):10521-30.
 41. Yu GL, Tang L, Yang Y, Geng ZX, Dong JX, Cao DD, et al. Enhanced resistance to heat stress in germinating rice seeds by calcium priming: Role of enhanced antioxidant system and accumulated heat shock proteins. Plant Physiol Biochem. 2024;204:108239.



Fake news detection through multi-level contextual and sequential modeling

Pimpa Cheewaprakobkit* and Rindra Razafinjatovo

Faculty of Information Technology, Asia-Pacific International University, Saraburi 18180, THAILAND

*Corresponding author: pimpa@apiu.edu

ABSTRACT

The spread of false information on digital platforms destabilizes democratic societies and undermines public trust. Conventional machine-learning models often fail to capture the subtle contextual cues and long-range linguistic dependencies required for accurate verification. To address these limitations, we propose a hybrid deep-learning architecture that integrates three modules: A Contextual Encoder Module (CEM) using RoBERTa to generate relational embeddings for deep semantic extraction; a Bidirectional Sequence Learner (BSL) to model long-range temporal dependencies; and a Context Refinement Layer (CRL) that employs attention mechanisms to highlight the most salient deceptive markers. We evaluated the model on the WELFake dataset, which contains 72,134 news articles. Our proposed method achieved 99% accuracy, precision, recall, and F1-score, significantly outperforming state-of-the-art baselines, including BERT (95%) and BiLSTM (97%). The CRL's attention outputs provide transparency into the model's decision-making process, an essential feature for applications in misinformation detection and automated content moderation. Beyond its technical contributions, this research supports society by enabling early and reliable detection of false information, reducing harm from misinformation, and promoting a safer, more trustworthy digital environment.

Keywords: Fake news detection, WELFake dataset, Hybrid deep learning, Text classification, RoBERTa

INTRODUCTION

The proliferation of online communication platforms has reshaped, often quite rapidly, how people encounter, circulate, and interact with news. While digital media has broadened participation and lowered barriers to information, it has simultaneously created a permissive environment for the swift diffusion of fake news and fabricated narratives intended to mislead audiences and steer public opinion. At this point, the social and political fallout is, by most accounts, far-reaching: such content corrodes trust in journalistic institutions, exacerbates societal polarization, distorts electoral outcomes, and, in the longer run, imperils democratic resilience. Consequently, responding to this challenge entails more than media literacy initiatives and policy reform; it also, notably, calls for computational infrastructures capable of reliably discriminating authentic reporting from deceptive material with a degree of precision commensurate with the speed and scale of contemporary platforms. [1].

Early lines of work in fake news detection were anchored in conventional machine-learning classifiers, Support Vector Machines, Logistic Regression, and Decision Trees, rather than more expressive neural architectures. These systems drew on hand-engineered textual descriptors, most notably TF and IDF [2], along with n-gram statistics and a range of readability indices.

Although such feature-driven pipelines yielded moderate success on many benchmarks, their accuracy seldom transferred well across domains, which can be inferred to stem from a shallow grasp of context and a limited capacity to encode linguistic nuance [3]. Put differently, and as others have documented, models keyed to surface cues struggle with pragmatics and semantics, thereby undermining robustness under domain shift.

With the advent of deep learning, performance advanced markedly. To better capture sequential dependencies and local regularities, researchers introduced architectures based on Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) [4], and Convolutional Neural Networks (CNNs) [5]. Yet, even as these models delivered stronger benchmarks, they arguably remained limited in two respects: they struggled to encode rich, long-range semantics, and they offered insufficient interpretability for use in real-world applications.

Transformer-based architectures, most notably BERT and RoBERTa, have arguably reshaped natural language processing (NLP) by formalizing self-attention, thereby enabling the modeling of long-range dependencies and yielding more expressive semantic representations [6, 7]. In turn, and not least for fake news detection, these advances have produced markedly improved performance [8].

Nonetheless, the same transformers carry well-known liabilities: they cope poorly with extended input sequences and, by tending to distribute attention relatively evenly across tokens, can blur the faint deceptive cues on which detection often hinges [9].

To address these gaps, researchers have turned to hybrid deep learning designs that fuse heterogeneous architectures to capture both local cues and broader discourse-level structure. Notably, CNN-BiLSTM combinations underscore how convolutional pattern detectors can be profitably paired with sequence-aware learners [10, 11]. Yet a recurring shortfall remains: many such hybrids forgo fine-grained attention that would highlight the most contextually salient tokens and spans. At this point, it is worth noting that transformer-based hybrids, while generally more robust, are computationally intensive and, arguably, less transparent in their decision-making.

To address these limitations, we propose a hybrid framework comprising three mutually interdependent components:

- 1) Contextual Encoder Module (CEM), producing deep, semantically nuanced contextual embeddings by leveraging RoBERTa;
- 2) Bidirectional Sequence Learner (BSL), learning temporal dependencies in both forward and backward directions; and
- 3) Context Refinement Layer (CRL) employs fine-grained attention to foreground semantically informative content.

Together, these modules enhance contextual understanding, strengthen sequence modeling, and improve interpretability, thereby increasing accuracy and robustness in fake news detection.

Overall, this study introduces an integrated architecture that effectively unites contextual encoding, bidirectional learning, and attention-based refinement. An empirical evaluation on the WELFake dataset demonstrates that the proposed model achieves near-perfect accuracy, precision, recall, and F1 scores, outperforming state-of-the-art methods.

MATERIALS AND METHODS

This study introduces a hybrid deep learning framework for fake news detection, illustrated in Figure 1. The model integrates contextual encoding, sequential learning, and attention-based refinement through three key components: The Contextual Encoder Module (CEM) for deep semantic extraction, the Bidirectional Sequence Learner (BSL) for capturing long-range temporal dependencies in news narratives, and the Context Refinement Layer (CRL), which employs attention mechanisms to highlight the most salient deceptive markers. Together, these modules enable a comprehensive understanding of linguistic, semantic, and sequential patterns within textual data.

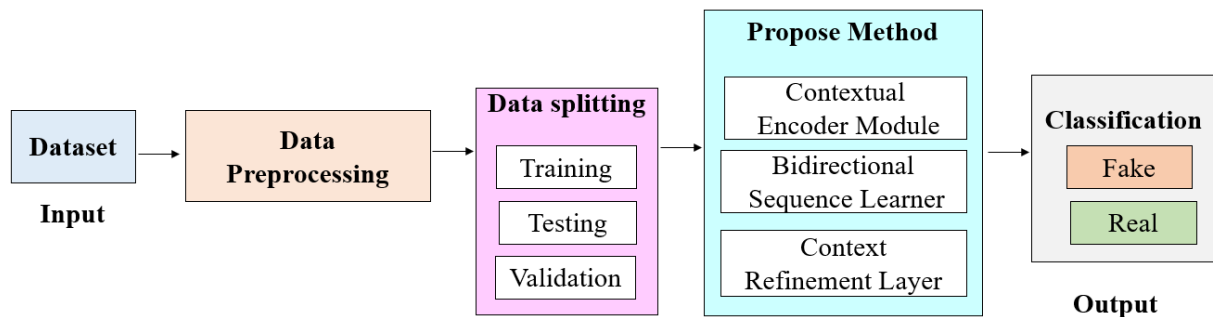


Figure 1 Our proposed architecture.

Data preprocessing

The WELFake dataset consists of raw textual content aggregated from diverse online sources, including news portals and social media platforms. Consequently, the raw data contains typical web-based artifacts and formatting irregularities that must be addressed prior to model training. To ensure high data quality while maintaining full compatibility with the RoBERTa encoder, we adopted a structured preprocessing strategy focused on noise reduction without sacrificing semantic integrity.

First, non-semantic artifacts introduced during web scraping, such as HTML tags, embedded hyperlinks, email addresses, and extraneous metadata, were removed, as they do not contribute meaningful linguistic information. Redundant whitespace was normalized

to ensure textual consistency. This preprocessing stage was explicitly designed to align with the RoBERTa tokenizer. Once cleaned, the text is tokenized and passed to the Contextual Encoder Module (CEM) for embedding generation. This design choice ensures that the architecture receives high-fidelity input, supporting robust semantic representation and contributing to the model's strong generalization performance [12].

Data splitting

To promote an equitable and unbiased assessment, the dataset was divided into training (70%), validation (15%), and test (15%) partitions via stratified sampling. Notably, this procedure, by design, preserves the class priors (fake versus real) within each subset. On the training split, parameter estimation

is performed; on the validation split, hyperparameters are tuned to temper overfitting; and on the holdout test set, out-of-sample generalization is assessed. Following the split, all documents are routed through the Contextual Encoder Module (CEM), which maps tokenized text to contextualized embeddings for downstream modules.

Contextual Encoder Module (CEM)

Within the CEM, as illustrated in Figure 2, we employ the RoBERTa-base model [13, 14], a bidirectional transformer, to generate context-conditioned semantic embeddings. In practical terms, both titles and the full body text are fed into the encoder, which maps them to dense vectors whose values are modulated by neighboring tokens and broader discourse cues. By contrast with static embeddings, these contextualized representations update with the local context, encoding syntactic dependencies and semantic relations in ways that arguably enable more fine-grained discrimination between authentic and fabricated content.

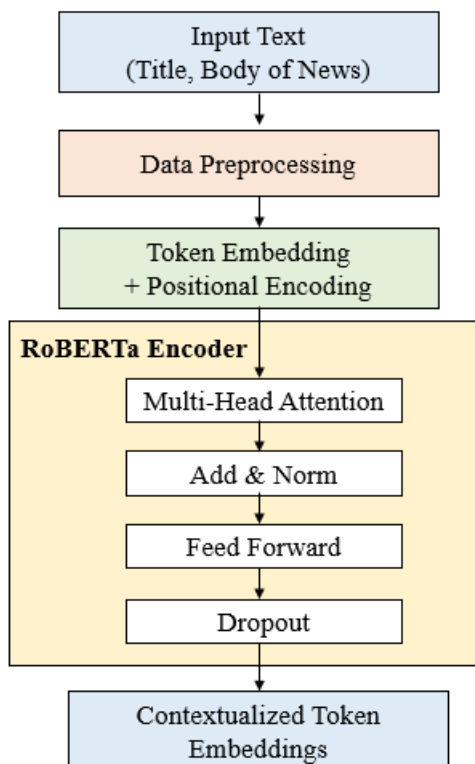


Figure 2 Contextual Encoder Module (CEM) flowchart.

At this stage, after cleaning and tokenization, the sequence is mapped to token embeddings and augmented with positional encodings, which together effectively encode lexical semantics while preserving token order. Notably, the core module is a RoBERTa encoder composed of twelve Transformer blocks that interleave multi-head self-attention with residual add-and-normalize steps, position-wise feed-forward sublayers, and dropout regularization. Accordingly, for downstream processing, these components produce high-dimensional token-level embeddings that, in

effect, preserve both syntactic structure and semantic content.

Among the hallmark contributions of the transformer is, arguably, the multi-head self-attention mechanism [8]. Rather than computing a single attention map, queries, keys, and values are first linearly projected into several learned subspaces, and within each head, attention is computed independently. This setup allows the model to attend, in parallel, to the task of distinguishing linguistic dependencies. The outputs from all heads are subsequently concatenated and passed through a final linear layer, thereby producing a composite attention representation.

Formally, the multi-head attention is defined as:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^o \quad (1)$$

where each attention head is computed as:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (2)$$

Here, W_i^Q , W_i^K , and W_i^V are the learned projection matrices for the queries, keys, and values, respectively, with dimensions $R^{d_{\text{model}} \times d_k}$, $R^{d_{\text{model}} \times d_k}$, $R^{d_{\text{model}} \times d_v}$. The final output projection matrix W^o has dimensions $R^{hd_v \times d_{\text{model}}}$.

Within this framework, self-attention serves as the core of the architecture. By design, it allows the model to assign relative salience to tokens anywhere in the sequence, regardless of distance. The mechanism proceeds by projecting the input into query (Q), key (K), and value (V) matrices through learned linear maps. Notably, each projection fulfills a distinct function: the query encodes what a given token seeks from its context, the key summarizes features of all tokens for matching, and the value carries the content to be pooled. Relevance across tokens is computed as the dot product between Q and K^T , yielding pairwise compatibility scores. To temper their magnitude and consequently stabilize gradient dynamics, the scores are rescaled by the square root of the key dimension (d_k). At this stage, the scaled similarity scores are fed into a softmax, which normalizes them into attention weights that quantify each token's influence on the current representation. Applied thereafter to the value matrix, these weights yield embeddings conditioned on context across the sequence. In this formulation, arguably the mechanism's fulcrum, the model can foreground the most pertinent portions of the input irrespective of position, thereby capturing long-range dependencies and refining semantic interpretation. Formally, the self-attention operation can be written as:

$$\text{Attention}(K, Q, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V_i \quad (3)$$

Through this mechanism, the model assigns differential salience to tokens across the entire sequence, regardless of how far apart they are, enabling it to model long-

range dependencies alongside fine-grained semantic nuances. In turn, the Contextual Encoding Module (CEM) emits a stream of context-enriched embeddings that, at this point, serve as input to the Bidirectional Sequence Learner (BSL).

Bidirectional Sequence Learner (BSL)

Once the CEM has produced contextualized token representations, those embeddings are, at this point, fed into a Bidirectional Sequence Learner (BSL) tasked with modeling the text's temporal structure and other sequential regularities. Scanning the sequence in both directions (left-to-right and right-to-left), the BSL exploits prior and forthcoming context alike, thereby capturing dependencies that hinge not only on preceding tokens but, notably, on tokens still to come, including long-range interactions. Consequently, this two-way treatment, arguably indispensable in practice, affords a more complete account of discourse-level dynamics, helping preserve narrative flow, maintain structural coherence, and uphold semantic intent. Notably, and this is especially pertinent to our use case, the bidirectional pass tends to surface faint inconsistencies and recurrent structural motifs that can be read as hallmarks of sophisticated disinformation.

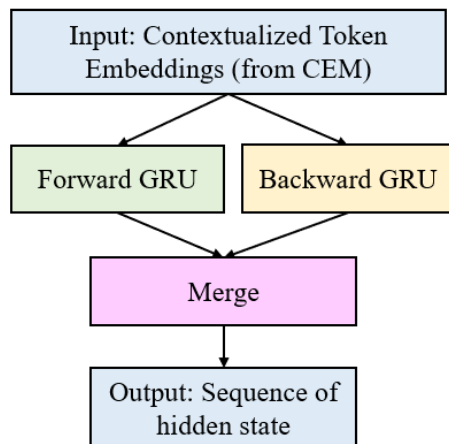


Figure 3 Bidirectional Sequence Learner (BSL) flowchart.

As shown in Figure 3, the BSL ingests the contextualized token embeddings produced by the CEM as its initial input. Instead of a single sweep, the module executes two passes in parallel: a forward pass that traverses the sequence from start to end, and a backward pass that traverses it in reverse. The activations from these two, stemming from distinct yet complementary subnetworks, are then concatenated to integrate context gathered in both temporal directions. Consequently, the BSL emits a sequence of hidden states, each state encoding a fused representation that blends forward and backward information at its corresponding position. Taken together, this output furnishes an arguably comprehensive, context-aware view of the entire sequence and is then handed off to the Context Refinement Layer (CRL).

Context Refinement Layer (CRL)

The Context Refinement Layer (CRL), positioned at the final stage of our architecture, is an attention-based module that focuses on the most informative cues for the downstream classifier. Into this layer, we pass the sequence of hidden states emitted by the Bidirectional Sequence Learner (BSL) without further transformation. Rather than treating all tokens uniformly, the CRL computes, for each hidden state, an attention coefficient that functions as an important score for its corresponding word in the sequence. Notably, words that are especially pertinent to disinformation detection, strong affective markers, concrete yet uncorroborated assertions, or propagandistic turns of phrase tend, consequently, to receive higher weights. Applying these weights yields a single, refined context vector, obtained by summing the hidden states. This vector, which can be inferred to provide a compact and task-tuned synopsis of the entire input, is optimized for the detection objective. In turn, the resulting representation serves as the final feature set and is forwarded to the classification layer to produce the prediction. The operational flow of the CRL is depicted in Figure 4.

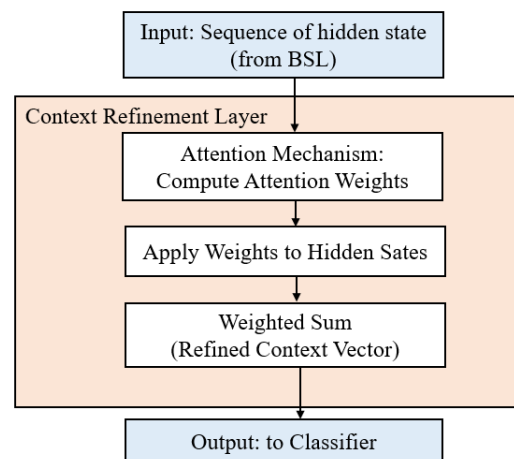


Figure 4 Context Refinement Layer (CRL) flowchart.

Classification

As depicted in Figure 5, the refined context vector is first subjected to dropout, a regularization technique intended to curb overfitting, and then routed to a dense (fully connected) layer that performs dimensionality compression and, arguably, feature consolidation. The resulting scalar is subsequently transformed by a sigmoid activation function [15]. A sigmoid activation function is defined as:

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (4)$$

where z denotes the weighted sum of features and bias, e is Euler's number (the base of the natural logarithm). The sigmoid function maps outputs to

the range [0,1], representing the probability that an input belongs to the “fake news” class (1). A threshold of 0.5 is used for binary classification: values above 0.5 are labeled as fake, and those below are labeled as real.

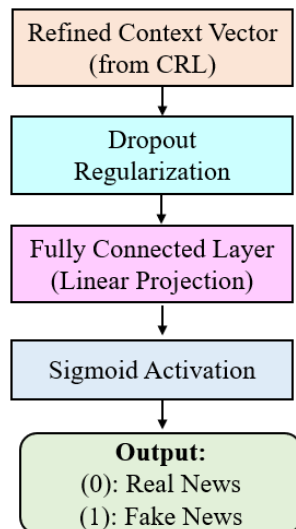


Figure 5 Classification flowchart.

Training and optimization

In practice, we optimized the Binary Cross-Entropy with Logits loss using the AdamW optimizer with a learning rate of 2×10^{-5} and a batch size of 64. Training was conducted on an NVIDIA GeForce RTX 4060 over 10 epochs. To ensure robust generalization, we monitored performance on a held-out validation set throughout training. The final model selected for evaluation is the checkpoint with the highest validation performance.

RESULTS AND DISCUSSION

Dataset

In this study, the WELFake dataset serves as the primary benchmark for evaluating fake news detection performance [16]. Publicly accessible and now widely used in the misinformation-detection literature, WELFake was curated specifically to support machine-learning models. Offering a balanced and comprehensive corpus of news articles, it is, consequently, well-suited to robust model training and evaluation. The dataset comprises 72,134 articles, 37,106 of which are labeled as fake and 35,028 as real. Drawn from Kaggle, the collection provides linguistic diversity alongside credible provenance. Notably, each entry contains three fields: Title, the article's headline; Text, the full body; and Label, a binary indicator in which 1 marks fake and 0 marks real news.

Figure 6 visualizes the distribution of article length for real versus fabricated news. Word count is plotted on the abscissa, and the ordinate records the number of articles in each bin. Fake items, by

and large, cluster at shorter lengths, whereas real news spans a broader, often longer, range. Even so, the two curves overlap appreciably, suggesting that length, taken in isolation, has limited diagnostic utility. Consequently, dependable classification requires deeper contextual and linguistically informed representations.

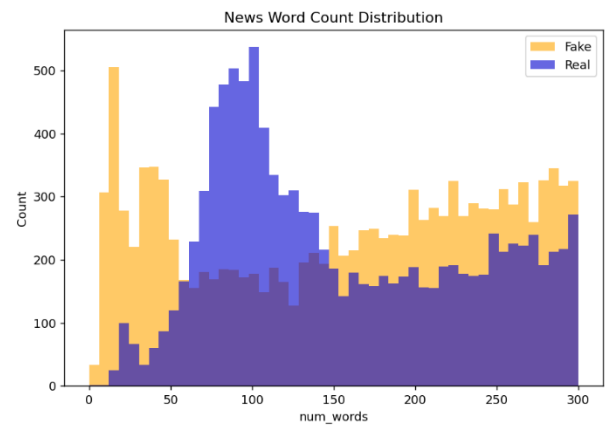


Figure 6 Word Count Distribution for Real and Fake News.

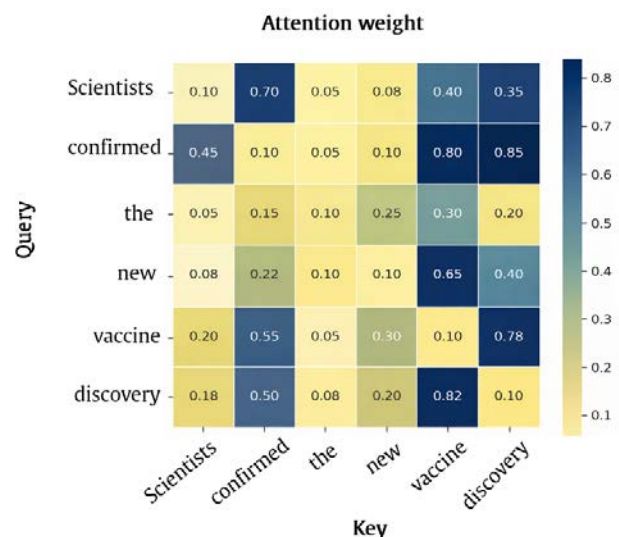


Figure 7 Attention weight heatmap.

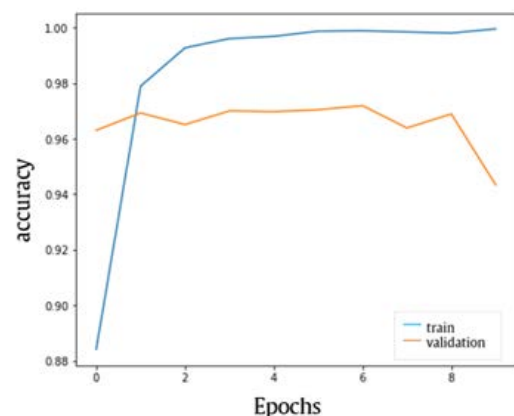


Figure 8 Training and validation accuracy curves.

Figure 7 visualizes attention weights from the Context Refinement Layer (CRL) to illustrate the

model's decision-making process. Darker blue shades indicate higher attention weights, highlighting tokens the model identifies as critical for classification. In contrast, lighter shades represent common function words or less meaningful terms that contribute less to the final representation. This visualization demonstrates how the CRL isolates contextually salient cues, such as affective markers or uncorroborated assertions, to refine the context used to detect fake news.

Figure 8 illustrates the classification performance of the proposed model throughout training, tracking its accuracy on both the training and validation sets. The curve monitors the overall proportion of correctly identified articles, both real and fake, as the model learns. A steady increase in these curves indicates that the hybrid architecture is effectively capturing the necessary contextual and sequential patterns to achieve its high final accuracy. Figure 9 displays the training and validation losses, which represent the error rates the model aims to minimize during optimization. A consistent decline in loss across both datasets demonstrates that the model is generalizing well and is not overfitting.

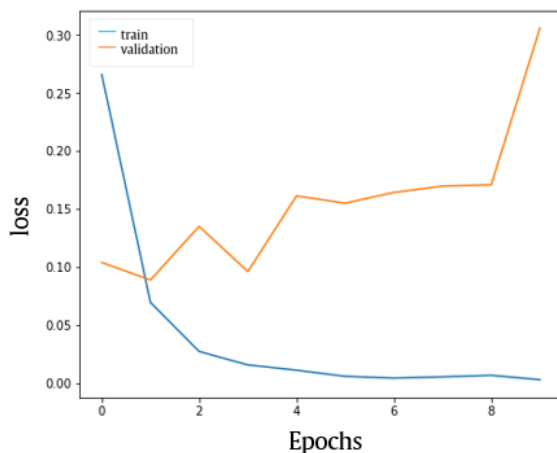


Figure 9 Training and validation loss curves.

Evaluation metrics

To rigorously assess model performance, four standard classification metrics were employed: accuracy, precision, recall, and F1-score [17].

Accuracy measures the overall proportion of correctly classified articles (both real and fake news). The formula is shown in (5).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

As shown in (6), precision, the fraction of truly fake articles within all items the model labels as fake (i.e., the positive predictive value), effectively gauges how accurately the model flags fake news.

$$Precision = \frac{TP}{FP+TP} \quad (6)$$

Recall measures the proportion of actual fake news articles that the model successfully detects, reflecting its ability to capture fake news cases, as shown in (7).

$$Recall = \frac{TP}{FN+TP} \quad (7)$$

The F1-score is the harmonic mean of precision and recall, providing a balanced measure of accuracy and detection capability, as shown in (8).

$$F1 = 2 * \frac{Precision * Recall}{Precision+Recall} \quad (8)$$

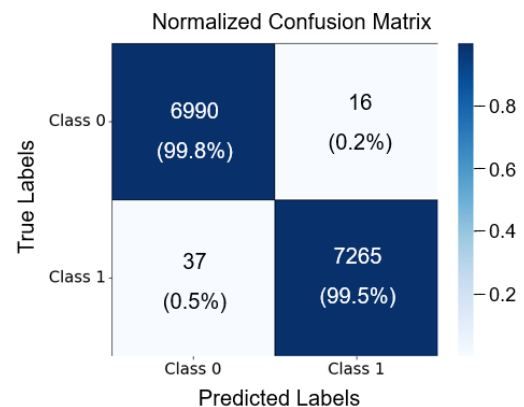


Figure 10 Normalized Confusion Matrix.

Figure 10 depicts the normalized confusion matrix for the proposed model. Considered together, the True Positive (TP) and True Negative (TN) entries are high; in aggregate, this points to robust discriminative performance. With the False Positive (FP) rate at only 0.2%, the system seldom tags legitimate news as fake, an aspect that, arguably, sustains public confidence in automated detection tools. By contrast, the False Negative (FN) rate is 0.5%, slightly higher yet still small, suggesting that only a handful of fake items slip through. Even so, such misses are nontrivial from a risk perspective, as undetected misinformation can propagate across digital networks. At this point, incorporating credibility scores or metadata-informed features could, arguably, further curtail this residual error.

On the WELFake corpus, Table 1 consolidates the side-by-side performance of the proposed framework against a suite of established machine- and deep-learning baselines. As an initial yardstick, the classical learners, CATBoost, Naïve Bayes, and SVM, appear to yield accuracies in the 89-92% range, which can be taken as a pragmatic baseline for automated fake news detection. Notably, logistic regression edges higher to 96%, whereas plain artificial neural networks (ANNs) hover around 95%.

Table 1 Result comparison of models on the WELFake dataset.

Algorithm	Accuracy	F1-Score	Precision	Recall	Reference
RoBERTa	0.86	0.85	0.85	0.86	[18]
CATBoost	0.89	0.89	0.89	0.89	[19]
GRU	0.90	0.90	0.90	0.90	[19]
SVM	0.92	0.92	0.88	0.92	[19]
BERT-Transformer-based Learning	0.95	0.95	0.95	0.95	[20]
Artificial Neural Network	0.95	0.95	0.94	0.95	[21]
Logistic regression	0.96	0.96	0.97	0.95	[21]
Naïve Bayes	0.89	0.89	-	-	[22]
LSTM	0.96	0.96	-	-	[22]
BiLSTM	0.97	0.97	0.98	0.98	[23]
CDLR	0.98	0.98	0.98	0.98	[24]
Proposed method (Our)	0.99	0.99	0.99	0.99	-

Further gains emerge with sequential architecture. GRU, LSTM, and BiLSTM improve accuracy by modeling temporal dependencies in the text; among them, BiLSTM reaches 97%, which arguably underscores the added value of bidirectional context.

Turning to transformer-based systems, their capacity for rich contextual inference is evident, though not uniformly realized. BERT posts 95% accuracy, while RoBERTa, despite its more advanced pretraining, registers a lower 86% on this dataset, likely reflecting limited task-specific fine-tuning or insufficient domain adaptation. These patterns, consequently, reinforce the need to tailor transformer models to the quirks of particular corpora and objectives.

Hybrid designs that blend contextual and sequential cues (e.g., CNN, BiLSTM, or CDLR) achieve up to 98% accuracy, attesting to the advantages of architectural fusion. At this point, the proposed hybrid framework exceeds all baselines, achieving 99% across all evaluation metrics: accuracy, precision, recall, and F1-score. This outperformance seems to arise from the synergistic integration of three components: CEM, which provides deep contextual embeddings through RoBERTa; BSL, operating bidirectionally, models the temporal and sequential dependencies; CRL, notably, recalibrates contextual relevance through attention-weighted scoring. Consequently, this layered design arguably allows the model to recover semantic, syntactic, and temporal dependencies in a unified manner.

Across all metrics, reaching 99% arguably signals not only the model's high accuracy but also its capacity to generalize and to be interpreted,

attributes that, notably, constitute core prerequisites for the operational deployment of misinformation detection systems.

Our results align with recent findings in the literature suggesting that hybrid models outperform standalone transformers. While previous studies using BERT achieved accuracies around 95%, our model's 99% performance on the WELFake dataset is due to the BSL and CRL modules, which capture nuances that standard encoders often overlook. This suggests that, for large-scale datasets, integrating sequential narrative tracking is as critical as capturing the text's deep meaning.

Ablation study

To evaluate the individual contribution of each module within our proposed architecture, we conducted an ablation study using the WELFake and FakeNewsNet datasets. We assessed three configurations: The Contextual Encoder Module (CEM) in isolation, the combination of CEM and the Bidirectional Sequence Learner (CEM+BSL), and our full hybrid model (CEM+BSL+CRL). The inclusion of the Context Refinement Layer (CRL), forming the full model, further elevated performance to 99% across all metrics. The CRL's attention mechanism identifies and weights contextually salient tokens, effectively filtering noise and sharpening the classifier's focus on deceptive markers. This progression demonstrates that while each module is beneficial on its own, their synergistic integration is critical for achieving near-perfect detection rates, as shown in Table 2.

Table 2 Ablation study results on the WELFake dataset.

Algorithm	Accuracy	F1-Score	Precision	Recall
CEM	0.86	0.88	0.82	0.95
CEM+BSL	0.97	0.97	0.97	0.98
CEM+BSL+CRL (Our)	0.99	0.99	0.99	0.99

To demonstrate the generalizability of our framework beyond the primary benchmark, we evaluated the model on the FakeNewsNet dataset. FakeNewsNet presents a distinct challenge, as it comprises 23,921 news items from Twitter, with diverse social context features and shorter text lengths than in standard news articles. As shown in Table 3, the proposed model maintained high performance

on this additional dataset, achieving an accuracy and F1-score of 0.95. Consistent with the ablation results, the full model outperformed the standalone CEM (85% accuracy) and the CEM+BSL configuration (92% accuracy). The architecture's ability to maintain high precision (0.95) and recall (0.95) on a social media dataset confirms its robustness across varying linguistic styles and data distributions.

Table 3 Comparison of model results on the FakeNewsNet dataset.

Algorithm	Accuracy	F1-Score	Precision	Recall
CEM	0.85	0.90	0.89	0.92
CEM+BSL	0.92	0.92	0.92	0.92
CEM+BSL+CRL (Our)	0.95	0.95	0.95	0.95

The 99% accuracy achieved on the WELFake dataset is attributed to the synergistic integration of the CEM, BSL, and CRL modules. While the standalone CEM model achieves 86% accuracy, our multi-level architecture sequentially captures deep semantics and temporal dependencies, thereby eliminating residual errors. Furthermore, the WELFake dataset is a large, balanced benchmark optimized for machine learning. To ensure these results were not due to overfitting, we validated the model on the FakeNewsNet dataset, where it maintained a robust accuracy of 95%, confirming the architecture's strong generalizability across different data distributions.

CONCLUSIONS

This study presents a new hybrid deep learning architecture for fake news detection that weaves together three components: a Contextual Encoder Module (CEM), built on RoBERTa to derive deep contextual embeddings; a Bidirectional Sequence Learner (BSL), which models temporal dependencies in both forward and backward directions; and a Context Refinement Layer (CRL) that applies fine-grained attention to surface semantically salient cues. Taken as a whole, these modules jointly remedy the shortcomings of standard machine learning and earlier neural models, approaches that, in practice, often fail to capture contextual nuance and sequential structure in tandem. On the WELFake dataset, the model attains 99% accuracy and surpasses all benchmark baselines. Notably, this performance appears to validate the multi-level integration strategy: contextual encoding fosters richer semantic understanding; bidirectional sequence modeling captures temporal relations that unidirectional pipelines tend to miss; and attention-based refinement sharpens interpretability while improving precision. The combination, consequently, suggests a deeper link between layered representation learning and robust detection in this domain.

Even so, several caveats deserve mention. First, although WELFake provides a widely used benchmark, it may not reflect the full linguistic heterogeneity or the evolving nature of misinformation observed in dynamic online settings (where concept drift and domain shift are common). Many social media items are multimodal, including images, videos, and memes, but the present model does not analyze them. Second, computational burden remains non-trivial; fine-tuning transformer encoders such as RoBERTa is resource-intensive and may impede real-time or large-scale deployment. Finally, the model currently assumes a binary label space (real vs. fake), whereas misinformation in the wild spans a continuum, from satire and propaganda to partially true content, that future work should more faithfully represent.

Taken together, the present results open several plausible extensions. To capture more nuanced patterns of deception, an obvious next step is a multimodal detector that fuses textual, visual, and metadata cues, potentially via co-attention or contrastive alignment, to leverage correlations among language, imagery, and provenance. Equally promising, though for a different reason, is domain adaptation and transfer learning aimed at robust cross-platform, cross-topic generalization with minimal retraining; under rapid topic drift and distributional shift, techniques such as unsupervised adaptation, parameter-efficient tuning (e.g., adapters or LoRA), or continual learning could be especially useful. At this point, efficiency also matters. Exploring lightweight architectures and model compression, pruning, distillation, quantization, and related low-rank or sparsity-inducing methods would improve throughput and latency for real-time inference, including in resource-constrained settings. Finally, integrating explainable AI tooling, whether through post hoc methods (saliency, concept activation, counterfactuals) or intrinsically interpretable modules, seems likely to deepen insight into model reasoning and to aid transparency, calibration, and user trust.

In summary, this study advances automated fake news detection by proposing an integrated deep learning architecture that, arguably, balances predictive performance with interpretability. While limitations remain, the evidence presented here lays a durable foundation for subsequent inquiry and, consequently, indicates strong potential for practical deployment against the accelerating spread of online misinformation.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

The authors used ChatGPT to improve the clarity and readability of the manuscript during its preparation. All generated text was carefully reviewed, revised, and validated by the authors after using this tool/service. The authors take full responsibility for the final content of the publication.

