

การตรวจจับพฤติกรรมต้องสงสัยระหว่างสอบออนไลน์ด้วยตัวกรองอนุภาค แบบราวแบล็คเวลล์ไลซ์

DETECTION OF SUSPICIOUS HUMAN BEHAVIOR DURING ONLINE EXAM WITH RAO-BLACKWELLIZED PARTICLE FILTER

นัศพ์ชาณัน ชินปัญชรณะ¹ และ เตจตัสสินปี เพี้ยชัย²

¹อาจารย์, วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิต
110/1-4 ถ.ประชาชื่น เขตหลักสี่ กรุงเทพฯ 10210, nutchanun.cha@dpu.ac.th

²อาจารย์, สาขาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช ถ.แจ้งวัฒนะ
เขตปากเกร็ด นนทบุรี 11120, tejtasin@gmail.com

Nutchanun Chinpanthana¹ and Tejtasin Phiasai²

¹Lecturer, College of Innovative Technology and Engineering, Dhurakij Pundit University,
110/1-4 Prachachuen rd. Laksi, Bangkok 10210, Thailand, nutchanun.cha@dpu.ac.th

²Lecturer, School of Science and Technology, Sukhothai Thammathirat Open University,
Chaengwattana rd. Bangpood, Pakkret, Nonthaburi 11120, Thailand, tejtasin@gmail.com

บทคัดย่อ

ปัจจุบันการจัดการเรียนการสอนรวมทั้งการประเมินผลหลังเรียนจะอยู่ในรูปแบบออนไลน์ การสอบออนไลน์ส่วนใหญ่จะเห็นเพียงใบหน้าของผู้เข้าสอบทำให้ประสิทธิภาพของผู้คุมสอบถูกจำกัด มีงานวิจัยในสาขาคอมพิวเตอร์วิชั่นที่แสดงการตรวจจับรูปแบบพฤติกรรมของการทุจริตแต่ยังคงเป็นปัญหาเนื่องจากรูปแบบของการแสดงออกทางพฤติกรรมมีหลากหลาย รวมทั้งมีนักวิจัยพยายามนำอัลกอริธึมเข้ามาใช้ในการตรวจจับรูปแบบของพฤติกรรมในการสอบ แต่ยังคงมีข้อจำกัดในหลายกรณี ดังนั้นในงานวิจัยนี้จึงได้นำเสนอ การตรวจจับพฤติกรรมต้องสงสัยระหว่างสอบออนไลน์ด้วยตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์ โดยใช้การสกัดคุณลักษณะตำแหน่งสำคัญจากเฟรมข้อมูลด้วย OpenPose เพื่อนำมาใช้ในการจำแนกตรวจหาพฤติกรรมต้องสงสัย วิธีการดำเนินงานประกอบด้วย 3 ขั้นตอนหลักดังนี้ (1) การเตรียมข้อมูล (2) การตรวจจับพฤติกรรมด้วยตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์ (3) การวัดประสิทธิภาพและประเมินผลจากการทดลอง จากผลการทดลองสามารถระบุได้ว่า การใช้แบบจำลองตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์เพื่อตรวจจับพฤติกรรมต้องสงสัยในการสอบออนไลน์สามารถเพิ่มประสิทธิภาพของการจำแนกได้ดียิ่งขึ้นถึงร้อยละ 79.63 ด้วยชุดข้อมูลที่ 2

คำสำคัญ: การประมวลผลภาพ, พฤติกรรมมนุษย์, ท่าทางมนุษย์, ตัวกรองอนุภาคแบบราวแบล็คเวลล์

ABSTRACT

Online exams have become widely used to evaluate students' performance in recent years. The ability to efficiently proctor is an important limiting for online exams due to the lack of face to face interaction. Prior research in computer vision has shown that online exams are more vulnerable to various cheating behaviors. The problem is challenging due to its various behavior activities in examination or by monitoring them. There are many algorithms attempt to analyze and detect in human behavior activities. But techniques are the limits in various conditions. To overcome this problem, this paper proposes a detection of suspicious human behavior with Rao-blackwellized particle filter. The proposed model extracts the keypoint features in an automated way with OpenPose. This paper presents main three steps including (1) data preprocessing, (2) detect behavior with Roa-blackwellized particle filter, and (3) efficiency measurement and evaluation of experimental results. The experimental results indicate that our framework offers performance improvements. The proposed model can achieve significant improvements for behaviour detection with maximum success rate of 79.63% with data set 2. The experimental results indicate that our proposed approach offers significant performance improvements in the detection of suspicious human behaviour.

KEYWORDS: image processing, human behaviour, human action, rao-blackwellized particle filter

1. บทนำ

ปัจจุบันเทคโนโลยีที่มีการพัฒนาอย่างรวดเร็วอีกทั้งสถานการณ์ได้มีการปรับเปลี่ยนรูปแบบการเรียนรู้มาเป็นออนไลน์มากขึ้นทำให้ผู้สอนต้องประเมินความรู้อของผู้เรียนในรูปแบบข้อสอบและแบบทดสอบที่อยู่บนออนไลน์เช่นกัน ดังนั้นเมื่อผู้เรียนมีการทดสอบผ่านทางออนไลน์ตามเกณฑ์ที่กำหนด สามารถรับใบรับรองหลักสูตรชั้นสุดท้ายได้หรือผ่านรายวิชาในหลักสูตรนั้น ๆ ในขณะเดียวกันโรงเรียนและมหาวิทยาลัยทั่วไปยังขยายหลักสูตรออนไลน์อย่างต่อเนื่อง ทำให้หลักสูตรออนไลน์ กำลังเป็นที่นิยมมากขึ้น [1] ตัวอย่างเช่น Coursera และ EdX เป็นต้น แต่อย่างไรก็ตามสิ่งที่สำคัญอย่างหนึ่งสำหรับการสอบออนไลน์คือวิธีการคุมสอบออนไลน์อย่างไรให้เกิดประสิทธิภาพ และมีความน่าเชื่อถือได้มากที่สุด จากงานวิจัยก่อนหน้า [2] แสดงให้เห็นว่าการสอบ

ออนไลน์นั้นเสี่ยงต่อพฤติกรรมการทุจริต จากแบบสำรวจมีถึง 74% ของผู้เข้าสอบออนไลน์บอกว่าการสอบออนไลน์ง่ายต่อการทุจริต และ เกือบ 29% ของผู้เข้าสอบออนไลน์ที่ระบุว่าทุจริตในการสอบออนไลน์ เนื่องจากความแตกต่างของการคุมสอบในห้องเรียนที่ผู้คุมสอบสามารถโต้ตอบและตรวจตราผู้เข้าสอบได้โดยตรง แต่การสอบออนไลน์นั้นไม่สามารถทำได้ ทำให้ต้องมีรูปแบบที่แตกต่างจากการสอบแบบเดิมๆ ด้วยรูปแบบพฤติกรรมการทุจริตแบบต่าง ๆ ที่อาจเกิดขึ้นในการสอบออนไลน์ [3, 4] เช่น การค้นหาข้อมูลบนอินเทอร์เน็ต การส่งข้อความไปยังบุคคลที่สาม การส่งข้อความไปทางแชทออนไลน์ หรือการใช้แอปพลิเคชันอื่นในการช่วยทำข้อสอบ ซึ่งทำให้ไม่ง่ายเลยที่จะตรวจจับพฤติกรรมในการทุจริต และเป็นผลให้เกิดการทุจริตได้ง่ายขึ้น ความน่าเชื่อถือของการสอบออนไลน์ลดลง

ปัญหาที่เกิดขึ้นจากการสอบออนไลน์ทำให้มีนักวิจัยหลายกลุ่มได้ Roger [5] ได้ทดลองติดตั้งโปรแกรมคุมสอบที่มีการรักษาความปลอดภัยลงในคอมพิวเตอร์เพื่อสามารถทำงานต่อเนื่องได้ด้วยการตรวจสอบหน้าจอคอมพิวเตอร์ของผู้เข้าสอบแต่ละคนในมุมมองการควบคุมผู้สอน ในทำนองเดียวกัน Cluskey et al [6] ได้นำเสนอแบบจำลองที่มีส่วนช่วยลดการกระทำทุจริตของผู้เข้าสอบ เช่น จำกัดระยะเวลาของการสอบ, จำกัดเวลาที่จัดสรรในการตอบแต่ละคำถาม เป็นต้น แต่อย่างไรก็ตามแบบจำลองนี้มีข้อจำกัดเนื่องจากต้องการให้ผู้สอบใช้เบราว์เซอร์เพื่อเข้าใช้งาน แอปพลิเคชันสอบอีกทั้งมีความจำเป็นที่ข้อสอบต้องเปลี่ยนคำถามทุกภาคการศึกษาเช่นกัน มีกลุ่มนักวิจัย Bella et al [7] ที่ได้ควบคุมความปลอดภัยในการเข้าถึงข้อมูลด้วยการเข้ารหัสผ่าน รวมทั้งการตรวจสอบการใช้งานโปรโตคอล และการถ่ายโอนข้อมูลไปยังบุคคลที่สาม เครื่องมือตรวจจับการลอกเลียน (Plagiarism detection) ได้รับความนิยมในการตรวจสอบเนื้อหา การคัดลอกส่วนโปรแกรม หรือความคล้ายคลึงของการคำนวณที่ไม่ได้รับอนุญาต เทคนิคการตรวจจับการลอกเลียนด้วยเทคนิคการวัดค่าความคล้ายกันของลำดับโทเค็น [8] และการใช้วิธีการสร้างแบบจำลองด้วยการจับคู่คุณลักษณะกราฟ [9, 10] ได้ป้องกันการทุจริตด้วยการตรวจสอบการลอกเลียนด้วยการเปรียบเทียบข้อสอบ โดยที่การวิเคราะห์คือการอ้างอิงจากการใช้โทเค็นและการหาค่าเฉลี่ยของคุณสมบัติต่าง ๆ ผลการศึกษาพบว่ามีข้อจำกัดในวิธีการ เนื่องจากการตรวจจับการลอกเลียนแบบนั้นขึ้นอยู่กับ การคำนวณเกณฑ์ระดับของความคล้ายคลึงกันในเนื้อหามาตรฐาน ทำให้เกิดข้อผิดพลาดได้ [11, 12]

