การแจกแจงผสมแบบเกาส์ด้วยค่าคาดหมายสูงสุดสำหรับการจำแนกภาพท่าทางกิจวัตรประจำวันของมนุษย์

GAUSSIAN MIXTURE MODELS WITH EXPECTATION MAXIMIZATION FOR HUMAN DAILY ACTIVITIES CLASSIFICATION

เตชค์ฐสิณป์ เพียซ้าย[1] และ นัศพ์ชาณัณ ชินปัญช์ธนะ[2]
[1]อาจารย์, สาขาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช
ถ.แจ้งวัฒนะ เขตปากเกร็ด นนทบุรี 11120, tejtasin.phi@stou.ac.th
[2]อาจารย์, วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์
110/1-4 ถ.ประชาชื่น เขตหลักสี่ กรุงเทพฯ 10210, nutchanun.cha@dpu.ac.th

Tejtasin Phiasai[1] and Nutchanun Chinpanthana[2]
[1]Lecturer, School of Science and Technology, Sukhothai Thammathirat Open University, Chaengwattana Rd., Bangpood, Pakkret, Nonthaburi 11120, Thailand, tejtasin@gmail.com
[2]Lecturer, College of Innovative Technology and Engineering, Dhurakij Pundit University, 110/1-4 Prachachuen rd. Laksi, Bangkok 10210, Thailand, nutchanun.cha@dpu.ac.th

บทคัดย่อ

ปัจจุบันการตรวจจับกิจกรรมของมนุษย์เป็นหัวข้อวิจัยที่มีความสำคัญอย่างมาก เนื่องจากความสามารถในการตรวจจับกิจวัตรประจำวันของมนุษย์จะช่วยเพิ่มขีดความสามารถของแอปพลิเคชันในคอมพิวเตอร์วิชัน มีหลายวิธีการที่พยายามรวมคำศัพท์จากเนื้อหาภายในภาพเข้ากับวิธีการทางสถิติเพื่อใช้ในการจำแนกรูปแบบท่าทางกิจกรรม และยังมีหลายวิธีการที่พยายามใช้การเรียนรู้แบบมีผู้สอนในการแยกแยะกิจกรรมของมนุษย์ ทำให้ยังมีข้อจำกัดที่ผูกติดกับเนื้อหาหรือคำอธิบายในภาพ แต่อย่างไรก็ตามยังคงเป็นหัวข้อที่ท้าทายเพื่อให้ได้ผลลัพธ์ที่ถูกต้องมากขึ้น ในงานวิจัยนี้ได้นำเสนอการจำแนกท่าทางกิจวัตรประจำวันของมนุษย์ด้วยการเรียนรู้แบบมีผู้สอนโดยใช้วิธีการแจกแจงแบบเกาส์ด้วยการประมาณค่าคาดหมายสูงสุด ได้แบ่งขั้นตอนการดำเนินงานวิจัย ออกเป็น 4 ส่วนหลัก ดังนี้ (1) การเตรียมข้อมูลภาพ (2) การกำหนดรูปแบบข้อมูลด้วยการแจกแจงแบบเกาส์ (3) การประมาณค่าคาดหมายสูงสุด และ (4) การวัดประสิทธิภาพการทำงาน ซึ่งได้ใช้ฐานข้อมูลมาตรฐานและจำแนกกิจกรรมออกเป็น 12 กลุ่มที่มีสภาพแวดล้อมพื้นหลังแตกต่างกัน จากผลการทดลองได้ค่าเฉลี่ยความถูกต้องสูงถึง 84.6% สำหรับการทดลองด้วยคำศัพท์ 150 คำ

คำสำคัญ: การประมวลผลภาพ, การแจกแจงผสมแบบเกาส์, ค่าคาดหมายสูงสุด, กิจวัตรประจำวันของมนุษย์

ABSTRACT

In recently, automatic detection of human activities has gained importance in a research topic due to the individual nature of the activities. Ability to monitor of human daily activities will enhance the capabilities of an application in computer vision. Combination of visual keyword embedding models and a statistical semantic prior model have been recently proposed in the task of mapping images to their contents. Many techniques were proposed to identify activities with supervised dictionary creation methods based on both supervised information with scene description, advantages and limits the discriminative power of the resulting visual words. However, it is a very challenging issue and none of the existing methods provides robust results. In this paper, we propose to classify a human daily activities supervised by learning algorithm based on a Gaussian Mixture model (GM) optimizing with an Expectation-Maximization-based (EM) approach. The approach is composed of four main phases: (1) data preprocessing (2) constructing the modeling with Gaussian distribution (3) Expectation-Maximization (4) measurement and evaluation. We test our model in a publicly available data-set that classify into twelve different daily activities performed by different environment background. This proved to be the case as GMEM with 150 keywords reached average accuracy of ≈ 84.6%. The experimental results indicate that our proposed approach offers significant performance improvements in the classification of human daily activities.

KEYWORDS: image processing, gussian mixture models, expectation maximization, human daily activities

1. บทนำ

ปัจจุบันการรู้จำท่าทางมนุษย์ (human activity recognition) เป็นวิธีการที่เข้ามาแทรกอยู่ในชีวิตประจำวัน ไม่ว่าจะเป็นการนำมาใช้ในการวิเคราะห์กิจกรรม ตรวจสอบลักษณะท่าทางของมนุษย์ที่เกิดขึ้นในกิจวัตรประจำวัน ไม่ว่าจะเป็นการจัดเก็บข้อมูลท่าทางสำหรับการรักษาผู้ป่วยหรือกิจกรรมในที่สาธารณะเพื่อรักษาความปลอดภัย สามารถนำมาประยุกต์เพื่อใช้ในการฝึกฝนสำหรับนักกีฬาให้อยู่ในท่าทางที่เหมาะกับกีฬาประเภทนั้น ๆ และจะพบอยู่ในหลายแอปพลิเคชัน