REFERENCES

1. Paliwal S, Bharti R. Reviewing the landscape: Fake news detection methods and social media user behavior. *Indian J Tech Educ.* 2025;48(2): 207-16.
2. Nuri Y, Senyurek E. Research Abstracts Similarity Implementation By Using TF-IDF Algorithm. *IOSR J Comput Eng.* 2025;27(1):4-10.
3. Shah S, Patel S. A Comprehensive Survey on Fake News Detection Using Machine Learning. *J Comput Sci.* 2025;21(4):982-90.
4. Padalko H, Chomko V, Chumachenko D. A novel approach to fake news classification using LSTM-based deep learning models. *Front Big Data.* 2023;6:1-18.
5. Haqi Al-Tai M, Nema BM, Al-Sherbaz A. Deep Learning for Fake News Detection: Literature Review. *Al-Mustansiriyah J Sci.* 2023;34(2):70-81.
6. Alghamdi J, Lin Y, Luo S. Towards COVID-19 fake news detection using transformer-based models. *Knowl Based Syst.* 2023;274:110642.
7. Aremu T, Akinwehinmi O, Nwagu C, Ahmed SI, Orji R, Amo PAD, et al. On the reliability of Large Language Models to misinformed and demographically informed prompts. *AI Mag.* 2025;46(1):1-15.
8. Vaswani A, Shazeer NM, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is All you Need. *Neural Inf Process Syst.* 2017:6000-10.
9. Boukouvalas Z, Shafer A. Role of Statistics in Detecting Misinformation: A Review of the State of the Art, Open Issues, and Future Research Directions. *Annu Rev Stat Appl.* 2024;11(1):27-50.
10. Alghamdi J, Lin Y, Luo S, editors. Modeling Fake News Detection Using BERT-CNN-BiLSTM Architecture. 2022 IEEE 5th International Conference on Multimedia Information Processing and Retrieval (MIPR). 2022; 354-357
11. Xiaoyan L, Raga RC, Xuemei S, Liguori R. GloVe-CNN-BiLSTM Model for Sentiment Analysis on Text Reviews. *J Sens.* 2022;1-12.
12. Sun W, Cai Z, Li Y, Liu F, Fang S, Wang G. Data Processing and Text Mining Technologies on Electronic Medical Records: A Review. *J Healthc Eng.* 2018;4302425.
13. Shuhai Zhang, Zeng You , Yaofo Chen , Zhiquan Wen, Qian Yue Wang , Zhijie Qiu , et al., editors. Curse of High Dimensionality Issue in Transformer for Long Context Modeling. Forty-second International Conference on Machine Learning; 2025; Vancouver, Canada: PMLR.
14. Tan KL, Lee CP, Anbananthan KSM, Lim KM. RoBERTa-LSTM: A Hybrid Model for Sentiment Analysis With Transformer and Recurrent Neural Network. *IEEE Access.* 2022;10:21517-25.
15. Xu S, Li X, Xie C, Chen H, Chen C, Song Z. A High-Precision Implementation of the Sigmoid Activation Function for Computing-in-Memory Architecture. *Micromachines.* 2021;12(10).
16. Verma PK, Agrawal P, Amorim I, Prodan R. WELFake: Word Embedding Over Linguistic Features for Fake News Detection. *IEEE Trans Computl Soc Syst.* 2021;8(4):881-93.
17. Salman T, Ghubaish A, Unal D, Jain R. Safety Score as an Evaluation Metric for Machine Learning Models of Security Applications. *IEEE Netw Lett.* 2020;2(4):207-11.
18. Rama Moorthy H, Avinash NJ, Krishnaraj Rao NS, Raghunandan KR, Dodmane R, Blum JJ, et al. Dual stream graph augmented transformer model integrating BERT and GNNs for context aware fake news detection. *Sci Rep.* 2025;15(1):25436.
19. Hashmi E, Yayilgan SY, Yamin MM, Ali S, Abomhara M. Advancing Fake News Detection: Hybrid Deep Learning With FastText and Explainable AI. *IEEE Access.* 2024;12:44462-80.
20. Raza N, Abdulkadir SJ, Abid YA, Albouq SS, Alwadain A, Rehman AU, et al. Enhancing fake news detection with transformer-based deep learning: A multidisciplinary approach. *PLoS One.* 2025;20(9):e0330954.

21. Trisna NJ, Sohan MFAA, Joy RM, Joy RD, editors. A Novel Machine Learning Approach for Detecting Fake News in WELFake Dataset. Proceedings of the International Conference on Industrial Engineering and Operations Management. 2024; p. 994-1002.
22. Al-Quayed F, Javed D, Jhanjhi NZ, Humayun M, Alnusairi TS. A Hybrid Transformer-Based Model for Optimizing Fake News Detection. IEEE Access. 2024;12:160822-160834.
23. Padalko H, Chomko V, Chumachenko D. A novel approach to fake news classification using LSTM-based deep learning models. Front Big Data. 2024;6(1):1320800.
24. Abbas F, Taeihagh A. A multi-level fusion-based framework for multimodal fake news classification using semantic feature extraction. Int J Mach Learn Cybern. 2025;16(1):6531–60.



Gamma shielding efficiency of sedimentary rock-based concrete blocks under Co-60 irradiation

Pitchpilai Khoonphunnarai^{1*}, Sutthisa Konruang² and Phayao Yongsiriwit¹

¹Department of Physics, Faculty of Science and Technology, Songkhla Rajabhat University, Songkhla Province 90000, THAILAND

²Department of Physics, Faculty of Science, Thaksin University, Phatthalung Province 93210, THAILAND

*Corresponding author: pitchpilai.kh@skru.ac.th

ABSTRACT

Ensuring human safety requires understanding the radiation shielding capabilities of building materials alongside their strength and durability. Common conventional radiation shielding materials include lead and concrete. However, lead is costly, toxic, and raises environmental concerns, prompting increased interest in alternative shielding materials. Therefore, this research investigates the potential of concrete made from two types of sedimentary rocks, shale and calcareous, for shielding against Co-60 irradiation. The sedimentary rocks were sourced from the Global Geopark region in Satun Province, Thailand. The shielding performance of sedimentary rock-based concrete is compared with standard construction concrete. All concrete blocks were 15 cm × 15 cm × 15 cm and were prepared with a cement:sand:stone ratio of 1:2:4. Concrete cube specimens were prepared to evaluate radiation attenuation properties. The linear attenuation coefficient (μ_l) was calculated from the reduction in count rate after transmission through the concrete block. Engineering properties, including density and compressive strength, were also analyzed. Radiation shielding efficiency is assessed using the linear attenuation coefficient (μ_l), mass attenuation coefficient (μ_m), mean free path (MFP), half-value layer (HVL), and tenth-value layer (TVL). All of these parameters were calculated from the measured count rates. The study found that standard concrete had the highest compressive strength at 25.45 MPa, followed by shale concrete at 13.59 MPa, and calcareous concrete at 5.21 MPa. However, shale concrete performed best in radiation shielding, showing a linear attenuation coefficient (μ_l) of 0.18 cm⁻¹, a mass attenuation coefficient (μ_m) of 0.09 cm²/g, a mean free path of 5.59 cm, a half-value layer of 3.87 cm, and a tenth-value layer of 12.86 cm. The improved shielding performance of shale concrete may be related to its mineral composition and internal structure. Analysis by XRD and XRF indicated that the main heavy metal compounds present were Al₂O₃ and Fe₂O₃.

Keywords: Radiation attenuation, Co-60 radiation, Radiation shielding concrete

INTRODUCTION

Gamma radiation, a form of ionizing radiation, poses significant risks to human health and the environment due to its high energy and strong penetrating ability [1]. Therefore, effective protection against gamma radiation is important in many fields, such as hospitals, nuclear power plants, and construction sites that deal with radiation [2]. In the past, radiation shielding has mostly used materials like lead and concrete, which are known for their ability to attenuate gamma radiation effectively [3]. But these traditional materials have several limitations, such as high costs, heavy weight, and environmental issues associated with their production, handling, and disposal [4].

Recent studies have focused on developing lightweight, economical, and environmentally sustainable alternative materials for radiation shielding. Sedimentary rocks are promising candidate materials

because they are naturally abundant and have varied chemical compositions. [6]. Mineral particles and organic matter build up, compact, and solidify over a long period of time to form sedimentary rocks. Sedimentary rocks are suitable for shielding against radiation because they exhibit a wide range of physical and chemical properties, including density, mineral composition, and porosity [7].

Natural rocks are widely used as construction materials because of their availability, durability, and mineral diversity. Concrete is a common composite material composed of cement and aggregate. Changing the aggregate type can affect the radiation attenuation performance of concrete [9]. People know that concrete composite materials are useful and can be used in many different ways. They work well to block radiation and can be used in nuclear power plants, handling radioactive waste, and nuclear medicine [10].

Mixing cement with sedimentary rock is an interesting way to improve radiation shielding [11]. Researchers have studied cementitious composite materials extensively to determine their mechanical strength, durability, and cost-effectiveness for building [12]. These composite materials can effectively reduce gamma radiation when combined with sedimentary rocks like sandstone, limestone, or shale [13]. Different kinds of composite materials can be shaped into different shapes without losing their strength. This makes them suitable for radiation shielding applications in hospitals, nuclear facilities, radioactive waste management, and industrial radiation areas.

The incorporation of sedimentary rock aggregates can affect the density, porosity, and microstructure of concrete [14], thereby influencing gamma-ray attenuation. Also, these composite materials are relatively inexpensive and can be made locally [15]. This makes them a suitable choice for places where it's difficult to locate or too expensive to buy normal shielding materials. There is a growing need for building materials that are environmentally sustainable and durable, and composite materials are a beneficial choice for that [15]. This is because they can be made and disposed of in an eco-friendly way.

Gamma-ray shielding materials have been widely investigated due to their important role in radiation protection in medical, industrial, and nuclear applications. Conventional shielding materials such as lead and heavy concrete are commonly used because of their high density and effective attenuation capability. However, the use of lead raises environmental and health concerns, while heavy concrete materials often involve higher production costs and limited availability of suitable aggregates.

In recent years, several studies have explored alternative materials for gamma-ray shielding, including natural rocks, industrial by-products, and modified concrete composites. Despite these efforts, few studies have focused on the use of sedimentary rock-based materials as a primary component in developing concrete blocks with enhanced gamma attenuation properties under Co-60 gamma irradiation.

Therefore, the present study aims to investigate the radiation shielding efficiency of concrete blocks produced from sedimentary rock. The rock's elemental composition was analyzed to identify metal oxides that may improve attenuation performance. The novelty of this work lies in the development and evaluation of sedimentary rock-based concrete blocks as a potential alternative shielding material, along with a detailed assessment of their linear attenuation coefficient and radiation protection efficiency under Co-60 irradiation.

This research aims to evaluate sedimentary rock-based concrete blocks as potential alternatives

or supplementary shielding materials for applications requiring moderate gamma-ray attenuation and construction compatibility.

MATERIALS AND METHODS

Sample preparation

Concrete blocks were prepared using two types of sedimentary rocks: shale and calcareous rock, in attenuating Co-60 gamma radiation. The study examines key physical properties, including density, compressive strength, and gamma-ray shielding performance. The study focused on two types of sedimentary rocks: shale and calcareous. These rocks were collected from the UNESCO Global Geopark region in Satun Province, Thailand. Then, the rocks were mixed with Portland cement in a 1:2:4 weight ratio (cement: sand: rock). This ratio is commonly used as a nominal concrete mix for general construction, as shown in Figure 1.



Figure 1 Preparation of sedimentary rock-based concrete blocks size 15 cm x 15 cm x 15 cm.

To ensure the mixture was even, it was mixed thoroughly in a concrete mixer. The mixture was homogenized using a concrete mixer, placed into 15 cm × 15 cm × 15 cm molds, and compacted. The specimens were then cured in water for 28-30 days before density, compressive strength, and gamma-ray shielding tests were performed. A Geiger-Muller detector and counter were used to measure the incident and transmitted count rates. The detector was operated at 960 V. A standard Co-60 source with an activity of 74 kBq was used as the gamma-ray source. The measurements were taken over a period of 300 seconds.

X-ray fluorescence (XRF) analysis was used to determine the concentration of each element in the shale and calcareous samples. This study was done to analyze the elemental composition and compounds of the samples. The Zetium X-ray fluorescence spectrometer (PANalytical, Netherlands) was used to do this analysis. The Empyrean X-ray diffractometer (PANalytical, Netherlands) was also used to analyze

the samples' crystallographic structures using X-ray diffraction (XRD). The elemental and mineralogical compositions are presented in Tables 1 and 2.

Property analysis

The density of the prepared concrete samples was evaluated by dividing the dry weight (M) of each specimen by its calculated volume (V). The mass measurements were obtained using a digital precision balance, while the sample volumes were derived from dimensional measurements of the concrete cubes. The density was then computed in accordance with Equation (1). Compressive strength is the maximum axial load that concrete can withstand before structural failure, such as cracking or crushing. It serves as a critical parameter in evaluating the material's mechanical performance under compressive stress. For the testing procedure, two steel bearing plates with hardened surfaces were mounted on the testing machine to ensure uniform load distribution. The compressive strength was calculated using Equation (2) [10].

$$\rho = M/V \quad (1)$$

$$F_c = P/A \quad (2)$$

where ρ , M , and V are density, mass and volume, respectively. F_c , P and A are compressive strength, load and loaded area, respectively.

Radiation shielding

The transmitted count rate, C , decreases when gamma radiation passes through a shielding material of thickness x . The reduction in the transmitted count rate is related to the incident count rate and the material's attenuation properties. The attenuation coefficient (μ) can be determined using the following equations (3) and (4) [8, 16]:

$$C = C_0 e^{-\mu x} \quad (3)$$

Alternatively, the equation can be expressed in terms of the mass attenuation coefficient as:

$$C = C_0 e^{(-\frac{\mu}{\rho})(\rho x)} \quad (4)$$

where:

C = transmitted count rate after shielding (counts/s),
 C_0 = initial count rate before shielding (counts/s),
 x = thickness of the shielding material (cm),
 ρ = density of the material (g/cm³),
 μ = linear attenuation coefficient (cm)⁻¹,
 μ/ρ = mass attenuation coefficient (cm²/g).

Several factors influence the attenuation coefficient (μ), which describes the ability of a material to attenuate gamma radiation.

The atomic number of a substance primarily determines its efficacy in attenuating radiation. Elements with higher atomic numbers generally

increase the probability of gamma-ray interactions within the material. Materials with lower atomic numbers are often less efficient in attenuating gamma photons.

Once the attenuation coefficient is determined, it can be used to calculate the Mean Free Path (MFP), Half Value Layer (HVL), and Tenth Value Layer (TVL) as follows:

MFP (Mean Free Path) refers to the average distance that radiation travels before interacting with the shielding material.

$$MFP = 1/\mu \quad (5)$$

HVL (Half Value Layer) is the thickness required to reduce the transmitted count rate to 50% of the initial count rate.

$$HVL = \ln 2 / \mu \quad (6)$$

TVL (Tenth Value Layer) is the thickness required to reduce the transmitted count rate to 10% of the initial count rate.

$$TVL = \ln 10 / \mu \quad (7)$$

To provide a quantitative evaluation of the radiation shielding performance, the radiation protection efficiency (RPE) of the developed concrete blocks was calculated using the relation:

$$RPE(\%) = \left(\frac{C_0 - C}{C_0} \right) \times 100 \quad (8)$$

where C_0 and C are the initial count rate before shielding and the transmitted count rate after shielding, respectively. The calculated RPE values indicate the effectiveness of the developed blocks in attenuating Co-60 gamma radiation. Higher RPE values correspond to greater shielding capability.

RESULTS AND DISCUSSION

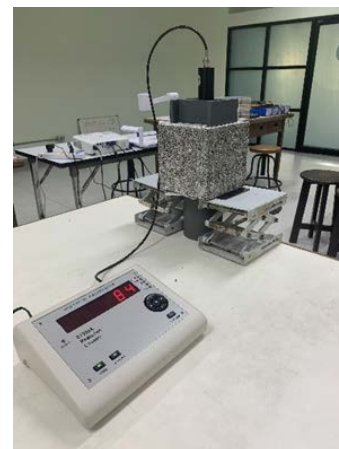


Figure 2 Experimental setup for gamma-ray shielding measurements using a Co-60 source and a Geiger-Muller detector. The system was used to measure incident and transmitted count rates for evaluating the shielding performance of the concrete samples.

Tables 1 and 2 present the elemental oxide composition and mineral phases of the shale and calcareous rocks. Both rock types contained similar major oxide compounds, but their concentrations differed. The five major oxide compounds identified in both rock types were SiO₂, CaO, Al₂O₃, K₂O and Fe₂O₃.

The physical and mechanical properties of concrete made from shale and calcareous rock were also investigated. This study examines the physical and mechanical characteristics of concrete produced from sedimentary rocks, specifically shale and calcareous rock. Two key properties were assessed:

1. The density of solid concrete samples made from each rock type, and
2. Their respective compressive strengths.

Table 3 summarizes the density and compressive strength of the concrete blocks. Conventional concrete showed the highest density and compressive strength, followed by shale concrete and calcareous concrete. The highest values were observed in typical construction concrete, with a density of 2.19 g/cm³ and a compressive strength of 25.45 MPa. Shale-based concrete blocks followed, with a density of 1.99 g/cm³ and a compressive strength of 13.59 MPa. The lowest values were recorded for concrete made from calcareous, at 1.88 g/cm³ and 5.21 MPa, respectively.

Table 1 Elemental and compound composition analysis of sedimentary rocks (XRF).

Calcareous		Shale	
compound	Concentration (%)	compound	Concentration (%)
Al ₂ O ₃	11.418	Al ₂ O ₃	10.823
SiO ₂	30.756	SiO ₂	28.461
K ₂ O	4.370	K ₂ O	3.892
CaO	22.826	CaO	24.965
Fe ₂ O ₃	3.081	Fe ₂ O ₃	3.472

Table 2 Types of compounds and chemical formula of sedimentary rocks (XRD).

Calcareous		Shale	
Mineral/compound name	Chemical Formular	Mineral/compound name	Chemical Formular
Quartz	SiO ₂	Calcite	Ca(CO ₃)
Calcite	Ca(CO ₃)	Quartz	SiO ₂
Pyrite	FeS ₂	Dolomite	CaMg(CO ₃) ₂
Orthoclase	K(AlSi ₃ O ₈)	Pyrite	FeS ₂
Muscovite	KAl ₂ Si ₃ AlO ₁₀ (OH) ₂	Orthoclase	K(AlSi ₃ O ₈)
Dolomite	CaMg(CO ₃) ₂	Muscovite	KAl ₂ Si ₃ AlO ₁₀ (OH) ₂

Table 3 Physical and mechanical properties of sedimentary rock-based concrete blocks.

Concrete type	m (kg)	Volume (m ³)	Density (kg/ m ³)	Compressive strength (Ksc)
Calcareous	6.375	0.0034	1,883.37	53.20
Shale	6.789	0.0034	1,993.72	138.72
concrete	7.526	0.0034	2,194.66	259.69

Table 4 Radiation shielding properties.

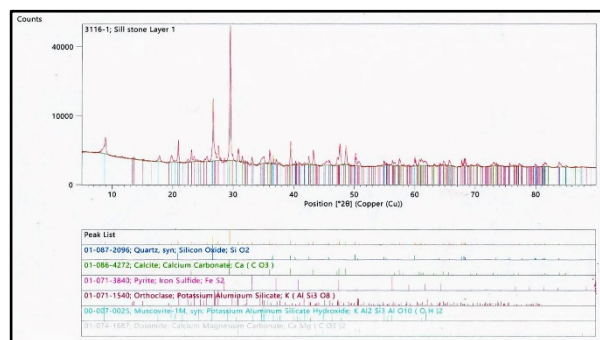
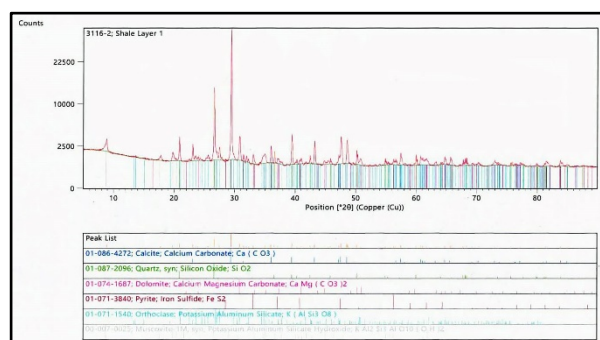
Operating Voltage (V)	Concrete type	μ_t (cm) ⁻¹	μ_m (cm ² /g)	MFP (cm)	HVL (cm)	TVL (cm)
960	Calcareous	0.13	0.07	7.59	5.26	17.47
	Shale	0.18	0.09	5.56	3.86	12.81
	Concrete	0.15	0.07	6.89	4.78	15.88

Table 5 presents the radiation protection efficiency (RPE) of the different concrete types. Shale concrete block is found to exhibit the highest RPE,

followed by conventional concrete and calcareous concrete block, respectively.

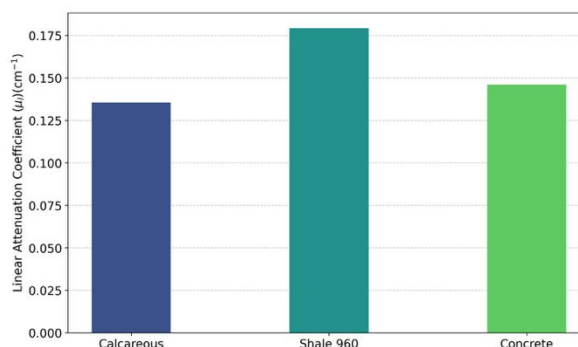
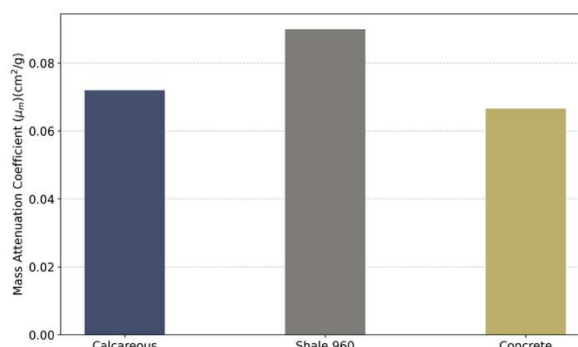
Table 5 Radiation shielding efficiency (RPE) of all concrete types.

Concrete type	C_0	C	RPE (%)
Calcareous	6,311.60	874.25	86.15
Shale	6,311.60	425.75	93.25
concrete	6,311.60	716.75	88.64

**Figure 3** XRD pattern of calcareous is shown, where quartz is identified as the dominant phase, with minor calcite, pyrite, orthoclase, and muscovite.**Figure 4** XRD pattern of shale is presented, in which calcite is identified as the dominant phase, with quartz, dolomite, pyrite, and orthoclase detected as minor constituents.

This research examines the efficacy of solid concrete as a gamma radiation shield by incorporating sedimentary rock samples with Portland cement. The analyzed sedimentary rocks consist of two varieties: shale and calcareous, combined with Portland cement, sand, and stone in a ratio of 1:2:4, respectively. The sedimentary rock samples used in the study were sourced from the La-Ngu District in Satun Province, which is part of the Satun Geopark. Next, the effectiveness of the dense concrete in blocking radiation was tested by measuring its ability to reduce radiation with a Geiger-Müller counter and a standard radiation source, Co-60 (shown in figure 2), which had an activity of 74 kBq and a half-life of 5.3 years. The measurement duration was 300 seconds to determine the attenuation coefficient of thick concrete from both varieties of sedimentary rock samples and to compare it with standard building concrete. The experimental findings are shown in Table 4. The Co-60 attenuation

measurements were done with the source at the bottom of the arrangement. It was set that the source and the concrete block would be 3 cm apart. The detector was just next to the 15 cm thick concrete block. In this setup, the source, sample, and detector were aligned vertically. So, the total distance from the source to the detector (C_0 measurement position) was 18 cm. The average value was utilized in the analysis after each measurement was taken three times for 300 seconds. Background radiation was measured separately and subtracted from all measured count values.

**Figure 5** Linear attenuation coefficient (LAC) values of calcareous concrete, shale concrete, and conventional concrete under Co-60 gamma irradiation.**Figure 6** Mass attenuation coefficient (MAC) values of calcareous concrete, shale concrete, and conventional concrete under Co-60 gamma irradiation.

From Figures 5 and 6, the graphs illustrate the relationships between the linear and mass attenuation coefficients for the three concrete types. The values of both the linear (μ_l) and mass (μ_m) attenuation coefficients presented in these graphs were calculated using Equation (4).

The results indicate that shale concrete blocks exhibited the highest linear attenuation coefficient, $0.18 (\text{cm})^{-1}$. Standard concrete showed a lower linear attenuation coefficient of $0.15 (\text{cm})^{-1}$, while the calcareous brick blocks had the lowest value of $0.13 (\text{cm})^{-1}$.

Regarding mass attenuation coefficients (μ_m), shale concrete blocks again demonstrated the highest

value at $0.09 \text{ cm}^2/\text{g}$. Both the calcareous concrete block and standard concrete had equal mass attenuation coefficients of $0.07 \text{ cm}^2/\text{g}$.

For the mass attenuation coefficient, shale concrete also showed the highest value of $0.09 \text{ cm}^2/\text{g}$, whereas calcareous concrete and conventional concrete showed values of $0.07 \text{ cm}^2/\text{g}$. These results indicate that shale concrete had the best shielding performance among the tested materials under the present Co-60 irradiation condition.

The radiation attenuation coefficient indicates how effectively a material reduces radiation count rate. Materials with higher values provide better shielding and are more suitable for radiation protection.

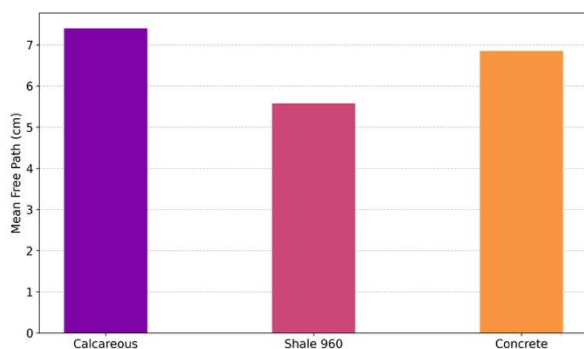


Figure 7 Mean free path (MFP) values of calcareous concrete, shale concrete, and conventional concrete calculated from the linear attenuation coefficient.

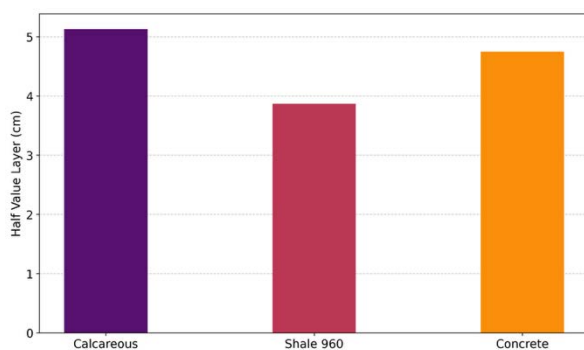


Figure 8 Half-value layer (HVL) values of calcareous concrete, shale concrete, and conventional concrete calculated from the linear attenuation coefficient.

Analysis of radiation shielding parameters in concrete blocks made from shale and calcareous under Co-60 irradiation. Figures 7 through 9 present the graphs of three key thickness parameters-Mean Free Path (MFP), Half Value Layer (HVL), and Tenth Value Layer (TVL). The MFP, HVL, and TVL values for shale concrete were 5.56 cm, 3.86 cm, and 12.81 cm, respectively. For calcareous concrete, the corresponding values were 7.59 cm, 5.26 cm, and 17.47 cm. Conventional concrete showed values of 6.89 cm, 4.78 cm, and 15.88 cm.

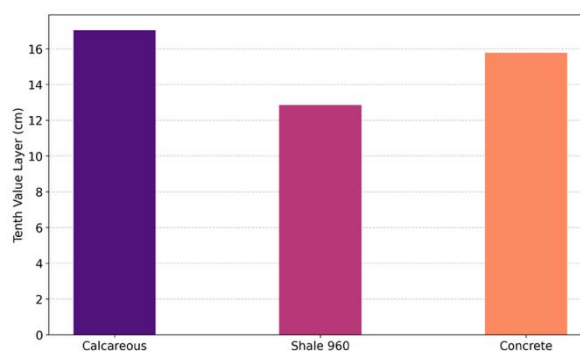


Figure 9 Tenth-value layer (TVL) values of calcareous concrete, shale concrete, and conventional concrete calculated from the linear attenuation coefficient.

All three parameters were calculated according to the formulas provided in equations (5) through (7). The results indicate that shale concrete had the lowest MFP, HVL, and TVL values, suggesting the best attenuation performance among the tested materials. Conventional concrete showed intermediate shielding performance, while calcareous concrete showed the lowest.

Lower MFP, HVL, and TVL values indicate that less material is required to reduce radiation count rate. Thus, materials with lower values are more efficient for Co-60 gamma-ray shielding, particularly in applications where space and material efficiency are important.

Discussion

1. Physical and Mechanical Properties: Density and Compressive Strength. This study examines how sedimentary rocks, specifically shale and calcareous rock, affect the basic properties of concrete. A mix ratio of 1:2:4 (cement: sand: aggregate by weight) was used, following previous studies [17] for general construction. The results show that conventional concrete had the highest density at 2.19 g/cm^3 . When shale was used, the density decreased to 1.99 g/cm^3 , and the lowest density, 1.88 g/cm^3 , was found in the mix with calcareous rock. Compressive strength showed the same pattern. Conventional concrete had a strength of 25.45 MPa, the shale mix had 13.59 MPa, and the calcareous mix had the lowest strength at 5.21 MPa. These results suggest that rock type affects both density and strength. Shale and calcareous rock are more porous than granite, which is commonly used in conventional concrete. Higher porosity leads to lower density and weaker structure. Overall, the results indicate that shale and calcareous concrete blocks have lower mechanical strength than conventional concrete. Therefore, their use should be limited to non-load-bearing or low-load shielding applications unless further mechanical optimization is performed.

2. Effect of Chemical Composition on Radiation Shielding. This study examines how the chemical

composition of calcareous and shale rocks affects the radiation shielding performance of concrete blocks. Table 1 shows that both rocks contain elements such as iron oxide (Fe_2O_3) and aluminum oxide (Al_2O_3), which are known to improve radiation attenuation. As a result, shale-based concrete may provide slightly better shielding than conventional concrete under the present experimental conditions, while calcareous concrete showed lower attenuation performance. Previous research [18] studied gamma-ray attenuation in several types of rock, including granite, sandstone, volcanic rocks, and feldspathic basalt, over an energy range of 122 to 1330 keV. Sandstone, another sedimentary rock, showed attenuation values similar to those reported in this study. The attenuation coefficients were 0.15 cm^{-1} for calcareous rock and 0.18 cm^{-1} for shale. These findings are consistent with previous studies, but this study focuses more clearly on the role of chemical composition. Calcareous rock collected in this study contains higher Al_2O_3 (11.419%) and Fe_2O_3 (10.823%) than shale (3.081% and 3.472%). This difference may explain its better shielding performance, especially at higher energies. Although lead is highly effective for radiation shielding, it is toxic and poses environmental risks. Therefore, safer alternative materials are needed. This study shows that using shale aggregate can slightly improve attenuation performance compared with conventional concrete, although this improvement cannot be attributed solely to density. These materials may serve as supplementary shielding materials in applications where moderate attenuation, low cost, and local availability are important. A related study [10] found that adding siltstone (0-40% by weight) increased concrete's shielding capacity. Higher siltstone content led to higher attenuation coefficients, mainly due to increased electron density, MnO content, density, and effective atomic number. Overall, sedimentary rocks can enhance the radiation shielding performance of concrete.

The results of this study and related literature show that radiation shielding depends on a few key factors. Density is generally an important factor in gamma-ray attenuation. However, the present results show that shale concrete had higher attenuation than conventional concrete despite its lower density. This suggests that chemical composition, mineral phases, aggregate distribution, and internal structure also contributed to the measured shielding performance. Compressive strength should be considered mainly as a mechanical suitability parameter rather than a direct indicator of radiation shielding ability. Chemical composition is also important because materials containing higher atomic-number elements generally exhibit improved gamma-ray attenuation. This study also suggests directions for future work. Adjusting the proportion of sedimentary rock in the mix could help better understand its effect on shielding. It would also be useful to compare different types of sedimentary

rocks with other common construction materials. In addition, studying other factors that affect shielding could help further improve radiation-resistant materials [19].

CONCLUSIONS

This study aims to evaluate the effectiveness of concrete blocks made from sedimentary rocks in La-ngu District, part of the Satun UNESCO Global Geopark, for radiation shielding. The study focuses on both physical and mechanical properties, as well as radiation attenuation performance. Concrete blocks in a cubic shape ($15 \times 15 \times 15 \text{ cm}$) were prepared using a 1:2:4 mix ratio of cement, sand, and rock by weight. Radiation levels were measured before and after placing shale and calcareous concrete blocks as shielding materials. Their performance was then compared with conventional concrete under Co-60 gamma irradiation. The results show that conventional concrete had the highest density and compressive strength, followed by shale blocks and then calcareous blocks. Radiation shielding was evaluated using the linear attenuation coefficient (μ_l), mass attenuation coefficient (μ_m), mean free path (MFP), half-value layer (HVL), and tenth-value layer (TVL). Shale blocks provided the best shielding, followed by conventional concrete and calcareous blocks.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this manuscript, the authors used ChatGPT and QuillBot for language editing and grammar improvement. The authors reviewed and edited the content and take full responsibility for the final manuscript.

REFERENCES

1. Sheekhoo WA, Abdulazeez KM. Assessing gamma-ray shield effectiveness in CuZnAl alloys: a comparative study. *J Next Energy*. 2025;9:100374.
2. Uzun Duran S, Küçükömeroglu B, Çiris A, Ersoy H. Determining the natural radiation level and gamma absorption coefficient of İkizdere obsidian. *Res Square* [internet]. 2022. Available from: <https://doi.org/10.21203/rs.3.rs-912918/v1>.
3. Farag HA, Draz WM, Ali FA, Arfa MM, Shazly RME, Sadawy MM, et al. Influence of engineering properties on radiation shielding characteristics of some Egyptian granitic rocks as potential materials. *Egypt J Chem*. 2025;68(6):67-78.
4. Abd El-Azeem SA, Harpy NM. Radioactive attenuation using different types of natural rocks. *Mater*. 2024;17(14):3462.