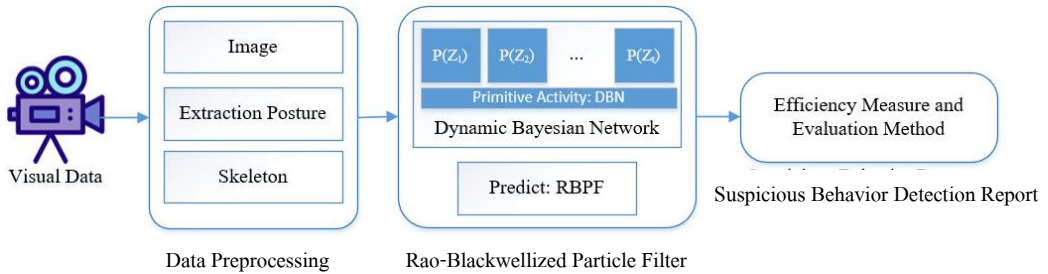
การตรวจจับรูปแบบเคลื่อนไหวของศีรษะของผู้เข้าสอบเป็นอีกวิธีหนึ่งที่มีกลุ่มนักวิจัย [13] ได้นำมาใช้ในการตรวจจับพฤติกรรมต้องสงสัยของผู้เข้าสอบจากตัวชี้วัดทางสถิติที่บันทึกได้ระหว่างการสอบ ด้วยการคาดการณ์จากรูปแบบการขยับศีรษะและการวางตำแหน่งของมุมศีรษะ เช่น การหัน เอียง ม้วน หรือดึงศีรษะ เป็นต้น Prathish et al [14] และ Narayanan et al [15] ได้ตรวจสอบความผันแปรของมุมเอียง การใช้เสียงและการจับภาพหน้าจอที่ใช้งานอยู่สามารถช่วยเฝ้าติดตามผู้เข้าสอบออนไลน์ได้ และ Wlodarczyk et al [16] ได้นำเสนอวิธีการตรวจจับท่าทางศีรษะร่วมกับ

จุดสังเกตใบหน้าที่จะช่วยในการระบุทิศทาง การจ้องมองของผู้ใช้ด้วยเป็นการจดจำใบหน้า ต่อมาในภายหลังมีหลายกลุ่มนักวิจัย Hu et al [17] Ghizlane et al [18-19] และ Garg et al [20] ได้พยายามนำรูปแบบการจดจำใบหน้าเข้ามาติดตามพฤติกรรมของระหว่างการประชุมออนไลน์มากขึ้น แต่อย่างไรก็ตามได้มีการพิจารณาในรูปแบบอย่างอื่นร่วมด้วย เช่น การสร้างความสัมพันธ์ที่เกี่ยวข้องของการกดแป้นพิมพ์ ระบบสามารถเก็บติดตามความเคลื่อนไหวภายในกรอบเวลาของการสอบ และ โครงข่ายประสาทเทียมตรวจสอบใบหน้าของผู้สอบ มีกลุ่มนักวิจัยบางกลุ่มพยายามวิเคราะห์พฤติกรรมเคลื่อนไหวของเมาส์เพื่อสร้างเป็นแบบจำลอง [21] ด้วยการใช้ความหนาแน่นเวลาในการขยับเมาส์ ความถี่ของการเคลื่อนไหวของเมาส์ และลำดับการเปลี่ยนตำแหน่งของเมาส์เพื่อนำมาใช้ในการตรวจจับกิจกรรมที่ผิดปกติในรูปแบบวิดีโอด้วยวิธีมาร์คอฟ (Markov) ในการติดตามการเคลื่อนที่ แต่จะเกิดปัญหาที่ไม่สามารถระบุรูปแบบทำงานได้เพราะรูปแบบที่ได้เป็นการจำเป็นระยะสั้นเท่านั้น มีคณะนักวิจัยของ Yang et al [22] ใช้ Hierarchical Dirichlet Process สำหรับการตรวจหากิจกรรมจากความสัมพันธ์ของกิจกรรมที่ถูกกละเลยไปของพฤติกรรมผู้เข้าสอบ แต่อย่างไรก็ตามการให้ความหมายของกลุ่มกิจกรรมยังไม่สามารถทำได้ทุกกรณี เนื่องจากบางกิจกรรมมีความซับซ้อนจนเกินไปทำให้ยังต้องมีข้อจำกัด สำหรับงานวิจัยนี้ได้นำเทคนิคเบย์เซียน (Bayesian) เป็นพื้นฐานของการหาค่าความน่าจะเป็นในสภาพแวดล้อมแปรผันตลอดเวลา ร่วมกับท่าทางภาษากาย (body language gestures) [23, 24] ที่สามารถสื่อถึงความหมายทางภาษา (sign language) ได้จากรูปแบบการเคลื่อนไหว เพื่อตรวจจับพฤติกรรมต้องสงสัยของผู้เข้าสอบในการสอบออนไลน์ด้วยวิธีการทำนายท่าทางจากตัวกรองอนุภาคแบบราวแบล็คเวลไลซ์ (Rao-Blackwellized Particle Filter) [25]

2. ขั้นตอนการดำเนินงานวิจัย

งานวิจัยนี้ได้พัฒนาวิธีการตรวจจับพฤติกรรมที่ต้องสงสัยทุจริตในการสอบออนไลน์ ประกอบด้วยขั้นตอนการดำเนินงานวิจัยคือ การเตรียมข้อมูล (Data Preprocessing) แบบจำลองตัวกรองอนุภาคแบบราวแบล็คเวลไลซ์ (Rao-Blackwellized Particle Filter) [25] และ รายการตรวจจับพฤติกรรมต้องสงสัย (Suspicious Behavior Detection Report) ดังแสดงในรูปที่ 1 โดยการทดลองได้นำข้อมูลจากฐานข้อมูลวิดีโอในระบบ Online Exam Proctoring (OEP) [26] ที่มีการกำหนดรูปแบบกิจกรรมที่ชัดเจนว่ากิจกรรมใดที่เป็นการส่อทุจริต และกิจกรรมใดที่อนุญาตให้สามารถกระทำได้สำหรับการสอบแบบเรียลไทม์ที่ประกอบด้วยหลายกลุ่มท่าทางรวมเข้ากันเป็นกิจกรรมหลักปกติ หรือท่าทางปกติ (Primitive activity) ในระหว่างการสอบประกอบด้วย กลุ่มท่าทางนั้น และขยับสลับเปลี่ยนไปเป็นรูปแบบท่าทางก้ำกัมนั้น เป็นต้น แต่ถ้ากิจกรรมที่ผิดปกติหรือกิจกรรมต้องสงสัย (Suspicious activity) ประกอบด้วย กลุ่มท่าทางนั้นและขยับสลับเปลี่ยนไปเป็นรูปแบบ

ท่าทางคดียุทธศาสตร์ เป็นต้น หรือกิจกรรมที่มีการสนทนากับบุคคลอื่น กิจกรรมที่มีการคดียุทธศาสตร์ หรือกิจกรรมที่มีการเดินไปมานอกเหนือจากการนั่งสอบที่โต๊ะ จะถือเป็นกิจกรรมที่ส่อทุจริต



รูปที่ 1 ขั้นตอนการตรวจจับพฤติกรรมต้องสงสัยด้วยตัวกรองอนุภาคแบบราวแบล็คเวลไลซ์

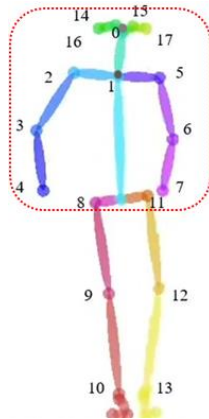
2.1 การสกัดข้อมูลตำแหน่งสำคัญบนโครงร่างกระดูก

