

การวิเคราะห์การมีส่วนร่วมของนักเรียนในห้องเรียนออนไลน์ โดยใช้ Convolutional Neural Networks (CNN)

Analysis of Student Engagement in Online Classroom using Convolutional Neural Networks (CNN)

พงษ์ศธร เชิดสม* และ วนิดา แก่นอากาศ

ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
มหาวิทยาลัยขอนแก่น
* ผู้รับผิดชอบบทความ
pongche@kku.ac.th

Received: 7 Sep 2023
Revised: 24 Dec 2023
Accepted: 26 Dec 2023

บทคัดย่อ

การระบาดของเชื้อไวรัสโคโรนา (COVID-19) ส่งผลกระทบต่อภาคการศึกษา เช่น การเรียนจากห้องเรียนปกติสู่ห้องเรียนออนไลน์ ทำให้การติดตามการมีส่วนร่วมในห้องเรียนออนไลน์เป็นไปด้วยความยากลำบากนอกจากจะส่งผลต่อประสิทธิภาพของผู้เรียนแล้วกรณีที่ร้ายแรงที่สุดที่อาจเกิดขึ้นคือการหลุดจากการศึกษาของผู้เรียน เพื่อให้ผู้สอนได้ทราบถึงการมีส่วนร่วมของผู้เรียนและสามารถปรับเปลี่ยนการเรียนการสอนให้เหมาะสมกับสภาพแวดล้อมในการเรียนออนไลน์ บทความนี้จึงได้นำเสนอการพัฒนาแบบจำลองที่ใช้ในการตรวจสอบการมีส่วนร่วมในชั้นเรียนออนไลน์โดยใช้โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Networks : CNN) ที่ใช้ใบหน้าของผู้เรียนในการตรวจจับอารมณ์และเชื่อมโยงเป็นระดับการมีส่วนร่วม 3 ระดับ คือ ไม่มีส่วนร่วม มีส่วนร่วมระดับปกติ มีส่วนร่วมระดับสูง ซึ่งได้ทำการปรับแต่งพารามิเตอร์ได้แก่ Learning Rate, Batch Size, Optimizer รวมทั้งการปรับแต่งในชั้นของการทำ Feature extraction และชั้นของการทำ Fully connected ซึ่งได้นำโครงสร้างแบบจำลองจากเว็บไซต์ Keras (Simple MNIST convnet) เป็นโครงสร้างแบบจำลองตั้งต้น ได้แบ่งการทดลองออกเป็น 3 กลุ่ม ได้แก่กลุ่มที่ 1 คือ การปรับแต่งโครงสร้างชั้นการสกัดคุณลักษณะเด่นของภาพ (Feature extraction) กลุ่มที่ 2 คือ การทดลองแต่ละพารามิเตอร์ ได้แก่ Learning Rate (LR), Batch Size (BZ), Optimizer (OTM), Fully connected (FC) กลุ่มที่ 3 คือ การทดลองจับคู่ 2 พารามิเตอร์ ซึ่งได้ทดลองกลุ่มที่ 1-3 ตามลำดับโดยใช้แบบจำลองจากกลุ่มการทดลองก่อนหน้าที่ได้ผลดีที่สุดมาใช้ในการทดลองกลุ่มถัดไป ผลการประเมินประสิทธิภาพแบบจำลองการใช้และการปรับแต่ง Optimizer Ranger ร่วมกับ Fully Connected (FC) จำนวนชั้น 2 ชั้นคือ 50 และ 100 ให้ประสิทธิภาพที่ดีที่สุดคือค่าความแม่นยำร้อยละ 82.30 (Accuracy) และค่าความสูญเสียร้อยละ 0.46 (Loss) ซึ่งมีค่าความแม่นยำเพิ่มขึ้นร้อยละ 16.51 และมีค่าความสูญเสียลดลงร้อยละ 0.31 เมื่อเทียบกับแบบจำลองตั้งต้น