5. Taalab SA, Saadawi DA, Faraj TK, Khandaker MU, Khattab MR, Hanfi MY. Radiological risk and mineralogical analysis of radionuclides in granitic rocks: environmental insights. *Nucl Eng Technol.* 2025;57(11):103759.
6. Akkurt I, Altindag R, Onargan T, Basyigit C, Kilincarslan S, Kun M, et al. The properties of various igneous rocks for γ -ray shielding. *Constr Build Mater.* 2007;21:2078-82.
7. Gümrükçüoğlu N, Uzun Duran S, Tuncay, Küçükömeroglu B, Yilmaz AH. Gamma ray shielding properties and natural radioactivity of some rocks of Trabzon district, Türkiye. *Egypt J Soil Sci.* 2023;63(3):381-90.
8. Günoğlu K, Akkurt I, Sayyed MI. Radiation shielding properties of some igneous rocks in Isparta province at different gamma energies: experimental and theoretical study. *J Radiat Res.* 2024;17:1-9.
9. Mahmoud KA, Sayyed MI, Tashlykou OL. Gamma ray shielding characteristics and exposure buildup factor for some natural rocks using MCNP-5 code. *Nucl Eng Technol.* 2019;51:1835-41.
10. Aita RS, Mahmoud KA, Abdel Ghany HA, Ibrahim EM, El-Feky MG, El Aassy IE. Impacts of siltstone rocks on the ordinary concrete's physical, mechanical and gamma-ray shielding properties: an experimental examination. *J Nucl Sci Technol.* 2024;56:2063-70.
11. Alorfi HS, Hussein MA, Tijani SA. The use of rocks in lieu of bricks and concrete as radiation shielding barrier at low gamma and nuclear medicine energies. *Constr Build Mater.* 2020;251(1):118908.
12. Al-Buriah MS, Alomayrah N, Alsaiari NS, Tehçi T, Çaliskan F. Effect of Si₃N₄ on metakaolin clay-based geopolymer: EDS, SEM, hardness and gamma radiation shielding properties. *Constr Build Mater* [internet]. 2025;492:142857. Available from: <https://doi.org/10.1016/j.conbuildmat.2025.142857>.
13. Camgöz YI, Camgöz B, Yaprak G. Investigation on radiation attenuation properties of natural stone samples traded in Turkey. *Sci Total Environ.* 2024;926:171452.
14. Tranler JK, Kudryavtsev VA, Scovell PR. Gamma-ray background from rock: studies for a next generation dark matter experiment based on liquid xenon. *Astropart Phys.* 2025;171:103119.
15. Al-Naggar TI, Khan A, Mohamed SE, Abdalla AM, Albargi HB, Ashraf O. Evaluation of granite tiles for gamma-ray and neutron shielding: a cost-effective and sustainable solution. *Radiat Phys Chem.* 2026;238:113179.
16. Phetsasithorn A, editor. Gamma and neutron attenuation coefficients of heavyweight concrete using domestic aggregates. *Proceedings of the 26th National Convention on Civil Engineering, Civil Engineering and Beyond the Limit Development*; 2021 Jun 23-25; Bangkok, Thailand. Bangkok: Department of Civil Engineering, King Mongkut's Institute of Technology Ladkrabang; 2021.
17. Chaisakulkiate U, editor. Concrete mix proportioning by nominal mix method. *Proceedings of the 4th National RMUTR Conference and the 1st International RMUTR Conference, Increasing Research to Sustainable Economic and Society*; 2019 Jun 26-28; Bangkok, Thailand. Bangkok: Rajamangala University of Technology Rattanakosin; 2019.
18. Obaid SS, Gaikwad DK, Pawa PP. Determination of gamma ray shielding parameters of rocks and concrete. *J Radioanal Nucl Chem.* 2017;144:356-60.
19. El-Arabi AM, Abbady AGE, Hussein AS. Gamma-ray measurements of natural radioactivity in sedimentary rocks from Egypt. *Nucl Sci Technol.* 2006;17(2):123-8.



A reinforcement learning model for route allocation optimization: A case study of C-130H transport aircraft in the Royal Thai Air Force

Patikorn Anchuen¹, Phummipat Daungklang², Nuntipat Phisutthangkoon², Nattawat Tanomchad² and Hatsadin Jantaboon^{1*}

¹Office of Graduate Studies, Navaminda Kasatriyadhiraj Royal Air Force Academy, Bangkok 10220, THAILAND

²Department of Computer Science, Navaminda Kasatriyadhiraj Royal Air Force Academy, Bangkok 10220, THAILAND

*Corresponding author: hatsadin_ja@rtaf.mi.th

ABSTRACT

This study aims to develop a route-allocation model for transport aircraft in the Royal Thai Air Force's (RTAF) air logistics operations. A Reinforcement Learning (RL) approach was applied to optimize resource allocation by determining the most suitable routes based on cargo capacity distribution, thereby reducing operational flight distances and, consequently, the frequency of aircraft maintenance. This research was conducted in a simulated environment using domestic air transport data as a reference for C-130H transport aircraft of Squadron 601, Wing 6, RTAF. Experimental results show that the developed model significantly improves air transport operational efficiency in support and service provision, facilitating network operations by reducing flight distances and increasing process continuity under varying cargo capacity conditions compared to current practices. Ultimately, this contributes to the improvement and sustainability of defense operations. The proposed work schedule also demonstrates adaptability to dynamic operational constraints and changing demand.

Keywords: Air logistics, Route allocation, Reinforcement learning, C-130H transport aircraft

INTRODUCTION

Tactical air transport and logistics play a vital role in military operations and national security, particularly in the Royal Thai Air Force, whose primary mission is to support personnel and equipment to enable effective operations. Under the Royal Thai Air Force's 20-Year Strategy (2018-2037) (Revised 2020), Strategic Issue 2 emphasizes strengthening national defense capabilities and readiness, including the development and improvement of the logistics system to meet international standards [1]. However, the air logistics process still faces several challenges, especially in reducing transport distances and times while increasing safety and operational efficiency. Although air transport is a fast and highly flexible option, it incurs high costs, particularly in fuel, which is a major factor affecting the overall operating costs [2]. The C-130H short-body tactical transport aircraft, based at Squadron 601, Wing 6, has a high fuel consumption rate due to its size and the volume of equipment it carries [3]. Therefore, efficient aircraft planning and resource allocation have become essential. Management is a key factor directly related to the Royal Thai Air Force's combat readiness, agility, and long-term sustainability, achieved by reducing fuel consumption and increasing operational efficiency.

At present, route and fleet management optimization have attracted considerable attention in both commercial transportation and security-related applications, as these problems are inherently complex and directly impact operational expenses and resource utilization efficiency [4, 5]. Many studies are currently being conducted on the original problem using the Vehicle Routing Problem (VRP) method, which aims to find the best way for multiple vehicles to deliver goods to different locations, taking into account time, distance, and load capacity constraints [6, 7]. The Pick-up and Delivery Problem (PDP) is a more modern version of the VRP. It can handle situations where a ministry has to pick up and drop off loads on the same route, which means moving government equipment and people and keeping track of changing needs [8, 9].

VRP and PDP in many research studies can efficiently solve problems within a limited time frame by applying mathematical methods such as Integer Linear Programming (ILP) and Mixed-Integer Programming (MIP) [10], as well as heuristics and metaheuristic methods such as Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO) [11, 12]. However, these methods still have limitations when problems are large or exhibit complex dynamics, such as real-

time-changing demands, resource constraints, and uncertain environments.

Mathematical optimization and heuristic algorithms provide efficient routing at a fundamental level, but they often have limitations when faced with adaptive transportation situations, such as fluctuating load demands, time constraints, and constantly changing resource conditions. The resulting decisions were not made promptly and were not adaptable in actual work situations. [13] In response to these limitations, recent research has turned to more learning approaches that leverage information and environmental interactions to enhance system adaptability, particularly through the application of machine learning (ML) and deep learning (DL) techniques for decision modeling [14]. This research, therefore, applies Deep Reinforcement Learning (DRL) technology to the air logistics planning process to improve operational efficiency and increase the speed and safety of transporting personnel and equipment from one point to another, reducing distance, time, and resource allocation in a bidirectional manner for simultaneous pick-up and delivery of goods.

MATERIALS AND METHODS

Route allocation for C-130H air transportation is a complex combinatorial optimization problem constrained by fuel capacity, payload, and flight range, and can be modeled as VRP or PDP. Traditional approaches have been widely applied but often struggle in dynamic and uncertain environments [7, 12]. As a result, reinforcement learning has gained attention as a promising alternative for achieving more adaptive and effective route optimization.

1. Deep reinforcement learning for routing problems

Reinforcement Learning (RL) enables an agent to learn optimal decision-making policies by interacting with an environment and maximizing long-term cumulative rewards. In route allocation problems, the agent performs sequential route selection under operational constraints. Due to the large state and action spaces involved, Deep Reinforcement Learning (DRL), particularly the Deep Q-Network (DQN), is well-suited for this task [15, 16]. By approximating state-action Q-values with neural networks, DQN enables adaptive, data-driven route optimization under dynamic operational conditions.

2. Data collection and preparation

The goal of this study was to create a deep reinforcement learning model capable of routing the C-130H aircraft. A Deep Q-Network (DQN) was trained and tested in a simulated environment incorporating historical flight data. The results obtained under different operating conditions provide insights into the model's performance, robustness, and adaptability,

while demonstrating its potential applicability to real-world air logistics missions.

Each table or figure must include a brief description of the results being presented and other necessary information in a legend. Number tables consecutively in accordance with their appearance in the text, and place any table notes below the table body. Be sparing in the use of tables and ensure that the data presented in them does not duplicate results described elsewhere in the article. Please avoid using vertical rules and shading in table cells. Figures can be sized between column and page widths if the author chooses; however, it is recommended that figures are not sized less than column width unless necessary. When texts are typed completely, authors can perform text formatting. Style names are summarized and tabulated below.

This study's data collection is based on flight operation statistics of the C-130H transport aircraft operated by Squadron 601, Wing 6, RTAF [1], and it analyzes domestic military airlift patterns using five years of historical data [2]. Flight routes, airspeed, altitude, flight hours, payload weight, and fuel consumption are among the information gathered. According to the analysis, the most popular routes start at Wing 6, Bangkok Don Mueang, and connect to several other air wings, such as Wing 7, Surat Thani; Wing 21, Ubon Ratchathani; Wing 23, Udon Thani; Wing 41, Chiang Mai; Wing 56, Songkhla; Wing 1, Nakhon Ratchasima; and Wing 46, Phitsanulok.

These routes serve as the foundational data for constructing the route network model, as illustrated in Figure 1, with inter-wing distances depicted in Figure 2.

The payload capacity of the C-130H transport aircraft is defined by a maximum gross weight of 80,000 pounds [3]. This weight, which includes about 32,000 pounds of fuel, gives military equipment an effective payload capacity of 48,000 pounds [4]. Five primary pallets (Pallets 1–5), each with a maximum capacity of 8,500 pounds, and an auxiliary pallet (Pallet 6) with a maximum capacity of 5,500 pounds make up the useful payload. The fuel load of 32,000 pounds is illustrated in Figure 3.

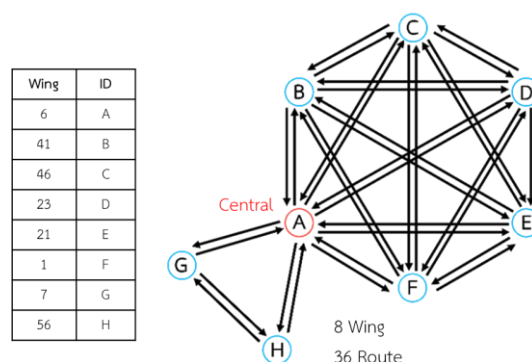


Figure 1 Domestic air wing flight route network.

Distance(Km)								
ID	A	B	C	D	E	F	G	H
A		570	320	450	480	195	555	775
B	570		260	500	780	540		
C	320	260		275	520	335		
D	450	500	275		325	280		
E	480	780	520	325		300		
F	195	540	335	280	300			
G	555							280
H	775						280	

Figure 2 Distances between air wings.

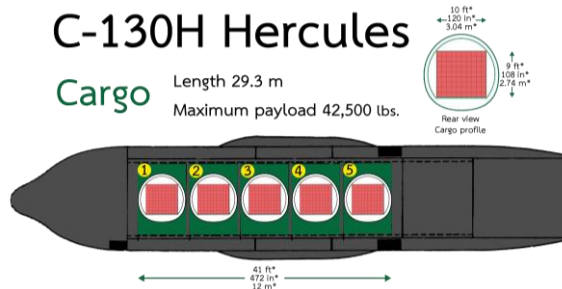


Figure 3 Number of pallets that can be carried on the C-130H aircraft.

The C-130H aircraft is capable of operating for a maximum flight duration of approximately 5 hours and 42 minutes, covering a distance of about 2,662 kilometers [6]. In addition, the aircraft can hold for approximately 1 additional hour in emergencies, in accordance with flight safety regulations [7]. The fuel consumption rate for JP-8 or Jet A-1 fuel is approximately 6,000 pounds during the first hour of operation, including engine start procedures, and approximately 4,500 pounds per hour during subsequent flight phases. An additional fuel reserve of 4,500 pounds is maintained for emergency contingencies [8].

3. Model development

Problem Formulation This study formulates the airlift route allocation problem as a sequential decision-making process under operational constraints. The objective is to minimize transportation cost while maximizing cargo delivery efficiency. The optimization objective is defined as:

$$\min J = \sum_{t=1}^T (\alpha D_t + \beta F_t - \lambda P_t) \quad (1)$$

D_t : total flight distance at time step t

F_t : fuel consumption

P_t : number of pallets delivered

α, β, λ : weighting coefficients

For reinforcement learning, the reward function is defined as:

$$r_t = \lambda P_t - \alpha D_t - \beta F_t - \eta C_t \quad (2)$$

C_t : penalty for constraint violations (e.g., payload overflow, range exceedance)

η : penalty coefficient

The system state at time step t is defined as:

$$s_t = [P_1, P_2, \dots, P_8, l_t, c_t, d_t] \quad (3)$$

P_i : number of pallets at air wing i

l_t : current aircraft location

c_t : remaining payload capacity

d_t : remaining allowable flight distance

The action $a_t \in A(s_t)$ represents a feasible round-trip airlift route sequence under operational constraints. Each action corresponds to a sequence of routes such as:

$$A \rightarrow B \rightarrow F \rightarrow C \rightarrow E \rightarrow A \quad (4)$$

The action space is dynamically constrained by payload capacity, flight range, and available pallets.

The state transition is governed by:

$$s_{t+1} = f(s_t, a_t) \quad (5)$$

where the function $f(\cdot)$ updates the system state based on cargo pick-up and delivery, aircraft movement, and resource consumption.

The reward function is defined as:

$$r_t = \beta_1 P_t - \beta_2 D_t - \beta_3 F_t - \beta_4 C_t \quad (6)$$

where C_t represents penalties for constraint violations such as payload overflow or exceeding flight range.

The model is subject to the following constraints:

- Payload capacity: $\sum P_t \leq 5$ pallets per flight
- Flight range constraint: $\sum D_t \leq D_{max}$
- Fuel constraint: $\sum F_t \leq F_{max}$

This formulation ensures that routing decisions remain feasible and aligned with real-world airlift operations.

Deep reinforcement learning is used to solve routing problems in air transport operations for C-130H aircraft, accounting for constraints such as flight distance, payload capacity, and fuel consumption [9]. This approach allows the model to learn how to make decisions by interacting with the working environment [10]. Deep Q-Network (DQN), which uses a deep neural network to estimate the Q-value function, enables the model to learn complex decision policies through iterative interactions with the environment [11]. Previous studies have demonstrated the effectiveness of this algorithm in solving routing problems and improving logistics [12, 13].

The data collected in the preceding section are transformed into a flight route network representation of air wings, consisting of a total of 8 nodes representing air wings responsible for pick-up and delivery operations within a single mission round. This configuration results in 36 possible flight routes, as illustrated in Figure 1 [14]. Wing 6 is designated as the central node, while the remaining air wings are treated as transit nodes to reflect the actual operational patterns of the squadron [17]. Route characteristics, including

flight distance, flight duration, and fuel consumption between air wings, are used as input variables to construct the simulated dataset for model training [15]. These variables play a critical role in evaluating the effectiveness of decision-making for each flight route [18].

Prior to model training, an initial data filtering

process based on reward scores is performed to select actions that are most appropriate and aligned with the research objectives. This process helps reduce problem complexity and improve the stability of the learning process [19]. As a result, the model is better able to identify efficient routes, as summarized in Table 1.

Table 1 C-130H flight operation data.

Route	Flight time (Minute)	Fuel Consumption (lbs.)	Distance (Km.)
BA	73	6975	570
CA	46	4600	320
CB	39	3900	260
DA	64	6300	450
DB	84	7800	500
DC	44	4400	275
EA	66	6450	480
EB	98	8850	780
EC	70	6750	520
ED	47	4700	325
FA	33	3300	195
FB	84	7800	540
FC	51	5100	335
FD	42	4200	280
FE	44	4400	300
GA	72	6900	555
HA	96	8700	775
HG	41	4100	280

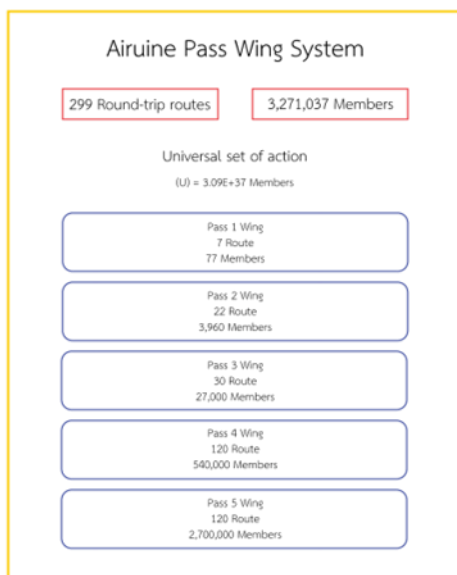


Figure 4 Route Classification Structure for DRL.

Prior to model training a data filtering process is conducted to select actions that correspond to flight routes yielding the highest rewards. The selection criteria are based on reward scores, as shown in Figure 4. A neural network architecture is used to estimate the Q-value function, which determines whether the action is optimal for each system state [21]. The loss function, which measures the difference between

predicted and actual rewards, is what the network's weights are always changing during training to minimize. This process enables the model to learn more accurate and effective decision-making policies under varying operational conditions.

After training with the DQN model, the Q value for each route can be automatically calculated, and the route yielding the highest overall reward score can be selected, taking into account factors such as payload capacity, flight distance, and fuel consumption. The results are then compared with those obtained from traditional route-allocation methods during the testing phase to verify the performance of the proposed model.

The proposed model employs a Deep Q-Network (DQN) to approximate the action-value function $Q(s,a)$. The neural network consists of two fully connected hidden layers, each containing 200 neurons. Rectified Linear Unit (ReLU) activation functions are applied to the hidden layers, while a linear activation function is used in the output layer.

The model is trained using experience replay with a memory buffer size of 10,000 transitions. A minibatch size of 36 is sampled uniformly from the replay memory at each training step. The discount factor is set to $\gamma=0.95$, and the learning process

follows an ϵ -greedy policy, where ϵ decreases linearly from 1.0 to 0.1 over the training period.

The network parameters are optimized using the Adam optimizer with a learning rate of 0.001. The model is trained for 100,000 iterations, and convergence is determined based on the stabilization of the Temporal Difference (TD) error.

A target network is employed to improve training stability, with parameters updated periodically every 1,000 iterations.

All experiments are conducted using the same dataset and configuration to ensure reproducibility.

In addition to the analysis of Temporal Difference (TD) error and reward trends, the performance of the proposed method was further evaluated using practical operational indicators. These include total flight distance, cargo throughput, and overall routing efficiency, which provide a clearer understanding of how the model performs in real-world air logistics operations.

The results show that the proposed method can significantly reduce the total flight distance compared to the conventional routing approach. In this study, the total distance decreases from 3,130 km to 2,445 km, representing a reduction of approximately

21.9%. This improvement suggests that the learned routing policy is able to eliminate unnecessary travel segments and produce a more efficient flight pattern.

Regarding cargo handling performance, the proposed method delivers 19 pallets in a single mission cycle. In contrast, the conventional approach is estimated to handle approximately 16 pallets due to repeated return trips and less efficient route planning. This corresponds to an increase of about 18% in cargo throughput, indicating better utilization of aircraft capacity.

Although fuel consumption is not directly modeled as an objective, reducing travel distance naturally leads to lower fuel use. This also implies reduced operating costs and aircraft usage, which are important factors in military air transport operations.

Another key observation is that the proposed method avoids repeated returns to the origin air wing, resulting in a more compact route structure. This allows multiple pick-up and delivery operations to be completed within a single mission, improving both time efficiency and resource utilization.

A summary of the comparison between the conventional method and the proposed approach is presented in Table 2.

Table 2 Performance comparison.

Method	Distance (km)	Reward	Pallets Delivered	Improvement
Conventional	3,130	27,870	16	baseline
Proposed (DQN)	2,445	16,555	19	-21.9% Distance and +18.8% Pallets Delivered

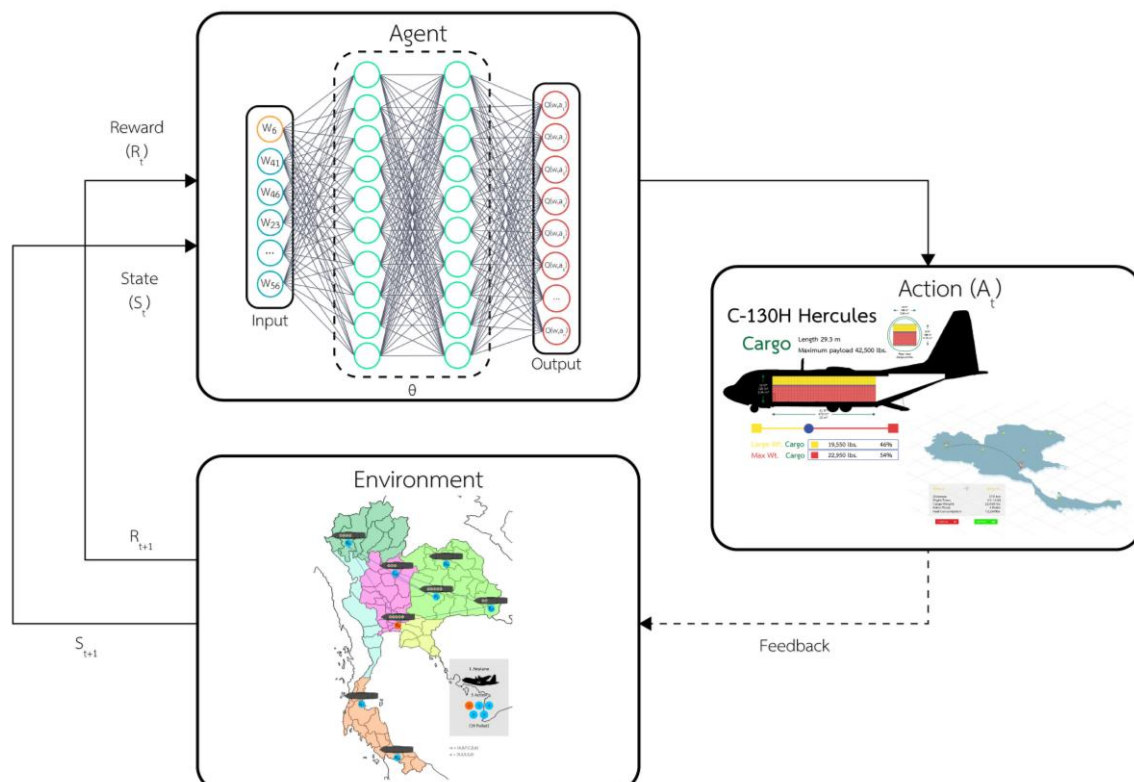


Figure 5 DQN architecture.

4. Model testing and performance evaluation

To assess the efficacy of deep reinforcement learning models, experiments were conducted in a simulated environment using operational data from either the 601st Wing or the 6th Wing of the Royal Thai Air Force. There were 36 air transport routes and 40 pallets of military gear. These data were employed as the testing dataset for both model training and performance evaluation of the decision-making process [22, 23]. During training, key neural network parameters within the DQN algorithm—including the number of training episodes, learning rate, and network depth—were configured to enable comparative analysis across three scenarios: Case I (learning rate = 0.05), Case II (learning rate of 0.07), and Case III (learning rate of 0.10) [24, 25].

The experimental results demonstrate that the model effectively learns to reduce the discrepancy between predicted and target Q-values. As the number of training episodes increases, the Temporal Difference (TD) error shows a steady decline [26], indicating improved learning stability. The DQN architecture employed in this study is illustrated in Figure 5, and the learning procedure follows the algorithm described below:

Algorithm: Deep Reinforcement Learning

1. Initialize replay memory D
2. Initialize Q-network with random weights θ
3. For each episode = 1 to N_{ep} do
 - Initialize starting state s_1
 - Preprocess input to get $\phi_1 = \Phi(s_1)$
 - For each time step $t = 1$ to T do
 - Select action a_t using epsilon-greedy policy based on $Q(\phi_t, a; \theta)$

```

Execute action  $a_t$  and observe reward  $r_t$  and
next state  $s_{t+1}$ 
Preprocess next state to get  $\phi_{t+1} = \Phi(s_{t+1})$ 
Store transition  $(\phi_t, a_t, r_t, \phi_{t+1})$  in replay
memory D
Sample random minibatch of transitions  $(\phi_j, a_j, r_j, \phi_{j+1})$  from D
If  $\phi_{j+1}$  is a terminal state then
   $y_j = r_j$ 
Else
   $y_j = r_j + \gamma * \max_{a'} Q(\phi_{j+1}, a'; \theta)$ 
End if
Perform gradient descent step on  $(y_j - Q(\phi_j, a_j; \theta))^2$ 
End for
End for
  
```

RESULTS AND DISCUSSION

The decision policy is constructed and its implementation is discussed. The DQN algorithm constructs a decision function that tells the system what to do based on its current state. This design aims to improve route allocation for C-130H operations, accounting for factors such as time and distance and fuel consumption. The proposed framework effectively simplifies the decision space while maintaining operational relevance and efficiency [27].

To make a synthetic dataset, you need to set important variables like the route ID, the action taken, the distance traveled, and the number of pallets moved. These variables are the inputs that the DQN model [20] uses to learn. The compiled data is organized and stored in a tabular format based on route sequences, as demonstrated in Table 3.

Table 3 Round-trip distances of routes under the selected actions.

Route	Action	Distance (Km.)
1	a_1 - a_{35}	1140
2	a_{36} - a_{70}	640
3	a_{71} - a_{105}	900
...	...	...
299	$a_{3248706}$ - $a_{3271205}$	1925

Table 4 Training configuration.

Parameters	Action
Data samples	10000
The number of links	36
Total Neurons	200
Total Output Layer	300
Threshold Error	0.0001
Total Iterations	100000
Learning rate	0.05/0.07/0.10

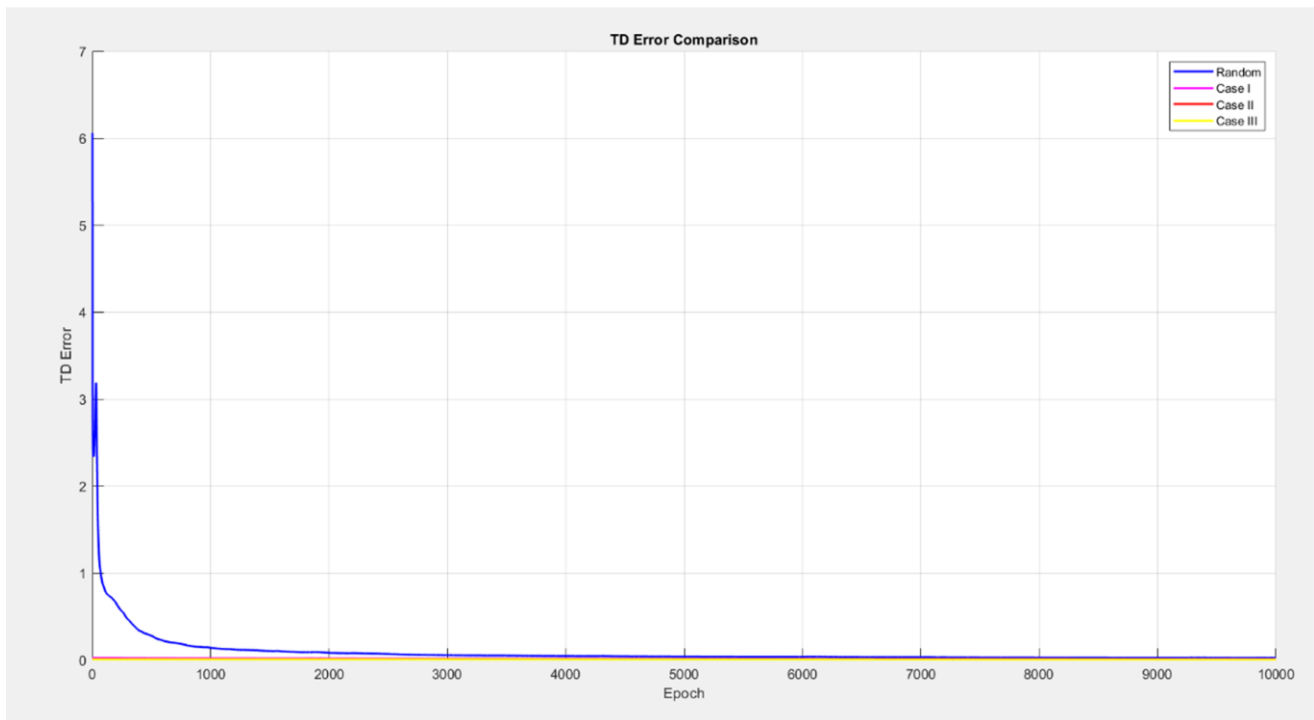


Figure 6 TD error between target values and q-values.

1. Policy construction

This stage constitutes a critical component of the deep reinforcement learning (DRL) process, in which the DQN algorithm is employed to derive an optimal decision function for allocating airlift routes for Squadron 601, Wing 6, RTAF. The primary objective is to minimize total flight distance while improving the efficiency of aviation resource utilization. The construction of a DRL framework aims to enhance various management efficiencies by enabling the algorithm to learn through interactions between an agent and the environment. This learning process is often formulated as a state-action pair, where each action receives a reward reflecting the quality of the decision at a given time step, though other formulations exist as well. The Q-value, $Q(s,a)$, representing the expected cumulative return obtained by taking action a in state s , is approximated using a neural network in accordance with the Deep Q-Network (DQN) paradigm [16, 19].

In this study, the state space consists of information on the number of pallets held at each air wing at time t , along with route distances and flight constraints. The action is to the set of available airlift routes (R1-R36) that the C-130H transport aircraft can select for operation. The DQN model incorporates an experience replay mechanism combined with the Adam optimizer to update network weights and to reduce the Temporal Difference (TD) error between predicted and target Q-values, thereby improving learning accuracy and stability [23].

The airlift activities of Squadron 601, Wing 6, comprising eight air wings, forty pallets of military equipment, and thirty-six potential routes between

origin and destination air wings, are the subject of the case study. The operational requirements of each air squadron will determine the reward for each flight route. The C-130H aircraft can accommodate up to 5 pallets per flight [28].

Given the complexity and uncertainty inherent in route allocation, this study adopts a trial-and-error learning process to identify optimal decision policies. For each experimental scenario, actions are randomly selected from the available route set and evaluated under three cases Case I (learning rate = 0.05), Case II (0.07), and Case III (0.10) to compare the learning speed and stability of the proposed model. A fully random action selection strategy is used as a baseline for performance comparison [22].

The learning rate is a key factor in how quickly the DQN neural network's weights are updated [22, 24]. To further improve the model's performance during retraining, parameters such as the number of hidden nodes, the number of training episodes, and the error threshold are set by parameters [25]. Figure 6 shows that the random baseline has a Temporal Difference (TD) error that stays pretty stable across training episodes. But this stability is due to the TD error converging to a value, not the policy converging to a value, because the random policy doesn't perform structured learning or policy improvement. So, even though the TD error in the random case looks low, the decisions and total rewards are still much lower than those from the DQN-based methods. The comparative results indicate that the random case, a conventional uninformed action selection strategy, exhibits slower learning while maintaining greater numerical stability due to the absence of policy updates. However, it does

not reach policy convergence or reward optimization [26]. On the other hand, Cases I–III show faster convergence in learning because they use different learning rates (0.05, 0.07, and 0.10, respectively). However, transient instability is observed during the initial training phase, evidenced by temporarily elevated TD errors resulting from frequent weight updates [27].

The final results reveal that as the number of training episodes increases, the algorithm consistently reduces the TD error and enhances the

effectiveness of the decision function in terms of cumulative Q-value rewards. As shown in Figure 7, the random baseline yields the lowest average reward across all states, confirming that the numerical stability observed in the TD error does not translate into effective route allocation or improved operational performance. Cases I–III ultimately achieve 100% convergence at the end of training. In contrast, the random baseline yields the lowest Q-values, as it does not account for structured route-based decision-making within the action space.

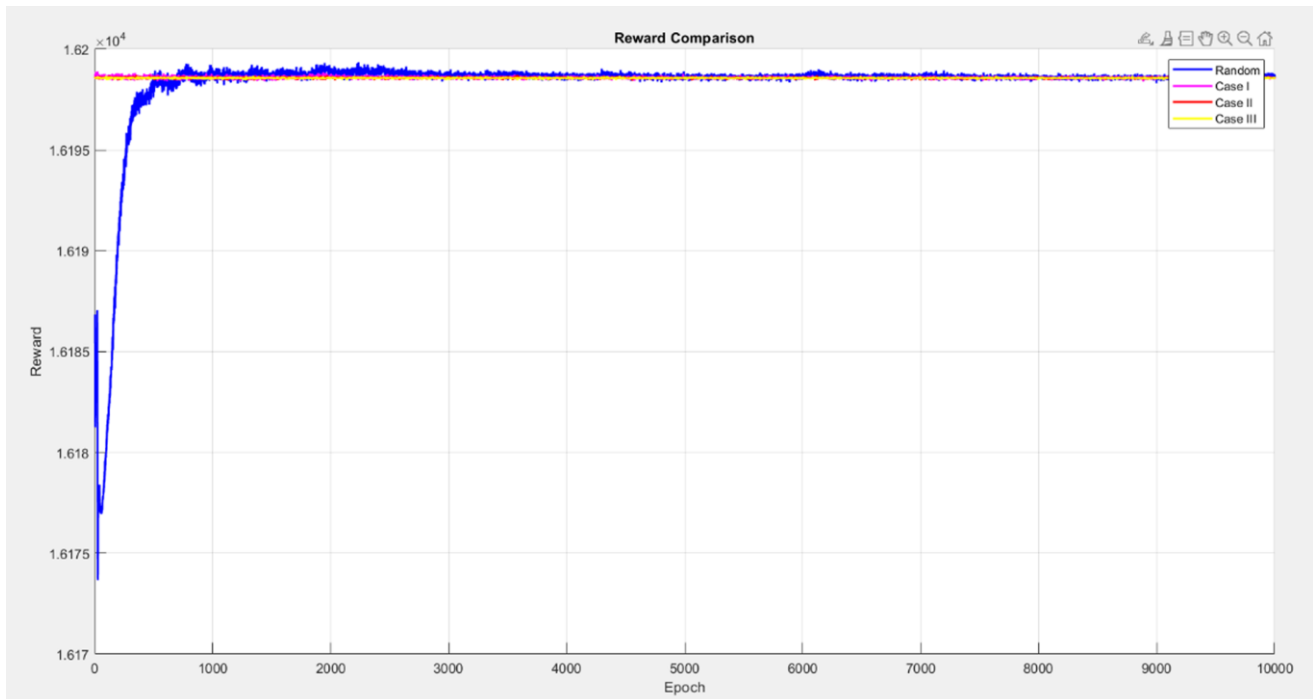


Figure 7 Average Reward for Each State.

2. Implementation

The decision function developed in the previous stage is applied to determine operational actions from the simulated system states for pallet balancing in the case study of Squadron 601, Wing 6, RTAF, which is responsible for managing C-130H airlift route allocation, as illustrated in Figure 8. The number of available pallets at each air wing is provided as input to the shared pallet management system, enabling the model to evaluate feasible transportation alternatives by computing Q-values and selecting the action that maximizes cumulative rewards under system constraints [22]. For example, the selected actions $R8 = 5$, $R29 = 4$, $R17 = 5$, $R25 = 3$, and $R4 = 2$ can be interpreted as an airlift route sequence $A \rightarrow B \rightarrow F \rightarrow C \rightarrow E \rightarrow A$. Within a single operational round, the C-130H aircraft performs pick-up and delivery operations totaling 19 pallets, covering five air wings with a total flight distance of 2,445 kilometers. The simulation results show that the proposed routing strategy significantly reduces the total flight distance compared to traditional fixed-route assignment

methods. This improves overall air transport efficiency. Moreover, the model's decision function explicitly accounts for operational constraints consistent with real-world deployment, including maximum aircraft payload capacity, flight range limitations, and mission profiles within a single airlift cycle. As a result, the proposed approach significantly enhances transportation completeness while reducing total distance-related expenses, as shown in Figure 9.