การเตรียมข้อมูลภาพเพื่อนำมาวิเคราะห์พฤติกรรมระหว่างการสอบจากท่าทางและรูปแบบการเคลื่อนไหว ด้วยการสกัดข้อมูลตำแหน่งสำคัญบนร่างกาย เพื่อนำมาวิเคราะห์ภาพรวมที่ได้จากท่าทางและความเร็วของการเคลื่อนไหว เป็นทฤษฎีพื้นฐานทางชีววิทยามนุษย์ทางด้านกลศาสตร์ ด้วยการสกัดข้อมูลและตำแหน่งจากการเคลื่อนที่ของร่างกาย (Human action movement) จากแต่ละตำแหน่งบนจุดสำคัญบนร่างกายด้วยมุมมองที่เกิดขึ้น (Movement trajectory) บนโครงร่างจำลองรูปแบบสองมิติที่เรียกว่า โครงร่างกระดูก แสดงถึงตำแหน่งข้อต่อที่เป็นจุดสนใจทั้งหมด 15 จุดเป็นตำแหน่งอ้างอิง (reference point) ส่วนร่างกายท่อนบน และส่วนร่างกายด้านล่าง ประกอบด้วย ส่วนหัว คอ ไหล่ แขนส่วนบน แขนส่วนล่าง สะโพก เข่า และ ข้อ ดังแสดงในรูปที่ 2 สำหรับการทดลองนี้ได้นำเพียงส่วนร่างกายท่อนบน (upper body) ประกอบด้วยส่วนหัว (head) คอ (neck) ไหล่ (shoulder) แขนส่วนบน (upper arms) แขนส่วนล่าง (forearms) ข้อ มาทำการวิเคราะห์รูปแบบการเคลื่อนไหวของตำแหน่งข้อต่อสำคัญบนร่างกาย สำหรับข้อมูลภาพในแต่ละเฟรมถูกสกัดด้วย OpenPose [27] ประกอบด้วย จุดเชื่อมต่อบนโครงร่างกายทั้งหมด N จุดที่ถูกสกัดออกมาซึ่งในบางครั้งการสกัดข้อมูลจะมีบางตำแหน่งที่ขาดหายไปในช่วงเฟรมแต่จะทำการเชื่อมต่อดูจุดเหล่านั้นเพื่อให้สามารถทำงานได้ด้วย OpenPose และ $\beta_k^i = (x^i, y^i)$. เมื่อ i แทนตำแหน่งของจุดบนโครงร่างกายที่ภาพ k และถ้า OpenPose ไม่สามารถตรวจจับตำแหน่งจุดข้อต่อที่ i ในภาพที่ k จะส่งค่ากลับเป็น $\beta_k^i = \emptyset$, และกำหนดให้ $\beta_k^i = \beta_{k-1}^i$, ถ้า $\beta_k^i = \emptyset$, และ $\beta_k^i \neq \emptyset, k = 1, \dots, k-1$ และ $j^i \leq j$ เมื่อ j^i แทนเป็นตำแหน่งข้อต่อที่ i โดยกำหนดให้ j มีค่าเป็น 7 และจะแสดงตำแหน่งแต่ละจุดบนโครงร่างกายเฉพาะท่อนบนที่มีทั้งหมด 8 จุด ดังนั้น β^0, \dots, β^7 , เมื่อ β^0 เป็นจุดเริ่มต้นในตำแหน่งข้อต่อที่ต้นคอ ดังแสดงในรูปที่ 2 และ OpenPose ปรับ

ด้วยการทำนอร์มอลไลซ์เพื่อพิจารณาความลึกจากตำแหน่งพิกัดข้อต่อที่ถูกเคลื่อนที่ตามความลึกของคอ d_n ในแต่ละภาพ ดังนั้นข้อต่อทั้งหมดจะถูกแทนที่ตามสามารถเขียนเป็นสมการได้

$$\beta^i \leftarrow (\beta^i - \beta^0)/\tilde{d}_n, \quad (1)$$

เมื่อ $\tilde{d}_n = (x_n, d_n; \theta_n)$, กำหนดให้ \tilde{d}_n แทนค่าประมาณของความลึก และ x_n แทนเวกเตอร์ท่าทาง (augmented pose vector) θ_n แทนค่ามุมระหว่างสองจุดที่เชื่อมต่อกันบนโครงร่างมนุษย์ และ n แทนค่าของตำแหน่งของข้อต่อด้านบน (upper-body joint coordinates) สำหรับการทดลองใช้เฉพาะส่วนร่างกายท่อนบนที่มี 8 ข้อต่อ ประกอบด้วยตำแหน่ง ต้นคอ คอ ไหล่ซ้าย ไหล่ขวา ต้นแขนซ้าย ต้นแขนขวา แขนซ้าย และ แขนขวา ดังแสดงในรูปที่ 2

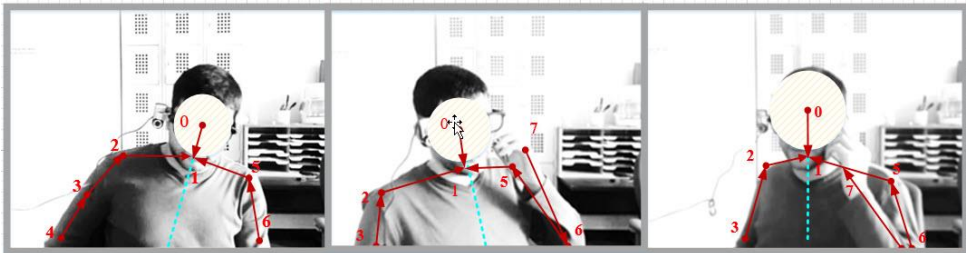


รูปที่ 2 ภาพจำลองโครงร่างกระดูก ตำแหน่งโครงร่างกระดูก OpenPose [23]

2.2 การวิเคราะห์พฤติกรรมต้องสงสัย

การวิเคราะห์พฤติกรรมต้องสงสัย เกิดจากการตีความหลายๆชุดกิจกรรมที่มีความต่อเนื่องกันตามเวลา จากการเรียนรู้ข้อมูลจำนวนหนึ่งจนสามารถแปลความหมายของชุดกิจกรรมได้ กิจกรรมเหล่านั้นมักมีลักษณะความสัมพันธ์ที่ซับซ้อน เช่น กิจกรรมคุยโทรศัพท์ -> การหยิบโทรศัพท์ขึ้นมา -> การถือโทรศัพท์แนบหู -> การพูดโทรศัพท์ -> การวางโทรศัพท์ รูปแบบความสัมพันธ์ของการเคลื่อนไหวท่าทางเหล่านี้สามารถแสดงถึงความหมายของกิจกรรมต้องสงสัย วิธีการไดนามิกเบย์เซียนเน็ตเวิร์ค (Dynamic Bayesian Network: DBN) [28] เป็นวิธีการหนึ่งที่สามารถคาดหมายความน่าจะเป็นจากเหตุการณ์ที่เคยเกิดขึ้นด้วยการประมาณค่า โดยจะทำการแทนข้อมูลความสัมพันธ์ด้วย กราฟที่แทนส่วนเวลาในระบบที่มีกราดบันทึกไว้ในรูปแบบของโหนด และค่าของตัวแปร สิ่งที่สามารถแทนความสัมพันธ์ในลักษณะแบบนี้ได้ดีที่สุดคือ การอนุมานแบบ

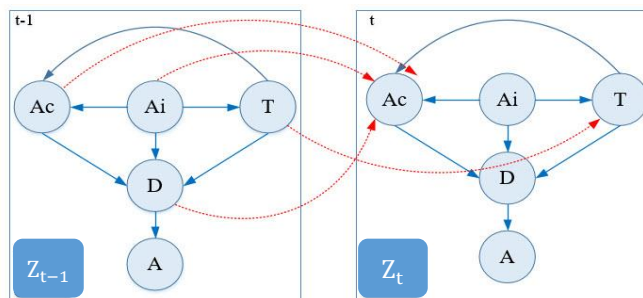
เบย์เซียน (Bayesian Inference) ประกอบด้วย ความสามารถในการให้เหตุผลกับความไม่แน่นอน และทำการประมาณค่าด้วยเบย์เซียนแบบวนซ้ำ การกำหนดโครงสร้างกิจกรรมต่าง ๆ ของมนุษย์ ลงในแบบจำลองกิจกรรมเพื่อให้สามารถรู้จำและดำเนินการตรวจสอบได้ ดังนั้นในงานวิจัยนี้ได้ เลือกวิธีการตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์ (Rao-Blackwellized Particle Filter: RBPF) [25] ที่มีพื้นฐานจากวิธีการเบย์เซียนเพื่อนำไปประยุกต์ในแบบจำลองการทำนายพฤติกรรมต้อง สงสัยในการสอบออนไลน์



รูปที่ 3 ตัวอย่างแบบจำลองโครงสร้างกระดูกถูกสกัดด้วย OpenPose [27] จากฐานข้อมูล ภาพ OEP [26]

2.2.1 ตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์

การจัดกลุ่มกิจกรรมโดยใช้แบบจำลองสถานะ DBN [15] ที่เกิดจากความสัมพันธ์ของท่าทาง กิจกรรม ($\beta_0, \beta_{\rightarrow}$) ระหว่างช่วงเวลา (t) โดยที่การเคลื่อนไหวของตำแหน่งข้อต่อสำคัญบนร่างกาย ถูกสกัดด้วย OpenPose [27] ได้เป็นชุดข้อมูล $\beta^i = \{\beta_0, \dots, \beta_n\}$ เมื่อ $\beta^i = (x^i, y^i)$, แทน ตำแหน่งจุดข้อต่อบนโครงร่างกายที่ i ในภาพที่ n และกำหนดให้ β_0 ถูกกำหนดให้เป็นค่าเริ่มต้น กิจกรรม ความน่าจะเป็นของชุดกิจกรรมในแบบจำลอง ท่าทางจากความน่าจะเป็นก่อนหน้าเป็น $P(Z_1)$ จากค่าเบย์เซียน เมื่อ กำหนดให้ Z_t เป็นค่าของโหนดกิจกรรมที่เกิดขึ้นในช่วง t ใด ๆ ดัง แสดงความสัมพันธ์ในรูปที่ 4 ดังนั้น สามารถกำหนดกิจกรรมที่มีความสัมพันธ์ได้