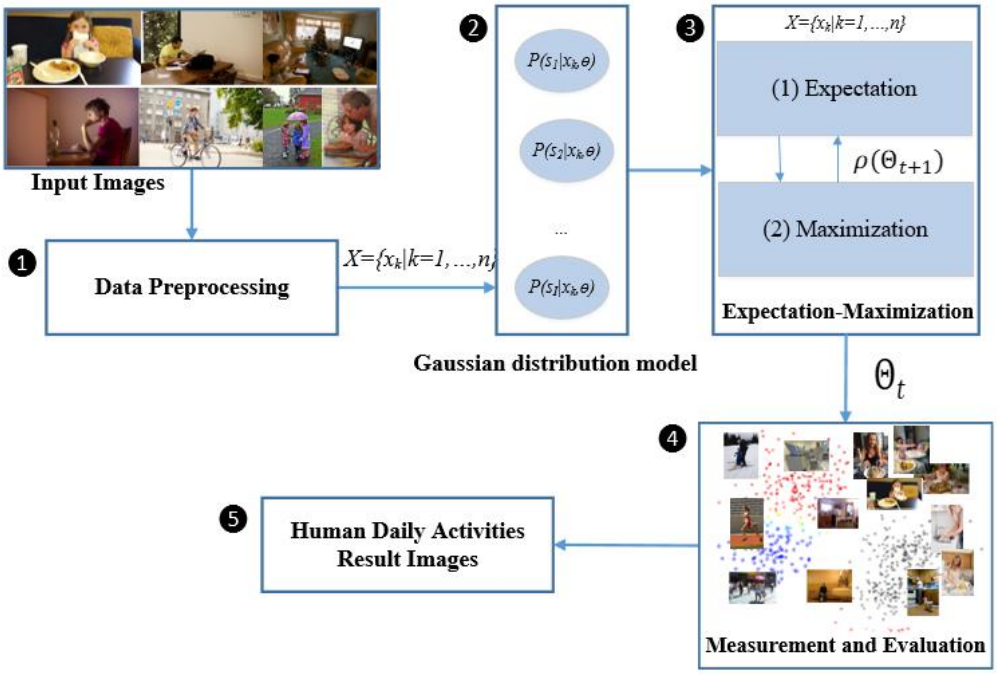
อย่างหลีกเลี่ยงไม่ได้ การวิเคราะห์กิจกรรมของมนุษย์เพื่อนำมาใช้ในการดำเนินการควบคุม ตรวจจับพฤติกรรมของมนุษย์โดยใช้ระบบคอมพิวเตอร์เพื่อนำมาใช้ในการวิเคราะห์ [1-3] เช่น การควบคุมความปลอดภัย (security control) กระบวนการเฝ้าระวัง (surveillance processes) การตรวจสอบกิจกรรม (monitor activity) หรือ การดูแลสุขภาพผู้ป่วย (processes of monitoring a patients' health) เป็นต้น ประกอบกับการพัฒนาทางเทคโนโลยีสารสนเทศและประสิทธิภาพการประมวลผลในระบบเวลาจริงที่รวดเร็ว อุปกรณ์และกล้องดิจิทัลที่มีราคาถูกลง กระบวนการรู้จำท่าทางมนุษย์เข้ามามีบทบาทต่อการดำรงชีวิตประจำวันและทำให้หลายองค์กรนำมาประยุกต์ใช้ในงานในหลากหลายส่วน จึงทำให้มีงานวิจัยหลายกลุ่มพัฒนาวิธีการประมวลผลเชิงลึกเพื่อเพิ่มประสิทธิภาพการทำงานด้วยการนำเทคโนโลยีของ การเรียนรู้เครื่อง (machine learning) และการจดจำรูปแบบ (pattern recognition) [4] เพื่อใช้ในการจำแนกกิจกรรมของมนุษย์ จะนิยมใช้คุณลักษณะที่ได้จาก ค่ากลาง (mean) ฟูเรียทรานสฟอร์ม (Fourier transforms) และการแทนด้วยสัญลักษณ์ หรือการแบ่งส่วนขอข้อมูลและจากการเรียนรู้เพื่อนำมาใช้ในการจำแนกต่อไป แต่อย่างไรก็ตามวิธีการในรูปแบบนี้ยังมีข้อจำกัดที่จำเป็นต้องพิจารณา สำหรับการประมวลผลการวิเคราะห์กิจกรรมของบุคคลจากภาพได้มีการนำข้อมูลเข้าสู่ระบบประมวลผลแบบอัตโนมัติแทนการบันทึกข้อมูลกิจกรรมด้วยระบบทำมือ (manual) จะเพิ่มประสิทธิภาพการทำงานมากกว่า และลดข้อผิดพลาดได้ด้วยการประมวลแบบอัตโนมัติ ดังนั้นเริ่มมีการสกัดข้อมูลด้วยเซ็นเซอร์บนร่างกาย [5] ที่เรียกว่า Body Sensor Networks ทำให้สามารถบันทึกข้อมูลได้จำนวนมากขึ้น และข้อมูลถูกนำมาใช้งาน การรู้จำกิจกรรมมนุษย์ (Human Activity Recognition) [6, 7] เป็นการสกัดข้อมูลจากตำแหน่งหรือมุมจากสัดส่วนของร่างกาย หรือข้อมูลแรงเฉื่อยจากการเคลื่อนไหวของมนุษย์ในการทำแต่ละกิจกรรมได้นำมาเรียนรู้และสร้างรูปแบบทางสถิติเพื่อการรู้จำกิจกรรมของมนุษย์ จากการหาส่วนที่เด่นของท่าทางจากลำดับท่าทางทั้งหมดด้วย การเรียนรู้สร้างความสัมพันธ์ของลำดับท่าทางจากเหตุการณ์ย่อย รูปแบบทางสถิติที่นิยมนำมาใช้ [8, 9]

ในปัจจุบันยังคงมีวิธีการต่าง ๆ เพื่อนำมาใช้แก้ไขปัญหาและเชื่อมโยงความหมายของภาพกิจกรรมที่เกิดขึ้นให้สอดคล้องกับกิจกรรมมนุษย์ ข้อมูลกิจกรรมได้มาจากการสกัดคุณลักษณะ [10-12] ประกอบด้วย การแปลงข้อมูลโดเมนความถี่ (time-frequency transformation) ข้อมูลเชิงสถิติ (statistical approaches) ที่ได้มาจากค่ากลางและความแปรปรวนของโดเมนตามลำดับเวลา และ กระบวนการวิธีการจำแนกข้อมูลทั่วไป เช่น ต้นไม้ตัดสินใจ (decision trees) การคำนวณเพื่อนบ้านใกล้ที่สุด (k-Nearest Neighbour) โครงข่ายประสาทเทียมเพอร์เซฟตรอนแบบหลายชั้น (multilayer perceptron) วิเคราะห์การถดถอยโลจิสติค (logistic regression) และ ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines) โดยข้อมูลจะได้จากการเรียนรู้และระบุกิจกรรมต่าง ๆ จากคุณลักษณะที่กำหนดไว้ การเรียนรู้แบบทั่วไปที่มีการสกัดข้อมูลโดยตรงจากข้อมูลที่ถูกรับเข้ามายังคงเป็นวิธีการที่ถูกนำเข้ามาใช้ในการเรียนรู้เชิงลึก (Deep learning approaches) [13, 14] เช่น

โครงข่ายความเชื่อแบบลึก (Deep Belief Networks) หรือ แมชชีนโบลทซ์มันน์เชิงลึก (Deep Boltzmann Machines) การเพิ่มลำดับชั้นของโหนดสำหรับการจำแนกและการเพิ่มความซับซ้อนของการประมวลผลยังสามารถช่วยทำให้ประสิทธิภาพดีขึ้น เช่น Zeng et al [15] ได้นำโครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Networks) และได้เพิ่มลำดับชั้นการกรองข้อมูล (max-pooling layers) เพื่อสกัดข้อมูลจากข้อมูลที่รับเข้า Alsheikh et al [16] ได้ปรับปรุงความสามารถของการรู้จำทำการรู้จำกิจกรรมมนุษย์ด้วยการใช้การเรียนรู้เชิงลึกจากหลายลำดับชั้นที่ซ่อนตัว (Hidden Layer) และยังได้มีการผสมผสานกับการเรียนรู้เชิงลึกและแบบจำลองมาร์คอฟซ่อนเร้น (hidden Markov model approach) แต่อย่างไรก็ตามยังไม่ได้แก้ไขประสิทธิภาพโดยรวมของระบบ Ordonez et al [17] ได้ทดลองการตรวจจับท่าทางด้วย หลักทางสถิติแบบเบย์บนพื้นฐานของการตรวจจับรูปแบบกิจกรรมในชีวิตประจำวันจากตัวเซ็นเซอร์ภายในบ้านตามช่วงเวลาต่าง ๆ บนท่าทางของมนุษย์ ที่เกิดขึ้นแต่อย่างไรก็ตามโดยส่วนใหญ่กิจกรรมเกิดการทับซ้อนกันของวัตถุจึงเป็นปัญหาที่ทำให้เกิดความตีความที่ผิดพลาดของค่าความถูกต้อง ดังนั้นในงานวิจัยนี้จึงได้นำเสนอการจัดกลุ่มภาพกิจกรรมของมนุษย์ตามคุณลักษณะของความหมายจากคำศัพท์ ด้วยหลักการความน่าจะเป็นโดยการสร้างกฎความเหมือนของคำจากพจนานุกรมคำศัพท์ที่สัมพันธ์กับหลักการทางสถิติ โดยใช้วิธีการแบ่งกลุ่มด้วยวิธีการแจกแจงผสมแบบเกาส์ (Gaussian Mixture Model: GM) และประมาณค่าพารามิเตอร์ด้วยกระบวนการค่าคาดหมายสูงสุด (Expectation-Maximization: EM)