3. Discussion

From the simulation and real-world results presented in the previous section, a Deep Reinforcement Learning (DRL) model developed using the Deep Q-Network (DQN) algorithm demonstrates a significant improvement in pallet handling within the C-130H air transport system. The model learns effective decision policies under resource and time constraints by estimating the Q-value function to select the action that maximizes the cumulative reward for each system state [16]. Although the DQN algorithm shows strong potential for handling large state and action spaces, increasing the action space size

inevitably increases computational complexity. This, in turn, may result in longer training times and increased computational resource consumption, particularly in airlift routing problems characterized by numerous operational constraints [26].

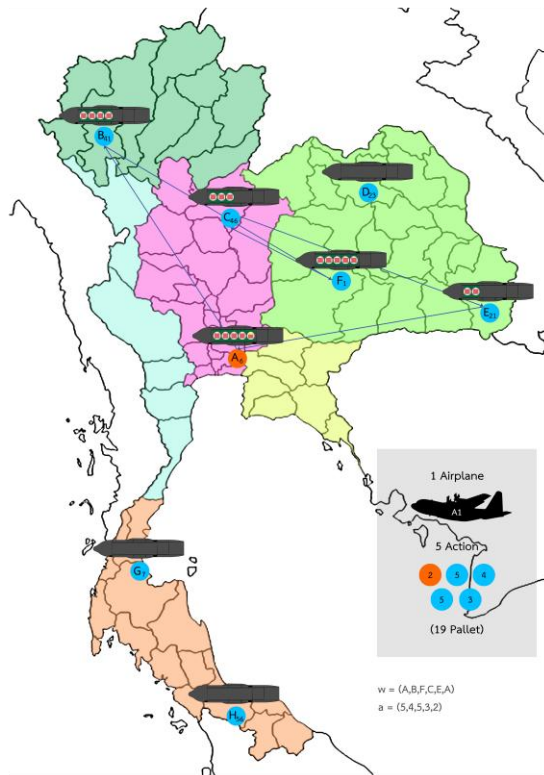


Figure 8 Application of the Decision Function for Military Airlift Operations Using C-130H Aircraft

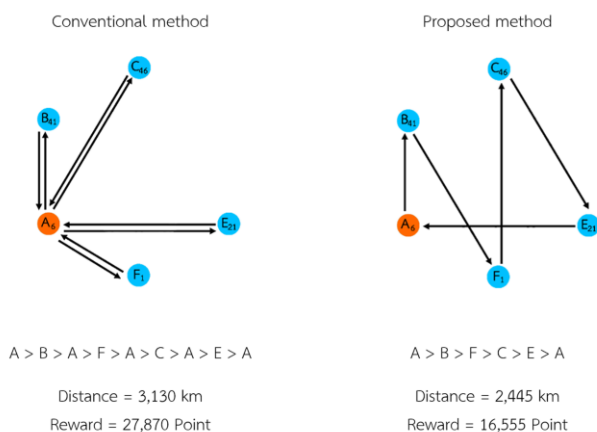


Figure 9 Comparison of Travel Distance and Reward Between the Traditional Method and the Proposed Approach for Military Airlift Operations Using C-130H Aircraft of Squadron 601, Wing 6, RTAF.

This study indicates that as the number of training episodes increases, the model's Q-values increase continuously, while the time difference error (TD) gradually decreases. This behavior reflects the stability and suitability of the DQN model's

internal parameter tuning [26]. These results are consistent with deep reinforcement learning, which aims to reduce the discrepancy between the estimated Q-value and the target Q-value once sufficient training has been completed [24]. Furthermore, applying the trained decision function to the Squadron 601, Wing 6, RTAF simulation environment has demonstrated that this model can efficiently select air transport routes that yield the highest Q-value, while maintaining operational constraints. This capability reduces operating expenses (OPEX) for military air transport missions compared to traditional routing methods.

However, an overly large action space, while potentially leading to higher Q-values, imposes a significant computational burden and slows down the learning process. Conversely, an overly limited action space can limit the model's ability to explore and identify optimal policies fully.

Therefore, designing an appropriate action-filtering strategy is crucial for balancing learning efficiency and computational complexity. This observation is consistent with accepted concepts in the field of function estimation and policy optimization within reinforcement learning frameworks [16]. Overall, the research confirms that the DQN algorithm, adapted to the operational characteristics of military air transport missions to reduce overall mission distance and increase pallet coverage under real-world operational constraints, shows promising potential for future application in military transport routing systems.

CONCLUSIONS

This study developed a pallet allocation and routing system for transportation between originating and destination air bases to address imbalances between bases. These imbalances arise from differing cargo demands, leading to overloaded aircraft or empty flights and negatively impacting the overall efficiency of the air logistics system. The proposed framework integrates a routing algorithm to appropriately filter and reduce the action space, in conjunction with a Deep Reinforcement Learning (DRL) approach—specifically the Deep Q-Network (DQN) to construct a decision function for pallet management and transportation operations. Experimental results using real operational data from Squadron 601, Wing 6, RTAF demonstrate that an optimally designed action space enables the model to learn and make decisions more efficiently. The proposed approach reduces overall flight distance, increases cargo volume, and lowers operating expenses (OPEX) compared to traditional methods. The key quantifiable performance improvements of the proposed method are summarized as follows: The total flight distance is reduced by approximately 21.9% compared to the conventional routing approach. Cargo throughput increases from an

estimated 16 pallets per mission to 19, representing an improvement of approximately 18.8%. The proposed method reduces redundant return trips to the origin air wing, resulting in a more compact, efficient routing structure. Furthermore, the learned decision-making functions can be applied to real-world logistics planning systems to support operators in making transportation decisions. Moreover, the proposed concept can be extended to other complex transportation management systems, such as land or sea logistics, which require real-time decision-making under resource constraints. Such an extension has the potential to increase operational resilience, reduce fuel consumption, and improve future transportation infrastructure. The proposed model relies on optimal factor analysis, algorithm selection, and the quality of training data. Therefore, practical implementation requires the systematic collection and management of C-130H operational data, including flight paths, speed, altitude, flight hours, and payload data. In addition, an automated, accurate data retrieval process is crucial for reducing operator workload and ensuring data continuity, enabling the model to adapt to real-world operating conditions effectively over time.

ACKNOWLEDGEMENT

This research was supported as part of the Master of Engineering program in Defense Technology, Office of Graduate Studies, Navaminda Kasatriyadhiraj Royal Air Force Academy (NKRAFA). The authors gratefully acknowledge Flt. Lt. Asst. Prof. Dr. Patikorn Anchuen for his valuable guidance, constructive comments, and continuous support throughout the research and thesis completion.

REFERENCES

1. Royal Thai Air Force. 20-Year Air Force Strategy (2018-2037). Revised 2020. Bangkok (Thailand): Royal Thai Air Force; 2020.
2. Royal Thai Air Force. Air Force Operational Plan 2023–2027. Bangkok (Thailand): Royal Thai Air Force; 2022.
3. Mouton CA, Powers JD, Romano DM, Guo C, Bednarz S, O'Connell C. Fuel reduction for the mobility air forces [Internet]. Santa Monica (CA): RAND Corporation; 2015 [cited 2026 Jan 26]. Available from: https://www.rand.org/content/dam/rand/pubs/research_reports/RR700/RR757/RAND_RR757.pdf.
4. Jayapal C, Devadharshini R, Praniis R, Praniis R. Enhancing fleet operations with a vehicle route management solution. In: Proceedings of the 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA); 2023; Coimbatore, India. p. 1-6.
5. Coppo M, Agostini M, De Matteis G, Bertolini M. Optimal management of commercial electric vehicle fleets with recharging stations and time-varying electricity prices. *Energies*. 2025;18(3): 453.
6. Bao Y. A summary and discussion on the current state of CVRP research. *Appl Comput Eng*. 2024; 115(1):35-42.
7. Reddy TS, Dhanush D, Krithin T, Sayan J. Supply chain logistics with hybrid optimization using ADMM and vehicle routing problem. In: Proceedings of the International Conference on; 2024.
8. Santos MJ, Jorge D, Bonomi V, Ramos TRP, Barbosa-Póvoa AP. Enhancing logistics through a vehicle routing problem with deliveries, pick-ups, and backhauls. *Int Trans Oper Res*. 2024.
9. Desticioğlu Taşdemir B, Özyörük B. Mathematical model for multi-depot simultaneously pick-up and delivery vehicle routing problem with stochastic pick-up demand. *Gazi Univ J Sci*. 2025;38(1): 219-35.
10. Jantaboon H, Anchuen P, Bua-ngam P. An optimization model for Air Force logistics: A case study of C-130H transport aircraft. In: Proceedings of the NCDT2025 Conference; 2025. p. 146-50.
11. Uthansakul P, Anchuen P, Uthansakul M, Khan AA. QoE-aware self-tuning of service priority factor for resource allocation optimization in LTE networks. *IEEE Trans Veh Technol*. 2020;69(1):887-900.
12. Sharma A, Srivastava S, Rautela PS, Joshi B. Metaheuristics-based routing optimization in on-chip network. In: Proceedings of the ACM International Conference; 2023.
13. Anchuen P, Uthansakul P, Uthansakul M, Poochaya S. Fleet optimization of smart electric motorcycle system using deep reinforcement learning. *Comput Mater Contin*. 2022;71(1):1925-43.
14. Poochaya S, Uthansakul P, Uthansakul M, Anchuen P, Thammakul K, Khan AA, et al. A multi-mode public transportation system using vehicular to network architecture. *Comput Mater Contin*. 2022;73(3).
15. Daungklang P, Mahavongvun A. Create a process for manually editing rapid 3D mapping datasets [Internet]. 2011 [cited 2026 Jan 26]. Available from: <https://www.diva-portal.org/smash/get/diva2:426325/FULLTEXT01.pdf>.

16. Sutton RS, Barto AG. Reinforcement learning: An introduction. 2nd ed. Cambridge (MA): MIT Press; 2018.
17. Silver D, Lever G, Heess N, Degris T, Wierstra D, Riedmiller M. Deterministic policy gradient algorithms. In: Proceedings of the International Conference on Machine Learning (ICML); 2014. p. 387-95.
18. Arulkumaran K, Deisenroth M, Brundage M, Bharath A. Deep reinforcement learning: A brief survey. IEEE Signal Process Mag. 2017;34(6):26-38.
19. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Marc G, et al. Human-level control through deep reinforcement learning. Nature. 2015;518(7540):529-33.
20. Busoniu L, Babuska R, De Schutter B, Ernst D. Reinforcement learning and dynamic programming using function approximators. Boca Raton (FL): CRC Press; 2017.
21. Watkins C, Dayan P. Q-learning. Mach Learn. 1992;8(3-4):279-92.
22. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, et al. Playing Atari with deep reinforcement learning. arXiv preprint. 2013;arXiv:1312.5602.
23. Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, et al. Continuous control with deep reinforcement learning. arXiv preprint. 2015; arXiv:1509.02971.
24. Hessel M, Modayil J, Van Hasselt H, Schaul T, Ostrovski G, Dabney W, et al. Rainbow: Combining improvements in deep reinforcement learning. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2018. p. 3215-22.
25. Wang Z, Schaul T, Hessel M, Hasselt H, Lanctot M, Freitas N. Dueling network architectures for deep reinforcement learning. In: Proceedings of the 33rd International Conference on Machine Learning (ICML); 2016. p. 1995-2003.
26. Van Hasselt H, Guez A, Silver D. Deep reinforcement learning with double Q-learning. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence; 2016. p. 2094-2100.
27. Baneshi F, Hernandez ML. Network-level aircraft trajectory planning via multi-agent deep reinforcement learning. Comput Ind Eng. 2025;175:109057.
28. Koh S, Zhou B, Fang H. Real-time deep reinforcement learning-based vehicle routing and navigation. Appl Soft Comput. 2020;96:106694.



Product development of mangosteen (*Garcinia mangostana* L.) with karonda (*Carissa carandas* L.) and rice protein sorbet

Jinussiga Monchaising¹, Sunisa Dorndi¹, Kamonchat Wanthongchan¹, Thiwaporn Jitdee¹, Amornrat Suwanposri², Warunya Nonmuang¹ and Kannikar Charoensuk^{1*}

¹Department of Food Innovation and Business, Faculty of Agro-industrial Technology, Rajamangala University of Technology Tawan-ok, Chanthaburi campus, Chanthaburi, THAILAND

²Department of Plant Production and Landscape Technology, Faculty of Agro-industrial Technology, Rajamangala University of Technology Tawan-ok, Chanthaburi campus, Chanthaburi, THAILAND

*Corresponding author: kannikar_ch@rmutto.ac.th

ABSTRACT

This study aimed to develop a functional sorbet formulation by combining mangosteen (*G. mangostana*) and karonda (*C. carandas*) juices, enriched with rice protein. Mixed fruits juice was prepared by substituting mangosteen juice with karonda juice at concentration of 0, 5, 10, 15 and 20% w/w. The physicochemical properties, including color, pH, total soluble solids (TSS), vitamin C content, and antioxidant activity, were evaluated. Results, indicated that increasing karonda juice concentration enhanced red coloration, DPPH radical scavenging activity, and vitamin C enrichment, the formulation containing 15% w/w karonda juice was selected for further development. Subsequently, sorbets enriched with rice protein at levels of 0, 5, and 10% w/w were prepared and characterized. Rice protein enrichment reduced the melting rate, demonstrating its role as natural stabilizing agent in sorbet formulations. Additionally, rice protein contributed functional and nutritional benefits, including increased energy and protein content, while improving textural properties. Despite these advantages, sensory attributes such as flavor and taste remained critical determinants of consumer acceptance and market success. Overall, the incorporation of karonda juice and rice protein into mangosteen sorbet enhances its nutritional profile and functional properties, highlighting the potential of plant-based proteins and fruit blends in the development of innovative frozen desserts.

Keywords: Mangosteen, Karonda, Sorbet, Plant-based protein, Functional food

INTRODUCTION

Mangosteen (*G. mangostana*) is a popular tropical fruit widely grown in Southeast Asia. June to August is the harvest season of mangosteen in Thailand. Recent updates on metabolite composition and medicinal benefits of mangosteen [1]. Due to it is highly nutritional value, thus, mangosteen juice was formulated. The juice made with 100% mangosteen juice exhibited higher levels of total phenolic content, antioxidant activity (measured by DPPH), and antimicrobial activity against both Gram-positive and Gram-negative pathogenic bacteria compared to the juice made 45% juice [2]. The fruit of Karonda (*C. carandas*) has white color to be pink-red until it is dark-purple too almost black when fully ripe with a sour have found the phytochemical constituents, including vitamin C, anthocyanin, flavonoids, glycosides, alkaloids, carbohydrates, sterols, terpenoids, tannins, and saponins taste, bitter, and astringent [3] A fruit drink enriched with protein is a beverage that combines real fruit flavors with a

protein boost, made for those needing extra protein or a healthier alternative to regular sugary juices [4], and the drinks are available as the commercial ready-to-drink options as mixes with plant-based protein, low-sugar, and free of fat and lactose [5]. Plant-based protein offers a nutritious, eco-friendly alternative to animal protein, supporting muscle repair, heart health, and longevity [6]. It's found in a wide variety of whole foods and can meet all dietary protein needs when consumed thoughtfully [7]. Rice protein is a naturally gluten-free protein and contains key protein fractions like prolamin, glutelin, globulin, and albumin, making it suitable for infants and those with gluten allergies [8].

Sorbets are a type of frozen dessert based on fruits, berries or vegetables with high water content and natural sugars. Unlike ice cream, they are free from milk proteins and fats, making them attractive to consumers adhering to principles of functional and healthy nutrition [9] or lactose intolerant [10]. The diversity of currently available functional food products also provides an opportunity for the use of

fruit-based sorbets as a carrier of functional ingredients [11]. Several studies are dealing with the protein enrichment of different products [12, 13], however there are only a few where sorbet are considered. Therefore, this study aims to develop a rice protein-enriched mangosteen-karonda mixed sorbet and evaluate its physicochemical properties, antioxidant activity and sensory acceptability.

MATERIALS AND METHODS

Preparation of fruits juice

The mangosteen and karonda used in this study was acquired from a local market in Chanthaburi Province, Thailand. The fruits were washed thrice, then manually dehusked (the rind was cut open). Afterward, a screw press conveyer (Zhengzhou Shuliy Machinery Co., Ltd., Model No. 008615093262873, China) was used to obtain a juice with a smooth texture [1].

Preparation of mangosteen-karonda mixed juice

Mangosteen-karonda mixed juice was prepared using the variation of karonda juice of 0, 5, 10, 15, and 20% w/w. The specific formulations for each variation can be found in Table 1.

Table 1 Different formulations of mangosteen- karonda mixed juice.

Formulas	Mangosteen juice (%w/w)	Karonda juice (%w/w)
T1	100	0
T2	95	5
T3	90	10
T4	85	15
T5	80	20

Subsequently, every mixture combined and subjected to batch pasteurization in a stainless-steel apparatus, wherein the juice's temperature was elevated to 70-80°C for a duration of 5 minutes. Following this, the heated juice was promptly filled into pasteurized bottles and securely sealed. To ensure the desired shelf life of the juice, the filled bottles underwent sterilization at a temperature of 100°C for a period of 15 minutes. Representative samples were then taken for physicochemical analysis, total phenolic content, antioxidant activity by using standard procedures, and sensory evaluation.

Physicochemical properties analysis

The color was examined using the color meter (Nippon Denshoku, ZE2000) and represented as L^* , a^* , and b^* . The pH level and total soluble solids (TSS) were assessed using pH meter (Extech PH100ExStik pH meter) and portable Otago Hand Refractometer (N1) with a range of 0-32%, respectively. The TSS was indicated as degrees Brix (°Brix) [14].

Total phenolic content

The quantification of total phenolic content was conducted using the Folin-Ciocalteu method. In this procedure, 0.2 mL of the juice sample, along with 0.2 mL of Folin-Ciocalteu reagent and 4 mL of 0.7 M Na_2CO_3 , were combined and mixed for a duration of 5 minutes. After 30 minutes, the absorbance was measured at 750 nm using a Thermo Scientific Orion AquaMate 7000 spectrophotometer, with gallic acid as the standard. The results were expressed as gallic acid equivalents ($\mu\text{g/mL}$) [15].

Antioxidant activity

The measurement of antioxidant capacity was conducted by employing 1,1-diphenyl-2-picrylhydrazyl (DPPH). To determine this, 15 μL of the juice was combined with 4 mL of DPPH dissolved in ethanol (100 mM). The absorbance was then determined at 515 nm using the Trolox reagent as a reference standard to express the Trolox equivalents per liter (TE)/L of the sample. The percentage of radical scavenging activity was calculated using the provided equation (1) [16].

$$\text{DPPH (\%)} = [(A_0 - A_1)/A_0] \times 100 \quad (1)$$

when A_0 = the absorbance of control reaction (containing all reagents except the sample).

A_1 = the absorbance of test compound.

Determination of vitamin C content

The vitamin C content was measured using the dichlorophenol indophenol titration as described previously [17].

Preparation of mangosteen-karonda juice sorbet enriched rice protein

Mixed fruit juice were supplemented with unflavored rice protein isolate (Wisamin, K.S. food product center Co., Ltd, Thailand) so that 0%, 5% and 10% w/w, respectively, without additional sugar and stabilizer. The sorbet processing method was modified from [18]. Briefly, the sorbets were prepared by first pasteurizing the mangosteen-karonda mixed juice at 65°C for 10 min. In the second step, the pasteurized mixed juice was blended with the rice protein and keep stirred in for another 10 min. Finally, sorbets were prepared by pour all the ingredients in a Gelato machine (OmniBlend V TM-800, Germany), with a speed of 5,000 rpm. at -25°C for 25 min. The sorbets were produced on the same day, then immediately stored in a blast freezer (SPEEDCRYO, Guangzhou Speed Refrigeration Equipment Co., Ltd, China) at -80°C for 45 min for hardening followed by storing in the freezer at -18°C before performing physicochemical and sensory analysis.

Determination of gross energy and protein content

The quantity of gross energy released from a burning of sample of mangosteen-karonda mixed juice enriched with rice protein was performed by using oxygen bomb calorimeter (Parr Instrument Company). The calories were measured using the in-house method TE-CH-169, which is based on the analysis method for nutrition labelling (1993). The protein content of samples was determined using the standard procedure outlined of the Central Laboratory Co. Ltd. Thailand [19].

Melting test

Melting rates were determined with a method modified from [18]. The sorbet sample, stored at -18°C , was taken out and placed on a stainless-steel mesh ($2\text{ mm} \times 2\text{ mm}$ square hole, 0.56 mm wire diameter) and a plastic plate at ambient temperature ($25^{\circ}\text{C} \pm 2^{\circ}\text{C}$) over 100 min. Then, the dripped of the sorbet was observed and recorded once every 20 min.

Structure analysis

The structure of mangosteen-karonda mixed juice sorbet samples was analyzed using a freeze dryer (Benchtop Freeze Dryer, Telstar, Spain) to remove the ice. About 50 g of each sample was placed into the chamber when temperature reached -35°C and the pressure was subsequently reduced to 4 mbar. After primary drying was finished, secondary drying began by slowly increasing the shelf temperature to $35\text{--}40^{\circ}\text{C}$. The mass and volume of the freeze-dried sorbet samples were then recorded following exposure to ambient room temperature ($25^{\circ}\text{C} \pm 2^{\circ}\text{C}$), during which the samples adsorbed moisture for over 100 min [19].

Sensory evaluation

For consumer acceptance, the sensory evaluation of all sorbet samples was conducted by 30 untrained panelists (students, instructors, staffs and ordinary citizens) from Rajamangala University of Technology Tawan-ok, Chanthaburi campus with an age of 20-70 years. For testing, the sorbet samples were randomized served at room temperature ($25^{\circ}\text{C} \pm 2^{\circ}\text{C}$) and the five or four parameters including color, odor, flavor, and overall acceptability were scored using a 9-point hedonic scale ranging from 1 (extremely dislike) to 9 (extremely like) [20].

Statistical analysis

All experimental treatments were done in three replications ($n = 3$). The data were analyzed with analysis of variance (ANOVA). Results are presented as means and standard deviations (SD). Differences between groups were assessed using Duncan's Multiple Range Test (DMRT) at a 95% confidence level ($p < 0.05$). Except, comparison of pigment assay using water versus ethanol was statistically analyzed by student's t-test.

RESULTS AND DISCUSSION

The physicochemical characteristics of the prepared mangosteen-karonda mixed juice were made analyzed using formulations containing 100% mangosteen juice and varying concentrations of karonda juice. Similar to other acidic fruit juices, mangosteen-karonda mixed juice exhibited high acidity with pH values ranging from 3.5 to 4.1. This level of acidity creates unfavorable conditions for the survival and growth of pathogenic bacteria [19]. Color analysis of the mixed juices revealed that increasing the concentration of karonda juice resulted in higher brightness (L^*) and redness (a^*) values, while yellowness (b^*) values decreased as shown in Table 2.

Table 2 Physiochemical properties of the mangosteen-karonda mixed juice.

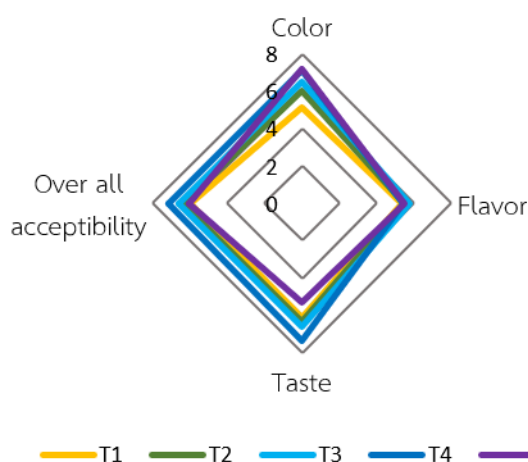
Formulas	Color			TSS ^{ns}	pH
	L^*	a^*	b^*		
T1	39.62 ^a	8.62 ^c	5.90 ^a	18	4.55 ^a
T2	38.90 ^a	9.31 ^{bc}	4.54 ^{ab}	17	4.00 ^b
T3	37.02 ^b	11.87 ^b	3.44 ^b	18	3.51 ^{bc}
T4	35.73 ^c	13.68 ^{ab}	3.28 ^{bc}	18	3.03 ^c
T5	34.86 ^c	15.20 ^a	2.80 ^c	18	3.00 ^c

The study examined the total phenolic content and antioxidant activity of mangosteen-karonda mixed juice. The results revealed that the formulation containing 20% (w/w) karonda juice exhibited the highest DPPH radical scavenging activity ($38.50 \pm 2.82\%$), whereas 100% mangosteen juice showed the lowest activity (Table 3). This observation was consistent with the levels of total phenolic compounds and vitamin C present in the mixed juices. These findings align

with previous study [2, 21], which initially reported the antioxidant properties of mangosteen juice [22]. Therefore, the biological activity of mangosteen-karonda mixed juice can be attributed to its total phenolic compounds, as well as the 1.5-2.0-fold increase in antioxidant activity and vitamin C content observed (Table 3). These bioactive components enhance free radical scavenging capacity, which may contribute to improve health benefits [23, 24].

Table 3 Physiochemical properties of the mangosteen-karonda mixed juice.

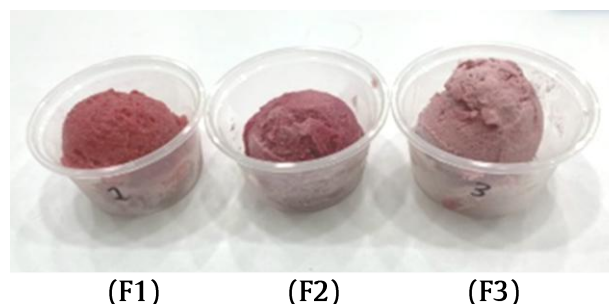
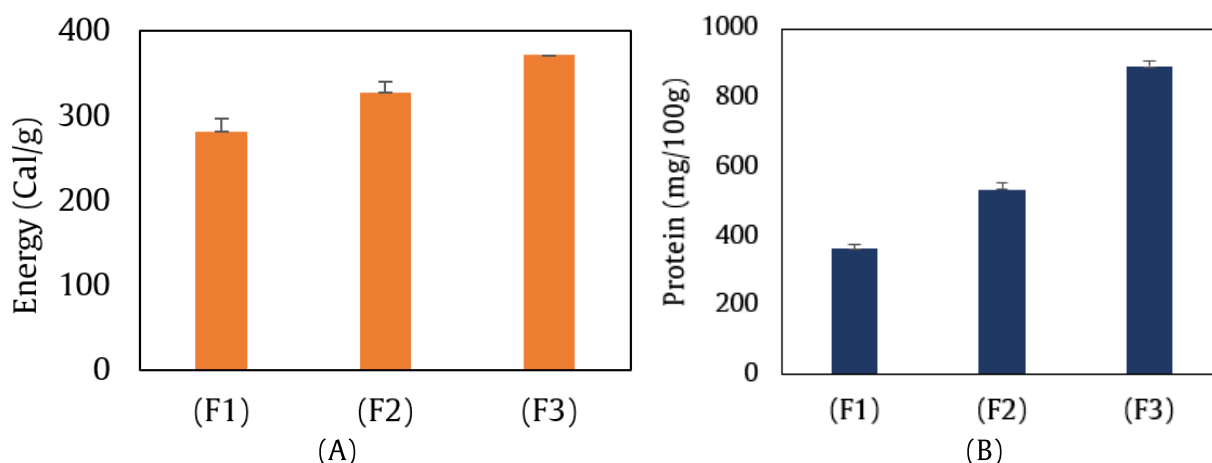
Formulas	Total phenolic Content (mg/ml)	DPPH radical Scavenging (%)	Vitamin C ($\mu\text{g/ml}$)
T1	13.40 $\pm$ 0.08 ^b	18.12 $\pm$ 3.21 ^c	24.50 $\pm$ 2.50 ^c
T2	10.50 $\pm$ 0.07 ^c	27.72 $\pm$ 2.50 ^{bc}	37.48 $\pm$ 2.50 ^{bc}
T3	11.10 $\pm$ 0.07 ^c	29.02 $\pm$ 1.50 ^b	39.22 $\pm$ 2.50 ^b
T4	12.80 $\pm$ 0.09 ^b	34.65 $\pm$ 3.50 ^{ab}	46.85 $\pm$ 2.50 ^{ab}
T5	15.00 $\pm$ 0.06 ^a	38.50 $\pm$ 2.82 ^a	52.05 $\pm$ 2.50 ^a

**Figure 1** Sensory scores of the mangosteen -karonda mixed juice; T1; 0%, T2; 5%, T3; 10%, T4; 15% and T5; 20% w/w, respectively.

The sensory evaluation of the mangosteen-karonda mixed juice is presented in Figure 1. In terms of color, the five formulations received scores ranging from 5.13 $\pm$ 0.60 to 7.15 $\pm$ 1.60, with T5 (7.15) achieving the highest score, though not statistically different from the T4 (7.12). For flavor, scores ranged from 5.35 $\pm$ 1.97 to 5.48 $\pm$ 1.67, with no statistically differences among the formulations. In terms of taste, the scores ranged from 6.12 $\pm$ 1.82 to 7.35 $\pm$ 1.23, with T4 (7.35) receiving the highest score, while T1, T2, and T3 were not statistically

different, and T5 (5.28) received the lowest score. For overall acceptability, the scores ranged from 5.77 $\pm$ 1.57 to 7.17 $\pm$ 1.40, with T4 (7.15) achieving the highest overall acceptability score, significantly higher than other formulations (Figure 1). The enhancement of red color and taste contributes to consumer acceptability, while the elevated antioxidant activity and vitamin C levels provide measurable health benefits. Therefore, T4 was selected for the development of mangosteen-karonda mixed juice sorbet enriched with rice protein, combining consumer acceptability with enhanced nutritional and functional properties.

Mangosteen-karonda mixed juice sorbets were enriched with rice protein that of 0% (F1), 5% (F2) and 10% (F3) w/w, respectively as shown in Figure 2.

**Figure 2** Mangosteen -karonda mixed juice sorbet enriched with rice protein; F1; 0%, F2; 5% and F3; 10% w/w, respectively.**Figure 3** Protein content (A) and energy content (B) of mangosteen-karonda mixed juice sorbet enriched with rice protein F1; 0%, F2; 5% and F3; 10% w/w, respectively.

The analysis of protein and energy content, as shown in Figure 3(A) and 3(B), revealed clear differences among the formulations. F3 exhibited the highest protein and energy values, followed by F2 and F1, respectively. This trend highlights the effectiveness of rice protein enrichment in enhancing the nutritional profile of the sorbet. The contribution of rice protein improved nutrition profiles and energy values has also been reported previously [25].

The ability of the mangosteen-karonda mixed juice sorbet enriched with rice protein variations to hold their shape during melting is presented in Figure 4. As the results, the F2 and F3 sample was completely melted for 80 min while F1 was 60 min, and its melting rate was lower than that of F1 (data not shown). The melting rate decreased when rice protein content was increased. Therefore, rice proteins stabilize structure and delay liquid separation from the sorbet matrix. Similar observations have been reported previously [26].

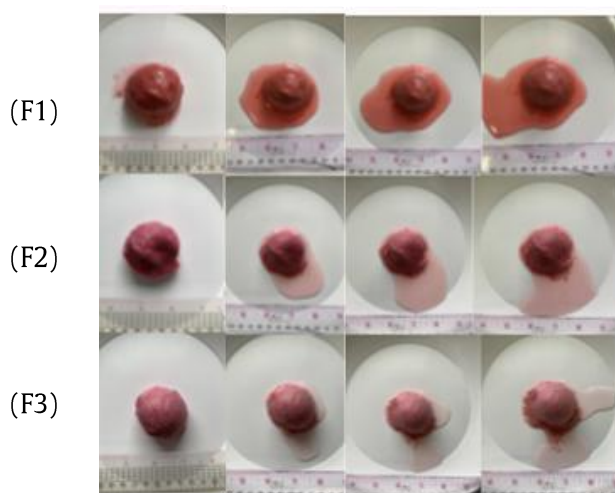


Figure 4 Melting behavior of mangosteen-karonda mixed juice sorbet enriched with rice protein; F1; 0%, F2; 5% and F3; 10% w/w, respectively.

Moreover, the addition of rice protein preparations influenced the physical stability of the mangosteen-karonda mixed juice sorbet, particularly its ability to retain shape following water removal via freeze-drying. It was observed that approximately 50% of the initial weight of all samples was removed. The volume of freeze-dried F2 and F3 were maintained, whereas the volume of freeze-dried F1 was reduced, as shown in Figure 5. Rice protein was explored as natural stabilizing agents in sorbet formulations, offering both functional and nutritional benefits. Unlike traditional hydrocolloids (e.g., guar gum, pectin), plant proteins can improve texture while simultaneously enriching the product with protein content. Milk proteins (caseins and whey protein) act as multifunctional stabilizers in ice cream, enhancing emulsion stability, air retention, viscosity, and ice crystal control [27, 28]. Similar to milk proteins, plant protein (e.g., pea

and soy protein isolates) confer structural stability to ice cream resulting in lower melting rates and improved shape retention was found and reported previously [18].



Figure 5 The freeze-dried product of the mangosteen-karonda mixed juice sorbet enriched with rice protein; F1; 0%, F2; 5% and F3; 10% w/w, respectively.

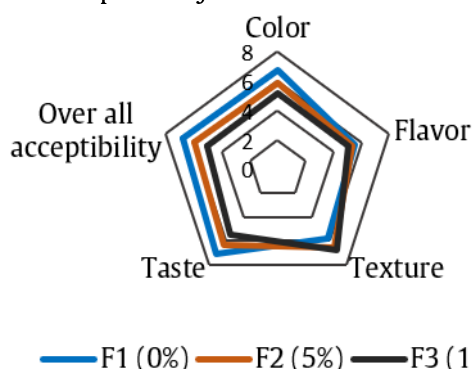


Figure 6 Sensory scores of the mangosteen-karonda mixed juice enriched with rice protein sorbet; F1; 0%, F2; 5% and F3; 10% w/w rice protein, respectively.

The sensory evaluation of the mangosteen-karonda mixed juice sorbet enriched with rice protein as shown in fig.6. In terms of color, the three formulations received scores ranging from 5.20 ± 0.98 to 6.75 ± 1.02 , with F1 receiving the highest score, followed by F2 and F3, respectively. For flavor, the scores ranged from 5.10 ± 1.40 to 5.50 ± 1.30 , with all formulators receiving not statistically different. In terms of taste, the scores ranged from 4.70 ± 1.05 to 7.15 ± 1.10 , with F1 received the highest overall score, while F3 was rated lowest. The overall acceptability decreased in the sensory analysis as rice protein content increased, primary due to reduced flavor and taste quality, which due to beany flavors and gritty mouthfeel. Similar findings were reported previously [29].