รูปที่ 4 ความสัมพันธ์ของกิจกรรมหลักตามเวลา Z_{t-1} และ Z_t ใน RBPF

$$P(Z_t|Z_{t-1}) = \prod_{i=1}^N P(Z_t^i|P_p(Z_t^i)) \quad (2)$$

เมื่อ $P_p(Z_t^i)$ แทนความน่าจะเป็นในกลุ่มกิจกรรมก่อนหน้าของ Z_t^i ที่เกิดขึ้นในช่วง t ได้ใด กำหนดตัวแปรสังเกตได้ (Observed variable) เป็น $V_t, V_{1:t} \triangleq \{v_1, v_2, \dots, v_t\}$ โดยกำหนดให้เป็นตัวแปรอิสระตามเงื่อนไขของกระบวนการ Z_t และเขียนเป็นสมการได้ $P(V_t|Z_t)$ และกำหนดให้ทุกเซตย่อยมีคุณสมบัติของแต่ละสถานะเป็น $Z_{0:t} \triangleq \{z_1, z_2, \dots, z_t\}$ จะมีความสัมพันธ์ของความน่าจะเป็นร่วมหลังปรับ (Joint posterior probability) เป็น $P(Z_{0:t}|V_{1:t})$ ดังนั้นการประมาณค่าเพื่อทำให้สามารถหาความน่าจะเป็นของ $P(Z_t|V_{1:t})$ สามารถเขียนสมการการแจกแจงได้ดังนี้

$$P(Z_{0:t}|V_{1:t}) = P(Z_{0:t-1}|V_{1:t-1}) \frac{P(V_t|Z_t)P(Z_t|Z_{t-1})}{P(V_t|V_{1:t-1})} \quad (3)$$

ถ้ากำหนดให้การกระจายค่าโดยใช้ชุดค่าน้ำหนัก N จำนวนสามารถเขียนเป็น $\{Z_{1:t}^i, \omega_t^i\}_{i=1}^N$, เมื่อ $Z_{1:t}^i$ แทนค่าที่ i ของตัวอย่าง $P(Z_{1:t})$ และสัมพันธ์กับค่าน้ำหนักของแต่ละ ω_t^i โดยที่

$$\sum_{i=1}^N \omega_t^i = 1 \quad (4)$$

$$\omega_t^i = \prod_{t=1}^T P(V_t|Z_t) \quad (5)$$

สามารถเขียนเป็นสมการแจกแจงการกรอง (Filtering distribution) ได้ดังนี้

$$P(Z_t|V_{1:t}) \approx \sum_{i=1}^N \omega_t^i \delta(Z_t, Z_t^i) \quad (6)$$

โดยทั่วไปแล้ววิธีการแบ่งการกรองออกเป็น ส่วน ๆ หรืออัลกอริทึมตัวกรองอนุภาค (Particle filtering algorithm) จะถูกอ้างอิงในแบบลำดับที่มีการสุ่มตัวอย่างเป็นหลัก แต่สำหรับการแบ่งเป็นส่วน ๆ จะถูกสร้างขึ้นด้วย $\{Z_t^i, \omega_t^i\}_{i=1}^N$ จากการกำหนดค่าน้ำหนักก่อนหน้า โดยกำหนดให้กลุ่มตัวอย่าง N ครั้ง ด้วยการแทนที่จากค่า $\{Z_t^i, \omega_{t-1}^i\}_{i=1}^N$, เมื่อแต่ละ Transition เป็น $Z_t^i \sim Q(Z_t^i|Z_{t-1}^i, V_t)$, ปรับค่าน้ำหนักด้วย

$$\omega_t^i = \omega_{t-1}^i \frac{P(V_t|Z_t)P(Z_t|Z_{t-1})}{P(Z_t|Z_{t-1}, V_t)} \quad (7)$$

เมื่อทำการสุ่มตัวอย่างใหม่ N ครั้งทำให้กลุ่มย่อยถูกสร้างใหม่เป็น $\{Z_t^i, \omega_t^i\}_{i=1}^N$ ด้วยการประมาณค่าความน่าจะเป็นของค่า $P(Z_t|V_{1:t})$ เมื่อ $P(Z_t^i = Z_t) = \omega_t^i$, ดังนั้นค่าน้ำหนักมีค่าเป็น $\{\omega_t^i = 1/N\}_{i=1}^N$ วิธีการราวแบล็คเวลไลซ์จะช่วยลดขนาดของกลุ่มตัวอย่างและเพิ่มประสิทธิภาพในการสร้างแบบจำลองที่มีโครงสร้างจากการวิเคราะห์ตัวแปรและการปรับเงื่อนไขจากตัวแปรอื่น ๆ ดังนั้นทำการแบ่งข้อมูลออกเป็นสองกลุ่มตัวอย่าง กำหนดให้เป็นตัวแปร $Z_t, Z_t = \{r_t, z_t\}, r_t = \{Ac, T, Ai\}_t, Z_t = \{D\}_t$

$$P(Z_t|Z_{t-1}) = P(x_t|r_{t-1}, z_{t-1})P(r_t|r_{t-1}) \quad (8)$$

สามารถแจกแจงในรูปแบบความน่าจะเป็น $P(Z_t|V_{1:t})$ เมื่อแต่ละกลุ่มตัวอย่างเป็น $\{r_t, z_t\}_t^i$ ที่มีความสัมพันธ์กับค่าน้ำหนัก ω_t^i โดยที่ $\sum_{i=1}^N \omega_t^i = 1$,

$$P(r_{0:t}, z_{0:t}|V_{1:t}) = P(z_{0:t}|V_{1:t}, r_{0:t})P(r_{0:t}|V_{1:t}) \quad (9)$$

การแจกแจงภายหลัง (Posterior distribution) ของค่า $P(r_{0:t}|V_{1:t})$ สามารถเขียนได้เป็น

$$P(r_{0:t}|V_{1:t}) = \frac{P(V_t|V_{1:t}, r_{0:t})P(r_t|r_{t-1})P(r_{0:t-1}|V_{1:t-1})}{P(Y_t|V_{1:t-1})} \quad (10)$$

2.2.2 การวิเคราะห์กิจกรรมด้วยตัวกรองอนุภาคแบบราวแบล็คเวลไลซ์

การกำหนดรูปแบบของกิจกรรม (Activity: A) ที่เกิดขึ้นในแต่ละช่วงเวลาบนเฟรมเพื่อสร้างความสัมพันธ์ระหว่างแต่ละกิจกรรมเพื่อนำมาใช้ในการกระจายตัวกรอง (filtering distribution) ในสมการ RBPF กำหนดให้โหนด T แทนกิจกรรมเป้าหมาย (target), โหนด Ac แทนโหนดปัจจุบัน (current activity)ที่กำลังสำรวจกิจกรรม และโหนด Ai แทนกิจกรรมที่ถูกขัดจังหวะ (interrupt activity) โครงสร้างของผลที่ได้จากแบบจำลองคือ D_t ค่ากิจกรรมที่ถูกตรวจจับได้ (desired activity) ของตัวแปรที่ได้จากเงื่อนไขของกลุ่มข้อมูลที่ประกอบด้วย $\{Ac, T, Ai\}_t$ เช่นเดียวกันกิจกรรมที่จะเกิดขึ้นถัดไปจะถูกคำนวณจากเงื่อนไขที่ขึ้นกับกลุ่มของข้อมูลกิจกรรมที่ประกอบด้วย $\{Ac, T, Ai\}_t$ เช่นเดียวกัน แทนค่ากิจกรรมลงในโครงสร้างของ RBPF ประกอบด้วย $r_t = \{Ac, T, Ai\}_t$ และ $Z_t = \{D\}_t$ ดังแสดงความสัมพันธ์ของกิจกรรมในรูปที่ 4 เมื่อกำหนดให้ RBPF ประกอบด้วย N เป็นจำนวนกลุ่มตัวอย่างข้อมูลแต่ละจุดมีค่าที่สัมพันธ์กับค่าน้ำหนักจะได้สมการดังนี้ $\{\{r, z\}_{1:t}^i, \omega_t^i\}_{i=1}^N = 1$ และ $P(Z_t|V_{1:t})$ ในแต่ละ $\{r, z\}_t^i$

$$P(Z_t|V_{1:t}) \approx \sum_{i=1}^N \omega_t^i (\{r, z\}_t, \{r, z\}_t^i) \quad (11)$$

$$P(Z_t|V_{1:t}) \approx \sum_{i=1}^N \omega^i(\{Ac, T, Ai, O\}_t, \{Ac, T, Ai, O\}_t^i) \quad (12)$$