2. ขั้นตอนการจำแนกท่าทางกิจวัตรประจำวันของมนุษย์

งานวิจัยนี้ได้พัฒนาวิธีการจำแนกข้อมูลกิจวัตรประจำวันของมนุษย์ ได้แบ่งขั้นตอนการดำเนินงานวิจัย ออกเป็น 4 ส่วนหลักดังนี้ (1) การเตรียมข้อมูลภาพ (data preprocessing) (2) การกำหนดรูปแบบข้อมูลด้วยการแจกแจงผสมแบบเกาส์ (generate data from Gaussian Mixture Model) (3) การประมาณค่าคาดหมายสูงสุด (Expectation-Maximization) และ (4) การวัดประสิทธิภาพการทำงาน (measurement and evaluation) ดังแสดงในรูปที่ 1



รูปที่ 1 ขั้นตอนการเรียนรู้ด้วยการแจกแจงผสมแบบเกาส์และประมาณค่าคาดหมายสูงสุด

2.1 การเตรียมข้อมูลภาพ

ปัจจุบันมีหลายกลุ่มวิจัยที่ได้จัดทำฐานข้อมูลภาพกิจกรรมมนุษย์ [18-23] ทั้งในรูปแบบข้อมูลภาพนิ่ง และภาพเคลื่อนไหว ข้อมูลที่จัดเก็บจะอยู่ในรูปแบบของโดเมนเวลาที่สัมพันธ์สภาพแวดล้อม และวัตถุในภาพโดยรอบในรูปแบบของคำศัพท์ บางชุดข้อมูลอาจจะมีความต่างกันในรูปแบบของอุปกรณ์ที่ใช้ในการบันทึกข้อมูล แต่ส่วนใหญ่จะอยู่บนท่าทางพื้นฐานของมนุษย์ เช่น ยืน (stand) เดิน (walk) นั่ง (sit) นอน (sleep) เป็นต้น และบางกลุ่มมีการปรับเปลี่ยนตามรูปแบบการเคลื่อนไหว เช่น วิ่ง (running) ปั่นจักรยาน (cycling) รีดผ้า (ironing) พายเรือ (rowing) วิ่งเหยาะ (jogging) กระโดด (jumping) เป็นต้น สำหรับการทดลองนี้จึงมีการคัดเลือกฐานข้อมูลภาพกิจกรรมต่าง ๆ เฉพาะภาพถ่ายบุคคลที่มีการทำกิจกรรมอย่างเด่นชัด และมีความสมบูรณ์ของวัตถุและคำศัพท์ชัดเจนตามกลุ่มที่ต้องการ ดังนั้นจะคัดเลือกข้อมูลภาพจากฐานข้อมูลกิจกรรมมาตรฐาน Pascal VOC [20] และ MS COCO [21, 22] ด้วยสภาพแวดล้อมแตกต่างกัน เช่น สถานที่ทำงาน (office) ห้องครัว (kitchen) ห้องนอน (bedroom) ห้องน้ำ (bathroom) กลางแจ้ง (outdoor) และห้องนั่งเล่น (living room) ดังแสดงในรูปที่ 2

รูปที่ 2 ตัวอย่างภาพกิจกรรมจากฐานข้อมูล Pascal VOC [23] และ MS COCO [21, 22]

2.2 กำหนดรูปแบบข้อมูลด้วยการแจกแจงแบบเกาส์

โดยทั่วไปวิธีที่นิยมสำหรับการแบ่งกลุ่มคือวิธี K-means clustering [23] เป็นการสร้างกลุ่มข้อมูลจากค่าเฉลี่ยที่ใกล้กันเป็นการกำหนดจุดสุ่มออกไปตามตำแหน่งศูนย์กลางและคำนวนค่าเฉลี่ยของจุดทั้งหมดเพื่อหาตำแหน่งกึ่งกลางตามจำนวนกลุ่มที่ต้องการ แต่ละกลุ่มข้อมูลได้จากการคำนวณเพื่อหาค่าใกล้ใดที่ใกล้กับจุดศูนย์กลางมากที่สุด ด้วยการทำซ้ำไปเรื่อย ๆ จนสามารถแยกกลุ่มข้อมูลออกมาได้ดีที่สุด แต่วิธีการนี้จะการจัดกลุ่มขึ้นกับการวัดค่าความต่างกันด้วยวิธีการ Euclidean Distance หรือ Manhattan Distance ของข้อมูล [24] ดังนั้นถ้าค่าความต่างของข้อมูลที่ถูกจัดไว้ภายในกลุ่มเดียวกันมีค่าที่น้อยหรือค่าที่ใกล้เคียงกันมาก แสดงถึงการจัดกลุ่มข้อมูลที่ดี ซึ่งจะเหมาะสมกับค่าข้อมูลที่มีการเปรียบเทียบค่าความหมายเป็นตัวเลข แต่สำหรับการแจกแจงผสมแบบเกาส์ (Gaussian Mixture Model: GM) [25, 26] สามารถจัดการแต่กลุ่มคัสเตอร์ด้วย ความน่าจะเป็นของข้อมูลที่เกิดขึ้น สำหรับงานวิจัยนี้มีการใช้คำศัพท์ในฐานข้อมูล ทำให้คำที่มีความใกล้เคียงกันจะไม่ถูกจำกัดเฉพาะหรือเฉพาะเจาะจงจนเกินไป และยังสามารถหลีกเลี่ยงการเกิดเหตุการณ์ overfitting ได้ด้วย รูปแบบการแจกแจงผสมแบบเกาส์ ถูกสร้างรูปแบบข้อมูล $i.i.d$ ด้วยการกำหนดข้อมูลจากฟังก์ชันความหนาแน่นของความน่าจะเป็น (Probability Density Function) การกระจายตัวของข้อมูลในแต่ละกลุ่มคัสเตอร์ด้วยการแจกแจงด้วยรูปแบบ GM โดยการกำหนด $\Theta = \{\theta_i, i = 1, \dots, I\}$ เมื่อ $\theta_i = \{\mu_i, \mathrm{E}_i, \omega_i\}$ เป็นรูปแบบพารามิเตอร์ที่ i ของเกาส์เซียนซึ่งสอดคล้องกับคัสเตอร์ S_i สามารถเขียนเป็นสมการของความน่าจะเป็น $x_k \in X$ ได้ดังนี้

$$p(x_k|\Theta) = \sum_{i=1}^{I} \omega_i \times N_{\mu_i,\mathrm{E}_i}(x_k), \quad (1)$$