CONCLUSIONS

Mangosteen (*G. mangostana*) mixed karonda (*C. carandas*) sorbet enriched with rice protein was developed and characterized. The results demonstrated that the intensity of the DPPH radical scavenging activity and the vitamin C content increased proportionally with the concentrations of the karonda juice incorporated. This finding highlights the improving functional qualities of the product. The enrichment of rice protein

had explored as natural stabilizing agents in sorbet formulations, offering both functional and nutritional benefits which enriching the product with energy and protein content. Furthermore, the results of the 9-point hedonic scale data highlight that mangosteen-karonda mixed enriched with 5% and 10% (w/w) of rice protein is liked moderately by consumers for texture. However, sensory attributes such as flavor, aroma, and taste remain critical determinants of market success. Future work should therefore focus on optimizing these sensory qualities through techniques such as enzyme hydrolysis, flavor masking, or blending rice protein with other plant-based proteins. Overall, the incorporation of karonda juice and rice protein into mangosteen sorbet improved its nutritional and functional properties, while highlighting the importance of sensory optimization for consumer acceptance and commercial viability.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

This work has benefited from the use of AI-assisted technologies. ChatGPT was employed to support language refinement during the drafting process which was critically reviewed, fact-checked, and edited by the author, who assumes full responsibility for the accuracy, originality, and integrity of the final manuscript.

ACKNOWLEDGMENT

This work was supported by Rajamangala University of Technology Tawan-ok (RMUTTO). The authors would like to express their gratitude to the researchers from the Department of Food Innovation and Business for their valuable guidance and assistance in the meticulous production of the product.

REFERENCES

1. Aizat WM, Jamil IN, Ahmad-Hashim FH, Normah MN. Recent updates on metabolite composition and medicinal benefits of mangosteen plant. *Peer J* [internet]. 2019;7(1):e6324. Available from: <https://doi.org/10.7717/peerj.6324>.
2. Dorndi S, Greedumrongsak S, Monchaising J, Srivilai D, Suwanposri A, Boonna S, et al. Formulation, physiochemical and anti-bacterial properties of Mangosteen (*Garcinia mangostana*) juice. In *Proceedings of the 12th International Conference on Integration of Science and Technology for Sustainable Development*. 2024 Nov 29-Dec 1, Chanthaburi, Thailand. 2024. p. 1-8.
3. Khuanekaphan M, Khobjai W, Noysang C, Wisidsri N, Thungmungmee S. Bioactivities of Karonda (*Carissa carandas* Linn.) fruit extracts for novel cosmeceutical applications. *J Adv Pharm Technol Res*. 2024;12(2):162-8.
4. Bäuerle L, Kühn S. Development of a protein supplemented fruit smoothie using pea protein isolate as a plant-based protein alternative. *Future food*. 2022;5:1-7.
5. Xie A, Dong Y, Liu Z, Li Z, Shao J, Li M, et al. A review of plant-based drinks addressing nutrients, flavor, and processing technologies. *Foods*. 2023;12(21):3952.
6. Faber I, Castellanos-Feijoo NA, Van de Sompel L, Davydova A, Perez-Cueto FJA. Attitudes and knowledge towards plant-based diets of young adults across four European countries. *Exploratory survey Appetite*. 2020;145:104498.
7. Langyan S, Yadava P, Khan FN, Dar ZA, Renu Singh R, Kumar A. Sustaining Protein Nutrition Through Plant-Based Foods. *Front Nutr*. 2022; 8:772573.
8. Jayaprakash G, Bains A, Chawla P, Fogarasi M, Fogarasi S. A narrative review on rice proteins: Current scenario and food industrial application. *Polymers*. 2022;14(15):3003.
9. Williams J, McKune AJ, Naumovski N. Sorbets as functional food products, unexplored food matrices, their challenges, and advancements. *Applied Science*. 2023;13(21):11945.
10. Szilagyi A, Ishayek N. Lactose intolerance, dairy avoidance, and treatment options. *Nutrients*. 2018;10:E1994.
11. Williams T, Kumar M, Singh R. Fruit-Based Sorbets as Functional Foods: A Review of Trends and Health Aspects. *Journal of Functional Foods*. 2023;102:105300.
12. Mohsen SM. Effect of substitution of soy protein isolate on aroma volatiles, chemical composition and sensory quality of wheat cookies. *Int J Food Sci Technol*. 2009;44:1705-12.
13. Brückner-Gühmann M, Vasil'eva E, Culetu A, Duta D, Sozer N, Drusch S. Oat protein concentrate as alternative ingredient for non-dairy yoghurt-type product. *J Sci Food Agric*. 2019;13:5852-7.
14. Boonna S, Phuangborisut S, Chanawanno T. Physical properties and nutritional value of durian juice produced from durian chips by-products. *International Journal of Agricultural Technology*. 2022;18(4):1837-46.
15. Mongkontanawat N, Boonna S, Wasikadilok N. Probiotic beverage from mangosteen juice fermented with *Lactobacillus* strains. *Trends in Sciences*. 2022;9(21):6305.

16. Suwanposri A, Khamphinit W, Sonthichai P, Maliyam R, Charoensuk K, Boonna S. Production of herbal kombucha with agarwood (*Aquilaria crassna* Pierre ex Lec.) leaves and its biological properties. *International Journal of Agricultural Technology*. 2025;21(3):1177-96.
17. Popova A. Comparison of Vitamin C Content of Commercially Available Fresh Fruits. *Asian Food Sci J*. 2019;13:1-6.
18. Ng CKZ, Leng WQ, Lim CH, Du J. Physicochemical property characterization, amino acid profiling, and sensory evaluation of plant-based ice cream incorporated with soy, pea, and milk proteins. *J Dairy Sci*. 2024;107:10268-79.
19. Nakagawa K, Kamisaki H, Suzuki T, Sano N. Model-based prediction of the moisture sorption kinetics and humidity-induced collapse for freeze-dried cakes. *Chemical Engineering Science*. 2022;248:117129.
20. Charoensuk K, Wongekalak L, Mongkontanawat N, Nonmoung W, Suwanposri A, Tanmanee P, et al. Formulation, sensory and pulp stability of Durian (*Durio zibethinus* Murr) juice. *International Journal of Agricultural Technology*. 2018;14(2): 215-24.
21. Atribst AAL, Sant'Ana SS, Massaguer PR. Review: Microbiological quality and safety of fruit juices past, present and future perspectives *Microbiology of fruit juices*. *Crit Rev Microbiol*. 2009;35(4):310-39.
22. Sukma M, Ngawhirunpat T, Opanasopit P, Siripong P. Antioxidant properties of squeezed mangosteen juice. *Thai Journal of Pharmaceutical Sciences*. 2011;35:176-87.
23. Lim YS, Lee SSH, Tan BC. Antioxidant capacity antibacterial activity of different parts of mangosteen (*Garcinia mangostana* Linn.) extract. *Fruits*. 2013; 68(6):83-9.
24. Busaid SAL. Phytochemical characterization and antioxidant potential of Karonda (*Carissa carandas*): Unlocking a neglected fruit for nutraceutical chemistry *International Journal of Advanced Chemistry Research*. 2025;7(7):51-6.
25. Tangjaidee P, Xiang J, Yin H, Wen X, Quek SY. Selenium, fiber, and protein enrichment of rice product: extrusion variables and product properties. *Food Quality and Safety*. 2019;3:40-51.
26. Khillari SA, Zanjad PN, Rathod KS, Raziuddin M. Effect of incorporation of whey protein concentrate as MSNF substitute on physico-chemical and sensory quality of low- fat ice cream. *Indian J Dairy Biosci*. 2005;16:26-9.
27. Goff HD. Formation and stabilization of structure in ice cream and related products. *Trends in Food Science and Technology*. 2002;13(3):60-8.
28. Sofjan RP, Hartel RW. Effects of overrun and storage on the microstructure of ice cream. *International Dairy Journal* 2004;14(7):655-62.
29. Kiełczewska K, Smoczyński M, Gutkowska M. The use of high-protein preparations in ice cream production. *Foods*. 2025;14(3):345.



Effective detection of Thai fake news using machine learning method

Nopphadol Sumleeroung¹, Komsan Kanjanasit² and Nithizethe Mhuadthongon^{1*}

¹Information and Communication Technology, School of Science and Technology, Sukhothai Thammathirat Open University, Nonthaburi 11120, THAILAND

²College of Computing, Prince of Songkla University, Phuket Campus, Phuket 83120, THAILAND

*Corresponding author: Nithizethe.Mhu@stou.ac.th

ABSTRACT

The rapid spread of misinformation on digital platforms has become a critical challenge affecting public trust, social stability, and policy communication in Thailand. The aim of this study is to develop and evaluate a reproducible machine learning framework for Thai fake news detection from real world verified data from the Anti-Fake News Center Thailand during 2019 - 2024. The data set comprises four key data groups including governmental policies, health products, financial and stocks as well as disaster news. The suggested research methodology entails thorough data cleaning and text normalization, sentiment and emotions analysis through the usage of AI for Thai, dimension reduction by selecting relevant features and dealing with unbalanced classes through Synthetic Minority Oversampling Technique (SMOTE). For evaluation, three supervised learning algorithms (Naïve Bayes, Support Vector Machine [SVM] and neural networks) were tested using 80:20 and 90:10 train-test split ratios, K-Fold Cross-Validation and hyperparameter optimization. Based on the experimental findings, the neural network classifier achieved better accuracy rates, that is, 92.39%, with other performance metrics including precision, recall, F1-score, ROC-AUC and AUC-PR at 92.64%, 92.39%, 92.38%, 96.79%, 96.07% respectively. Therefore, the methodological framework was effective in identifying false news stories in the Thai language due to the use of linguistic features and oversampling strategies. The main contribution of this study is an interpretable baseline framework that integrates linguistic, affective, and metadata features for Thai fake news detection in a low-resource language setting.

Keywords: Fake news, Machine learning, Sentiment and emotion analysis

INTRODUCTION

The tremendous growth of internet-based platforms for information gathering has resulted in a huge shift in how information is produced and shared. These sources have been effective in improving the availability of information, yet they have at the same time facilitated the circulation of misinformation that is not backed up by evidence. In today's world, misinformation spreads rapidly due to algorithms designed to promote sensationalist or emotional information [1]. This means that besides being a communications issue, fake news has become a growing menace in society that causes misunderstanding, distrust of institutions, economic harm, and even risks to health and safety.

Fake news is not a recent phenomenon. However, due to the digitization and rapid mobile penetration, the impact of fake news has become significant in society. The findings of previous natural language processing (NLP) and machine learning (ML) studies show promise in detecting fake news automatically using linguistic, semantic,

and statistical textual features [2-4]. As early as 2015, the main approach to detect hate speech from online data sources was traditional text classification-based approaches. However, recent work has utilized advanced machine-learning and deep-learning architectures to enhance detection accuracy and robustness [5, 6]. Yet, there remains a limitation. Most of the existing in-house LLMs are developed and tested on English datasets, thus limiting their usability in non-English and low-resource language settings.

The phenomenon of fake news also poses problems for Thailand since people are now more inclined to access information through social media and messaging apps. Issues related to public health, economic policy, and disaster mitigation are among those being affected. Unfounded claims that involve government activities, dubious health products, scams, and others are frequently shared on popular platforms and sometimes lead to public hysteria, thus reducing public trust in communication from the government side. To address this problem, the Thailand Anti-Fake News Center was established.

The center is responsible for monitoring, verifying, and correcting fake news. As part of its operations, a systematic Thai-language verification dataset has been created.

There are some issues associated with fake news detection within the Thai language. Word boundaries are not easily distinguishable, and word order is quite flexible in this language; moreover, Thai is a highly contextual language. This adds up to the complexity of its automatic processing and feature extraction [7]. Segmentation mistakes can cause errors at the stage of feature extraction, leading to poor feature representation and performance. Moreover, this will add an extra layer of difficulty during feature engineering, which may potentially affect the model's performance and capacity to generalize knowledge. Another issue worth mentioning is the lack of annotated Thai corpora that consist of a substantial number of instances, and there is currently no benchmark or framework for evaluation.

Despite advancements in fake news detection in popular languages, the field of fake news detection in the Thai language still requires much attention and remains fragmented. A number of solutions depend on the task of text classification and are created specifically for small datasets, which does not fit well with the requirements of an online environment in Thailand. At the same time, there is no use of features related to the sentiments and emotions involved in posts by just a few researchers, even though the importance of the use of the content with affective properties was proven long ago. There have also been no systematic efforts to integrate national natural language processing tools like PyThaiNLP and AI for Thai into a unified model for fake news detection.

The suggested work introduces a machine learning approach to detecting fake news in Thai language, combining several kinds of features' representation with structural and contextual awareness. A fully developed and validated Thai dataset of fake news is used in the experiments; this database was created using data obtained from the Thailand Anti-Fake News Center. This database stands out because it is one of the first to consider public policy topics along with common ones. In addition to the database, a universal data processing pipeline is introduced and includes creating TF-IDF representations as well as integrating NLP models specific to Thai, for sentiment and emotion analysis, and contextual features like publication features. Moreover, a series of experiments on comparing the performance of Naïve Bayes, SVM, and Neural Network algorithms in various conditions, including feature selection and data balance.

The objective of this study is to develop and evaluate the machine learning framework for the Thai fake news detection based on linguistics, affective,

and metadata features extracted from the verified news data provided by the Anti-Fake News Center Thailand. In particular, the study compares the performance of Naïve Bayes, Support Vector Machine, and Neural Network classifiers under different train-test split configurations and examines the contribution of feature selection and class balancing to classification performance [8].

MATERIALS AND METHODS

Conceptual framework

The present research study design was constructed in such a way that various feature domains can be united into one modeling pipeline for fake news detection among Thai texts using an integrative approach. According to Figure 1, the framework used in the study can be described in three phases (i.e., Data Preparation, Processing, and Evaluation).

First of all, during the Data Preparation phase, textual data was obtained from a reliable source to assure the high quality of the data and its relevance to the real world. Noise reduction and data cleaning took place to enhance the overall quality. Moreover, affective features were also incorporated into the data set as feature representations. The structured data can be used for further processing and analysis.

The main objective of the second phase (i.e., Processing) is to transform the already structured data into feature representations suitable for classification purposes. During the initial stage of data processing, standardization of the data took place. Furthermore, after feature selection, the problem of class imbalance in the data set was solved by using oversampling methods. Thus, the framework was capable of facilitating the work of various classifiers.

The performance of the model was evaluated in the Evaluation stage with the validation strategy. The evaluation process involved splitting the data into training and testing sets, performing K-fold cross-validation, and conducting hyperparameter tuning. The system used evaluation criteria including accuracy, precision, recall, and F1-score to evaluate the model. The evaluation outputs of this stage were produced according to the model results in the framework.

Dataset

The dataset employed for this study was obtained from open data provided by the Anti-Fake News Centre Thailand (AFNC), which contained confirmed misinformation cases that were gathered by authoritative organizations based in Thailand [9]. The data retrieved contain news that has been published between the year 2019 and 2024. All these news items were individually labelled for their verification status, which is either real or fake. This allowed the use of supervised learning methodologies for carrying out the classification task.

The datasets covered various fields of public interest and were further divided into four main categories, namely, government policy and public

affairs, health products, finance and stock, and disaster-related information, as summarized in the Table 1.

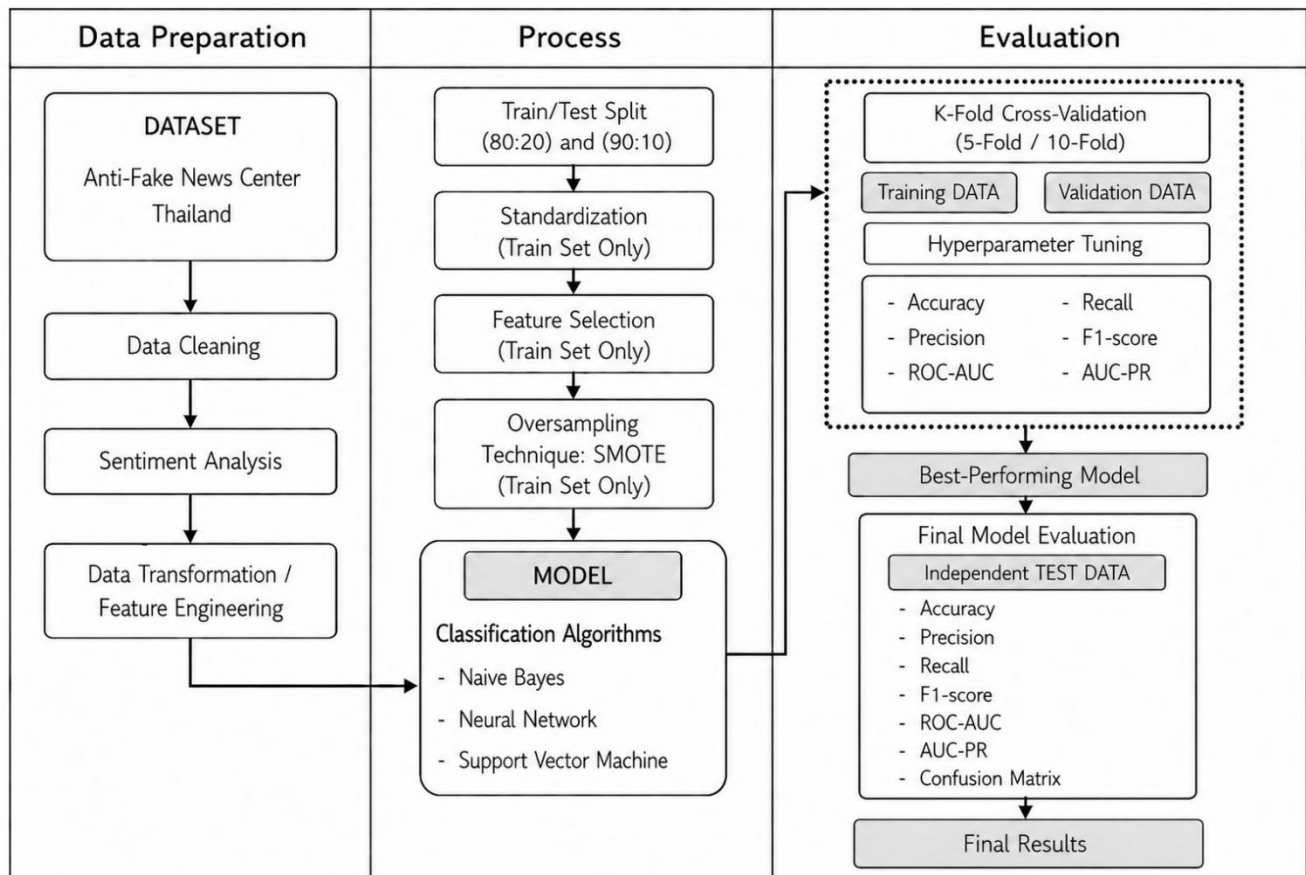


Figure 1 Proposed research framework.

Table 1 Number of news items by category.

Category	Number of Records
Government Policy and Public Affairs	2,737
Health Products	2,636
Finance and Stock	1,385
Disaster	256
Total	7,014

To ensure the quality of the data and reliability of the analyses, the datasets were initially screened to remove any incomplete, inconsistent, or irrelevant records. Moreover, any category with too few cases was excluded to avoid bias and instability in model training and assessment. Through filtering, a total of 7,014 news items were selected for model development.

Along with the textual data, the dataset also contains other attributes such as the date-time of publication, news URL, news headline, full news content, verifying agency, class label (real/fake), news category, and news hashtags, as given in Table 2 [9].

Data preprocessing and feature generation

The acquired dataset underwent data preprocessing and feature generation techniques to enhance its quality and suitability for subsequent analysis. First, the original text data underwent cleaning procedures through removal of unnecessary characters, solving any inconsistencies in the text format, and eliminating noise from the online news articles. Considering that the Thai language does not have distinct word boundaries, word segmentation was performed on the text by the PyThaiNLP library [10].

In addition, to incorporate semantic and affective aspects of the text, sentiment analysis was conducted using the AI for Thai web service through the use of its publicly available API [11, 12]. The output from the API included multiple variables such as sentiment polarity, sentiment score, and number of positive and negative words, which were employed as input features.

Sentiment analysis was followed by the conversion of the textual data to numerical form through the implementation of the Term Frequency-Inverse Document Frequency (TF-IDF) technique [13-15].

Furthermore, the sentiment values and text features were converted to structured numeric inputs ready for use in machine learning algorithms. Specifically, the sentiment values were mapped to three possible classes as follows: -1 for negative class, 0 for neutral class, and +1 for positive class. The number of positive

words and the number of negative words were captured as numeric values, while token-based features were summarized using numeric representation. A comprehensive list of the engineered features used in this work is described in Table 3.

Table 2 Information extracted from the Thailand Anti-Fake News Center website.

No	Extracted Data	Description
1	Publication Date and Time	The date and time when the news item was created or published
2	News URL	The web link where the news item is published on the Thailand Anti-Fake News Center website
3	Headline	The title of the news article
4	News Content	The full textual content of the news article
5	Verifying Agency	The organization responsible for verifying the news item
6	News Classification	The classification of the news item (real or fake)
7	News Category	The assigned category of the news item
8	Hashtag	Hashtags associated with the news item

Table 3 Description of engineered features used in model development.

No.	Attribute	Description	Data Type	Values / Definition
1	create_day	Day of publication	Polynomial	Sun = 1, Mon = 2, Tue = 3, Wed = 4, Thu = 5, Friday = 6, Sat = 7
2	create_time	Time period of publication	Polynomial	Midnight = 1, Morning = 2, Afternoon = 3, Evening = 4
3	isVerified	Verification status	Binomial	Verified by = 1, Not verified by agency = 2
4	Label	News classification	Binomial	Real news = 1, Fake news = 0
5	Category	News category	Polynomial	Government policy and public affairs = 1, Health products = 2, Finance and stock = 3, Disaster = 4
6	has_hashtag	Presence of hashtag	Binomial	Present = 1, Not present = 2
7	count_hashtags	Number of hashtags	Numeric	Integer ≥ 0
8	Sentiment	Sentiment polarity	Polynomial	Positive = 1, Neutral = 0, Negative = -1
9	sentiment_score	Sentiment score	Numeric	Range: 0-100
10	count_pos	Number of positive	Numeric	Integer ≥ 0
11	count_neg	Number of negative	Numeric	Integer ≥ 0
12	content_length	News content length	Numeric	Integer ≥ 0

Train-test data splitting

To evaluate the performance of the model, the dataset was divided into two subsets a training set and a test set. The training set was used to fit the model, while the test set was kept separate and used only for final evaluation. This approach helps ensure that the model is assessed on data it has never seen before, providing a more reliable estimate of its real-world performance.

The split was performed carefully to avoid data leakage between training and testing. In cases where the dataset was imbalanced, resampling techniques was applied only to the training set, while the test set remained unchanged. This allowed the

evaluation results to reflect the model's true ability to generalize to previously unseen data

Standardization and feature selection

After constructing the features for the dataset, feature scaling and feature selection processes were employed for consistency in measurements within the features and for stable training of models.

Among some features in the dataset were the count_hashtags, sentiment_score, count_pos, count_neg, and content_length features which were all numerical and whose units were on various scales and ranges. This can be seen in Table 4 below, showing the need for data standardization to avoid disparity in scales among the features.

Table 4 Statistical summary of numerical features before standardization.

Attribute	Max	Min	Mean	SD.
count_hashtags	10	0	3.16	1.03
sentiment_score	99.98	0	34.65	36.78
count_pos	5	0	0.24	0.50
count_neg	9	0	0.54	0.91
content_length	44	3	13.44	4.77

Table 5 Statistical summary of numerical features after standardization.

Attribute	Max	Min	Mean	SD.
count_hashtags	6.67	-3.09	0	1
sentiment_score	1.78	-9.42	0	1
count_pos	9.54	-4.76	0	1
count_neg	9.30	-5.89	0	1
content_length	6.40	-2.19	0	1

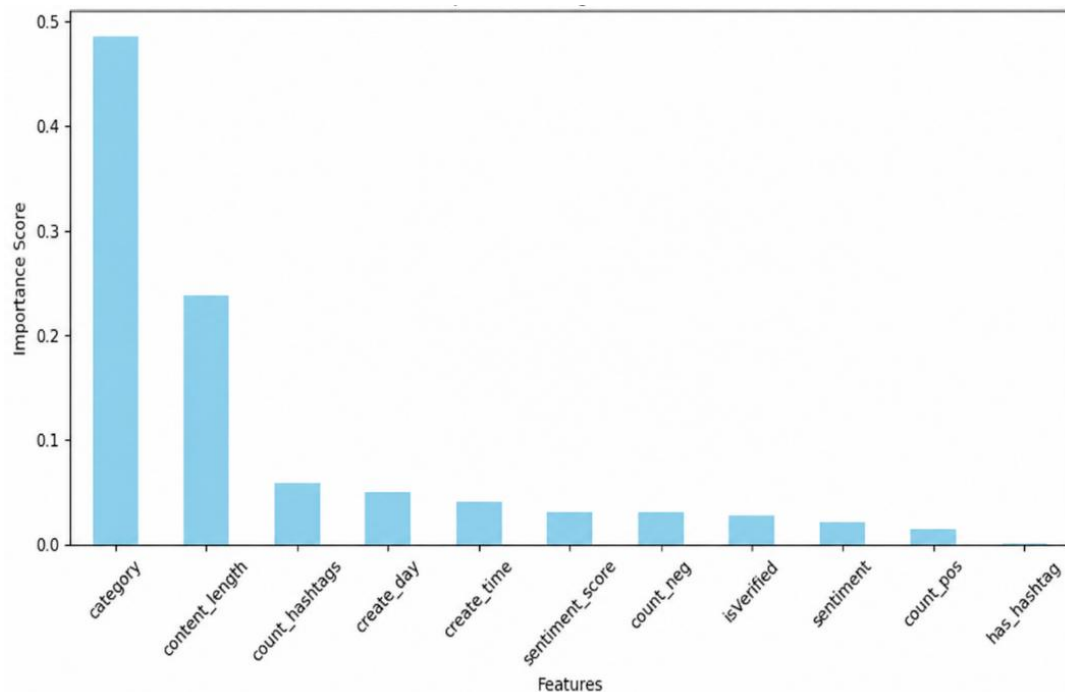


Figure 2 Feature importance scores obtained from the extra tree's classifier.

The StandardScaler technique was applied to transform the numerical features into a common scale with zero mean and unit variance [16]. This normalization process was applied to standardize feature distributions prior to model training. The transformed feature distributions after standardization are presented in Table 5.

Apart from the scaling, feature selection was done to select the most useful features and minimize redundancy among the features. Feature importance was calculated by the Extra Trees Classifier that calculates the feature's importance from impurity decrease while building trees [16].

The feature importance values obtained were then used to rank the importance of the features, as shown in Figure 2.

Moreover, correlation analysis was applied to analyze the relationship between the features and the target variable through a correlation matrix. Correlation analysis was applied to study the relationship between the features and the target variable, as indicated in Figure 3.

In addition to the above, Backward Feature Elimination technique was applied to determine the significance of each feature while training the model [17]. This procedure involves a systematic elimination of features in order to study their impact on the model. The outcomes of this analysis are outlined in Table 6. From the above processes, the features were selected for model development. Comparative analysis of feature selection techniques is presented in Table 7.

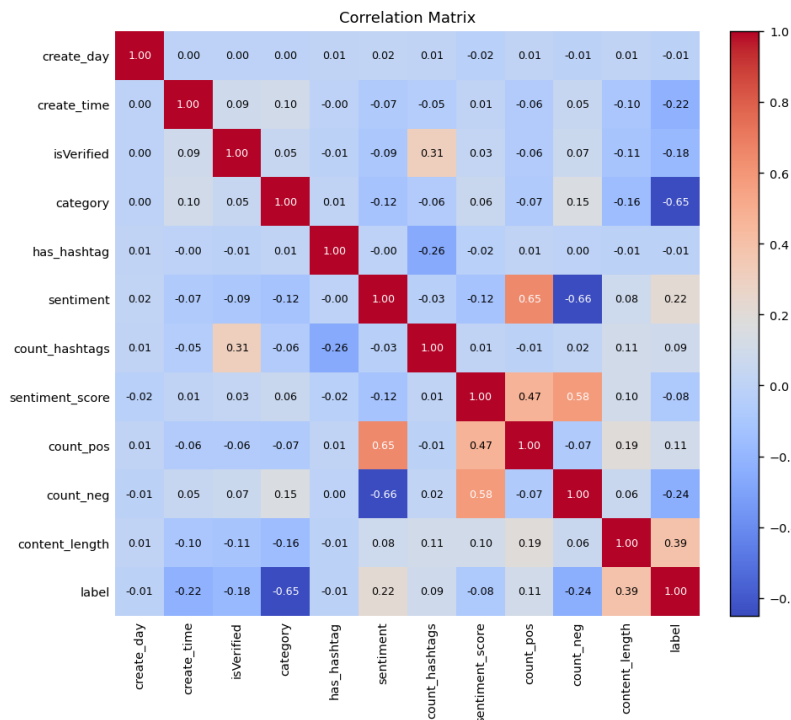


Figure 3 Correlation matrix showing relationships between features and the target variable.

Table 6 Impact of backward feature elimination on model accuracy.

Feature	NN	NB	SVM	Comparison with Baseline			
				NN	NB	SVM	Summary
Baseline Model	0.8978	0.8579	0.8308	-	-	-	-
Drop create_day	0.8978	0.8574	0.8308	-	Decreased	-	Decreased
Drop create_time	0.8971	0.8487	0.8308	Decreased	Decreased	-	Decreased
Drop isVerified	0.8962	0.8566	0.8308	Decreased	Decreased	-	Decreased
Drop category	0.8359	0.8111	0.8308	Decreased	Decreased	-	Decreased
Drop has_hashtag	0.8980	0.8509	0.8308	Increased	Decreased	-	-
Drop sentiment	0.9003	0.8637	0.8308	Increased	Increased	-	Increased
Drop count_hashtags	0.8942	0.8514	0.8308	Decreased	Decreased	-	Decreased
Drop sentiment_score	0.9019	0.8657	0.8948	Increased	Increased	Increased	Increased
Drop count_pos	0.8956	0.8603	0.8308	Decreased	Increased	-	-
Drop count_neg	0.8909	0.8774	0.8308	Decreased	Increased	-	-
Drop content_length	0.8613	0.8065	0.8308	Decreased	Decreased	-	Decreased

Table 7 Comparative results of feature selection methods.

Feature	Rank (Feature Importance)	Rank (Correlation Matrix)	Backward Feature Elimination	Final Selection
create_day	4	9	✓	✓
create_time	5	4	✓	✓
isVerified	8	5	✓	✓
Category	1	1	✓	✓
has_hashtag	11	9	-	✗
Sentiment	9	4	✗	✓
count_hashtags	3	7	✓	✓
sentiment_score	6	8	✗	✓

Feature	Rank (Feature Importance)	Rank (Correlation Matrix)	Backward Feature Elimination	Final Selection
count_pos	10	6	-	✓
count_neg	7	3	-	✓
content_length	2	2	✓	✓

Class imbalance

SMOTE was applied to the dataset to resolve the class imbalance issue in the dataset. In real-world fake news datasets, it is common for class distributions to be imbalanced, where one class can be much larger than the other.

The SMOTE was used to create synthetic observations of the minority class by interpolation of feature space observations [18]. The method was utilized to enhance the presence of the minority class without duplicating of existing representations.

Before developing the model for Thai fake news detection, the training data were resampled. Table 8 presents the distribution of the classes before and after using SMOTE.

Table 8 Comparison of training dataset distribution before and after SMOTE-based oversampling.

Split	Class	Before Over-sampling	After Over-sampling
80:20	Real	950	4,661
	Fake	4,661	4,661
	news		
90:10	Real	1,068	5,244
	Fake	5,244	5,244
	news		

This procedure was applied prior to model training to address class imbalance in the training dataset. SMOTE was applied only to the training set, while the test dataset was left unchanged.

Classification algorithms

The following machine learning methods were selected for implementation of the above-mentioned detection approach: Naïve Bayes (NB), Support Vector Machine (SVM), and Neural Networks (NN). Each algorithm was used to present alternative approaches to classifying data. All algorithms were trained based on the feature vector extracted from textual content and sentiment-emotion attributes.

Naïve Bayes (NB) is a probabilistic model based on Bayes' theorem with an additional conditional independence assumption of attributes [19]. In this study, Gaussian Naïve Bayes was used as the classification model.

Support Vector Machine (SVM) is a discriminative classifier that aims at choosing the best possible hyperplane by finding a maximal margin between

classes [20]. Thus, some transformations, including kernels, are needed for better separation of nonlinear patterns. In this study, the SVM model was configured with an RBF kernel, $C = 100$, and $\gamma = \text{auto}$, with probability estimation enabled.

Neural Network (NN) is a feedforward network model, which helps to find a nonlinear relation between feature vectors. Parameters of NN are optimized iteratively based on backpropagation [21, 22]. In this study, the NN model was implemented as a multilayer perceptron with two hidden layers containing 128 and 64 neurons, respectively. The tanh activation function and the Adam optimizer were used, and the model was trained for a maximum of 1,000 iterations.

Experimental setup and model evaluation

To evaluate the proposed Thai fake news detection framework, a systematic experimental setup was established. The dataset was partitioned using two train-test split strategies, namely 80:20 and 90:10. We also used K-fold cross-validation with $k = 5$ to check the model [23].

Table 9 Model configuration and hyperparameters.

Model	Key Parameters
NB	Gaussian Naïve Bayes
SVM	RBF kernel, $C = 100$, gamma = auto, probability = True
NN	Hidden layers = 128 and 64, tanh activation, Adam optimizer, max_iter = 1,000

Table 10 Experimental configuration.

Component	Description
Train-Test Split	80:20, 90:10
Cross-Validation	5-Fold, 10-Fold
Evaluation Metrics	Accuracy, Precision, Recall, F1-score, ROC-AUC, AUC-PR
Oversampling Method	SMOTE
Tuning Method	Grid Search

We changed the hyperparameters for each model that did the classification. A grid search method was used to carefully look at different combinations of important parameters. Table 9 gives a brief overview of this. We used four standard evaluation metrics

accuracy, precision, recall, and F1-score to see how well the model worked. All of the experiments were done in a controlled computer environment.

All experiments were conducted within a controlled computational setting. The overall experimental configuration, including data splitting

strategies, validation methods, evaluation metrics, and tuning procedures, is summarized in Table 10.

The complete workflow of preprocessing, feature construction, model training, and evaluation is illustrated in Figure 4.

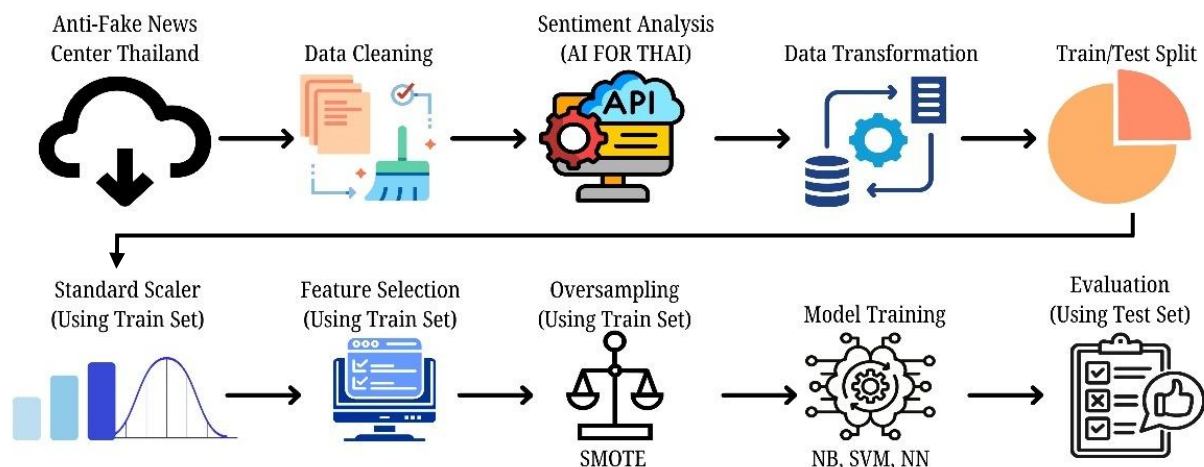


Figure 4 System architecture.