จากสมการที่ (5) ปรับค่าน้ำหนักสำหรับสองกิจกรรม Ac และ T ดังนั้นสามารถแทนความสัมพันธ์ที่เกิดขึ้นได้ดังนี้

$$\omega_t^i = \omega_{t-1}^i \times P(Z_t^i|V_t) \times \frac{|Ac_t^i|+1}{|T_t^i|+1} \quad (13)$$

เพื่อเป็นการป้องกันการเกิดข้อผิดพลาดที่ทำให้ค่าน้ำหนักมีค่าเข้าใกล้ศูนย์ในสมการที่ (10) และมีผลทำให้สมการที่ (12) เช่นเดียวกันอาจจะทำให้เกิดค่าของ $P(A_t^i|\{Ac, T, Ai\}_t^i, A_{1:t-1}^i) \approx 0$ และถ้าค่าของ $1 - P(D_t|A_t)$ เกิดมีค่าเป็นลบ (false positive) ขึ้นมาจะทำการกำหนดค่าน้ำหนักที่ได้เป็น 0.01 โดยจะอธิบายการทำงานของอัลกอริทึม RBPF ดังรูปที่ 5

Algorithm 1 : Rao-Blackwellized particle filter (RBPF)

Input : Ac : current activity

T : target activity

Ai : interrupt activity

ω : weight vector for the feature attributes

N : number of random sample

Resampling: $\{\{Ac, T, Ai\}_0^i, \omega_0^i\}_{i=1}^N$ and ω_1

Output : Predict D_t^i from set activity $\{Ac, T, Ai\}_t^i$

Begin

For $t = 1$ to T do

for $i = 1$ to N do

Sample : $\{Ac, T, Ai, D\}_{t-1}^i$

Transition : $Z_t^i \sim \{Ac, T, Ai\}_{t-1}^i$

If $P(A_t^i|\{Ac, T, Ai\}_t^i == 0)$ && $(Ac_t^k == T_t^k)$

Reset activity

end if

Compute: $P(D_t^i|\{Ac, T, Ai\}_t^i, A_{1:t}^i)$

Predict D_t^i

Calculate ω_t^i Equation (13)

end for

Normalize: ω_t and Resample Z_t

End for

End

รูปที่ 5 อัลกอริทึม Rao-Blackwellized particle filter

3. การวัดประสิทธิภาพการทำงาน

การวัดประสิทธิภาพเพื่อทดสอบวิธีการว่ามีประสิทธิภาพมากหรือน้อย สอดคล้องกับความต้องการเพียงใดเมื่อนำมาใช้ในงานจริง สำหรับในงานวิจัยนี้ได้นำเสนอ วิธีการตรวจหาพฤติกรรมที่ต้องสงสัยในการสอบออนไลน์จากรูปแบบของการเคลื่อนไหวด้วยตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์ เพื่อใช้ในการตรวจสอบพฤติกรรมทดสอบ การสกัดข้อมูลจากภาพวิดีโอ ด้วย OpenPose [22] ร่วมกับข้อมูลจากระบบ Online Exam Proctoring (OEP) [26] เป็นระบบคุมสอบออนไลน์แบบอัตโนมัติที่ควบคุมและดูแลการสอบด้วยการบันทึกวิดีโอลงในระบบฐานข้อมูลแบบเรียลไทม์มีขนาด 640x480 พิกเซลด้วยอัตราความเร็ว 25 เฟรมต่อวินาที เป็นระยะเวลาประมาณ 20-25 นาที รวมทั้งหมด 24 ชุดข้อมูล โดยใช้อุปกรณ์เซ็นเซอร์ในการตรวจจับกล้องและไมโครโฟน ซึ่งข้อมูลของผู้สอบที่ได้รับจากการจัดเก็บข้อมูลบนระบบ OEP ประกอบด้วย ข้อมูลตัวอักษร ข้อมูลเสียง ข้อมูลหน้าจอคอมพิวเตอร์ และ ข้อมูลภาพเคลื่อนไหวระหว่างการสอบ ซึ่งในการทดลองนี้จะนำข้อมูลมาเพียงส่วนเดียว คือข้อมูลของระหว่างการสอบที่ประกอบด้วยท่าทางการเคลื่อนไหวลักษณะต่าง ๆ ข้อมูลตำแหน่งร่างกาย และ ค่าประมาณของความถี่รวม 8 จุด ดังแสดงภาพตัวอย่างในรูปที่ 3 สำหรับการทดลองได้กำหนดรูปแบบพฤติกรรมอ้างอิงตาม Baseline จากการกำหนดรูปแบบมาตรฐานกลุ่มพฤติกรรมทดสอบจากระบบ OEP ที่แบ่งออกเป็นผู้เข้าสอบแสดงพฤติกรรมทุจริต โดยการสอบถามผู้อื่นระหว่างที่มีการสอบ ผู้เข้าสอบแสดงพฤติกรรมต้องสงสัยด้วยการหยิบจับอุปกรณ์สื่อสารต่าง ๆ ขณะที่มีการสอบ ผู้เข้าสอบมีการเปิดหนังสือเอกสารต่าง ๆ ระหว่างที่สอบ และ ผู้เข้าสอบที่ไม่มีพฤติกรรมเด่นชัดหรือมีท่าทางปกติในการสอบ โดยการบันทึกข้อมูลมีการควบคุมสภาพแวดล้อมเดียวกันทั้งห้องและอุปกรณ์ที่ใช้ในการบันทึก ดังนั้นสามารถแบ่งกลุ่มตามกิจกรรมที่มีการตรวจจับได้ 4 กลุ่มดังนี้ (1) พฤติกรรมปกติ (Normal behaviour: normal) และพฤติกรรมต้องสงสัย (Suspicious behaviour) [26] ซึ่งมีรูปแบบสามารถจำแนกพฤติกรรมต้องสงสัยออกเป็นกลุ่มย่อยดังนี้ (2) การพูดคุยกับเพื่อน (talking to the person in the room :Talking) (3) การใช้โทรศัพท์ (using a phone: Phone) และ (4) นำเอกสารต่าง ๆ ขึ้นมา (cheating from a book or notes: Note) ดังแสดงภาพตัวอย่างกิจกรรมในรูปที่ 6 สำหรับการทดลองได้ทำการเปรียบเทียบกับวิธีการที่นำเสนอตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์ (Rao-Blackwellized Particle Filter: RBPF) ด้วย 4 วิธีดังนี้

3.1 ซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines: SVM) [29] ใช้หลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งสำหรับแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกกลุ่มข้อมูลได้ดีที่สุด กำหนดให้ชุดข้อมูลการเรียนรู้ N ที่มี

ข้อมูลเป็น $\{y_i, x_i\}_{i=1}^m$ เมื่อ $x_i \in R^n$ เป็นรูปแบบข้อมูลเข้าที่ i และ $y_i \in R^n$ เป็นรูปแบบข้อมูลออกที่ i และมีรูปแบบของการจำแนกด้วยวิธีการ SVM ดังสมการ $y(x) = \text{sign}(\sum_{i=1}^m \alpha_i y_i K(x_i, x) + b)$ เมื่อ α_i เป็นค่าบวกคงที่และ b เป็นค่าคงที่จำนวนจริง ข้อมูลที่นำมาใช้ในระบบการสอนให้ระบบเรียนรู้ส่วนใหญ่มักเป็นข้อมูลแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาดังกล่าวด้วย Kernel Function มาใช้จำแนกข้อมูลบนระนาบหลายมิติ สามารถเขียนสมการของการลดทอน Cost function ได้ดังนี้

$$\text{minimize}_{w, \xi, b} \{ \|w\|^2 / 2 + C \sum_{i=1}^m \xi_i \} \quad (14)$$

ภายใต้ข้อจำกัดของ $y_i(w \cdot x_i - b) \geq 1 - \xi_i$ และ $\xi_i \geq 0$, สำหรับ $i = \{1, 2, \dots, m\}$ เมื่อ ξ_i เป็นข้อมูลใช้วัดดีกรีการจำแนกข้อมูลที่ผิดพลาดของ x_i และตัวแปรควบคุม C จะควบคุมระหว่างค่าผิดพลาดและค่าสูงสุดของข้อมูลการเรียนรู้ โดยที่ถ้า $C = \infty$ จะนำพาไปยัง Margin

3.2 โครงข่ายประสาทเทียมแบบสังวัตนาการ

โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks: CNN) [30] เป็นเพอร์เซ็ปตรอนหลายชั้น (Multi-layer Perceptron) ประกอบด้วย (1) Convolution Layer: ลำดับชั้นจะทำการกรองข้อมูลแบบ 2 มิติ ระหว่างข้อมูล x ตัวกรอง w และ h เป็นข้อมูลภาพอีกชุดที่ถูกสร้างขึ้น และให้ CT แทนความสัมพันธ์ของข้อมูลเข้าและออก เมื่อตัวกรองตอบสนองจากข้อมูลเข้าที่เชื่อมต่อกับข้อมูลออกตัวเดียวกันจะถูกเชื่อมต่อกันเขียนเป็นสมการ

$$h_j = \sum_{i, k \in CT_{i, k, j}} (x_i * W_k) \quad (15)$$

(2) Sub-sampling Layer: ลำดับชั้นเพื่อดำเนินการ “Max-Pooling” เป็นการเปลี่ยนแปลงและลดทอนให้สามารถทำงานได้อย่างรวดเร็วขึ้น (3) Fully-connected Layer: ลำดับชั้นที่รับข้อมูลเข้าเพื่อทำการรวมเข้าเป็นข้อมูลออก แบบ 1 มิติสำหรับลำดับชั้นถัดไป (4) Output Layer: ลำดับชั้นข้อมูลออกที่ถูกจำแนกเพียงคลาสเดียวแบบ 1 มิติจากลำดับชั้น Fully-connected Layer