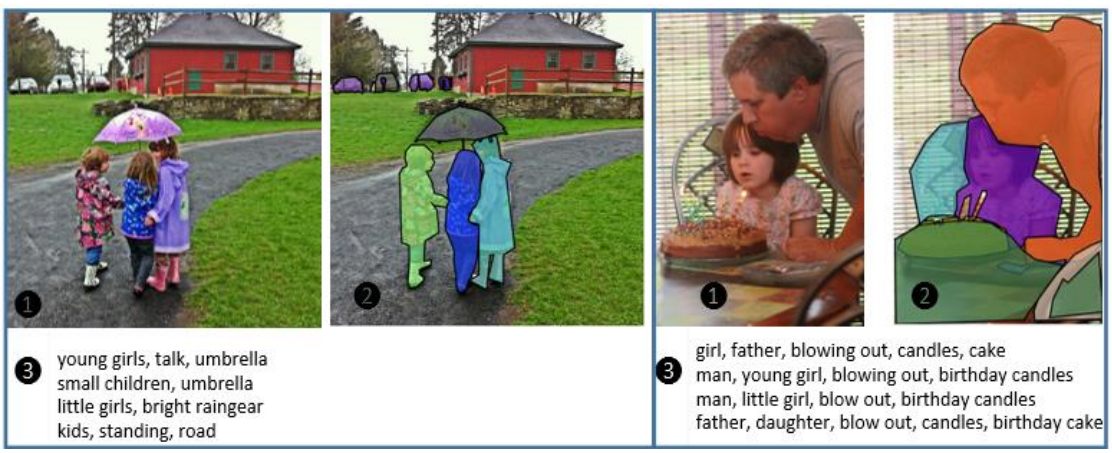
โดยกำหนดให้ข้อมูลในฐานข้อมูลภาพเป็น $X = \{x_k|k = 1, \dots, n, x_k \in \mathbb{R}^D\}$, เมื่อ n เป็นจำนวนข้อมูล และ D แทนมิติของข้อมูล $C = \{c_j|j = 1, \dots, R\}$ แทนกลุ่มข้อมูล เมื่อข้อมูล X จะถูกป้ายกำกับ (labeled) เป็นคำศัพท์ที่ประกอบด้วย ท่าทางพื้นฐาน ท่าทางกิจวัตรประจำวันเวลา สถานที่ชื่อวัตถุ และสภาพแวดล้อมภายนอก รวมทั้งคุณลักษณะอื่นที่ไม่ใช่คำศัพท์ดังแสดงในรูปที่ 3 และ $S = \{s_i|i = 1, \dots, I\}$ แทนชุดข้อมูลคัสเตอร์ เมื่อให้ $I > 1$ และ μ_i เป็นค่ากลาง (mean) E_i เป็นค่าความแปรปรวน (variance) และ ω_i เป็นค่าน้ำหนักของตำแหน่งชุดข้อมูลที่ i และการกระจายแบบเกาเสี้ยน (Gaussian distribution) ถูกแทนด้วย $G_{\mu,\mathrm{E}}(x)$

$$G_{\mu,\mathrm{E}}(x) = \frac{1}{\sqrt{(2\pi)^{D/2}|\mathrm{E}|}} exp\left\{-\frac{1}{2}(x-\mu)' \sum^{-1}(x-\mu)\right\} \quad (2)$$

สำหรับโมเดล GM ค่าของความน่าจะเป็นภายหลัง (Posterior Probability) $p(s_i|x_k, \Theta)$ สามารถเขียนเป็นสมการได้ดังนี้

$$p(is_i|x_k, \Theta) = \frac{\omega_i \times p(x_k|s_i,\theta_t)}{\sum_t \omega_t \times p(x_k|s_t,\theta_t)}, \quad (3)$$

เมื่อ $\sum_i \omega_t = 1$, เมื่อกำหนดให้ค่าความน่าจะเป็น $p(x_k|\mathrm{s_i}, \theta_i)$ สำหรับ x_kในตำแหน่งที่ i ของเกาส์เซียนมีค่าเป็น $G_{\mu_i,\mathrm{E_i}}(x_k)$ จากสมการที่ (2)



รูปที่ 3 ตัวอย่างป้ายกำกับบนภาพกิจกรรม ประกอบด้วยคำศัพท์จากฐานข้อมูลภาพ MS COCO [21, 22]

2.3 การเรียนรู้แบบมีผู้สอนด้วยการแจกแจงผสมแบบเกาส์

รูปแบบของการเรียนรู้มีทั้งการเรียนรู้แบบมีผู้สอน (unsupervised learning) และไม่มีผู้สอน (supervised learning) [24] สำหรับการแบ่งกลุ่มแบบไม่มีผู้สอนนั้นจะทำให้เกิดหลายกลุ่มย่อยและบางครั้งกลุ่มที่ได้เป็นกลุ่มที่ไม่ได้สนใจทำให้ยากแก่การควบคุม โดยเฉพาะการหาความคล้ายกันในความหมายในรูปแบบของคำศัพท์ที่มีความต่างกันของพยัญชนะ (stone และ rock มีความคล้ายกัน) ดังนั้นในงานวิจัยนี้ได้เลือกการเรียนรู้แบบมีผู้สอน ด้วยการนำคำศัพท์ที่ได้จากพจนานุกรมฐานข้อมูลเข้ามาเรียนรู้เพื่อจัดกลุ่มตามความหมายที่ต้องการ ดังแสดงในรูปที่ 3 เป็นภาพตัวอย่างจากฐานข้อมูลหมายเลข ❶ เป็นภาพต้นฉบับ หมายเลข ❷ เป็นวัตถุบนภาพที่ถูกให้ความหมายแสดงบริเวณการแบ่งส่วน คำศัพท์ที่ถูกใช้ในภาพแสดงในหมายเลข❸ และวัดค่าความคล้าย (similarity measure) ด้วยการพิจารณาค่าที่ได้จากการจัดกลุ่มและค่าภาวะน่าจะเป็น (likelihood) ของข้อมูลร่วมกันเพื่อให้เกิดความเที่ยงตรง และป้องกันการเกิด overfitting ได้ด้วย ดังนั้นในงานวิจัยนี้ใช้วิธีการ GM ที่ถูกเรียนรู้ด้วยวิธีการหาค่าคาดหมายสูงสุด [23, 24, 26] เป็นการหาค่าพารามิเตอร์ของภาวะน่าจะเป็นสูงสุด (Maximum likelihood: ML) จะได้ค่าของฟังก์ชันล็อกภาวะน่าจะเป็น (Log-likelihood) ของพารามิเตอร์ Θ ของระบบและชุดข้อมูล X เขียนเป็นสมการของพารามิเตอร์แบบลอการิทึมได้ดังนี้

$$\mathcal{L}(X) = \sum_{i=1}^{n} \log\left(p(x_i|\Theta)\right) \tag{4}$$

โดยที่ $p(s_i|x_k,\Theta)$ และ $p(x_k|\Theta)$ เป็นค่าใดใดที่ถูกประมาณค่าด้วย GM ดังนั้นการพิจารณาค่าความคล้ายกันจากความหมายของคำเท่านั้น สามารถเขียนเป็นค่าประมาณของความน่าจะเป็นได้ $\hat{p}(s_i|x_k,\Theta)$ และ $\hat{p}(x_k|\Theta)$