RESULTS AND DISCUSSION

Results

Neural Network (NN) model showed the highest values for all the evaluation parameters according to the experimental results. With the 90:10 train-test configuration, NN achieved an accuracy of 92.39%, precision of 92.64%, recall of 92.39%, F1-score of 92.38%, ROC-AUC of 96.79%, and AUC-PR of 96.07%, respectively (Table 12). According to Table 11, with the 80:20 split, the NN had 92.38% accuracy and 96.77% ROC-AUC, while SVM and NB had 92.09% and 87.31% accuracy, and 95.63% and 92.88% ROC-AUC, respectively. In terms of AUC-PR, the NN also achieved the highest value at 96.05%, followed by SVM at 93.46% and NB at 91.74%.

On the various train-test configurations, the NN model generated the highest values for all evaluation metrics, SVM generated the second highest

value, and NB generated the lowest value among the three models. The AUC-PR results further confirm that NN provided the best precision-recall balance in both experimental settings.

As indicated by the information in Table 11, the three models were evaluated using a 80:20 train-test configuration. NB achieved 87.31% accuracy and 92.88% ROC-AUC, while SVM and NN achieved 92.09% and 92.38% accuracy, and 95.63% and 96.77% ROC-AUC, respectively. All models exhibited uniformity in terms of precision, recall, F1-score, ROC-AUC, and AUC-PR.

The outcomes of the 90:10 train-test configuration, in other words, are summarized in Table 12. The NN model achieved 92.39% accuracy and 96.79% ROC-AUC, SVM followed with 92.01% accuracy and 95.71% ROC-AUC, and NB achieved 87.28% accuracy and 92.86% ROC-AUC, respectively. The precision, recall, F1-scores, ROC-AUC, and AUC-PR values are consistently distributed across the three models.

Table 11 Model performance on the independent test set using a 80:20 train-test split.

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	AUC-PR
NB	87.31	88.26	87.31	87.23	92.88	91.74
SVM	92.09	92.63	92.09	92.06	95.63	93.46
NN	92.38	92.64	92.38	92.37	96.77	96.05

Table 12 Model performance on the independent test set using a 90:10 train-test split.

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	AUC-PR
NB	87.28	88.24	87.28	87.20	92.86	91.74
SVM	92.01	92.55	92.01	91.98	95.71	93.67
NN	92.39	92.64	92.39	92.38	96.79	96.07

In this study, the Real News class was defined as the positive class for performance evaluation. Tables 13 and 14 present the confusion matrices of the NN model under the 80:20 and 90:10 train-test split configuration on the testing dataset. In both configurations, the numbers of True Positives and True Negatives were higher than those of False Positives and False Negatives.

Table 13 Confusion matrix of the best-performing neural network model using a 80:20 train-test split configuration on the testing dataset.

	Predicted Real	Predicted Fake
Real	209 (TP)	28 (FN)
Fake news	79 (FP)	1,087 (TN)

Table 14 Confusion matrix of the best-performing neural network model using a 90:10 train-test split configuration on the testing dataset.

	Predicted Real	Predicted Fake
Real	105 (TP)	14 (FN)
Fake news	39 (FP)	544 (TN)

However, although high classification performance was achieved by the NN model, some false positives and false negatives were still observed in the confusion matrices. False positives can be produced when fake news articles are written in a neutral, formal, or official style of language, and therefore linguistically closer to verified news. On the other hand, real news articles that contain emotional expressions, warnings and sensational wording, particularly in health- or disaster-related content, can increase false negatives. Moreover, the short content length and the Thai word segmentation errors might deteriorate the quality of the feature representation that can influence model decisions in some cases.

Discussion

Based on the results of this paper, it can be stated that non-linear analysis of the interaction between features is one of the factors in fake news detection in Thai language. The NN model was better than NB and SVM for both train-test splits; thus, the model was able to discover the correlation between TF-IDF, sentimental and metadata features. This is consistent with existing literature, which proves the efficiency of using neural networks for text classification in high-dimensional spaces [4, 21, 22].

The inclusion of affective and emotional features positively impacted classification performance by detecting affective features within the content of fake news. Therefore, it can be argued that affective features increase the separability between fake and real news as stated before in other works [2, 11].

The use of SMOTE as the balancing technique increased classification stability by improving the

representation of minority samples. This result is consistent with previous findings proving the benefits of oversampling [15].

Conclusions derived from feature selection include category, content length, time-based variables, and sentiment score as significant features affecting classification decisions. The removal of less relevant features does not impair accuracy, demonstrating that fewer but informative features will yield good predictive performance.

This conclusion follows several other investigations reporting the capacity of ensemble tree-based learning to detect significant predictors in highly dimensional feature spaces [24, 25].

The conclusions drawn from the confusion matrix output also reveal that the NN classifier performs equally well in predicting classes in either of the two categories. Nonetheless, the presence of a few classification errors implies the existence of a trade-off between precision and recall that ought to be modified according to the demands of a particular task.

The affective features involved in the classification process included sentiment_score, count_pos and count_neg, which resulted in capturing emotional cues in news content. Fake news often employs persuasive, emotional or attention-grabbing phrases, whereas verified news is typically written in neutral and factual language. Thus, these affective features helped to increase the separability between real and fake news, especially when combined with linguistic and metadata features.

Nevertheless, the analysis is constrained to a small number of machine-learning models and a single dataset. This study focused on an interpretable conventional machine-learning baseline. therefore, transformer-based models such as WangchanBERTa and mBERT were not included in the current experiments. Future research should compare the proposed framework with these models to evaluate technical competitiveness and potential performance improvement.

CONCLUSIONS

This research paper has presented a machine learning system aimed at identifying Thai-language fake-news using multiple approaches such as content, sentiment, and context analysis. The results of the study indicate that a neural network system has demonstrated the best classification performance in terms of the test results with the maximum accuracy score of 92.39% in the 90:10 setting.

The importance of the use of both content and sentiment features in Thai-language settings cannot be understated in terms of improving the accuracy rate. Specifically, the inclusion of affective attributes leads to higher class separation in terms

of fake vs. real news while SMOTE helps to boost the stability of the classifier.

This work has delivered one of the early structured frameworks used for detecting misinformation in low-resource languages. This research delivers an initial structured framework for creating a model capable of identifying Thai-language misinformation. As far as the practical implications of the study are concerned, the proposed approach allows for the creation of monitoring systems for fake news detection purposes.

REFERENCES

- Vosoughi S, Roy D, Aral S. The spread of true and false news online. *Science*. 2018;359(6380): 1146-51.
- Shu K, Sliva A, Wang S, Tang J, Liu H. Fake News Detection on Social Media. *ACM SIGKDD Explorations Newsletter*. 2017;19(1):22-36.
- Conroy NJ, Rubin VL, Chen Y. Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*. 2015;52(1):1-4.
- Jurafsky D, Martin JH. *Speech and Language Processing*. 3rd ed. Hoboken (NJ): Pearson; 2020.
- Zhou X, Zafarani R. A Survey of Fake News. *ACM Computing Surveys*. 2020;53(5):1-40.
- Kaliyar RK, Goswami A, Narang P. FakeBERT: fake news detection in social media with a BERT-based deep learning approach. *Multimed Tools Appl*. 2021;80(8):11765-88.
- Mookdarsanit P, Mookdarsanit L. The COVID-19 fake news detection in Thai social texts. *Bulletin of Electrical Engineering and Informatics*. 2021; 10(2):988-98. doi:10.11591/eei.v10i2.2745.
- Bird S, Klein E, Loper E. *Natural Language Processing with Python*. Sebastopol (CA): O'Reilly Media; 2009.
- Ministry of Digital Economy and Society. Thailand Anti-Fake News Center [Internet]. Bangkok: Ministry of Digital Economy and Society; 2025 [cited 2026 Mar 16]. Available from: <https://www.antifakenewscenter.com/>.
- Phatthiyaphaibun W, Chaovavanich K, Polpanumas C, Suriyawongkul A, Lowphansirikul L, Chormai P, et al. PyThaiNLP: Thai natural language processing in Python. In: *Proceedings of the 3rd Workshop for Natural Language Processing Open Source Software (NLP-OSS 2023)*; 2023. p. 25-36.
- Pang B, Lee L. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*. 2008;2(1-2):1-135.
- NECTEC. AI for Thai: natural language processing services for Thai language [Internet]. Pathum Thani: National Electronics and Computer Technology Center; 2025 [cited 2026 Mar 16]. Available from: <https://aiforthai.in.th/>.
- Zhou JL, Ye Z, Zhang S, Zhao G, Han N, Yang T. Investigating response behavior through TF-IDF and Word2vec text analysis: A case study of PISA 2012 problem-solving process data. *Heliyon*. 2024;10(16):e35945.
- Khamphakdee N, Seresangtakul P. An efficient deep learning for Thai sentiment analysis. *Data*. 2023;8(5):90. doi:10.3390/data8050090.
- Dhiman P, Kaur A, Gupta D, Juneja S, Nauman A, Muhammad G. GBERT: A hybrid deep learning model based on GPT-BERT for fake news detection. *Heliyon*. 2024 Aug 30;10(16).
- Géron A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. 3rd ed. Sebastopol (CA): O'Reilly Media; 2022.
- Meesad P. Thai fake news detection based on information retrieval, natural language processing and machine learning. *SN Computer Science*. 2021 Nov;2(6):425.
- Dablain D, Krawczyk B, Chawla NV. DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data. *IEEE Transactions on Neural Networks and Learning Systems*. 2022;34(9):6390-404.
- McCallum A, Nigam K. A comparison of event models for Naïve Bayes text classification. In: *Proceedings of AAAI-98 workshop on learning for text categorization*; 1998 Jul 27; Madison (WI). Menlo Park (CA): AAAI Press; 1998. p. 41-8.
- Cortes C, Vapnik V. Support-vector networks. *Machine Learning*. 1995;20(3):273-97.
- Goodfellow I, Bengio Y, Courville A. *Deep learning*. Cambridge (MA): MIT Press; 2016.
- Rumelhart DE, Hinton GE, Williams RB. Learning representations by back-propagating errors. *Nature*. 1988;323(6088):533-6.
- Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*; 1995. p. 1137-43.
- Breiman L. Random forests. *Machine Learn*. 2001; 45(1):5-32.
- Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Machine Learning*. 2006;63(1):3-42.



Physical property of 925 sterling silver metal injection molding and morphology of brown parts under varying solvent-based debinding conditions

Montri Kawsuk¹, Witthaya Daodon² and Surat Wannasri^{2*}

¹Department of Agricultural Engineering for Industry, Faculty of Technology and Industrial Management, King Mongkut's University of Technology North Bangkok, Prachinburi Campus, THAILAND

²Department of Industrial Engineering, Faculty of Engineering and Technology, Rajamangala University of Technology Isan, Nakhon Ratchasima 30000, THAILAND

*Corresponding author: surat.wn@rmuti.ac.th

ABSTRACT

This research aims to study the physical properties of raw parts produced using Metal Injection Molding (MIM), an advanced manufacturing process suitable for the industrial-scale production of highly complex, small, and high-precision three-dimensional metal parts. However, applying this technique to precious metals, especially 925 sterling silver, still has significant limitations and research gaps, particularly regarding the use of atomized metal powders, which are a byproduct and have lower costs than general-purpose specialized metal powders. 925 silver feedstock used to produce the MIM samples consisted of 90% 925 silver powder and 10% binder. Solvent-based binder removal was performed in water for 5 hours at three temperatures: 30°C, 40°C, and 50°C, respectively. Microstructure and particle morphology were examined using a scanning electron microscope (SEM). The goal was to identify the optimal conditions for forming an open porous structure conducive to sintering and shape retention in subsequent industrial processes. The experimental results showed that the optimal material ratio consisted of 90% silver powder and 10% binder by weight, resulting in an average material density of 4.507 g/cm³ and an average material tensile strength of 17.70 MPa. The results indicated that the binder removal process using distilled water at 50°C for 5 hours was the most efficient method.

Keywords: Metal injection molding, Sterling silver, Debinding kinetics, Morphology, Physical properties

INTRODUCTION

Powder metallurgy (P/M) processes involve the production of metal powders, property analysis, and extrusion (or molding) [1-3]. These processes are used to produce metal powders, analyze their properties, and then sinter them to produce finished components. The application of powder metallurgy technology in the Thai gem and jewelry industry shows relatively limited use in manufacturing processes. Thai manufacturers often use it as a powder solder or binder for joining jewelry components [4-7]. However, this process can be applied to a variety of manufacturing processes, such as powder injection molding (PIM), metal injection molding (MIM), and additive manufacturing (AM) [8-10].

In the context of metal injection molding (MIM), which involves the preparation of the injection molding material, part forming, and sintering, the flow behavior of the feedstock, which is a mixture of silver powder and binder, is studied during the material preparation. The rotor speed and preheating temperature are then calculated to achieve the optimal

viscosity. This ensures that the metal flows fully into the mold without causing component segregation. Sintering is a thermal process that bonds individual powder particles together. This is a crucial step in increasing the strength, rigidity, and toughness of the particles for various applications (e.g., compression, tensile, and other conditions). Sintering involves heating the injection-molded part to a temperature close to the melting point of the metal. This sintering temperature causes the internal structure of the particles to move at the atomic level. This results in the dispersion and interconnection of particles. The resulting interparticle adhesion is removed by the binder (Debinding Kinetics): The temperature and time of preheating are usually derived from the melting point and thermal decomposition point of the binder, which must be consistent with the TGA analysis results. This can occur either in the solid state or through the formation of a liquid phase from the material mixture under sintering conditions. Various methods for binder removal exist, including catalytic debinding using a chemical agent, solvent

debinding with a liquid solvent or water, and thermal debinding using heat [11-14].

To study the properties of the various components present in silver alloys in the MIM process, the main process parameters were systematically analyzed to understand the strength required for forming while maintaining the silver luster and to understand better the processing behavior of 925 sterling silver powder [15-19]. The leftover powder from the research of Montri et al. [20], with particle sizes in the range of 75-100 μm , was used as the raw material. This research aims to evaluate the physical properties of 925 sterling silver parts produced by the MIM process. Process parameters: type of solvent, temperature, time, solid-to-liquid ratio and stirring method, binder removal rate, percentage weight loss versus time and temperature, and determining the optimal separation conditions for dissolving water-soluble binders from MIM raw parts, thus strengthening Thailand's potential to be a regional hub for the production and trade of high-value precious metals.

MATERIALS AND METHODS

Materials and production of metal powder

The material used in this study was 925 sterling silver alloy prepared according to the Thai Industrial Standard (TIS 21-2515). The alloy composition was obtained by mixing 99.99% pure silver with 99.9% pure copper in a ratio of 92.5:7.5 by weight. Metal powder production was done using a prototype gas atomizer at Rajamangala University of Technology Isan, as shown in Figure 1. This gas atomizer consists of various components and operates using nitrogen gas as a medium to atomize the molten metal into fine droplets. The nozzle used in the experiment was a close-fitting nozzle. After mixing the metal for the specified time, the nozzle was lifted to allow the molten metal to flow into a water tank located below the melting chamber. Samples were collected for further testing.

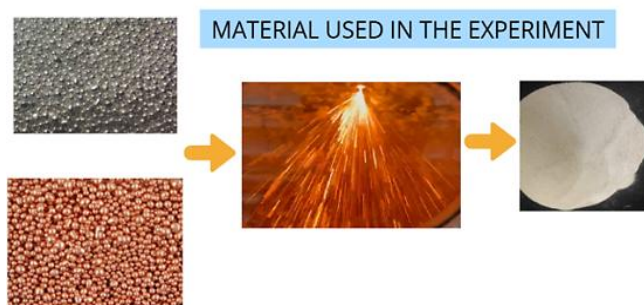


Figure 1 Material used in this study 925 sterling silver alloy powder.

Binder composition, feedstock preparation and metal injection molding

The 925 silver powder, with a particle size of 75-100 μm obtained from gas atomization, was used as the base material. The binder system was formulated to ensure adequate flowability during injection molding, shape retention during debinding, and ease of removal prior to sintering. Based on our previous work [21], the binder consisted of high-density polyethylene (HDPE) as the backbone polymer, polyvinyl alcohol (PVA) and polyethylene glycol (PEG) as water-soluble binders, paraffin wax (PW) as a flow-promoting agent, and stearic acid (SA) as a lubricant and surfactant. The binder composition by weight was 38 wt.% HDPE, 3.75 wt.% PVA, 1.25 wt.% PEG, 54 wt.% PW, and 3 wt.% SA. Feedstock preparation was performed using an internal mixer. The 925 sterling powder was thoroughly mixed with the binder materials in various proportions ranging from 1 wt.% to 10 wt.% binder, with the remainder being silver powder. First, the 925 silver powder was preheated to 60°C for 15 min at a rotor speed of 10 Hz, except when using HDPE, which required 15 Hz. Subsequently, the binder components SA, PW, and HDPE were added sequentially, with each addition lasting 10 min. PVA and PEG were introduced simultaneously at 150°C with a rotor speed of 15 Hz, followed by continued mixing for 20 min to ensure homogeneous dispersion of the powder within the binder matrix. Finally, the heater was turned off while the mixing process continued for 90 min, after which the feedstock was cooled to room temperature and granulated for ease of feeding into the injection molding machine.

Metal injection molding was carried out using a ram injection machine under constant injection pressure of 150 bar and a barrel temperature at the extruder of 190°C, with a holding time of 30 s and a cooling time of 150 s. After cooling, the mold was opened, and the parts were ejected. The injection pressure must be high enough to fill the mold cavities, and the as-molded parts, referred to as "green parts," must have sufficient green strength to maintain their shape and withstand handling loads throughout the debinding and sintering processes, as shown in Figure 2.

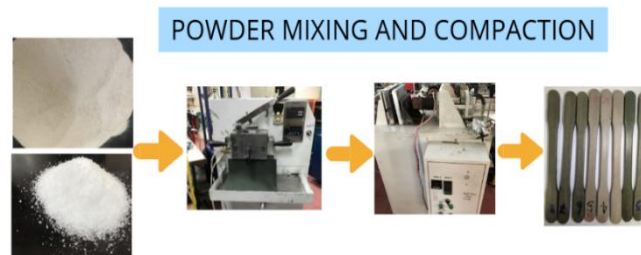


Figure 2 Gas atomized silver powder and binder material, feedstock preparation using an internal mixer, and metal injection molding.

Particle shape analysis and physical property measurement of debound parts

1. Particle shape analysis by Scanning Electron Microscope (SEM), Thermo Scientific, Model: Apreo 2 SEM

The shapes of the metal particles were examined using a scanning electron microscope (SEM). The obtained gas-atomized metal powders have a spherical or near-spherical shape, as shown in Figure 3, which is advantageous for powder injection molding because of improved packing density and flowability. High-magnification observations revealed that the particle surface exhibited a dendritic structure.

2. Physical property evaluation

The density of the MIM green parts was determined using Archimedes' principle, which states that the buoyant force acting on an object immersed in fluid is equal to the weight of the fluid displaced by the object. The mass of the samples was weighed in air and distilled water using a balance with 0.0001 g resolution. The density ρ (g/cm³) can be calculated using the following equation [22]:

$$\rho = \left(\frac{A}{A-B} \right) (\rho_0 - \rho_l) + \rho_l, \quad (1)$$

where ρ is the density of the sample, A is the mass of the sample in air, B is the mass of the sample in the auxiliary liquid, ρ_0 is the density of the auxiliary liquid, and ρ_l is the density of air (0.0012 g/cm³). For this study, distilled water at room temperature was used as the auxiliary liquid, with a density of 0.998 g/cm³.

Solvent debinding

Metal injection molding consists of four main steps: feedstock preparation, injection molding, debinding, and sintering. This study focused on determining the optimal conditions for solvent-based binder removal of 925 sterling silver components. To determine the optimal debinding conditions for 925 sterling silver components, the study used three temperature levels and one time level, as shown in Table 1. Debinding was conducted by immersing the molded "green" parts in distilled water to dissolve the water-soluble binder components. Solvent debinding was performed for 5 h at three different temperatures: 30°C, 40°C, and 50°C. These parameters were selected to evaluate the effect of the solvent temperature on the binder removal efficiency. The experiment was repeated 3 times at different temperatures.

Table 1 Conditions for solvent debinding.

Conditions	Temperature (°C)	Time (hours)
1	30	5
2	40	5
3	50	5

RESULTS AND DISCUSSION

Particle shape of 925 sterling silver metal powder

Figure 3 shows SEM microimages of gas-atomized 925 sterling silver powder under different overheating conditions. The atomized powders exhibited similar characteristics: when gas-atomized, the small metal droplets easily transformed and dispersed better than in highly viscous metal solutions. The effects of overheating the molten metal and peak angles could be observed using scanning electron microscopy (SEM). The size of the gas-atomized silver particles ranged from 75-100 μ m. The resulting silver powders had a variety of shapes [23, 24]. Shapes observed with scanning electron microscopy (SEM) included: plate-like or thin sheet, fibrous, spherical, and irregular. The produced particles were spherical to nearly spherical, with a wide range of sizes. In addition, slight variations in particle shape were found, such as slight irregularity or fibrousness, but these were much less frequent compared to spherical particles. These observations are consistent with the findings of [25,26]. Different experimental conditions resulted in spherical or nearly spherical metal powder particles.

PARTICLE SHAPE OF 925 STERLING SILVER METAL POWDER

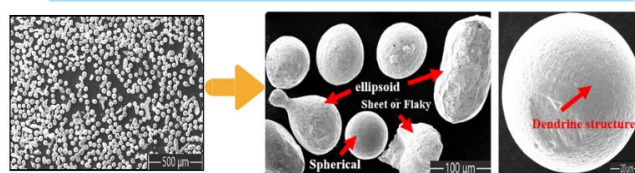


Figure 3 SEM micrographs of 925 sterling silver particles.

Physical property of MIM green parts

The green density of the injection-molded specimens was determined using Archimedes' principle, according to equation 1: the mass of the object lost in water is equal to the mass of water displaced by the object. Therefore, the density of each specimen was calculated from its measured mass in air and in water, as described in the previous section. Figure 5 presents the average green density values obtained from five specimens for each feedstock composition. The results indicate that the green density varied inversely with binder content, which is consistent with the findings of Sasitorn Pramakhmo et al [21]. The feedstock with the lowest binder content, 2 % binder (corresponding to 98 % 925 sterling silver powder), exhibited the highest green density of 6.558 g/cm³. This is attributed to the higher volume fraction of the high-density silver powder ($\rho_{925Ag} = 10.37$ g/cm³) in this composition. As the binder content increased from 2 % to 30 %, the green density progressively decreased. [27] This trend is due to the lower density of the main binder constituents,

paraffin wax ($\rho_{PW} = 0.91 \text{ g/cm}^3$) and high-density polyethylene ($\rho_{HDPE} = 0.95 \text{ g/cm}^3$), compared to that of the silver alloy. Higher binder fractions therefore reduce the overall packing density of the feedstock, leading to lower green densities after molding, as shown in Figure 4.

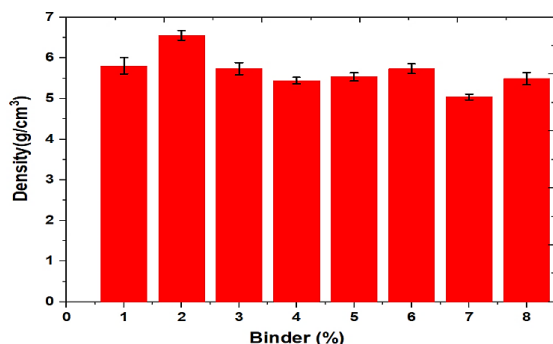


Figure 4 Green density of MIM 925 sterling silver specimens as a function of binder content.

MIM parts and binder removal behavior

The most commonly used method is extrusion molding, with standard test specimens conforming to ASTM E8/E8M-11 standards selected. [28] 925 silver behaves similarly to industrial heavy metals in the MIM process as it is an “alloy” that relies on atomic diffusion during annealing, has similar shrinkage levels, and, importantly, the copper content in this alloy requires oxygen control. [29-30] Solvent debonding involves immersing the specimen in a solvent for a period of time to allow the binder to dissolve and evaporate from the specimen (debonding process), which is controlled by gas diffusion and heating rate. Selected test specimens are shown in Figure 5.

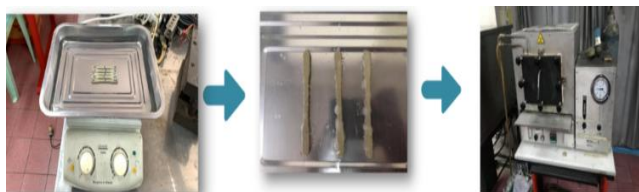


Figure 5 Sample of solution removal and atmosphere metal oven.

In powder injection molding (PIM) processes, multi-component binder systems typically consist of a water-soluble main binder such as polyethylene glycol (PEG) and an insoluble main binder such as polyvinyl butyral (PVB) or polymethyl methacrylate (PMMA). Research by Hwang et al. [31] showed that a binder system containing PEG, when removed at temperatures of around 40-50 °C, reduces the surface tension and viscosity of the solution within the porous structure, resulting in orderly particle shrinkage and preventing the formation of large air gaps (pore clustering). Thus, experiments showed that increasing the temperature from 30 °C (condition 1) to 40 °C (condition 2) and up to 50 °C (condition 3) improved

efficiency. However, temperature increases in this range often resulted in a decrease in the average pore size or a more uniform pore fineness. These results are shown in Figure 6.

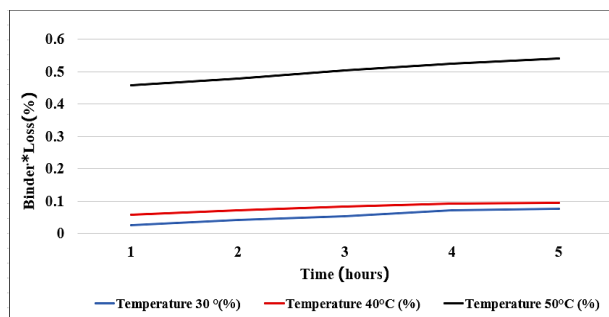


Figure 6 Percentage of binder dissolution as a function of debinding time for MIM 925 sterling silver specimens under three solvent debinding conditions.

Microstructure of the test piece after solvent debonding

The internal microstructures of the green specimens after solvent debinding were investigated to evaluate the influence of debinding temperature on the porosity and pore distribution of the fast-solidifying alloy powder compacts. Distinct solidification structures evolved, governed primarily by the alloy composition and specific solidification parameters. The specimens were extracted by immersion in distilled water for 5 h at various temperatures, followed by drying at 100°C for 1 h to remove residual moisture. Microstructural observations (Figure 7) reveal that the progressive dissolution and diffusion of the water-soluble binder generate continuous pore networks along the extraction pathways. This interconnected porosity acts as critical escape routes, allowing gaseous degradation products to safely permeate and vent during subsequent thermal debinding without causing structural defects such as bloating, blistering, or cracking. Significant morphological differences in the pore structure were observed across the three debinding temperatures [32-33]:

In conclusion, the solvent debinding temperature plays a critical role in determining the pore architecture and binder extraction efficiency of fast-solidifying alloy powder specimens. Increasing the water bath temperature from 30°C to 50°C transitions the pore structure from isolated, large voids to a highly uniform and interconnected network of fine pores. This structural transformation is attributed to the thermally activated diffusion kinetics of the soluble binder. The formation of a well-distributed, interconnected pore network at 50°C successfully establishes low-resistance pathways for gas transport. [34] Consequently, this optimization of the solvent debinding stage ensures safe, defect-free gas venting during subsequent thermal processing, representing a critical parameter

for maintaining structural integrity in high-quality powder metallurgy applications.

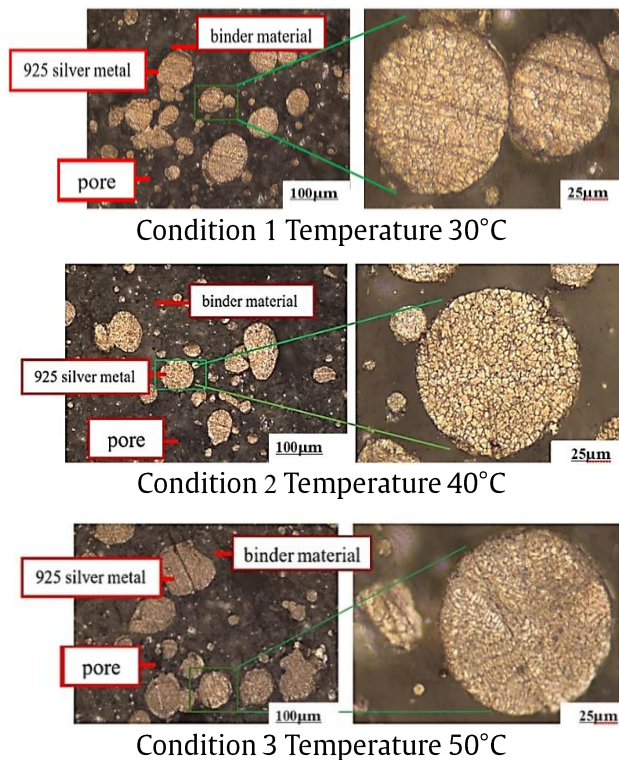


Figure 7 Optical micrograph of the cross-section of the MIM 925 sterling silver sample after solvent separation under conditions.

CONCLUSIONS

In this study, waste 925 sterling silver powder (75–100 µm) produced by gas atomization was used to fabricate parts via metal injection molding (MIM). This study aimed to examine the physical properties of MIM-fabricated 925 silver parts and to investigate the influence of solvent-based debinding temperature on binder removal efficiency and molded-part morphology. Three debinding conditions (30°C, 40°C, and 50°C) were evaluated over a 5 h solvent immersion period. The green density, binder loss, morphology, and pore formation were analyzed to determine the optimal debinding conditions. Based on the experimental results, the following conclusions can be drawn. (1) Scanning electron microscopy (SEM) analysis of the particle morphology revealed that the gas-atomized powders were predominantly spherical or near-spherical in shape, with a fine dendritic surface substructure. This is because the metal droplets cool faster than under equilibrium conditions under the different experimental conditions. As the pouring temperature increased, the viscosity of the metal solution decreased, increasing the mass flow rate. (2) Binder removal behavior was examined through solvent debinding in distilled water at three temperatures: 30°C, 40°C, and 50°C for 5 h. The water-soluble binder components, polyvinyl alcohol (PVA) and polyethylene

glycol (PEG) were progressively removed as the debinding temperature increased. The test samples exhibited varying porosities, with 50°C being the optimum temperature for binder removal. (3) The morphological and microstructural features observed with optical microscopy after debinding show that solvent debinding at 50°C results in fine pores uniformly distributed across the cross-section of the sample, which is conducive to subsequent thermal debinding and sintering. Conversely, lower temperatures resulted in fewer and larger pores, indicating less efficient debinding. Although the binder was completely dissolved at 50°C, some binder residue remains, highlighting the need for further thermal debinding to ensure its removal. Further temperature increases might lead to better results.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author used ChatGPT in order to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

ACKNOWLEDGEMENT

The authors would like to express their appreciation to Department of Industrial Engineering, Faculty of Engineering and Technology, Rajamangala University of Technology Isan, and Department of Agricultural Engineering for Industry, Faculty of Industrial Technology and Management, King Mongkut's University of Technology North Bangkok, Prachinburi Campus for supporting this research and Suntara Refiner Co., Ltd. and Pata Chemical Co., Ltd are kindly acknowledged for providing the materials used in this work.

REFERENCES

- Georgarakis K, Kvashnin VI. Metallic glass-reinforced metal matrix composites: Design, interfaces and properties. *Materials*. 2022;15(23):8278.
- Georgarakis K. Core-shell particle reinforcements-A new trend in the design and development of metal matrix composites. *Materials*. 2022;15(7):2629.
- Chorney MP, Downey JP, Tripathy PK. On the sintering behavior of Nb₂O₅ and Ta₂O₅ mixed oxide powders. *Materials*. 2022;15(14):5036.
- Kazior J. Influence of sintering atmosphere, temperature and the solution-annealing treatment on the properties of precipitation-hardening sintered 17-4 PH stainless steel. *Materials*. 2023;16(2):760.

5. Song YQ, Savvakini D, Xu X, Stasiuk O, Ivasishin O, Cheng T. In situ Ti₆Al₄V/TiB composites prepared by hydrogen-assisted sintering of blends containing TiH₂ and ball-milled Ti+TiB₂ powders. *Materials*. 2022;15(3):1049.
6. Moser S, Vrel D, Perrière L, Pirès-Brazuna R, Couque H, Bernard F. Elaboration and characterization of WMoTaNb high entropy alloy prepared by powder metallurgy processes. *Materials*. 2022;15(15):5416.
7. Edosa OO, Gupta K. A review on the influence of process parameters on powder metallurgy parts. *Eng Appl Sci Res*. 2021;49(3):433-43.
8. Fayyazi S, Kasraei M. Optimizing high-velocity oxygen fuel-sprayed WC-17Co coating using Taguchi experimental design to improve tribological properties. *Trans Indian Inst Met*. 2018;71(12):3045-62.
9. Li YY, Zhang DT, Long Y, Xia W. Effect of die wall lubrication on warm compaction powder metallurgy. *J Mater Process Technol*. 2002;129(1-3):354-8.
10. Ward M, Billington JC. Effect of zinc stearate on apparent density, mixing, and compaction/ejection of iron powder compacts. *Powder Metall*. 1979;22(4):201-8.
11. Kurgan N. Effects of sintering atmosphere on microstructure and mechanical property of sintered powder metallurgy 316L stainless steel. *Mater Des*. 2013;52:995-8.
12. Huang X, Wang G, Alexandrov S. Effect of powder size on microstructure and mechanical properties of 2A12Al compacts fabricated by hot isostatic pressing. *Adv Mater Sci Eng*. 2018;2018:1-7.
13. Jamaludin SB. The effect of sintering on the properties of powder metallurgy (PM) F-75 alloy. *Adv Mater Res*. 2013;795:573-7.
14. Guo RQ, Rohatgi PK, Nath D. Preparation of aluminium-fly ash particulate composite by powder metallurgy technique. *J Mater Sci*. 1997;32(15):3971-4.
15. Doremus P, Imbault D, Puente G. High-velocity compaction and conventional compaction of metallic powders; comparison of process parameters and green compact properties. *Proc Inst Mech Eng E J Process Mech Eng*. 2010;224(3):177-85.
16. El-Bassuony AAH, Abdelsalam HK. Impact of different magnetic materials added to silver-magnetite nanoparticles on the structural, magnetic and antimicrobial properties. *Eur Phys J Spec Top*. 2023;232:1339-51.
17. Hamad A, Hadi A. Silver nanoparticles and silver ions as potential antibacterial agents. *J Inorg Organomet Polym Mater*. 2020;30(12):4811-28.
18. Niculescu AG, Grumezescu AM. Magnetite nanoparticles: Synthesis methods - a comparative review. *Methods*. 2022;199:16-27.
19. Vasiliev G, Kubo AL, Vija H, Kahru A, Bondar D, Karpichev Y, Bondarenko O. Synergistic antibacterial effect of copper and silver nanoparticles and their mechanism of action. *Sci Rep*. 2023;13(1):9202. doi: 10.1038/s41598-023-36460-2.
20. Montri K, Sirichai T. Morphology of 925 silver powder particles produced from gas atomization. *J Appl Eng Sci Technol (JAETS)*. 2024;5(2):966-76.
21. Pramakhamo S, Chaekrathok S, Wannasri S, Wannarumon Kielarova S, Budchar N, Kraiklang R, et al. Investigating of mixture ratio of 925 sterling silver compound for the metal injection molding process. *Journal of Manufacturing and Management Technology*. 2022;1(2):63-71.
22. Cubberly WH. *Metals Handbook Powder Metallurgy*. Vol. 9. Metals Park, OH: American Society for Metals; 1984.
23. Sharp DH. An overview of Rayleigh-Taylor instability. *Physica D Nonlinear Phenom*. 1984;12(1-3):3-18.
24. German RM. *Powder Metallurgy Science*. Princeton, NJ: Metal Powder Industries Federation; 1994.
25. Dai Y, Song C, Han Q, Zhai Q. Solidification structure of C2.08Cr25.43Si1.19Mn0.43Fe70.87. *Mater Charact*. 2010;61(1):116-22.
26. Waters C, Ajinola S. Porosity comparative analysis of porous copper and OOF modelling. *J Porous Mater*. 2015;22(4):989-95.
27. Seerane M, Ndlangamandla P, Machaka R. Influence of particle size distribution on the properties of metal-injection-moulded 17-4 PH stainless steel. *Journal of the Southern African Institute of Mining and Metallurgy*. 2016;116(11):1045-1050.
28. ASTM International. *ASTM E8/E8M-11: Standard Test Methods for Tension Testing of Metallic Materials*. West Conshohocken, PA, 2011
29. German RM. *Metal Injection Molding: A Handbook of Handbook of Powder Technology*. Elsevier, 2013.
30. Ji CH, Loh NH, Khor KA, Tor SB. Study on sintering behavior of 925 sterling silver powder by metal injection molding. *Journal of Materials Processing Technology*. 2001;111(1-3):103-7.