3.3 โครงข่ายประสาทเทียมแบบสังวัตนาการ-ซัพพอร์ตเวกเตอร์แมชชีน

โครงข่ายประสาทเทียมแบบสังวัตนาการ-ซัพพอร์ตเวกเตอร์แมชชีน (Convolutional Neural Network - Support Vector Machines: CNNSVM) [29, 30] เป็นการนำสองวิธีการมาผสมรวมกัน สามารถเขียนฟังก์ชันการตัดสินใจและสมการทั่วไปของ SVM เป็น $f(x) = (w \cdot \phi(x) + b)$, เมื่อ

w เป็นเวกเตอร์ของค่าน้ำหนัก, b เป็นค่าอคติ (Bias) และทุกพารามิเตอร์จะเพิ่มค่า \emptyset แทน Arbitrary Function ดังนั้นวัตถุประสงค์หลักของการจำแนกด้วย CNNSVM เพื่อทำการหาค่าเส้นแบ่งที่เหมาะสมที่สุดสำหรับข้อมูล $f(x)$ ดังนั้นกำหนดให้ข้อมูลการเรียนรู้เป็น $S = \{(x_i, y_i)\}_{i=1}^m$ เมื่อ $x_i \in R^n$ และ $y_i \in \{+1, -1\}$ สำหรับการจำแนกสองกลุ่ม และค่า Cost function มีสมการเป็น

$$\text{minimize}_w \{ \lambda/2 \|w\|^2 + 1/m \sum_{i=1}^m l(x_i, y_i; w) \}, \quad (16)$$

เมื่อ $\lambda \geq 0$, และ $l(x_i, y_i; w) = \max\{0, 1 - y_i f(x_i)\}$. สำหรับเทอม $\max\{0, 1 - y_i f(x_i)\}$ จะถูกแทนด้วย Hinge-loss และกำหนดให้ $f(x_i)$ มีความสัมพันธ์กับ Kernel $K(\dots)$, และกำหนดให้ $\lambda/2 \|w\|^2$ ถูกใช้เป็นตัวปรับค่าของ Margin เมื่อเทอม $1/m \sum_{i=1}^m l(x_i, y_i; w)$, เป็นค่าต่ำสุดของค่าความคลาดเคลื่อนในการเรียนรู้และให้ ค่า C เป็นพารามิเตอร์ของการจำแนกคลาดเคลื่อน โดยที่มีความสัมพันธ์กับพารามิเตอร์ $\lambda = 1/mC$ โดยทั่วไปแล้วการเรียนรู้ในชุดของข้อมูลของ CNN และ SVM จะต้องทำให้ไปถึงค่าต่ำที่สุด $R(f(x_i), y_i)$

3.4 โครงข่ายประสาทเทียมแบบหน่วยความจำระยะสั้นแบบยาว

โครงข่ายประสาทเทียมแบบหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory: LSTM) [31] เป็นระบบโครงข่ายประสาทเทียมแบบวนซ้ำที่มีวิธีการสามารถจดจำข้อมูลได้ในช่วงระยะเวลายาวนานกว่าในรูปแบบอื่น มีระยะเวลาในแต่ละ gate ดังนี้

- 1) Input gate (i) เป็นตัวควบคุมข้อมูลเข้าเพื่อทำการพิจารณาชุดข้อมูล (d_i)

$$i_t = \sigma(\omega_{di}d_t + \omega_{hi}h_{t-1}) \quad (17)$$

กำหนดให้ σ แทน ฟังก์ชันซิกมอยด์ (sigmoidal function), ω_{di} แทนข้อมูลซ่อน (hidden input) h_t แทน hidden state

- 2) Forget gate (f) เป็นตัวที่อนุญาตให้ LSTM ลืมความจำก่อนหน้า (c_{t-1}) เพื่อเตรียมพื้นที่สำหรับรับข้อมูลใหม่ ด้วยการตัดสินใจจากข้อมูลนำเข้ร่วมกับ hidden state (h_t) ก่อนหน้า

$$f_t = \sigma(\omega_{df}d_t + \omega_{hf}h_{t-1}) \quad (18)$$

- 3) Output gate (o) เป็นตัวที่ทำการตัดสินใจว่าข้อมูลในหน่วยความจำจะถูกส่งไปยัง hidden state (h_t)

$$o_t = \sigma(\omega_{do}d_t + \omega_{ho}h_{t-1}) \quad (19)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \phi(\omega_{dc}d_t + \omega_{hc}h_{t-1}) \quad (20)$$

$$h_t = o_t \odot \phi(c_t) \quad (21)$$

กำหนดให้ σ แทน ฟังก์ชันซิกมอยด์ (sigmoidal function), ϕ แทน ไฮเปอร์โบลิกแทนเจนท์ (hyperbolic tangent), และ c แทน ฟังก์ชันกระตุ้น (activation vector) สำหรับข้อมูลค่านำหน้า ระหว่างข้อมูลนำเข้าหรือข้อมูลซ่อน (hidden input) ถูกแทนด้วย $\omega_{di}, \omega_{hi}, \omega_{df}, \omega_{hf}, \omega_{do}, \omega_{ho}, \omega_{dc}$ และ ω_{hc} ดังนั้นการปรับปรุงค่าระหว่างขั้นตอนการทำงานจะเป็นการรวมข้อมูลจาก ขั้นตอนต่างเพื่อประกอบการพิจารณาในการปรับปรุงค่าของเซลล์ ซึ่งข้อดีของการใช้ LSTM คือ สามารถบอกได้ว่าเมื่อใดจะทำการปรับปรุงเซลล์ อ่านเซลล์หรือเขียนข้อมูลเข้ามาซึ่งตัว gate จะเป็น ตัวควบคุมการไหลของข้อมูล

4. ผลการทดลอง

การทดลองเพื่อตรวจหาพฤติกรรมที่ต้องสงสัยในการสอบจากระบบของการเคลื่อนไหวด้วย ตัวกรองอนุภาคแบบราวแบบลึกลับ โดยทำการแบ่งข้อมูลภาพการสอบออนไลน์ในระบบ OEP โดยสุ่มข้อมูลจากระบบ OEP (primary dataset) จากชุดข้อมูลเดิมที่มีการตัดแบ่ง ออกเป็น 3 ช่วง ข้อมูลความยาวสำหรับการทดลอง 10 นาที ดังนั้นข้อมูลเริ่มต้นเป็น 72 ชุดข้อมูล (primitive dataset) และข้อมูลจะถูกทำการจำลองเพื่อเพิ่มปริมาณข้อมูลภาพด้วยการทำ data augmentation [32] ประกอบด้วย การหมุน การตัดบางส่วน การปรับเปลี่ยนความสว่าง เพื่อเพิ่มจำนวนข้อมูลในการทดลองทั้งหมดเป็น 576 ชุดข้อมูล ด้วยการสุ่มข้อมูลสำหรับการทดลองที่ 1 และ 2 ตามลำดับ การประเมินประสิทธิภาพด้วยค่าความถูกต้องของแต่ละกลุ่มข้อมูล ประกอบด้วย ค่าความแม่นยำ (False positive rate / Precision: Pre.) เป็นอัตราส่วนของการตรวจพบพฤติกรรมที่ถูกต้องจาก จำนวนท่าทางทั้งหมด, ค่าความระลึก (True positive rate / Recall: Re) เป็นอัตราส่วนของการ ตรวจพบพฤติกรรมที่ถูกต้องจากจำนวนข้อมูลที่ถูกต้องทั้งหมด และ ค่าประสิทธิภาพโดยรวม (F-measure: F1) เป็นการวัดค่าความสัมพันธ์ระหว่างค่าความระลึกและค่าความแม่นยำในเชิง ฮาร์โมนิคแสดงผลการทดลองเป็นค่าเฉลี่ยความถูกต้องรวม (total average accuracy) ค่า ความแม่นยำ ค่าความระลึก และค่าประสิทธิภาพโดยรวม (F_1) ด้วยการเปรียบเทียบกับ ดังแสดงใน ตารางที่ 1 และ 2 เป็นชุดข้อมูลการทดลองที่ 1 และ 2 ตามลำดับ วิธี SVM กำหนดค่า Kernel sigmoid กำหนดค่า $\xi = 1$, สำหรับวิธี CNN เป็นเครือข่ายแบบการเรียนรู้เชิงลึก ประกอบด้วย Convolutional layer1 มี 10 kernel ที่มีขนาด 11×11 และ Convolutional layer2 มี 6 kernel ที่มี