โดยทั่วไปพารามิเตอร์ของ GM จะได้มาจาก ฟังก์ชันล็อกภาวะน่าจะเป็นของข้อมูลที่มีการใช้วิธีการค่าคาดหมายสูงสุด ดังนั้นในงานวิจัยนี้จะต้องมีการหาค่าพารามิเตอร์ Θ จากการหาจุดสมดุลอาจจะไม่ใช้แต่เพียงการภาวะน่าจะเป็นจากสมการที่ 4 ดังนั้นจึงมีการแบ่งค่าของเป็นสองส่วนคือ $\hat{p}(c_j|x_k)$ และ $\hat{p}(s_i|x_k)$เพื่อใช้ในการประมาณค่าดังนี้

$$\rho(\Theta) = (1-\alpha) \times \sum_{x_k \in X} \log\left(\hat{p}(x_k|\Theta) + \alpha \times \sum_i^I \log\left(\varepsilon(s_i)\right)\right), \tag{5}$$

เมื่อ $0 \le \alpha \le 1$, ถ้า $\alpha = 0$ เป็นการหาค่าที่ไม่ได้เกิดจากการเรียนรู้ แต่ถ้า $\alpha = 1$ แล้วนั้นการหาจุดสมดุลนี้จะเสี่ยงต่อการเกิดเหตุการณ์ overfitting สำหรับการกระจายตัวของ GM จะมีการเลือกเป็นแบบ negative logarithm likelihood และมีการกำหนดค่าการกระจายตัว $\varepsilon(s_i|\theta_i)$ จากค่า entropy ของแต่ละคัสเตอร์ s_i เพื่อให้เกิดความคล้ายกันของข้อมูลภายในคัสเตอร์มากขึ้น

$$\varepsilon(s_i|\theta_i) = -\log\left(-\sum_j \hat{p}(c_j|s_i,\theta_i) \times \log\left(\hat{p}(c_j|s_i,\theta_i)\right)\right) + \phi, \quad (6)$$

เมื่อ $\hat{p}(c_j|s_i,\theta_i)$ เป็นค่าประมาณความน่าจะเป็นของคลาส c_j สำหรับคัสเตอร์ s_j จะขึ้นกับค่า θ_i ที่เป็นพารามิเตอร์ของ GM ของคัสเตอร์ s_i^2 เมื่อ $\phi = \log(\log(R))$, เมื่อ $R > 2$, และอื่นๆ $\phi = 0$ โดยที่ค่าของ $\varepsilon(s_i)$ จะมั่นใจว่ามีค่าฟังก์ชันบวกเสมอ และการกำหนดการแบ่งส่วนของคลาส c_j เพื่อให้กับคัสเตอร์ s_i

$$\hat{p}(c_j|s_i) \propto \frac{\sum_k \hat{p}(c_j,s_i,x_k)}{\hat{p}(s_i)}, \quad (7)$$

$$\hat{p}(c_j|s_i) = \frac{\sum_k \hat{p}(x_k|c_j)\times\hat{p}(x_k|s_i)}{\sum_t\sum_k \hat{p}(x_k|c_t)\times\hat{p}(x_k|s_i)}, \quad (8)$$

จากสมการที่ (5) การปรับค่าให้พอเหมาะและเกิดสมดุลขึ้นสามารถเขียนพารามิเตอร์ฟังก์ชันของการปรับค่าให้เหมาะสมเป็น

$$\Theta^* = argmax_\Theta \rho(\Theta) \quad (9)$$

การแทนค่าจากพารามิเตอร์ข้างต้นด้วยข้อมูลเข้าดังนี้ μ_i ,E_i และ ω_i ด้วยตำแหน่งชุดข้อมูลที่ i เพื่อใช้ Lagrange multiplierแล้ว λ สามารถเขียนเป็นสมการได้ดังนี้

$$\tilde{\rho}(\Theta) = \rho(\Theta) + \lambda(1 - \sum_i \omega_i) \quad (10)$$

$$\frac{\partial\tilde{\rho}}{\partial_{\mu_i,\mathrm{E}_i}} = \sum_k \{(1-\alpha)\hat{p}(s_i|x_k) + \alpha\kappa_i \times \hat{p}(x_k|c_j) \times \hat{p}(x_k|s_i)\} \times \frac{\partial}{\partial(\mu_i,\mathrm{E}_i)} \log\left(\hat{p}(x_k|s_i)\right) \quad (11)$$

เมื่อกำหนดให้

$$\kappa_i = \left(\frac{-1}{F(s_i)[\sum_j \hat{p}(c_j|s_i)\times\log\left(\hat{p}(c_j|s_i)\right)]} \times \frac{1}{\sum_t\sum_k \hat{p}(x_k|c_t)\times\hat{p}(x_k|s_i)}\right) \sum_j (1 + \log\left(\hat{p}(c_j|s_i)\right)) \quad (12)$$

ปรับพารามิเตอร์และเขียนเป็นสมการที่ได้มาแทนลงใน GM ด้วยการเรียนรู้เพื่อประมาณค่าด้วยการทำซ้ำต่อไป

2.4 การประมาณค่าคาดหมายสูงสุด

กระบวนการแจกแจงผสมแบบเกาส์ด้วยการประมาณภาวะน่าจะเป็นสูงสุดโดยใช้ค่าคาดหมายสูงสุด (Expectation-Maximization: EM) [26] เป็นการวิเคราะห์ความสัมพันธ์ของพารามิเตอร์ภายในของแต่ละกลุ่ม ซึ่งเริ่มจากการกำหนดค่าเริ่มต้น $t = 0$ และ การกำหนดค่า Θ_0 ด้วย K-means และใช้ข้อมูลจากกลุ่มพารามิเตอร์ที่ได้จากสมการด้านบน เข้ากระบวนการทำซ้ำจะแบ่งการทำงานเป็น 2 ขั้นตอน

(1) ขั้นตอนการประมาณการ (Expectation) เป็นขั้นตอนคาดคะเนพารามิเตอร์เบื้องต้นจากการเรียนรู้ค่าคาดหวัง ประกอบด้วย $\hat{p}(x_k|s_i), \hat{p}(s_j|x_k), \hat{p}(x_k|c_j), \kappa_i$ และ $\{\forall_i, j, k\}$ ข้อมูลของพารามิเตอร์เพื่อนำไปใช้ในขั้นตอนถัดไป

(2) ขั้นตอนการปรับปรุงค่า (Maximization) จะส่งค่ากลับไปให้ขั้นตอนแรกทำงานกระบวนการทั้งหมดจะหยุดทำเมื่อขั้นตอนในรอบที่ผ่านมากับรอบปัจจุบันมีค่าใกล้เคียงกันมาก จะมีการปรับปรุงค่าดังนี้ Θ_{t+1} และ จากสมการที่ 11 สามารถเขียนเป็นสมการได้ดังนี้

$$\mu_i = \frac{\sum_k\{(1-\alpha)\hat{p}(s_i|x_k)+\alpha B_i \sum_j a_i^j \times \hat{p}(x_k|s_j)\times\hat{p}(x_k|c_i)\}x_k}{\sum_k\{(1-\alpha)\hat{p}(s_i|x_k)+\alpha B_i \sum_j a_i^j \hat{p}(x_k|s_j)\times\hat{p}(x_k|c_i)\}} \tag{13}$$

$$\mathrm{E}_i = \frac{\sum_k\{(1-\alpha)\hat{p}(s_i|x_k)+\alpha B_i \sum_j a_i^j \times \hat{p}(x_k|s_j)\times\hat{p}(x_k|c_i)\}(\mu_i-x_k)(\mu_i-x_k)^t}{\sum_k\{(1-\alpha)\hat{p}(s_i|x_k)+\alpha B_i \sum_j a_i^j \hat{p}(x_k|s_j)\times\hat{p}(x_k|c_i)\}} \tag{14}$$