31. Hwang KS, Hu SC. Analysis of pore evolution and swelling during solvent extraction of powder injection molded parts. *Metallurgical and Materials Transactions A*. 2005;36(5):1293-301.
32. Enneti RK, Park SJ, Atre SV, German RM. Effect of solvent debinding parameters on the diffusion coefficient of polyethylene glycol (PEG) in metal injection molded (MIM) parts. *Materials Letters*. 2012;71:105-7.
33. Lin HP, Samal S. Study on solvent debinding kinetics of water-soluble binder systems in metal injection molding. *Journal of Materials Processing Technology*. 2018;252:452-9.
34. Omar MA, Subuki I, Ismail MH, Muhamad N. Influence of solvent debinding temperature on the defect formation and properties of metal injection molded compacts. *Journal of Solid State Science and Technology*. 2010;18(1):12-21.



Isolation and selection of *Lactiplantibacillus plantarum* CLT-L01 (TISTR P088) from traditional fermented vegetables for solid-state fermentation of soybean meal (SBM)

Phongsupha Chanthachaiyaphum^{*}, Wang Yue and Wang Wenbo

Research and Development Department, Colotex bio-technology (Thailand) co.,Ltd., Chonburi 20220. THAILAND

Corresponding author: Pongsupa18@gmail.com

ABSTRACT

Lactic acid bacteria are widely used as beneficial microorganisms in fermentation processes for animal feed production. In soybean meal, which is an important protein source in animal feed, lactic acid bacteria fermentation can improve nutritional quality by promoting acidification and reducing antinutritional factors. This study aimed to isolate lactic acid bacteria from traditional fermented vegetables and evaluate their potential as starter cultures for the solid-state fermentation of soybean meal. Traditional fermented vegetables were selected as the isolation source because they are locally available, inexpensive, and rich in naturally occurring lactic acid bacteria. Nine acid-producing bacterial isolates were obtained from traditional fermented vegetable samples using MRS agar supplemented with calcium carbonate. The isolates were preliminarily characterized by Gram staining and catalase testing. Six isolates, namely CLT-35, CLT-51, CLT-33, CLT-52, CLT-53, and CLT-L01, were selected for the evaluation of acid production efficiency in MRS broth. Among them, CLT-L01 showed the fastest acid production, with the pH decreasing to 3.21 ± 0.02 after 24 h of incubation. Therefore, CLT-L01 was selected for further evaluation in solid-state soybean meal fermentation. The results showed that soybean meal inoculated with CLT-L01 exhibited a rapid increase in lactic acid content, exceeding 10% after 96 h of fermentation. The highest lactic acid content was $11.13 \pm 0.06\%$ at 127 h, whereas the uninoculated control reached only $2.43 \pm 0.12\%$ at 151 h. These findings indicate that CLT-L01 enhanced acid production and improved fermentation efficiency compared with natural fermentation. Molecular identification based on 16S rDNA sequencing revealed that CLT-L01 showed more than 99% sequence similarity to the reference strain *Lactiplantibacillus plantarum*. Overall, CLT-L01 has potential as a starter culture for promoting lactic acid production and acidification during soybean meal fermentation.

Keywords: *Lactiplantibacillus plantarum*, Lactic acid bacteria, Solid-state fermentation, 16S rDNA

INTRODUCTION

Soybean meal (SBM) is a widely used plant protein ingredient in animal feed because of its high protein content, availability, and relatively low cost. However, its nutritional value can be limited by antinutritional factors (ANFs), such as trypsin inhibitors, phytates, oligosaccharides, glycinin, and β -conglycinin, which reduce nutrient digestion, particularly in young or monogastric animals. Therefore, suitable processing methods are needed to improve SBM quality. Microbial fermentation is an effective approach because it can reduce ANFs, produce organic acids and enzymes, and promote the breakdown of complex proteins into smaller peptides, thereby improving digestibility and nutrient utilization in animals [1]. In addition, fermentation supports current interest in the sustainable use of biological resources by adding value to locally available agricultural materials [2-4].

Lactic acid bacteria (LAB) are attractive starter cultures for fermentation because they utilize

carbohydrates in plant-based materials as energy sources and convert them into organic acids, particularly lactic acid. This acidification lowers the pH of the fermentation system and creates unfavorable conditions for spoilage and pathogenic microorganisms. In soybean meal fermentation, LAB may contribute to controlled acidification, fermentation stability, and improvement of feed quality. Previous studies have reported that fermented soybean meal can show increased lactic acid production, lower pH, improved protein degradation, increased small peptides and free amino acids, and reduced oligosaccharides or other antinutritional substances [5-7].

Traditional fermented vegetables are widely consumed, readily available locally, inexpensive, and rich in naturally occurring acid-producing microorganisms. Vegetable fermentation generally relies on microorganisms naturally present on vegetable surfaces and plant tissues. Brining and oxygen-restricted conditions favor salt- and acid-tolerant microorganisms, particularly lactic acid

bacteria, which often become dominant during fermentation [8-11]. Therefore, traditional fermented vegetables provide a suitable natural habitat for isolating LAB that are adapted to plant-based fermentation environments and may have good acid-producing potential for application in soybean meal fermentation.

During soybean meal fermentation, LAB can use available carbohydrates and added sugar as substrates for lactic acid production, resulting in rapid acidification of the fermentation matrix. The resulting acidic environment may enhance fermentation stability and contribute to the partial degradation of soybean proteins, including glycinin and β -conglycinin, into smaller peptides and free amino acids. These biochemical changes are important because they are associated with improved digestibility and reduced effects of some antinutritional components in fermented soybean meal [7-10]. Therefore, this study aimed to isolate LAB from traditional fermented vegetables, screen the isolates for acid-producing ability, identify the selected isolate, and evaluate its potential for improving acid production during the solid-state fermentation of soybean meal.

MATERIALS AND METHODS

Sample collection

Traditional fermented vegetable samples were randomly collected from 2 local markets in Chonburi Province, Thailand. The samples were aseptically collected in sterile containers and transported to the laboratory under refrigerated conditions. All samples were processed within 24 h after collection for the isolation of lactic acid bacteria.

Culture media

De Man, Rogosa and Sharpe (MRS) agar and MRS broth were used for the isolation, purification, activation, and propagation of lactic acid bacteria [11]. MRS agar supplemented with CaCO_3 was used for the preliminary screening of acid-producing colonies (clear zones indicate acid dissolution of CaCO_3). [12]. Sterile saline solution was used for serial dilution of the samples.

Isolation and screening of lactic acid bacteria

Traditional fermented vegetable samples were serially diluted from 10^{-1} to 10^{-6} using sterile saline solution. Appropriate dilutions were transferred onto solid MRS agar supplemented with CaCO_3 and then incubated at 35°C for 48 h. After incubation, colonies with clear zones around them colonies were selected and purified by transferring them to MRS agar using the streak plate method. The purified cultures were then tested for lactic acid bacteria by Gram stain and catalase testing [13].

Acid production efficiency test in MRS broth

Each strain of LAB bacteria was grown in MRS broth and incubated at 30°C for 24 h to use as the starting cell inoculum. The initial cell density of each isolate was adjusted to the same 1.0 McFarland standard, then transferred into MRS broth and incubated at 30°C . The pH was monitored at 0, 6, 12, 24, 36, and 48 h to assess the acidogenic capacity of each isolate. The isolate showing the greatest reduction in pH was considered to have the highest acid-producing ability and was selected as a representative strain for subsequent soybean meal fermentation experiments.

Solid-state fermentation of soybean meal

1. Starter Culture Preparation: The selected LAB isolate was cultured in MRS broth and incubated at 30°C for 24 h to activate the culture prior to fermentation. Then the culture was then adjusted to the 1.0 McFarland standard using sterile distilled water to be used as a starter culture for the solid fermentation of soybean meal.

2. Preparation of Soybean Meal Fermentation: The experiment consisted of two sample groups: a LAB inoculum group and a control group without inoculum. The fermentation formula consisted of 50% SBM, 10% sugar, and 1% lactic acid bacteria for the LAB inoculum group. DI water was then added to reach 100% (volume by weight of the total substrate). The soybean meal and sugar were thoroughly mixed, and a separate liquid phase was prepared. For the LAB inoculum sample, the bacteria were mixed with water, and the liquid phase was gradually added to the dry mixture, mixing until homogeneous. The prepared samples were then placed into fermentation containers, sealed, and labeled with the sample group, experimental group number, and fermentation start time. Fermentation was carried out at room temperature. Each sample group had three experimental sets. Samples were collected separately at 0, 7, 24, 31, 48, 56, 72, 79, 96, 103, 120, 127, 144, 151, and 168 h to determine the total acid content.

Total titratable acidity analysis

The total titratable acid (TTA) was determined by titration with 0.1 N NaOH and is expressed as percentage of lactic acid (% lactic acid; %LA). The fermented soybean meal sample was mixed with distilled water in a 1:10 ratio. Then, the mixture was homogenized and filtered to obtain the filtrate. An aliquot of 10 mL filtrate was titrated with standardized 0.1 N NaOH using phenolphthalein as an indicator until a pale pink endpoint persisted for approximately 30 s. A blank titration was performed using distilled water under the same procedure. The net volume of NaOH used was calculated as follows:

$$V_{\text{net}} = V_{\text{sample}} - V_{\text{blank}}$$

Total acidity was calculated and expressed as % lactic acid for solid samples using the following equation:

$$\%LA = \frac{V_{net} \times N \times 90.08 \times V_{total} \times 100}{1000 \times V_{aliquot} \times W}$$

where:

V_{net} = net volume used for titration.

V_{sample} = volume used for the sample.

V_{blank} = volume used for the blank.

N = normality of NaOH.

V_{total} = final volume after sample dilution.

$V_{aliquot}$ = volume of filtrate used for titration.

W = sample weight.

The results were expressed as % Lactic Acid (%LA, w/w) for comparison of acid production between the LAB-inoculated samples and the non-inoculated control samples.

Molecular identification of the selected isolate

1. Extraction of Bacterial DNA: Genomic DNA of selected isolates was extracted using a bacterial genomic DNA extraction kit according to the manufacturer's protocol.

2. PCR Amplification of 16S rDNA: The primers and amplification program required for PCR amplification of bacterial 16S rDNA referred to reference [14].

3. Sequencing of Amplified Products: After purification, the 16S rDNA amplification product was sequenced with universal 16S rDNA primers by Shanghai Sangon Biotech Co., Ltd and the obtained sequence was compared with reference sequences available in the NCBI GenBank database using BLAST analysis.

4. Phylogenetic analysis was performed based on the 16S rRNA gene sequence using MEGA12.1.2 software. A phylogenetic tree was constructed to determine the taxonomic relationship between the selected isolate and reference strains.

5. Statistical analysis was performed using SPSS software. All experiments were conducted in triplicate, and data are expressed as mean $\pm$ standard deviation (SD, $n=3$). Differences in pH values among strains were analyzed using one-way ANOVA followed by Games-Howell post hoc test. The effects of treatment, fermentation time, and their interaction on total lactic acid content were analyzed using two-way ANOVA. Statistical significance was set at $p < 0.05$.

RESULTS AND DISCUSSION

Description of samples

Two types of traditional fermented vegetables, pickled mustard greens and spider weed, were used as sources for bacterial isolation. Visual observation showed that the samples ranged from light to dark

green and consisted of whole stems and leaves immersed in cloudy brine. The samples also had a characteristic sour odor associated with natural fermentation (Figure 1).



Figure 1 Traditional fermented vegetable samples used for bacterial isolation: (A) pickled mustard greens and (B) spider weed.

Isolation and screening of lactic acid bacteria

After isolating bacteria on MRS medium supplemented with CaCO_3 , acid-producing colonies can be identified by the presence of a clear ring around the colony. This ring is formed when organic acids produced by lactic acid bacteria dissolve CaCO_3 in the medium, creating an observable ring around the colony. Based on the presence of clear zones, a total of 9 preliminary isolates were obtained. Six bacterial strains, namely CLT-33, CLT-35, CLT-51, CLT-52, CLT-53, and CLT-L01, were selected after Gram staining and microscopic examination, as well as catalase reaction testing. These strains were selected based on their Gram-positive, catalase-negative, and non-spore-forming characteristics, consistent with the features of lactic acid bacteria. Subsequently, the six strains were tested for their acid-producing ability in MRS broth. The results showed that strain CLT-L01 had a high acid-producing rate. This is an important factor affecting the duration of the fermentation process and serves as a key indicator for evaluating the fermentation ability or robustness of the strain, as shown in Table 1. From Table 1, CLT-L01 showed the fastest acid production rate, with the lowest pH decreasing to 3.21 ± 0.02 after 24 h of incubation. This was followed by CLT-35, which had a pH of 3.46 ± 0.01 , and CLT-51, CLT-33, CLT-53 and CLT-52 which had a lowest pH of 3.94 ± 0.03 , 4.15 ± 0.02 , 4.53 ± 0.02 and 4.85 ± 0.01 after 24 h of incubation. Therefore, based on the acid production rate, CLT-L01 was selected as the strain with rapid acid production for use in subsequent experiments.

Morphological, colony and physiological and biochemical characteristics

1. Cell Morphology: The microscopic observations of strain CLT-L01 revealed that it consisted of short, rod-shaped or nearly spherical cells arranged singly or

in short chains. These cells were non-motile, non-spore-forming, and measured approximately $0.5\text{--}0.8\ \mu\text{m} \times 1.0\text{--}1.2\ \mu\text{m}$, as shown in Figure 2A. These morphological features are consistent with the general description of members of the genus *Lactobacillus*, particularly *L. plantarum*, which are generally rod-shaped and non-spore-forming [15,16]. Morphological testing combined with Gram staining and the catalase test are the primary screening criteria in the microbiology of fermented foods, being particularly important in identifying gram-positive and catalase-negative lactic acid bacteria [17].

2. Colony Characteristics: The colonies of strain CLT-L01 grown on MRS agar are opaque, spherical, and homogeneous, with a center of approximately 1.2–1.5 mm, slightly convex, smooth surface and have a slight viscosity as shown in Figure 2B. Previous reports indicate that lactic acid bacteria isolated from common

fermentation media produce small, smooth, and convex colonies on MRS agar due to acid metabolism and adaptation to a carbohydrate-rich environment [15,17].

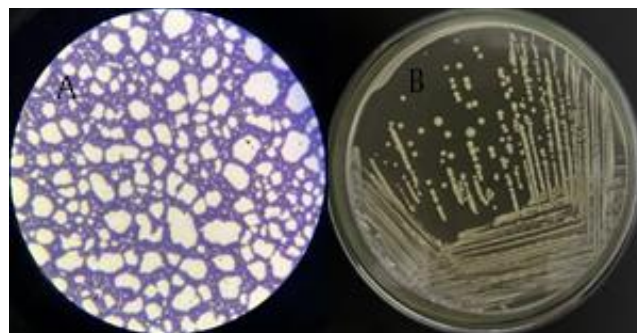


Figure 2 The microscopic appearance of cells (A) and colonies of strain CLT-L01 grown on MRS agar (B).

Table 1 The lowest pH values after culturing 24 hours in MRS broth by six strains of lactic acid bacteria.

Strain	CLT-33	CLT-L01	CLT-35	CLT-51	CLT-52	CLT-53
pH value	4.15 ± 0.01^d	3.22 ± 0.02^a	3.46 ± 0.01^b	3.92 ± 0.03^c	4.86 ± 0.01^f	4.54 ± 0.02^e

* Values are mean $\pm$ SD (n = 3), Different superscript letters indicate significant differences among strains (Games-Howell post hoc test, $p < 0.05$).

Solid-state fermentation testing of soybean meal using *L. plantarum* CLT-L01

The results of the soybean meal test showed that the percentage of total lactic acid (%LA) tended to increase in both the samples inoculated with CLT-L01 and the uninoculated control samples. However, the samples inoculated with CLT-L01 showed a significantly higher effective increase in total lactic acid compared to the control samples ($p < 0.05$), as shown in Figure 3. From figure 3 shows that the total lactic acid inoculated with CLT-L01 increased continuously, rising to $1.16 \pm 0.02\%$, $3.50 \pm 0.01\%$, $6.00 \pm 0.21\%$, and $8.49 \pm 0.15\%$ after 7, 24, 48, and 72 h of fermentation, respectively. It then increased to $10.20 \pm 0.25\%$ after 96 hours of fermentation and peaked at $11.13 \pm 0.06\%$ at 127 hours of fermentation. It then slightly decreased towards the end, remaining at $10.20 \pm 0.15\%$, $8.85 \pm 0.30\%$, and $8.40 \pm 0.15\%$ at 144, 151, and 168 h, respectively. While the uninoculated samples showed a gradual increase. The total lactic acid increased to $1.32 \pm 0.03\%$ after 48 h of fermentation and peaked at only $2.43 \pm 0.12\%$ after 151 h. This demonstrates that the addition of CLT-L01 inoculum promotes lactic acid production during SBM fermentation more efficiently than natural fermentation as seen in the uninoculated group. This is consistent with previous research reporting similar findings on soybean meal fermentation using *Lactobacillus* spp. inoculum, which increases lactic acid production and improves SBM fermentation quality [5,6]. Higher total lactic acid content after fermentation is a desirable characteristic for animal feed products,

which can help improve growth performance and balance of beneficial gut microorganisms, comparable to probiotic supplementation [18]. It was found that there was a slight decrease in acid content towards the end of fermentation, which may be related to lower nutrient levels and accumulation of acidic final products, as well as reduced lactic acid metabolism during prolonged fermentation. This is consistent with previous reports concerning the growth of lactic acid bacteria, which indicate that nutrient deficiencies and the accumulation of lactic acid or higher production products can inhibit bacterial growth and limit bacterial metabolic activity [19].

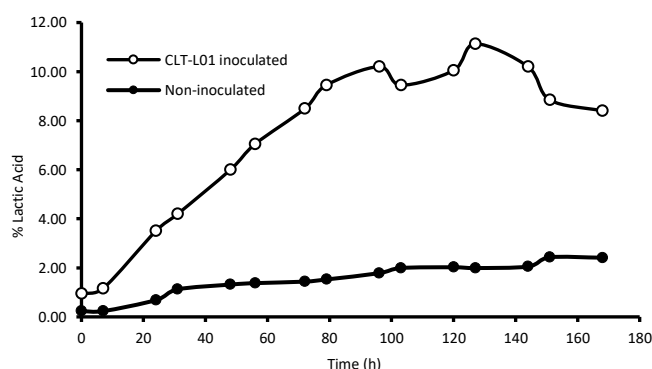


Figure 3 Changes in total lactic acid content in fermented SBM inoculated with CLT-L01 compared with the non-inoculated control during 168 h of fermentation.

Soybean meal served as the solid substrate, and added sugar provided an available carbon source

for CLT-L01. Under sealed solid-state fermentation, CLT-L01 produced lactic acid from available carbohydrates, acidifying the soybean meal matrix and improving fermentation stability. Thus, the higher lactic acid content in the inoculated treatment reflected the acid-producing ability of CLT-L01 and its compatibility with the sugar-supplemented soybean meal system [5, 6, 19].

However, the total lactic acid content of samples inoculated with CLT-L01 remained higher than that of controls throughout the fermentation period. These findings indicate that CLT-L01 has potential as a starter culture for controlled SBM fermentation and can improve acid production efficiency during fermentation.

Sequence, homology comparison and phylogenetic analysis

To determine the taxonomic position of strain CLT-L01, the 16S rRNA gene was amplified using universal bacterial primers, resulting in a PCR of approximately 1.5 kilobases [14]. The purified amplicons were then subjected to sequencing. The obtained sequences were compared with reference sequences in the GenBank database using BLAST analysis [20]. The results of sequence analysis showed that strain CLT-L01 had more than 99% sequence similarity to the reference strain *L. plantarum*. In addition, phylogenetic analysis based on 16S rRNA sequences showed that strain CLT-L01 belongs to the phylogenetic group of *L. plantarum* and has branching that is consistent with the reference strain. As shown

in Figure 4, in the taxonomy of bacteria 16S rRNA sequence similarity has long been used as a key criterion for taxonomic classification. The approximately 97% similarity can be used as a reference for species-level relationships [21,17]. Therefore, the sequence similarity and phylogenetic grouping found in this study can indicate that strain CLT-L01 is closely related to *L. plantarum*. This result is consistent with previous reports that fermented fruits and vegetables are a significant source of the lactic acid bacteria [7-8], particularly *L. plantarum*, a bacterium that is highly adaptable to different fermentation environments and substrates, supporting its potential as a starter culture for plant-based fermentation systems [22-23]. The CLT-L01 strain was then deposited in the culture archive with the access number TISTR P088 to prepare for future development and application.

This study's findings contribute to the development of knowledge and industrial applications by utilizing readily available bio-resources. In this study, low-cost and locally available traditional fermented vegetables were used to isolate potential lactic acid bacteria starter cultures. The selected strain, CLT-L01 (TISTR P088), effectively promoted lactic acid production and acid formation during soybean meal fermentation, suggesting a practical and cost-effective strategy for improving plant-based animal feed formulations. However, further studies on optimizing fermentation conditions, starter culture stability, and product efficacy are needed to support future industrial-scale applications.

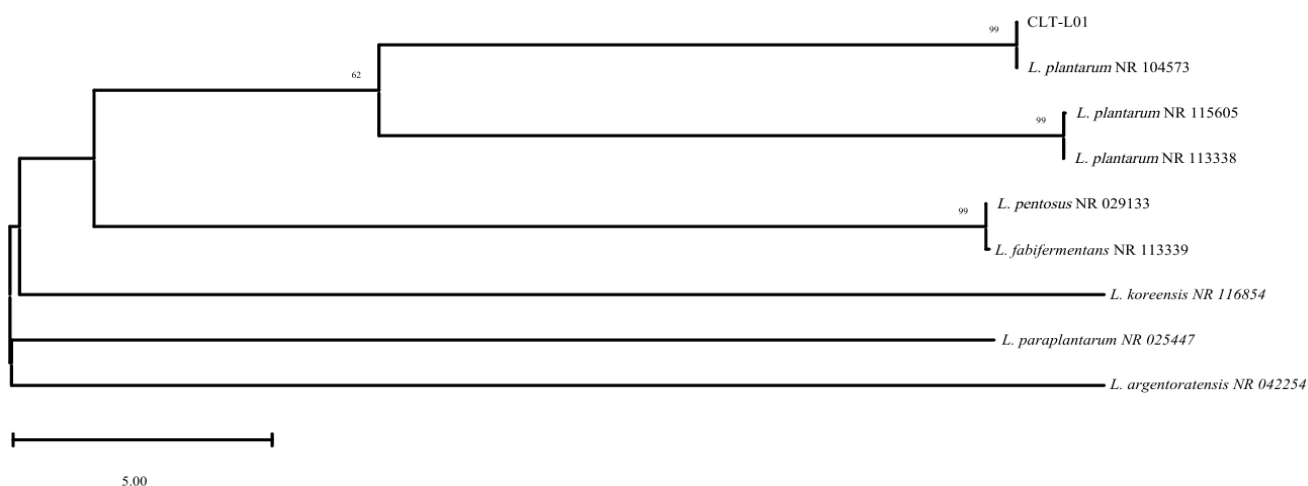


Figure 4 Phylogenetic chart of CLT-L01 isolated from traditional fermented vegetables based on 16S rDNA sequence.

CONCLUSIONS

The selected inoculum CLT-L01 exhibited high 16S rDNA sequence similarity to the reference strain *L. plantarum* and demonstrated efficient acid production capabilities during solid-state fermentation of soybean meal. The lactic acid percentage increased by more than 10% from 96 h onwards, peaking at $11.13 \pm 0.06\%$ at 127 hours. These results indicate that

CLT-L01 promotes acid production and improves fermentation efficiency compared to natural fermentation. Therefore, CLT-L01 (deposited as TISTR P088) is considered a standout choice as a starter culture for controlling solid-state fermentation of soybean meal.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

The authors used Gemini to improve the legibility and clarity of the manuscript during its preparation. The tool was used solely for language editing. All experimental design, data analysis, interpretation and conclusions were carried out by the authors. The authors reviewed and edited the manuscript after using this tool and assume full responsibility for the content of this publication.

ACKNOWLEDGEMENT

The authors would like to express their sincere appreciation to Colotex Bio-Technology (Thailand) Co., Ltd. for providing laboratory facilities, technical support and materials used in this study. Their support and cooperation were essential for the successful completion of this research.

REFERENCES

1. Mukherjee R, Chakraborty R, Dutta A. Role of fermentation in improving nutritional quality of soybean meal-a review. *Anim Biosci.* 2016;29(11): 1523-29.
2. Murtaza G, Usman M, Ahmed Z, Rizwan M, Iqbal R. Non-wood-based biochars as promising and eco-friendly adsorbents for chromium hexavalent Cr(VI) removal from aquatic systems: state-of-the-art, limitations, and potential future directions. *Environ Pollut Bioavailab.* 2024;36(1):2387680.
3. Murtaza G, Deng G, Usman M, Hyder S, Ahmed Z, Iqbal J, et al. Potential of cotton sticks-derived biochar to alleviate aluminium toxicity in cotton. *Ind Crops Prod.* 2025;229:121022.
4. Murtaza G, Hassan NE, Usman M, Deng G, Ahmed Z, Iqbal J, et al. Synergistic effects of allantoin and *Achyranthes japonica*-biochar profoundly alleviate lead toxicity during barley growth. *Ecotoxicol Environ Saf.* 2025;290:117784.
5. Su LW, Cheng YH, Hsiao FSH, Han JC, Yu YH. Optimization of mixed solid-state fermentation of soybean meal by *Lactobacillus* species and *Clostridium butyricum*. *Pol J Microbiol.* 2018;67(3):297-305.
6. Chen L, Zhao Z, Yu W, Zheng L, Li L, Gu W, et al. Nutritional quality improvement of soybean meal by *Bacillus velezensis* and *L. plantarum* during two-stage solid-state fermentation. *AMB Express.* 2021;11:23.
7. Tamang JP, Watanabe K, Holzapfel WH. Review: diversity of microorganisms in global fermented foods and beverages. *Front Microbiol.* 2016;7:377.
8. Swain R, Anandharaj M, Ray RC, Parveen Rani R. Fermented fruits and vegetables of Asia: a potential source of probiotics. *Biotechnol Res Int.* 2014;2014: 250424.
9. Armenova N, Tsigoriyna L, Arsov A, Petrov K, Petrova P. Microbial detoxification of residual pesticides in fermented foods: current status and prospects. *Foods.* 2023;12(6):1163.
10. Baralić K, Živančević K, Božić D, Đukić-Čosić D. Probiotic cultures as a potential protective strategy against the toxicity of environmentally relevant chemicals: state-of-the-art knowledge. *Food Chem Toxicol.* 2023;172:113582.
11. de Man JC, Rogosa M, Sharpe ME. A medium for the cultivation of lactobacilli. *J Appl Bacteriol.* 1960;23(1):130-5.
12. Hwanhlem N, Buradaleng S, Wattanachant S, Benjakul S, Tani A, Maneerat S. Isolation and screening of lactic acid bacteria from Thai traditional fermented fish (Plasom) and production of Plasom from selected strains. *Food Control.* 2011;22(3-4), 401-7.
13. Mokoena MP. Lactic acid bacteria and their bacteriocins: classification, biosynthesis and applications against uropathogens: a mini-review. *Molecules.* 2017;22(8):1255.
14. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol.* 1991;173(2):697-703.
15. Yang J, Tan H, Cai Y. Characteristics of lactic acid bacteria isolates and their effect on silage fermentation of fruit residues. *J Dairy Sci.* 2016; 99(7):5325-34.
16. Jeyagowri N, Rameshkumar MR, Kumar V, Anandham R. Phenotypic characterisation and molecular identification of lactic acid bacteria isolated from fermented food. *Fermentation.* 2023;9(9):807.
17. Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol.* 2007;45(9):2761-4.
18. Jazi V, Ashayerizadeh A, Toghyani M, Shabani A, Tellez G, Toghyani M. Fermented soybean meal exhibits probiotic properties when included in Japanese quail diet in replacement of soybean meal. *Poult Sci.* 2018;97(6):2113-22.
19. Leroy F, De Vuyst L. Growth of the bacteriocin-producing *Lactobacillus sakei* strain CTC 494 in MRS broth is strongly reduced due to nutrient depletion: a nutrient depletion model for the growth of lactic acid bacteria. *Appl Environ Microbiol.* 2001;67(10):4407-13.

20. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389-402.
21. Stackebrandt E, Goebel BM. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol.* 1994;44(4):846-9.
22. de Vries MC, Vaughan EE, Kleerebezem M, de Vos WM. *L. plantarum*-survival, functional and potential probiotic properties in the human intestinal tract. *Int Dairy J.* 2006;16(9):1018-28.
23. Fidanza M, Panigrahi P, Kollmann TR. *L. plantarum*-nomad and ideal probiotic. *Front Microbiol.* 2021;12:712236.



Enhancing lung cancer screening accuracy through a dual-pipeline deep learning and machine learning framework for integrated risk assessment

Patthanan Uthaitipitak*

Ruamrudee International School, Bangkok 10510, THAILAND

*Corresponding author: patthanan.utha@gmail.com

ABSTRACT

Lung cancer is a leading cause of death worldwide, with delayed or misdiagnosis further hindering professionals' ability to lower lung cancer-related mortality rates. Misdiagnosis could look like false positive results, where benign lung diseases are mistaken for malignant lung carcinomas, or more life-threatening false negative results, where harmful eosinophilic pneumonia is misinterpreted as a benign lung hamartoma. The objective of this study consists of the following goals: (1) to develop CNN-based classification models for lung nodule analysis, (2) to compare the performance of multiple deep learning models under consistent controlled environments, (3) to evaluate and present the interpretability of the models through a clinically grounded probability equation, and (4) to propose an integrated risk-scoring framework that combines the imaging and clinical pipelines into a single interpretable measure of lung cancer risk. However this framework is only theoretical as imaging and clinical models were developed using separate datasets; future studies should note that a practical development will require validation on a unified multimodal patient cohort. In regard to the creation of the models, the deep learning convolutional neural network, which utilized a YOLOv8n backbone trained on the IQ-OTH/NCCD dataset of 649 annotated CT images, achieved a mean average precision (mAP50) of 0.687 when evaluated against expert-verified annotations. The machine learning branch utilizes a structured clinical dataset of 309 patient records with symptom- and risk-based features, including age, smoking status, chronic disease, fatigue, wheezing, and shortness of breath, all evaluated under standardized conditions. These machine learning models, showed similarly promising results, with the Logistic Regression (Log) achieving the highest mean accuracy of 0.929, while CatBoost (CB) demonstrated the strongest overall balance with a Mean F1 score of 0.953 at optimal hyperparameter settings of 1,000 iterations and a tree depth of 8. Interpretability is further enhanced through a clinically grounded cancer probability equation. This framework that showcases both deep and machine learning models will offer insight into the effectiveness of incorporating artificial intelligence in the screen accuracy of lung cancer classification.

Keywords: Lung cancer detection, Deep learning, Machine learning, CT scan, Pulmonary nodule classification

INTRODUCTION

Cancer is the leading cause of death worldwide, accounting for nearly 10 million deaths in 2020 [1, 2], with lung cancer, together with prostate and colorectal cancers, accounting for nearly half of all incident cancer cases [3]. Delays in diagnosis and treatment may lead to an uptick in advanced-stage case and mortality specifically [4]. CT imaging is often used in clinical practice to detect and classify pulmonary nodules at early stages [5], potentially allowing for the early classification of malignant tumors and therefore improving patient survivability. With this technology in mind, differentiating between the subtle features of malignant and benign tumors remains a challenge, especially when considering

how radiologist interpretation may be influenced by intrapersonal factors such as workload; where some studies found radiologists made 21% higher than their normal productivity [6]. With the combined concerns of both alarming cancer mortality rates and the possible human-related risks linked to tumor classification, the application of artificial intelligence (AI) in the practice looks increasingly promising. Developments in deep learning technology, especially convolutional neural networks (CNNs) have demonstrated many advantages. Most crucially, CNNs are capable of automatically learning hierarchical features from large volumes of imaging data [7], which is unique to traditional computerized detection systems, as they can process data without external feature engineering. These two factors allow for efficient

recognition and classification of benign and malignant tumors. On top of efficient deep learning models, AI-based diagnostics have proven high accuracy in identifying tumors [8]. Ardila et al. developed a deep learning system that outperformed radiologists in detecting lung cancers in CT scans [9]. Nam et al. similarly reported that their deep learning algorithm outperformed physicians in “radiograph classification and nodule detection performance for malignant pulmonary nodules on chest radiographs” [10]. The combination of classification efficiency, elimination of human error, and accurate pattern recognition indicates effective serviceability of AI in supporting radiologic decision-making.

Despite AI's promising to support clinical decision-making many considerations must first be accounted for as well. In particular, more research must be conducted into the generalizability of models in using datasets from various institutions [11]. Aside from the practical aspects of AI integration into clinical environments, professionals must also answer questions related to gaps in AI-assisted lung tumor classification. Firstly, many studies rely on only a single dataset, with one study finding that only around 23% of ML-based healthcare papers used multiple datasets [12]. Secondly, effective models should be transparent with their decisions, yet many act as black boxes and do not provide justifications that clinicians require for informed decision-making. Thirdly, regarding deep learning models, direct systematic comparisons between multiple deep learning architectures under identical experimental conditions must also be addressed. These gaps in key information will be investigated to gauge the effectiveness of AI algorithms in applying them to lung cancer diagnosis.