ขนาด 5x5 Convolutional layer3 มี 23 kernel ที่มีขนาด 5x5 และในการทดลองจะมีการใช้โครงสร้างของ CNN ร่วมกับ SVM และ LSTM จากตารางผลการทดลองที่ 1 ในชุดข้อมูลที่ 1 ได้ค่าเฉลี่ยความถูกต้องรวมของการตรวจพบพฤติกรรมต้องสงสัยด้วยวิธีการ SVM, CNN, CNN-SVM, CNN-LSTM และ RBPF มีค่าความถูกต้องเฉลี่ยรวมอยู่ที่ 50.47% 52.99% 69.43% 72.13% และ 76.98% ตามลำดับ จะเห็นว่าการใช้วิธีการ RBPF มีการจำลองรูปแบบของความสัมพันธ์ทำทางกิจกรรมตามช่วงเวลาในรูปแบบของค่าการอนุมานเบย์เซียน จึงทำให้มีค่าความถูกต้องที่เพิ่มขึ้น

ตารางที่ 1 ประสิทธิภาพการตรวจจับพฤติกรรมต้องสงสัยระหว่างสอบออนไลน์ ด้วยชุดข้อมูลที่ 1

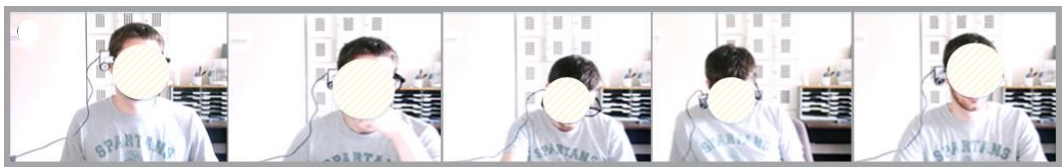
Categories		Performance (%)				Average Accuracy (%)
		Normal	Talking	Phone	Note	
SVM	Pre.	60.22	50.96	44.07	57.89	50.47
	Recall	58.33	56.99	59.09	36.42	
	F ₁	59.26	53.81	50.49	44.72	
CNN	Pre.	50.98	54.78	54.39	68.13	52.99
	Recall	56.52	61.17	63.27	39.24	
	F ₁	53.61	57.80	58.49	49.80	
CNN-SVM	Pre.	75.00	60.58	73.12	72.45	69.43
	Recall	64.08	72.41	65.38	77.17	
	F ₁	69.11	65.97	69.04	74.74	
CNN-LSTM	Pre.	72.45	75.82	71.00	70.87	75.13
	Recall	78.02	75.82	71.72	75.26	
	F ₁	75.13	75.82	71.36	73.00	
RBPF	Pre.	82.29	84.27	75.73	75.96	76.98
	Recall	78.22	71.43	78.00	80.61	
	F ₁	80.20	77.32	76.85	78.22	

ตารางที่ 2 ประสิทธิภาพการตรวจจับพฤติกรรมต้องสงสัยระหว่างสอบออนไลน์ ด้วยชุดข้อมูลที่ 2

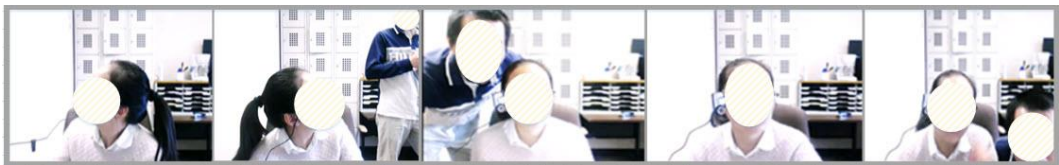
Categories		Performance (%)				Average Accuracy (%)
		Normal	Talking	Phone	Note	
SVM	Pre.	59.14	52.88	44.07	57.89	52.93
	Recall	56.12	56.12	58.43	44.00	
	F ₁	57.59	54.46	50.24	50.00	
CNN	Pre.	47.06	58.26	53.51	58.24	54.27
	Recall	52.17	62.62	62.89	42.06	
	F ₁	49.48	60.36	57.82	48.85	
CNN-SVM	Pre.	73.86	65.38	67.74	64.29	67.62
	Recall	63.73	73.12	69.23	64.95	
	F ₁	68.42	69.04	68.48	64.62	
CNN-LSTM	Pre.	75.51	76.92	73.00	72.82	74.49
	Recall	80.43	70.00	72.28	75.76	
	F ₁	77.89	73.30	72.64	74.26	
RBPF	Pre.	81.72	85.39	78.22	74.29	79.63
	Recall	79.17	73.79	82.29	83.87	
	F ₁	80.42	79.17	80.20	78.79	

จะเห็นว่าการใช้วิธีการ RBPF สามารถตรวจจับพฤติกรรมได้สูงที่สุดโดยเฉพาะ Talking ด้วยค่าความแม่นยำ 84.27% และ ค่าความแม่นยำสำหรับการสอบที่มีท่าทางปกติ Normal อยู่ที่ 82.29% แต่การจำแนกพฤติกรรมด้วย CNN และ SVM ในท่าทางปกติ Normal จะได้ค่าความแม่นยำอยู่ที่ 50.98% และ 60.22% แต่วิธีการ CNN-SVM สามารถจำแนกพฤติกรรม Phone มีค่าความแม่นยำสูงถึง 73.12% แต่ CNN-LSTM ได้ค่าความแม่นยำเพียง 71% และ RBPF สามารถตรวจจับพฤติกรรมได้ค่าความแม่นยำมากถึง 75.73% ดังแสดงภาพเฟรมตัวอย่างการทดลองในรูปที่ 6 จากการทดลองในชุดข้อมูลที่ 2 เพื่อทดสอบประสิทธิภาพการตรวจจับพฤติกรรมระหว่างการสอบออนไลน์ สามารถแสดงผลในตารางที่ 2 ได้ค่าเฉลี่ยความถูกต้องรวมของวิธีการ SVM, CNN, CNN-SVM, CNN-LSTM และ RBPF ที่นำเสนอผลปรากฏว่ามีค่าความถูกต้องเฉลี่ยรวมอยู่ที่ 52.93% 54.27% 67.62% 74.49% และ 79.63% ตามลำดับ สำหรับการใช้วิธี SVM มีค่าความ

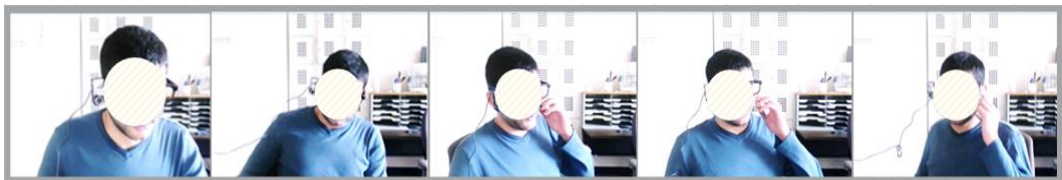
แม่นยำของ Normal. สูงถึง 59.14% แต่ CNN จะได้เพียง 47.06% และ CNN-SVM สามารถจำแนกพฤติกรรม Normal. ได้สูงถึง 73.86% แต่เมื่อใช้วิธี RBPF สามารถบอกพฤติกรรมจากท่าทางกิจกรรมที่มีความสัมพันธ์กันได้ ถึง 81.72% เนื่องจากการตรวจจับเป็นไปตามกิจกรรมลำดับการเคลื่อนไหวของท่าทางตามกิจกรรมที่เกิดขึ้น แต่อย่างไรก็ตามวิธี RBPF สามารถจำแนกท่าทาง Talking ได้ค่าความถูกต้องสูงถึง 85.39% สำหรับชุดข้อมูลในชุดที่ 2 ซึ่งพฤติกรรมจากท่าทางในแต่ละเฟรมสามารถเห็นได้เด่นชัดจากการเคลื่อนไหวตัวของผู้เข้าสอบทำให้ผลที่ได้จากค่าความแม่นยำค่อนข้างสูง ทั้งสองชุดการทดลอง



ก. ตัวอย่างกลุ่มเฟรมท่าทางการสอบปกติ (Normal)



ข. ตัวอย่างกลุ่มเฟรมท่าทางการคุย (Talking)



ค. ตัวอย่างกลุ่มเฟรมท่าทางการคุยโทรศัพท์ (Phone)



ง. ตัวอย่างกลุ่มเฟรมท่าทางการนำเอกสารขึ้นมา (Note)

รูปที่ 6 ตัวอย่างเฟรมการสอบออนไลน์จากฐานข้อมูลภาพ OEP [26]

5. สรุปผลการทดลอง

งานวิจัยนี้ได้นำเสนอแบบจำลองตัวกรองอนุภาคแบบราวแบบลิคเวลไลซ์เพื่อตรวจจับพฤติกรรมต้องสงสัยในการสอบออนไลน์ โดยจะแบ่งกลุ่มตามรูปแบบของพฤติกรรมที่ส่วนใหญ่จะ