จากสมการที่ 10 เมื่อแทนค่า $\rho(\Theta) = 0$ เพื่อหาค่า ω_i แล้วสามารถเขียนเป็นสมการได้ดังนี้

$$\omega_i = \frac{\sum_k \hat{p}(s_i|x_k)}{n} \tag{15}$$

จากสมการที่ (13) ถึง (15) ทำการปรับปรุงค่า $\rho(\Theta_{t+1})$ จากสมการ (10) ผลลัพธ์ที่ได้จะเป็นการประมาณค่าความน่าจะเป็นสูงสุดของพารามิเตอร์ Θ_t

3. การวัดประสิทธิภาพการทำงาน

ในงานวิจัยนี้ได้นำเสนอวิธีการแจกแจงผสมแบบเกาส์ด้วยการประมาณภาวะน่าจะเป็นสูงสุดด้วยค่า EM ด้วยเครื่องมือ Bayes Net Toolbox ในโปรแกรม Matlab [27] เพื่อทำการวัดและประเมินผลการทำงานที่สามารถเลือกใช้ค่า EM สำหรับประมาณค่าที่เหมาะสมบนการแจกแจงผสมแบบเกาส์ได้ และในการทดลองนี้ได้คัดเลือกข้อมูลภาพกิจกรรมจาก Pascal VOC [20] และ MS COCO [21, 22] โดยภาพทั้งหมดเป็นกิจวัตรประจำวันของมนุษย์ (Activities of Daily Living)

[28, 29] จะมีความสัมพันธ์กับดำรงชีวิตและการอยู่ในสังคม ตามสภาวะพื้นฐานของกิจกรรม เช่น การแต่งตัวการรับประทานอาหาร การเคลื่อนที่ การเดินทาง และเมื่อนำมาพิจารณาร่วมกันกับอุปกรณ์เครื่องมือร่วมกันกับท่าทางในกิจกรรม สามารถจัดกลุ่มได้ดังนี้ (1) การติดต่อสื่อสาร เช่น มือถืออีเมล อินเทอร์เน็ท เป็นต้น (2) การคมนาคม เช่น รถยนต์ รถไฟ รถสาธารณะ เป็นต้น (3) การประกอบอาหาร เช่น อาหาร อุปกรณ์ครัว เป็นต้น (4) การซื้อของ เช่น อาหาร เสื้อผ้า เป็นต้น (5) งานบ้าน เช่น เครื่องซักผ้า เครื่องดูดฝุ่น เป็นต้น (6) การรักษาพยาบาล และ (7) การจัดการทางด้านการเงิน ดังแสดงภาพตัวอย่างในรูปที่ 2 ประกอบด้วยข้อมูล 4,800 ภาพสำหรับทดลอง และ 3,000 สำหรับการเรียนรู้ ข้อมูลจะประกอบด้วยกลุ่มคำศัพท์ที่ถูกให้ความหมายแล้ว 1,500 คำศัพท์ ใช้ภาพที่ได้จากการคัดเลือกแบบสุ่ม 3,000 ภาพ สำหรับการทดลองในแต่ละชุดข้อมูล ทำการเปรียบเทียบกับวิธีการที่นำเสนอการแจกแจงผสมแบบเกาส์ด้วยวิธีค่าคาดหมายสูงสุด (Gaussian Mixture Model with Expectation-Maximization: GMEM) ด้วย 4 วิธีดังนี้ (1) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines: SVM) [25] (2) เคมีน (K-means) [25] (3) คอร์เนลเคมีน (Kernel K-means: KK-means) [25] และ (4) Supervised Dictionary Learning Model (SDLM) [30, 31] ซึ่ง วิธี K-means กำหนดให้เป็น baseline เพื่อเปรียบเทียบกับการเรียนรู้ด้วยฐานข้อมูลคำศัพท์ด้วยวิธี SDLM จะเป็นการรวมกันระหว่าง รูปแบบการเรียนรู้แบบไม่มีผู้สอน และการเรียนรู้แบบมีผู้สอนด้วยค่าถดถอยโลจิสติค (logistic regression model) จะเปรียบเทียบกับวิธี SVM โดยมีการวัดเป็นค่าความถูกต้องของแต่ละกลุ่มข้อมูล ประกอบด้วย ค่าความแม่นยำ (Precision: Pr.) ค่าความระลึก (Recall: Re.) และ ค่าประสิทธิภาพโดยรวม (F-measure: F_1) สำหรับการทดลองจะแบ่งกลุ่มตามกิจวัตรประจำวันของมนุษย์ ออกเป็น 12 กลุ่ม ดังนี้ การอาบน้ำ (bathing) การแต่งตัว (dressing) การกิน (eating) การซื้อของ (shopping) การทำอาหาร (cooking) งานบ้าน (housework) ซักรีด (Laundry) ขับรถ (driving) ออกกำลังกาย (exercise) การพักผ่อน (leisure) ทำงาน (working) เดิน (walking)

4. ผลการทดลอง

การทดลองจำแนกกิจวัตรประจำวันสำหรับวิธีการ GMEM ได้กำหนดค่าควบคุมการทำงานของฟังก์ชัน $\rho(\Theta)$ ด้วยค่า $\lambda = 0.5$ เป็นค่าที่เหมาะสมสำหรับการทำงานด้วยการเรียนรู้ของวิธีการ GMEM มากที่สุด และสุ่มข้อมูลภาพกิจกรรมออกเป็น 2 ชุด ชุดละ 3,000 ภาพ จำนวนคำศัพท์ที่ใช้ในการทดลอง 250 และ 150 คำ สำหรับการทดลองที่ 1 และ 2 ตามลำดับ ทำการทดลองด้วยการเปรียบเทียบกับวิธี SVM K-Means KK-Means SDLM และ GMEM แสดงผลการทดลองเป็นค่าเฉลี่ยความถูกต้องรวม (total average accuracy) ค่าความแม่นยำ (Pr.) ค่าความระลึก (Re.) และค่าประสิทธิภาพโดยรวม (F_1) แสดงในตารางที่ 1 และ 2 เป็นชุดข้อมูลการทดลองที่ 1 และ 2 ตามลำดับ

ตารางที่ 1 เปรียบเทียบประสิทธิภาพของการจำแนกข้อมูลด้วยชุดข้อมูลที่ 1 (λ =0.5)