To address these gaps, this study proposes a dual-pipeline framework that, to the authors' knowledge, is among the first to integrate a convolutional object detection model operating on CT scan images with a gradient-boosting classifier trained on structured clinical and symptom data, combining their outputs into a single interpretable risk score. This integration is novel in three respects: it leverages complementary information modalities-imaging and clinical data-within a unified evaluation framework; it enables systematic, controlled comparison of multiple deep learning architectures; and it produces a clinically grounded, interpretable risk score rather than an opaque black-box prediction, directly addressing the transparency gap identified above.

This study aims to evaluate the effectiveness of multiple deep learning and machine learning architectures in differentiating between benign and malignant tumors on CT scans. To break it down into clear objectives, this study aims to: (1) develop CNN-based classification models for lung nodule analysis, (2) compare the performance of multiple deep learning models under similar controlled environments, (3)

evaluate and present the interpretability of the models through a clinically grounded probability equation, and (4) propose an integrated risk-scoring framework that combines the imaging and clinical pipelines into a single interpretable measure of lung cancer risk. The findings of this study contribute to the evidence base on AI-assisted lung cancer diagnostics and provide practical insights for future integration into clinical practice.

MATERIALS AND METHODS

Overall research design

This study follows a two-branch parallel pipeline for developing AI-based models to differentiate between benign and malignant tumors in CT scans, as illustrated in Figure 1. The overall workflow begins with planning for the AI model and then divides into two parallel tracks: (1) an Object Detection model branch and (2) a Machine Learning model branch. Both branches converge at a shared model evaluation stage, after which a decision is made on whether the model meets performance criteria. If the criteria are not met, the model is retrained with adjusted parameters. Once the criteria are satisfied, the model is built and deployed for clinical applications.

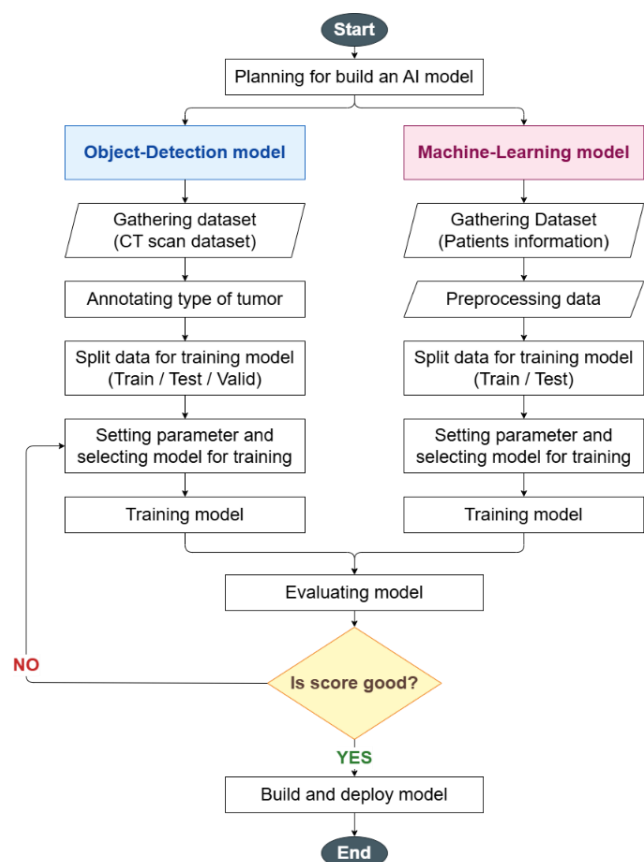


Figure 1 Overall research design flowchart for developing the AI-based tumor classification model.

Tumor and Nodule Identification

Pulmonary nodules are defined as small, well-defined, and typically spherical areas of abnormal tissue growth which could be benign or malignant, whereas tumors represent a broader classification encompassing abnormal growths of various sizes and morphologies. Accurately classifying these anomalies as either benign or malignant is fundamental to the development and training of a high-performance Object Detection Model. The differentiation between benignity and malignancy involves several physiological and radiological factors. Key indicators include calcification patterns, nodule size, growth rate, overall shape, and margin characteristics (such as spiculation). Additionally, location within the lungs is also crucial for the classification of the aforementioned anomalies, hence why the inclusion of multiple datasets in model training is vital to recording variation between the locations of the CT scan images. This information was primarily used to annotate the type of tumor under the Object Detection pathway. See below a comprehensive summary of these variables in determining the probability of whether a nodule is benign or malignant, presented in Table 1.

Table 1 summary of qualities used differentiating benign and malignant nodules on CT scans [13 – 18].

Qualities	Benign	Malignant
Calcification	Diffuse, central, laminated, or popcorn calcifications	Expect other calcification patterns
Size	Typically less than 3cm in size	Typically more than 3cm in size
Growth	Relative comparisons with prior imaging studies	
Shape	Three-dimensional ratio > 1.78	
Margin	Smooth margins (unless metastatic in origin)	Corona radiata signs
Location	Hamartoma: Peripheral areas	Adenocarcinoma: Peripheral areas
	Chondroma: Location varies	Squamous Cell Carcinoma: Central areas
	Carcinoid Tumors: Often central but may appear in other areas	Large Cell Carcinoma: Peripheral areas
		Small Cell Lung Cancer: Central areas

Probability of lung cancer

In addition to the symptom-based clinical data used to train machine learning models in this study, the established Mayo Clinic malignancy prediction model [34] considers several key radiological features for estimating the probability of malignancy. (1), and one of the most vital characteristics of tumors is the Nodule Diameter. Measured in millimeters at the widest point across all views, Nodule Diameter serves as a baseline indicator of growth, with tumors larger than 3 cm in diameter considered as malignant. (2) Spiculation, star-like patterns, are clinically recognized as a sign of malignancy. (3) The location of the tumor, specifically a position in the upper lobe, may indicate that the tumor is malignant. (4) Reported patient data, such as Age, Smoking History, and Family Cancer history, provides invaluable context for malignancy.

$$x = -6.8272 + (0.0391 * Age) \quad (1)$$

$$+ (0.7917 * SH) \\ + (1.3388 * CH) \\ + (0.1274 * DN) \\ + (1.0407 * SPICULATION) \\ + (0.7838 * UPPER LOBE)$$

$$PC = (e^x) / (1 + e^x) \quad (2)$$

- Age: In years.
- Smoke history (SH): 1 if current/former smoker, 0 if never.
- Cancer history (CH): 1 if history of cancer 5+ years ago, 0 if not.
- Diameter of nodules (DN): Largest diameter of the nodule in mm.
- Spiculation: 1 if spiculation (ragged edge) is present, 0 if not.
- Upper Lobe: 1 if the nodule is in the upper lobe, 0 if not.
- Probability of cancer (PC)

Multi-view morphological and positional analysis

While multi-view analysis provides a more complete perspective into tumors, this study emphasizes a focus on Axial View CT scans for Object Detection CT scan analysis. Although this may cause a small number of tumors to be missed due to the difficulty of precisely identifying tumor shapes during the labeling phase, this study is still willing to move forward with this blueprint. Seeing as a key focus of this study is on evaluating the performance of numerous AI models in screening for tumors, by focusing on one perspective, this study better enhances computational efficiency and reduces complexity, something which is essential for real-time or high-throughput processing in clinical environments. This streamlined approach allows the models to maintain high sensitivity in identifying benign and malignant nodules without compromising processing speed or accuracy. For

informative purposes, see below for the types of views considered for the model to built on:

- **Axial View:** A perspective that slices the body into top and bottom sections. Axial views can be used for initial nodule detection and horizontal localization. The images are presented in Figure 2 (a)
- **Coronal View:** This view images the body in slices from front to back; coronal views are used to determine the vertical position of the nodule relative to landmark organs like the heart and diaphragm. The images are presented in Figure 2 (b)
- **Sagittal View:** This view is critical for identifying the specific lung lobe, allowing radiologists to determine whether the nodule is in the upper or lower lobe and figure out the possibility of malignancy. The images are presented in Figure 2 (c).

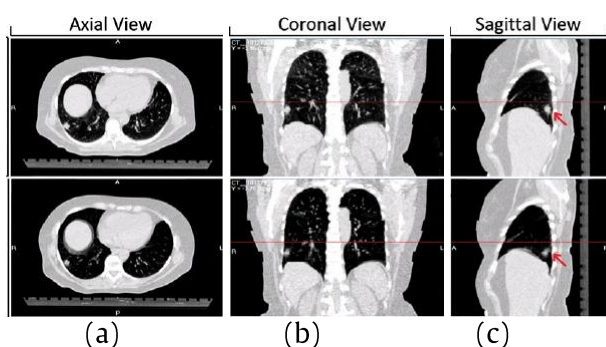


Figure 2 Morphology (Source: Li et al. 2012).

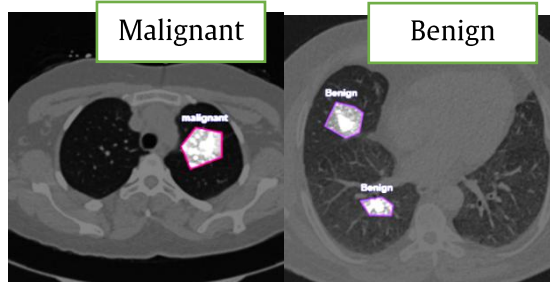


Figure 3 Annotating malignant and benign.

Dataset preparation

Two separate datasets are used for each model branch in this study. The Object Detection branch is built on CT scan images from the publicly available IQ-OTH/NCCD dataset, which contains annotated low-dose CT scans with labelled pulmonary nodules [19]. Each identified nodule is annotated as either benign or malignant based on radiologist consensus following established annotation protocols for pulmonary nodule classification. Samples of labelled CT scan images are presented in Figure 3. The Machine Learning branch, on the other hand, incorporates patients' demographic and symptom-based clinical information, such as age, smoking status, and respiratory symptoms, to compile a structured tabular dataset.

Both datasets undergo preprocessing steps, including normalization, removal of incomplete records, and class balancing using the Synthetic Minority Oversampling Technique (SMOTE) before model training [20].

Object detection model branch

The Object Detection branch is designed to identify and localize tumors directly within CT scan images. After gathering the CT scan dataset, each nodule is manually annotated with a bounding box and labeled as benign or malignant. The dataset is then split at the image level into training (note that these per-image patient identifiers were not exposed in the public release), validation, and test sets at a ratio of 70:15:15 [21]. The detection dataset comprised 649 axial CT images derived from the IQ-OTH/NCCD lung cancer dataset and annotated using Roboflow, partitioned into 456 training, 96 validation, and 97 test images and containing 277 benign and 1,173 malignant annotated nodules. Although this study acknowledges potential leakage risk within this internal split during training. However, final model performance was independently evaluated on a separate, Memorial Sloan Kettering Cancer Center (MSKCC) verified dataset [22] with no patient overlap with the training data, providing an external generalization check unaffected by the internal split's leakage risk.

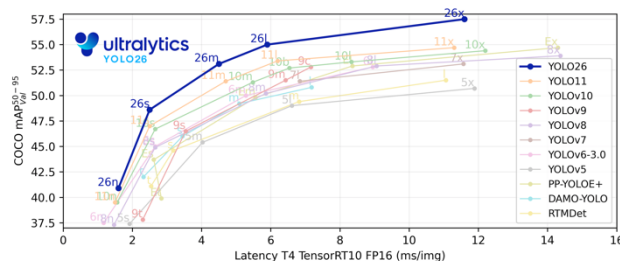


Figure 4 YOLO models (Source: Ultralytics, 2024).



Figure 5 Data augmentation techniques.

In the first stage, the bounding boxes were annotated by the researchers themselves, who identified nodule regions based on radiological characteristics based on radiological characteristics, including nodule size and location within the lung, using class labels provided by the IQ-OTH/NCCD dataset. The models were then trained and tested on these annotations. In the second stage, the same model predictions were compared against a separate dataset whose annotations had been confirmed by the MSKCC verified dataset [22], which provided a more reliable ground truth for

evaluating performance. Model parameters such as the number of training epochs, batch size, and learning rate are configured prior to training. See Figure 4 for the deep Malignant Benign learning frameworks based on YOLO models, which have been fine-tuned on the annotated CT scan images. Random flipping, rotation, and brightness adjustments are shown in Figure 5 in order to improve model generalizability in training.

1	GENDER	AGE	SMOKING	YELLOW_FINGERS	ANXIETY
2	M	69	1	2	2
3	M	74	2	1	1
4	F	59	1	1	1
5	M	63	2	2	2
6	F	63	1	2	1
7	F	75	1	2	1
8	M	52	2	1	1
9	F	51	2	2	2
10	F	68	2	1	2
PEER_PRESSURE	CHRONIC_DISEASE	FATIGUE	ALLERGY	WHEEZING	
1		1	2	1	2
1		2	2	2	1
2		1	2	1	2
1		1	1	1	1
1		1	1	1	2
ALCOHOL_CONSUMING	COUGHING	SHORTNESS_OF_BREATH			
2		2			2
1		1			2
1		2			2
2		1			1
1		2			2
1		2			2
2		2			2
1		1			2
SWALLOWING_DIFFICULTY	CHEST_PAIN	LUNG_CANCER			
2		2 YES			
2		2 YES			
1		2 NO			
2		2 NO			
1		1 NO			
1		1 YES			

Figure 6 Example of patient information.

Machine learning model branch

The Machine Learning branch uses structured patient data rather than raw images. Patient information including clinical features is shown in Figure 6 such as patient age, gender, smoking history, anxiety, peer pressure, chronic disease, fatigue, allergy, wheezing, alcohol consumption, coughing, shortness of breath, swallowing difficulty, and chest pain are gathered and preprocessed. Feature engineering steps include encoding categorical variables, handling missing values, and applying standardization. The dataset is evaluated using 10-fold stratified cross-validation, with SMOTE applied exclusively within each training fold to prevent data leakage; reported metrics represent the mean across all folds. Although the 10-fold stratified cross-validation is somewhat practical in preventing data leakage, it must be stated that the lack of a nested cross-validation may act as a limitation when it comes to ensuring that the final test set remains unseen. The clinical dataset comprised 309 patient records,

with 270 positive and 39 negative cases, reflecting a class imbalance of approximately 6.9:1 that motivated the use of SMOTE. Multiple machine learning classifiers are trained and compared, including Logistic Regression (Log), XGBoost (XGB), LightGBM (LGBM), CatBoost (CB), Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Naive Bayes (NB) [23]. A summary of the performance comparison across all models is presented in Table 4. Hyperparameter tuning is performed using grid search with cross-validation to identify the optimal model configuration for each algorithm [24, 25].

Model evaluation and deployment

Both model branches feed into a unified evaluation stage. Model performance is assessed using a comprehensive set of classification metrics derived from the confusion matrix. These metrics include accuracy (Eq.3), precision (Eq.4), recall or sensitivity (Eq.5), and F1 score (Eq.6) [26]. In these equations, TP denotes true positives, TN denotes true negatives, FP denotes false positives, and FN denotes false negatives.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

Since k-fold cross-validation is applied, the mean value of each metric across all k folds is used as the final reported performance indicator, where k denotes the number of cross-validation folds. The mean accuracy (MA), mean precision (MP), mean recall (MR), and mean F1 score (MF1) are defined in Equations 7 through 10 respectively.

$$MA = \frac{1}{k} \sum_{i=1}^k Accuracy_i \quad (7)$$

$$MP = \frac{1}{k} \sum_{i=1}^k Precision_i \quad (8)$$

$$MR = \frac{1}{k} \sum_{i=1}^k Recall_i \quad (9)$$

$$MF1 = \frac{1}{k} \sum_{i=1}^k F1\ Score_i \quad (10)$$



Figure 7 Representative CT scan images with model-predicted malignancy labels and confidence scores.

A model is considered to pass evaluation if it achieved an AUC-ROC of at least 0.85 and a mean sensitivity of at least 0.80 on the held-out test set, thresholds selected based on clinically accepted standards for diagnostic screening tools. If a model

fails to meet these criteria, the training parameters are adjusted and the model is retrained. Once the performance criteria are met by either or both branches, the final model is built and packaged for deployment as a decision-support tool that can be integrated into a clinical CT image review workflow, as shown in Figure 7.

Integrated Risk Scoring

To translate the complementary outputs of the two pipelines into a single, clinically interpretable measure, this study proposes an integrated risk score. Although the object detection and machine learning pipelines are trained independently on domain-appropriate datasets, both can be applied to the same patient at inference time: the object detection model contributes the predicted probability of malignancy, $P(\text{malignant})$, from the detected nodule, while the machine learning model contributes the predicted cancer probability, $P(\text{cancer})$, from the patient's clinical and symptom data. The two probabilities are weighted equally and combined into a composite score on a 0-100 scale, as defined in Equation 11:

$$\text{Risk Score} = (P(\text{malignant}) \times 50) + (P(\text{cancer}) \times 50) \quad (11)$$

Table 2 Representative examples of the proposed integrated risk score. $P(\text{mal})$: imaging-derived probability of malignancy; $P(\text{cancer})$: clinical probability of cancer; Score: composite value on a 0-100 scale.

Case	$P(\text{mal})$	$P(\text{cancer})$	Score	Risk
A	0.00	0.70	35.0	Moderate
B	0.85	0.60	72.5	High
C	0.20	0.15	17.5	Low

Here $P(\text{malignant})$ and $P(\text{cancer})$ each range from 0 to 1, yielding a final score between 0 and 100, where higher values indicate greater estimated risk. The equal 50:50 weighting reflects the absence of prior evidence favouring either modality and can be recalibrated once validation data become available. This formulation allows the framework to remain informative even when one modality is uncertain; for example, when no nodule is detected ($P(\text{malignant}) = 0$), the clinical branch still contributes a risk estimate, mirroring evidence that clinical predictors retain value in the absence of visible nodules. The use of complementary imaging and clinical predictors for risk stratification is consistent with established nodule-malignancy models [27, 28] and standardized risk-reporting systems such as Lung-RADS [29] and recent multimodal fusion approaches [30]. It should be emphasised that, because the two branches were trained on separate datasets, the integrated risk score is presented here as a proposed framework; its empirical validation requires a multimodal cohort

containing CT scans, clinical records, and confirmed diagnoses for the same patients, as noted in the limitations. Table 2 illustrates representative cases showing how the score behaves across differing imaging and clinical inputs.

RESULTS AND DISCUSSION

Machine learning model comparison

Table 3 and Figure 8 summarize the machine learning model performance across each evaluated metric, including accuracy, precision, recall, and F1-score. Among the evaluated models, logistic regression emerged as the top performer in terms of overall predictive accuracy, achieving the highest mean accuracy of 0.929. The CB model followed closely with an accuracy of 0.906 and demonstrated superior balance, yielding a strong F1-score of 0.945. Upon closer analysis of the models, it has become apparent that different models stand out in different areas of strength, in particular between efficiency and accuracy. Comparing the top machine learning architectures, the logistic regression model excels in its simplicity and efficiency while the CB model offers robust predictive power through its advanced gradient-boosting framework.

Table 3 Performance comparison of various machine learning models (default parameter).

Model	MA	MP	MR	MF1
Log	0.929	0.960	0.959	0.959
XGB	0.900	0.938	0.948	0.942
LGBM	0.900	0.938	0.948	0.942
CB	0.906	0.948	0.944	0.945
RF	0.903	0.941	0.948	0.944
DT	0.887	0.945	0.926	0.934
SVM	0.479	0.925	0.441	0.588
KNN	0.871	0.947	0.904	0.924
NB	0.900	0.920	0.970	0.944

Table 4 AUC comparison of various machine learning models (Default parameter).

Model	Mean AUC	Std AUC
Log	0.937	0.0535
XGB	0.9006	0.0847
LGBM	0.9043	0.0904
CB	0.9145	0.0817
RF	0.9224	0.0669
DT	0.7949	0.1423
SVM	0.6781	0.1346
KNN	0.8559	0.1439
NB	0.8986	0.0888

Following up on concerns of false negative evaluations, high recall models are vital to ensure that all potentially malignant nodules are accounted for, even if it results in a less harmful false positive outcome. The NB model fits this precautionary criterion, achieving the highest recall (MR) of 0.970. The NB model favors high sensitivity; however, it achieves lower precision than the ensemble methods. These results suggest that while NB is highly effective at minimizing false negatives, ensemble-based models like CB and RF offer a greater balance of metrics; making them highly suitable for robust clinical support systems.

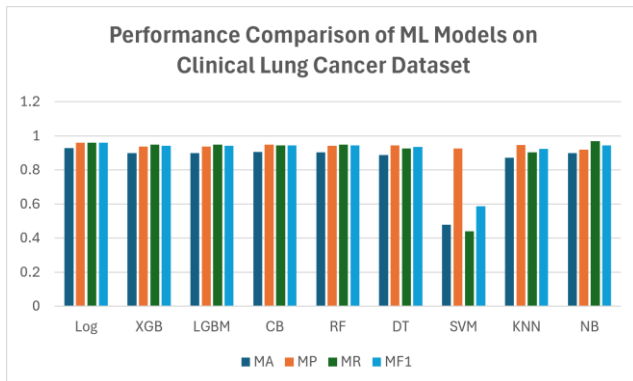


Figure 8 Performance Comparison of ML Models on Clinical Lung Cancer Dataset.

Table 5 Evaluation of CB Model Efficiency with Varying Iteration Values.

Iteration	MA	MP	MR	MF1
100	0.913	0.951	0.948	0.949
200	0.910	0.948	0.948	0.947
500	0.913	0.949	0.952	0.95
1000	0.919	0.956	0.952	0.953
2000	0.919	0.956	0.952	0.953

Table 6 Evaluation of CB Model Efficiency with Varying Eta Values.

Eta	MA	MP	MR	MF1
0.01	0.9100	0.948	0.948	0.947
0.05	0.919	0.956	0.952	0.953
0.15	0.916	0.955	0.948	0.951
0.20	0.919	0.959	0.948	0.953
0.30	0.916	0.955	0.948	0.951

Table 7 Evaluation of CB Model Efficiency with Varying Depth Values.

Depth	MA	MP	MR	MF1
2	0.900	0.938	0.948	0.942
4	0.919	0.953	0.956	0.953
6	0.913	0.949	0.952	0.950
8	0.919	0.956	0.952	0.953
10	0.916	0.956	0.948	0.951

Deep learning analysis

The experimental results summarized in Tables 8 and 9 provide a comparative evaluation of the

CB hyperparameter analysis

Further investigation of the CB model in Tables 5-7 shows that adjusting certain hyperparameters can directly affect diagnostic efficiency [31]. The number of iterations played a crucial role in model convergence. Increasing the iterations to 1,000 gave the optimal mean F1-score (MF1) of 0.953. Past this point, however, the model's performance slightly decreased and plateaued, indicating that beyond the 1,000 iteration threshold, no significant additional predictive gains are achieved. Another critical factor in the CB model is the depth of the decision trees in the dataset, with tree depths of 4 and 8 both yielding the highest recorded accuracy of 0.919 and an F1-score of 0.953; depth 8 was selected for its greater representational capacity.

Within the training process, Learning Rate (eta) was identified as key to ensuring model stability. Decrease yielded comparable results; however, 0.05 was selected for its stability in preventing overfitting. Furthermore, higher learning rates lead to a slight decrease in overall accuracy. Collectively, these findings highlight how systematic hyperparameter tuning is fundamental in maximizing model effectiveness in tumor classification.

performance of various YOLO architectures, emphasizing the influence of ground truth quality on model evaluation. Out of the YOLO architectures, YOLOv8n

is shown to be the most effective architecture for the dataset (see Table 7). Despite newer architectures such as YOLOv9 to YOLOv11, YOLOv8n outperformed these models and achieve the highest mAP50 of 0.3857 [32].

To further support YOLOv8n's effectiveness over additional models, the YOLOv6n performance must be addressed. While YOLOv6n exhibited the highest precision at 0.694, its significantly lower recall of 0.456 and an mAP50 of 0.232 may prove ineffective in real clinical settings. The aforementioned metrics in particular indicate that YOLOv6n operates with an overly conservative detection threshold, potentially leading to a higher rate of false negatives and therefore a likelihood that the model will miss positive instances. Conversely, YOLOv8n remains balanced on factors of precision and recall, allowing for high sensitivity in diagnostics. Additionally, the longitudinal performance of the YOLO models demonstrated a consistent upward trend in the number of epochs directly correlated with improved model convergence and detection accuracy (see in Figure 9), highlighting the importance of training duration in the optimization of deep learning performance in medical imaging.

Table 8 Performance comparison of various deep learning models.

Model	mAP50	Precision	Recall
YOLOv6n	0.232	0.694	0.456
YOLOv8n	0.385	0.647	0.435
YOLOv9c	0.346	0.657	0.429
YOLOv10n	0.311	0.602	0.432
YOLOv11n	0.372	0.664	0.446

Table 9 Performance evaluation of YOLOv8n with the 'MSKCC verified dataset' [22].

Model	mAP50 (Original Val)	mAP50 (MSKCC Verified Dataset)
YOLOv8n	0.385	0.687

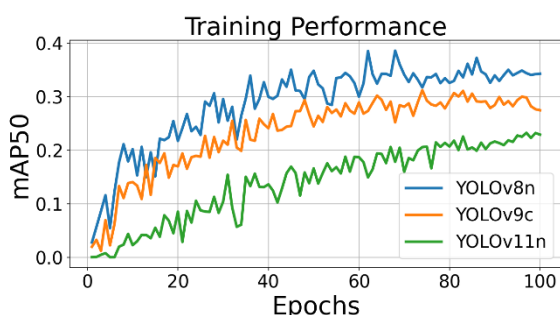


Figure 9 Training Performance (mAP50/Epochs).

A key observation across the YOLO architectures is the relatively low recall, with YOLOv8n achieving a recall of 0.435 against the researcher-annotated

test set. In a cancer screening context, recall is a critical metric, since missed malignant nodules (false negatives) carry the greatest clinical risk. Several factors likely contributed to this limitation. First, the relatively small size of the annotated dataset (approximately 1,000 images) constrains the ability of the model to generalize across the morphological variability of pulmonary nodules. Second, the first-stage bounding boxes were annotated by the researchers rather than qualified radiologists, introducing label noise that disproportionately reduces recall. The increase in mAP50 from 0.385 to 0.687 under the MSKCC verified dataset (Table 9) supports this interpretation, indicating that a substantial portion of the apparent error originated from ground-truth quality rather than model capacity. Future work should therefore prioritise expanding the dataset, obtaining expert annotations, and applying nodule-focused augmentation to raise sensitivity to a clinically acceptable level.

Comparison with previous studies and limitations

When situated against prior work, the present findings are consistent with the broader literature in which deep learning systems for lung nodule analysis report strong performance under expert-curated ground truth while remaining sensitive to annotation quality and dataset scale [10, 11]; the improvement observed under MSKCC verified dataset [22] aligns with this pattern. Nevertheless, several limitations should be acknowledged. First, the deep learning and machine learning pipelines were trained and evaluated on separate datasets, and the proposed integrated risk score is therefore presented as a conceptual framework rather than an empirically validated outcome. Second, the object detection dataset is limited in size and was partially annotated by non-experts, which constrains generalisability. Third, the detection data were split at the image level rather than the patient level, and no external multi-institutional validation was performed, limiting evidence for cross-population robustness. Finally, the clinical machine learning dataset is symptom-based and does not include the radiological measurements used in established models such as the Mayo Clinic equation [27]; integrating these complementary feature sets within a single multimodal cohort represents an important direction for future work. These limitations temper any claim of immediate clinical deployment and instead position the study as evidence supporting the feasibility of AI-assisted lung cancer screening.

CONCLUSIONS

This research evaluated the effectiveness of deep learning and machine learning techniques in differentiating between benign and malignant lung nodules in CT scans. Through the use of a two-branch

parallel pipeline consisting of an object detection branch for image analysis and a Machine Learning branch for clinical data, the study provides a comprehensive approach to automated lung cancer screening.

Performance of machine learning models

This study has demonstrated the effectiveness of machine learning algorithms for classification tasks by incorporating contextual patient data and radiological features. The log models showed excellent accuracy with a mean accuracy of 0.929, making them most optimal for screening accuracy. Additionally, hyperparameter optimization of the CB model yielded a mean F1 score of 0.953 with 1,000 iterations and a tree depth of 8. These findings demonstrate the above adjustments will best improve the performance of the CB model.

Performance of object detection models

The object detection framework was able to classify benign and malignant tumors based on CT scan images. The YOLO architectures were compared, YOLOv8n was found to be the best model in terms of mAP50, precision and recall. Notably, the performance of YOLOv8n was heightened to mAP50 of 0.6875 when tested against datasets verified by medical experts. These results highlight the potential of deep learning architectures for tumor classification and the crucial need for high-quality expert annotated data to build trustworthy AI tools for the clinic. The low recall of 0.435 is likely influenced by the relatively small size of the IQ-OTH/NCCD dataset (110 cases, 1,190 images) and the use of single-annotator labeling. While bounding boxes were drawn based on established radiological criteria, the absence of multi-expert consensus annotation may have introduced variability in ground truth quality, contributing to missed detections. Future work should prioritize expanding the dataset with direct physician-verified annotations, which, as demonstrated in this study through the MSKCC verified dataset [22], improved mAP50 from 0.385 to 0.687."

Clinical significance and future work

The findings of this study underscore the potential of AI-driven systems to support clinical decision-making tools in clinical workflows. Through the incorporation of key identifying features such as nodule diameter, spiculation, and anatomical location into a predictive probability equation, along with deep learning object detection models designed to differentiate between benign and malignant tumors in CT scan images; dual-pipeline frameworks such as the one in this study show great promise towards early detection and improving survival outcomes for lung cancer patients. Future research should focus on enhancing model generalizability across diverse patient populations and further improving the

transparency of black box algorithms for to facilitate their potential clinical application, future studies may also develop a more practical integrated risk scoring system using a unified multimodal cohort.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

The author utilized ChatGPT to improve the legibility and clarity of the text during the preparation of this work.

ACKNOWLEDGEMENT

The author would like to thank Mr. Pichayapazz Naewthong and Mr. Dechanin Tengcharoenkul for their support and facilities provided throughout this research.

REFERENCES

1. World Health Organization. Cancer Son Erişim Tarihi. 2021;24. Available from: <https://www.who.int/news-room/fact-sheets/detail/cancer>.
2. Neacşu F, Vârban AŞ, Simion G, Şurghie R, Pătraşcu OM, Sajin M, et al. Lung cancer mimickers—a case series of seven patients and review of the literature. *Rom J Morphol Embryol*. 2021;62(3):697-704.
3. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA: a cancer journal for clinicians*. 2023;73(1):17-48.
4. Ellis PM, Vandermeer R. Delays in the Diagnosis of Lung Cancer. *J Thorac Dis*. 2011;3(3):183-8.
5. Ahmed S, Subhan F, Mohd Su'ud M, Alam MM, Waheed A. A Survey of Lung Nodules Detection and Classification from CT Scan Images. *Comput Syst Sci Eng*. 2024;48(6):1483–511
6. Hanna TN, Lamoureux C, Krupinski EA, Weber S, Johnson JO. Effect of Shift, Schedule, and Volume on Interpretive Accuracy: A Retrospective Analysis of 2.9 Million Radiologic Examinations. *Radiology*. 2018;287(1):205-12.
7. Mienye ID, Swart TG, Obaido G, Jordan M, Ilono P. Deep Convolutional Neural Networks in Medical Image Analysis: A Review. *Information*. 2025;16(3):195.
8. Wang M, Chang W, Zhang Y. Artificial Intelligence for the Diagnosis and Management of Cancers: Potentials and Challenges. *MedComm*. 2025;6(11):e70460.
9. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature medicine*. 2019;25(6):954-61.

10. Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, et al. Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. *Radiology*. 2019;290(1):218-28.
11. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS medicine*. 2018;15(11):e1002686.
12. Yang J, Soltan AA S, Clifton DA. Machine Learning Generalizability across Healthcare Settings: Insights from Multi-Site COVID-19 Screening. *Npj Digital Medicine*. 2022;5(1). doi: 10.1038/s41746-022-00614-9.
13. Zhen F, Schraeder R. The Changing Pathology of Lung Cancer. *Surg Oncol Clin N Am*. 2011;20(4): 637-53.
14. Constantin AA, Cotea AA, Mihăițan FD. The Multiple Faces of Lung Adenocarcinoma: Challenges in Diagnosis. *Internal Medicine*. 2023;20(3):69-80.
15. Rudin CM, Brambilla E, Faivre-Finn C, Sage J. Small-cell lung cancer. *Nature Reviews Disease Primers*. 2021;7(1):3.
16. Tian D, Wen H, Zhou Y, Fu M. Pulmonary chondroma: A clinicopathological study of 29 cases and a review of the literature. *Mol Clin Oncol*. 2016;5(3): 211-5.
17. Ann L, Smithuis R. The Radiology Assistant: Benign versus Malignant. Available from: <https://radiologyassistant.nl/chest/solitary-pulmonary-nodule/benign-versus-malignant>.
18. Onno M, Smithuis R. The Radiology Assistant: Fleischner 2017 Guideline. Available from: <https://radiologyassistant.nl/chest/plumonary-nodules/fleischner-2017-guideline>.
19. Al-Yasriy HF, AL-Husieny MS, Mohsen FY, Khalil EA, Hassan ZS. Diagnosis of Lung Cancer Based on CT Scans Using CNN. *IOP Conference Series: Materials Science and Engineering*. 202;928 (2):022035.
20. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321-57.
21. Joseph VR. Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*. 2022;15(4):531-8.
22. The Coffee-break lung CT collection with scan images reconstructed at multiple imaging parameters [Internet]. The Cancer Imaging Archive (TCIA). 2025. Available from: <https://doi.org/10.7937/k9/tcia.2015.u1x8a5nr>.
23. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data*. 2019;6(1):1-48.
24. Breiman L. Random forests. *Machine learning*. 2001;45(1):5-32.
25. Bergstra J, Bengio Y. Random search for hyperparameter optimization. *J Mach Learn Res*. 2012 Feb 1;13(2):281 - 305.
26. Fawcett T. An introduction to ROC analysis. *Pattern recognition letters*. 2006;27(8):861-74.
27. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Archives of Internal Medicine*. 1997;157(8):849-55.
28. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of cancer in pulmonary nodules detected on first screening CT. *New England Journal of Medicine*. 2013;369(10):910-9.
29. Lung CT. Screening Reporting and Data System (Lung-RADS) version 1.1. American College of Radiology Web site [Internet]. 2019.
30. Huang H, Zheng D, Chen H, Wang Y, Chen C, Xu L, et al. Fusion of CT images and clinical variables based on deep learning for predicting invasiveness risk of stage I lung adenocarcinoma. *Medical Physics*. 2022;49(10):6384-94.
31. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*. 2018;31. Available from: <https://doi.org/10.48550/arXiv.1706.09516>
32. Jocher G, Chaurasia A, Qiu J. Ultralytics YOLOv8. 2023. Available from: <https://github.com/ultralytics/ultralytics>.



Institute of Research and Development
Rajamangala University of Technology Thanyaburi

39 M.1 Klong 6, Thanyaburi, Pathumthani 12120, Thailand
Tel. (02) 549-4492 Fax. (02) 577-5038, (02) 549-4680

Website : <https://ird.rmUTT.ac.th>



THAIJO