ตรวจพบในการสอบทั่วไป สำหรับการทดลองนี้ได้แบ่งออกเป็น 4 กลุ่ม ประกอบด้วย พฤติกรรมปกติ และ พฤติกรรมต้องสงสัยที่แบ่งกลุ่มเป็น พฤติกรรมพุดคุยกับเพื่อน พฤติกรรมใช้โทรศัพท์ และ พฤติกรรมนำเอกสารขึ้นมา โดยใช้วิธีการสกัดข้อมูลในตำแหน่งที่สำคัญที่เป็นข้อต่อเฉพาะท่อนบนของร่างกายเพื่อตรวจจับหาความสัมพันธ์ของการเคลื่อนไหวในรูปแบบของความน่าจะเป็นในกลุ่มของกิจกรรมเพื่อทำให้เกิดความสัมพันธ์ที่กลายเป็นพฤติกรรมในแต่ละกิจกรรมของผู้เข้าสอบ จากผลการทดลองที่ทำการเปรียบเทียบประสิทธิภาพจะเห็นว่าวิธีการจำแนกด้วยการใช้ตัวกรองอนุภาคแบบราวแบล็คเวลล์ไลซ์มีค่าเฉลี่ยความถูกต้องสูงถึง 79.63% สำหรับการทดลองด้วยชุดข้อมูลที่ 2 ซึ่งรูปแบบการทำงานของอัลกอริธึมจะมีความสัมพันธ์กับท่าทางกิจกรรมจากค่าเบี่ยงเบนแบบวนซ้ำจึงทำให้มีการทำงานที่แม่นยำมากกว่าวิธีการอื่น แต่อย่างไรก็ตามข้อมูลที่นำมาใช้ในการทดลองนี้เป็นการทดลองในห้องที่กำหนดสภาพแวดล้อมและมีการตั้งค่ากล้องที่แน่นอน เพื่อให้เหมาะสมกับรูปแบบวิธีการมากที่สุด รวมทั้งยังไม่สามารถตรวจสอบการส่งข้อความหรือการแชร์หน้าจอให้กับผู้อื่นได้ซึ่งเป็นข้อจำกัดของการทดลองนี้ ดังนั้นควรจะเพิ่มวิธีการตรวจจับหน้าจอที่ผู้สอบใช้งานอยู่ รวมทั้งการใช้ตัวอย่างข้อมูลการสอบที่หลากหลายและแปรเปลี่ยนสภาพแวดล้อมได้เพื่อความถูกต้องและแม่นยำของการทดลองที่เพิ่มมากขึ้นสำหรับการทดลองต่อไป

References

- [1] Prince DJ, Fulton RA, Garsombke TW. Comparisons of proctored versus non-proctored testing strategies in graduate distance education curriculum. *Journal of College Teaching & Learning* 2009;6(7):51-62.
- [2] King DL, Case CJ. E-cheating: Incidence and trends among college students. *Issues in Information Systems* 2014;15(1):20-7.
- [3] Ullah A, Xiao H, Barker T. A classification of threats to remote online examinations. In: 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON); 2016. p.1-7.
- [4] Wei H, Li H, Xia M, Wang Y, Qu H. Predicting student performance in interactive online question pools using mouse interaction features. In: the 10th International Conference on Learning Analytics and Knowledge; 2020. p.645-54.
- [5] Rogers CF. Faculty perceptions about e-cheating during online testing. *Journal of Computing Sciences in Colleges* 2006;22(2):206-12.
- [6] Cluskey GR, Ehlen CR, Raiborn MH. Thwarting online exam cheating without proctor supervision. *Journal of Academic and Business Ethics* 2011;4(1):1-7.

- [7] Bella G, Giustolisi R, Lenzini G, Ryan PYA. A secure exam protocol without trusted parties. In: Proceedings International Conference on ICT Systems Security and Privacy Protection; 2015. p.495-509.
- [8] Mariani L, Micucci D. Audentes: Automatic detection of tentative plagiarism according to a reference solution. ACM Transactions on Computing Education; 2012. p.1-26.
- [9] Chen K, Liu P, Zhang Y. Achieving accuracy and scalability simultaneously in detecting application clones on android markets. In: Proceedings 36th International Conference on Software Engineering; 2014. p.175-86.
- [10] Ullah A, Xiao H, Barker T. A dynamic profile questions approach to mitigate impersonation in online examinations. Journal of Grid Computing 2018;17:209-23.
- [11] Ullah A, Xiao H, Barker T. A classification of threats to remote online examinations. In: Proceedings IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON); 2016. p.1-7.
- [12] Kuhnke F, Ostermann J. Deep head pose estimation using synthetic images and partial adversarial domain adaption for continuous label spaces. In: 2019 IEEE/CVF International Conference on Computer Vision; 2019. p.10163-72.
- [13] Ruiz N, Chong E, Rehg JM. Fine-grained head pose estimation without keypoints. In: 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops; 2018. p.2074-83.
- [14] Prathish S, Narayanan SA, Bijlani K. An intelligent system for online exam monitoring. In: Proceedings International Conference on Information Science (ICIS); 2016. p.138-43.
- [15] Narayanan A, Kaimal RM, Bijlani K. Yaw estimation using cylindrical and ellipsoidal face models. IEEE Transactions on Intelligent Transportation Systems 2014;15(5):2308-20.
- [16] Wlodarczyk M, Kacperski D, Krotewicz P, Grabowski K. Evaluation of head pose estimation methods for a noncooperative biometric system. In: Proceedings 23rd International Conference Mixed Design of Integrated Circuits and Systems (MIXDES); 2016. p.394-8.
- [17] Hu S, Jia X, Fu Y. Research on abnormal behavior detection of online examination based on image information. In: Proceedings 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC); 2018. p.88-91.

- [18] Ghizlane M, Reda FH, Hicham B. A smart card digital identity check model for university services access. In: Proceedings The 2nd International Conference on Networking, Information Systems and Security; 2019. p.1-4.
- [19] Ghizlane M, Hicham B, Reda FH. A new model of automatic and continuous online exam monitoring. In: Proceedings 2019 International Conference on Systems of Collaboration Big Data, Internet of Things Security (SysCoBloTS); 2019. p.1-5.
- [20] Garg K, Verma K, Patidar K, Tejra N, K. Patidar. Convolutional neural network based virtual exam controller. In: Proceedings 4th International Conference on Intelligent Computing and Control Systems (ICICCS); 2020. p.895-9.
- [21] Xia M, Wei H, Xu M, L YL, Wang Y, Zhang R, et al. Visual analytics of student learning behaviors on K-12 mathematics E-learning platforms. Computer Science, Education; 2019. p.1-2.
- [22] Yang T, Chen Y C, Lin YY, Chuang YY. FSA-Net: Learning fine-grained structure aggregation for head pose estimation from a single image. In: 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2019. p.1087-96.
- [23] Wan J, Zhao Y, Zhou S, Guyon I, Escalera S, Li SZ. Chalearn looking at people RGB-D isolated and continuous datasets for gesture recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops; 2016. p.56-64.
- [24] Negin F, Rodriguez P, Koperski M, Kerboua A, González J, Bourgeois J, et al. PRAXIS: Towards automatic cognitive assessment using gesture recognition. Expert Systems with Applications 2018;106:21-35.
- [25] Doucet A, Freitas N, Murphy K, Russell S. Rao-Blackwellised particle filtering for dynamic Bayesian networks. In: Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence; 2000. p.176-83.
- [26] Atoum Y, Chen L, Liu AX, Hsu SDH, Liu X. Automated online exam proctoring. IEEE Transactions on Multimedia 2017;19(7):1609-24.
- [27] Cao Z, Hidalgo G, Simon T, Wei SE, Sheik Y. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence 2018;43(1):172-186.
- [28] Murphy KP. Dynamic Bayesian networks: representation, inference and learning [Ph.D. thesis in Computer Science]. Berkeley, CA, USA: University of California, Berkeley; 2002.

- [29] Bishop CM. Pattern recognition and machine learning. New York, USA: Springer; 2006.
- [30] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. In: International Conference on Learning Representations;2015.
- [31] Donahue J, Hendricks LA, Sergio G, Marcus R, Subhashini V, Kate S, et al. Long-term recurrent convolutional networks for visual recognition and description. Computer Vision and Pattern Recognition 2014;39(4):677-91.
- [32] Yun S, Oh SJ, Heo B, Han D, Kim J. VideoMix: Rethinking data augmentation for video classification. Computer Vision and Pattern Recognition. 2020. arXiv:2012.03457.

ประวัติผู้เขียนบทความ



นัทพ์ชาณัน ชินปัญช์ธนะ อาจารย์ประจำวิทยาลัยนวัตกรรมการด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์ งานวิจัยทางด้านการประมวลผลภาพดิจิทัลทางด้านความหมายภาพ งานประมวลผลภาพทางด้านโรงงานอุตสาหกรรมการผลิต การจำแนกภาพดิจิทัล การรู้จำภาพดิจิทัล การประมวลผลภาพวิดีโอแบบเรียลไทม์ อีเมล nutchanun.cha@dpu.ac.th โทรศัพท์ 0-2954-7300



เดชค์ฐสิณป์ เพี้ยชัย อาจารย์ประจำสาขาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช งานวิจัยทางด้านโทรคมนาคมไฟฟ้า อิเล็กทรอนิกส์ การประมวลด้วยเสียง IoT การประมวลผลแบบ 3D การประมวลผลภาพดิจิทัล การประมวลผลภาพวิดีโอแบบเรียลไทม์ อีเมล tejtasin@gmail.com

Article History:

Received: March 7, 2022

Revised: August 22, 2022

Accepted: August 24, 2022