Activity	Method														
	SVM			K-means			KK-means			SDLM			GMEM		
	Pr.	Re.	F_1	Pr.	Re.	F_1	Pr.	Re.	F_1	Pr.	Re.	F_1	Pr.	Re.	F_1
Bathing	53.5	58.1	55.7	66.7	55.2	60.4	55.4	56.0	55.7	74.7	66.3	70.3	79.8	89.3	84.3
Relaxin.	43.0	55.3	48.4	62.9	63.6	63.3	52.0	53.0	52.5	66.3	68.3	67.3	79.6	78.7	79.1
Cooking	54.3	52.1	53.2	58.8	61.0	59.9	56.6	53.3	54.9	76.5	75.8	76.1	89.2	75.5	81.8
Dressin.	49.1	44.5	46.7	60.0	61.4	60.7	53.3	60.6	56.7	71.6	70.9	71.2	96.0	82.8	88.9
Shopin.	44.8	46.7	45.7	51.6	57.1	54.2	56.3	63.7	59.8	71.7	80.9	76.0	93.2	75.8	83.6
Eating	42.7	45.6	44.1	50.5	50.5	50.5	54.6	61.5	57.8	70.8	72.1	71.4	73.5	83.3	78.1
Workin	57.3	52.0	54.5	54.7	58.0	56.3	60.2	54.9	57.4	70.1	75.0	72.5	71.6	80.4	75.7
HousW.	41.1	42.2	41.6	49.5	50.0	49.7	60.0	63.6	61.8	68.5	75.5	71.8	63.2	87.8	73.5
Exercise	75.0	58.7	65.9	59.6	59.6	59.6	64.6	65.3	64.9	84.0	79.0	81.4	93.3	83.8	88.3
Driving	50.4	55.3	52.8	56.3	61.1	58.6	60.2	65.3	62.6	75.5	74.0	74.7	80.4	85.4	82.8
Walking	51.4	50.5	50.9	56.4	61.4	58.8	59.4	55.3	57.3	70.3	68.9	69.6	83.9	72.9	78.0
Laundry	74.1	64.3	68.9	65.5	52.8	58.5	84.4	59.6	69.9	84.3	75.0	79.4	91.3	85.7	88.4
Ave. Acc.	**51.8**			**57.5**			**59.2**			**73.4**			**81.6**		

จากตารางที่ 1 ในชุดข้อมูลที่ 1 ใช้คำศัพท์รวมทั้งหมด 200 คำ ได้ค่าเฉลี่ยความถูกต้องรวมของวิธีการ SVM K-Means KK-Means SDLM และ GMEM อยู่ที่ 51.8% 57.5% 59.2% 73.4% และ 81.6% ตามลำดับ จะเห็นว่าการจัดกลุ่มด้วย GMEM ในกลุ่มกิจกรรม Dressing และ Exercise มีค่าความแม่นยำสูงสุดอยู่ที่ 96% และ 93.3% แต่การจัดกลุ่มด้วยวิธี SVM แบบมีผู้สอนสามารถจัดกลุ่มของ Eating และ Housework ด้วยค่าความแม่นยำเพียง 42.7% 41.1% ในทางกลับกัน GMEM สามารถจัดกลุ่มได้ถึง 73.5% 63.2% และเมื่อทำการทดลองโดยลดจำนวนการใช้คำศัพท์ลงให้เหลือเพียง 150 คำด้วย ชุดข้อมูลที่ 2 ดังแสดงผลในตารางที่ 2 ได้ค่าเฉลี่ยความถูกต้องรวมของวิธีการ SVM K-Means KK-Means SDLM และ GMEM อยู่ที่ 61.0% 64.6% 68.51% 74.45% และ 84.6% ตามลำดับ จะเห็นว่าวิธี GMEM สามารถจัดกลุ่ม Bathing Relaxing และ Housework ได้ค่าความแม่นยำเพิ่มขึ้นจากเดิมถึง 84.4% 83.2% และ 74.7% และเช่นเดียวกันวิธี SDLM สามารถจำแนกในกลุ่ม Exercise ได้ค่าความแม่นยำสูงถึง 84.9% และกลุ่ม Laundry สูงถึง 85.9% แต่วิธี GMEM สามารถจำแนกได้ 87.9% และ 91.2% จะเห็นว่าการจำแนกด้วยวิธี GMEM สามารถได้ผลลัพธ์ที่มีค่าเฉลี่ยความถูกต้องสูงถึง 84.6% สำหรับการทดลองโดยใช้คำศัพท์ 150 คำ และจำแนกได้ค่าเฉลี่ยความถูกต้อง 81.6% สำหรับการทดลองโดยใช้คำศัพท์ 200 คำ มากกว่าวิธีการอื่น ๆ

ตารางที่ 2 เปรียบเทียบประสิทธิภาพของการจำแนกข้อมูลด้วยชุดข้อมูลที่ 2 (λ =0.5)

Activity	Method														
	SVM			K-means			KK-means			SDLM			GMEM		
	Pr.	Re.	F_1	Pr.	Re.	F_1	Pr.	Re.	F_1	Pr.	Re.	F_1	Pr.	Re.	F_1
Bathing	62.9	64.2	63.5	63.0	67.0	64.9	72.0	60.4	65.7	73.7	77.7	75.6	84.4	88.0	86.2
Relaxin.	66.3	57.6	61.6	62.0	57.0	59.4	67.4	56.9	61.7	70.8	78.9	74.6	83.2	83.2	83.2
Cooking	53.7	53.7	53.7	68.3	57.1	62.2	68.0	58.1	62.7	69.9	76.6	73.1	86.4	77.6	81.7
Dressin.	53.1	57.1	55.0	68.7	55.9	61.6	67.0	69.7	68.3	70.1	68.8	69.4	95.1	85.7	90.2
Shopin.	58.6	52.6	55.4	62.0	58.8	60.3	62.7	64.0	63.4	70.8	73.5	72.1	90.4	81.5	85.7
Eating	55.8	63.0	59.2	63.6	63.6	63.6	56.7	73.5	64.0	74.4	68.4	71.3	78.8	86.7	82.5
Workin	62.9	68.0	65.3	57.8	67.7	62.4	67.3	70.4	68.8	72.4	73.2	72.8	82.3	78.2	80.2
HousW.	63.9	66.0	64.9	60.9	63.8	62.3	70.1	78.9	74.3	71.6	69.5	70.5	74.7	81.3	77.9
Exercise	65.5	76.0	70.4	67.9	79.6	73.3	70.9	83.0	76.5	84.9	80.6	82.7	87.9	87.9	87.9
Driving	65.5	57.3	61.1	69.1	65.0	67.0	71.0	71.7	71.4	77.3	78.9	78.1	82.4	93.7	87.7
Walking	61.4	62.0	61.7	61.1	70.4	65.4	70.2	70.9	70.5	75.0	76.5	75.7	82.5	84.2	83.3
Laundry	63.4	53.6	58.1	75.0	71.6	73.3	85.2	69.7	76.7	85.9	72.3	78.5	91.2	88.3	89.7
Ave. Acc.	61.0			64.60			68.51			74.45			84.6		

5. สรุปผลการทดลอง

งานวิจัยนี้ได้นำเสนอรูปแบบการจำแนกข้อมูลภาพที่เป็นกิจกรรมของมนุษย์โดยจะแบ่งกลุ่มตามกิจวัตรประจำวันที่เกิดขึ้นเป็น 12 กลุ่มทั่วไป และใช้วิธีการแจกแจงผสมแบบเกาส์ด้วยวิธีค่าคาดหมายสูงสุด ซึ่งมีการหาความสัมพันธ์ของคำศัพท์ และค่าเหมาะสมของพารามิเตอร์ภายในกลุ่มข้อมูลที่ใช้ในการจำแนก เพื่อให้ได้ผลลัพธ์ตามกลุ่มความหมายที่ดีที่สุด จากผลการทดลองจะเห็นว่า มีค่าเฉลี่ยความถูกต้องสูงถึง 84.6% สำหรับการทดลองโดยใช้คำศัพท์ 150 คำ และค่าเฉลี่ยความถูกต้อง 81.6% สำหรับการทดลองโดยใช้คำศัพท์ 200 คำมากกว่าวิธีการอื่น ๆ แต่อย่างไรก็ตามเมื่อทำการทดลองโดยเพิ่มคำศัพท์จะเห็นว่ามีค่าเฉลี่ยความถูกต้องที่ลดต่ำลงซึ่งยังคงเป็นสิ่งที่ต้องแก้ไขและปรับจำนวนกลุ่มให้มีความหลากหลายมากขึ้นสำหรับการทดลองต่อไป

References

[1] Ranasinghe DC, Torres RLS, Wickramasinghe A. Automated activity recognition and monitoring of elderly using wireless sensors: Research challenges. 5th IEEE International Workshop on Advances in Sensors and Interfaces; 2013. p. 224-7.

[2] Ann OC, Theng LB. Human activity recognition: A Review. 2014 IEEE International Conference on Control System, Computing and Engineering; 2014. p. 389-93.

[3] Kay W, et al. The kinetics human action video dataset. Computer Vision and Pattern Recognition. 2017.

[4] Zainudin MNS, Sulaiman MN, Mustapha N, Perumal T. Activity recognition based on accelerometer sensor using combinational classifiers. 2015 IEEE Conference on Open Systems (ICOS); 2016. p. 68-73.

[5] Bulling A, Blanke U, Schiele B. A tutorial on human activity recognition using body-worn inertial sensors. ACM Computing Surveys (CSUR) 2014;46(3):1-33.

[6] Yang JB, Nguyen MN, San PP, Li XL, Krishnaswamy S. Deep convolutional neural networks on multichannel time series for human activity recognition. Proceedings of the 24th International Conference on Artificial Intelligence. AAAI Press; 2015. p. 3995-4001.

[7] Chinpanthana N, Phiasai T. Deep textual Searching for Visual Semantics of Personal Photo Collections with a Hybrid Similarity Measure. International Symposium on Computer Science and Intelligent Controls; 2017. p. 124-28.

[8] Fine S, Singer Y, Tishby N. The hierarchical hidden Markov model: analysis and applications. Machine Learning 1998;32:41-62.

[9] Natarajan P, Nevatia R. Coupled hidden semi Markov models for activity recognition. 2007 IEEE Workshop on Motion and Video Computing; 2007.

[10] Jindong W, Yiqiang C, Shuji H, Xiaohui P, Lisha H. Deep Learning for Sensor-based Activity Recognition: A Survey Pattern Recognition Letters. 2019;119: 3-11.

[11] Catal C,Tufekci S, Pirmit E, Kocabag G. On the use of ensemble of classifiers for accelerometer-based activity recognition. Applied Soft Computing 2015;37:1018-22.

[12] Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. International Workshop of Ambient Assisted Living (IWAAL 2012); 2012. p. 216-23

[13] Lopes N, Ribeiro B. Deep belief networks (DBNs). Machine Learning for Adaptive Many-Core Machines - A Practical Approach. Springer International Publishing; 2015. p. 155-86.

[14] Chen Y, Xue Y. A deep learning approach to human activity recognition based on single accelerometer. 2015 IEEE International Conference on Systems, Man, and Cybernetics; 2015. p. 1488-92.

[15] Zeng M, Nguyen LT, Yu B, Mengshoel OJ, Zhu J, Wu P, et al. Convolutional Neural Networks for human activity recognition using mobile sensors. 6th International Conference on Mobile Computing, Applications and Services; 2014. p. 197-205.

[16] Alsheikh MA, Selim A, Niyato D, Doyle L, Lin S, Tan HP. Deep activity recognition models with triaxial accelerometers. AAAI workshop; 2016.

[17] Ordonez FJ, Toledo P, Sanchis A. Sensor-based bayesian detection of anomalous living patterns in a home setting. Personal Ubiquitous Comput 2015;19(2):259-70.

[18] Jennifer RK, Gary MW, Samuel AM. Activity recognition using cell phone accelerometers. ACM SigKDD Explorations Newsletter 2011;12(2):74-82.

[19] Reyes-Ortiz JL, Oneto L, Samà A, Parra X, Anguita D. Transition-aware human activity recognition using smartphones. Neurocomputing 2016;171:754-67.

[20] Everingham M, Gool L, Williams CK, Winn J, Zisserman A. The PASCAL visual object classes (VOC) challenge. International Journal of Computer Vision 2010;88:303-38.

[21] Caesar H, Uijlings J, Ferrari V. COCO-Stuff: Thing and stuff classes in context. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018. p. 1209-18.

[22] Lin T, Maire M, Belongie SJ, Hays J, Perona P, Ramanan D, et al. Microsoft COCO: common objects in context. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. Springer, Cham. p. 740-55.

[23] Dhillon IS, Guan Y, Kulis B. Kernel K-means: spectral clustering and normalized cuts. Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2004. p. 551-6.

[24] Frey PW, Slate DJ. Letter recognition using Holland-style adaptive classifiers. Machine Learning.1991;6:161-82.

[25] Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006.

[26] Banfield JD, Raftery AE. Model-based Gaussian and non-Gaussian clustering. Biometrics 1993;49:803-21.

[27] Chen M. EM algorithm for Gaussian mixture model (EM GMM) [Internet]. MATLAB Central File Exchange; 2021 [cited 2021 January 8]. Available from: https://www.mathworks.com/matlabcentral/fileexchange/26184-em-algorithm-for-gaussian-mixture-model-em-gmm

[28] Drumm-Boyd C. Activities & instrumental activities of daily living - definitions, importance and assessments [Internet]. 2020. Available from: https://www.payingforseniorcare.com/activities-of-daily-living.

[29] Costenoble A, Knoop V, Vermeiren S, Vella RA, Debain A, Rossi G, et al. A comprehensive overview of activities of daily living in existing frailty instruments: a systematic literature search. The Gerontologist; 2019.

[30] Lian X, Li Z, Wang C, Lu B, Zhang L. Probabilistic models for supervised dictionary learning. Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition; 2010. p. 2305-12.

[31] Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A. (2008). Discriminative learned dictionaries for local image analysis. IEEE Conference on Computer Vision and Pattern Recognition; 2008.

ประวัติผู้เขียนบทความ

เตชค์ฐสิณป์ เพียซ้าย อาจารย์ประจำสาขาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช งานวิจัยทางด้านโทรคมนาคมไฟฟ้าอิเลคทรอนิค การประมวลด้วยเสียง IoT การประมวลผลแบบ 3D การประมวลผลภาพดิจิทัล การประมวลผลภาพวิดีโอแบบเรียลไทม์ อีเมล tejtasin.phi@stou.ac.th

นัศพ์ชาณัณ ชินปัญช์ธนะ อาจารย์ประจำวิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์ งานวิจัยทางด้านการประมวลผลภาพดิจิทัลทางด้านความหมายภาพ งานประมวลผลภาพทางด้านโรงงานอุตสาหกรรมการผลิต การจำแนกภาพดิจิทัล การรู้จำภาพดิจิทัล การประมวลผลภาพวิดีโอแบบเรียลไทม์ อีเมล nutchanun.cha@dpu.ac.th

Article History:

Received: February 17, 2021

Revised: April 1, 2021

Accepted: April 17, 2021