คำสำคัญ: การตรวจจับการมีส่วนร่วม การจดจำอารมณ์ การมีส่วนร่วมในชั้นเรียน

Abstract

The pandemic of the coronavirus (COVID-19) has affected the education system, disrupting all from traditional classrooms to online classrooms. This complicates tracking involvement in online classes even further. A student's separation from education is the most serious scenario that can occur, in addition to damaging the efficiency of the learners. Teachers should indeed be regularly

informed about student participation so that they may customize their instruction to the online learning environment. This paper presents a model for analyzing and tracking student participation in online classrooms using neural networks. The approach utilizes Convolutional Neural Networks (CNN) to analyze learners' faces, employing a pre-trained base model obtained from the Keras website. The model categorizes student engagement into three levels: disengagement, normal engagement, and high engagement. The experiments were divided into three groups: adjusting the image feature extraction layer, analyzing individual parameters (Learning Rate, Batch Size, Optimizer, Fully connected), and conducting a sequential two-parameter matching test. The best-performing models from each test were applied to subsequent trials, leading to progressive improvements in the model's performance for monitoring online class participation. The evaluation of the model's performance yielded promising results. By incorporating the Optimizer Ranger and employing a Fully Connected (FC) layer with a configuration of 50-100 units, the accuracy of the model experienced a significant boost, reaching an impressive 82.30%. Simultaneously, the loss was notably reduced to 0.46. These improvements were substantial compared to the baseline model, with accuracy seeing a remarkable enhancement of 16.51% and a reduction in loss by 0.31. These findings showcased the effectiveness of the proposed approach in accurately monitoring and categorizing learner engagement levels based on facial analysis.

Keywords: Engagement detection, Emotion recognition, Class engagement

1. บทนำ

เมื่อเดือนธันวาคม 2562 เกิดการแพร่ระบาดของเชื้อไวรัสโคโรนา 2019 (COVID-19) ซึ่งมีต้นกำเนิดการระบาดจากเมืองอู่ฮั่น ประเทศจีน [1] และมีการแพร่ระบาดไปทั่วโลกมากกว่า 200 ประเทศ [2] ส่งผลกระทบในทุก ๆ ภาคส่วน รวมทั้งภาคการศึกษาที่ได้มีการปรับเปลี่ยนการเรียนการสอน จากห้องเรียนปกติสู่ห้องเรียนออนไลน์ทำให้การติดตามพฤติกรรมการณ์มีส่วนร่วมเป็นไปด้วยความยากลำบากและยากลำบากมากกว่าการเรียนในห้องเรียนแบบปกติหลายเท่าตัวหรือไม่สามารถติดตามได้เลย

ในการวิเคราะห์และการติดตามการมีส่วนร่วม จาก [3] ได้แบ่งวิธีในการวิเคราะห์และติดตามการมีส่วนร่วมของผู้เรียนออกเป็น 3 แบบคือ 1) แบบอัตโนมัติ ที่เป็นการใช้เครื่องมือหรืออุปกรณ์ เข้ามาช่วยในการวิเคราะห์และประเมินผล เช่นการใช้ข้อมูลจากเซ็นเซอร์ ข้อมูลการเข้าใช้งานเว็บไซต์ รูปภาพหรือวิดีโอที่ถูกบันทึกในการเรียนการสอนโดยมีการใช้ใบหน้า การเคลื่อนไหวและท่าทางของร่างกาย การเคลื่อนไหวของดวงตา 2) แบบกึ่งอัตโนมัติ ที่เป็นการถามตอบในชั้นเรียน 3) แบบดำเนินการเองโดยมนุษย์ ที่ได้รับความนิยมมากที่สุดในการประเมินการเรียนและประเมินระดับการมีส่วนร่วมโดยการใช้แบบประเมินจากผู้สอนหรือตัวผู้เรียนเองถึงแม้ว่าการดำเนินการเองโดยมนุษย์จะถูกใช้นำมาใช้มากที่สุดในการเรียนในห้องเรียนปกติในสถานการณ์ที่ไม่ปกติและการเปลี่ยนแปลงวิธีการเรียนการสอนให้มีความเหมาะสมจึงมีการใช้เทคโนโลยีและการติดตามการมีส่วนร่วมแบบอัตโนมัติเข้ามาช่วยมากยิ่งขึ้น

ซึ่งในปัจจุบันมีการใช้ศาสตร์ทางด้านคอมพิวเตอร์ที่เรียกว่าปัญญาประดิษฐ์ (AI : Artificial Intelligence) เข้ามาช่วยวิเคราะห์และติดตามการมีส่วนร่วมของผู้เรียนที่มีการใช้เทคโนโลยีในการประมวลผลภาพหรือวิดีโอที่ให้ความแม่นยำและมีความน่าเชื่อถือสูงที่เรียกว่าคอมพิวเตอร์วิทัศน์ (Computer Vision) เป็นการใช้ความสามารถจากการเรียนรู้ของเครื่อง (Machine learning) และการเรียนรู้เชิงลึก (Deep Learning) ในการวิเคราะห์การมีส่วนร่วมของผู้เรียน จาก [4] ใช้อัลกอริทึมที่เป็นที่นิยมอย่าง Convolutional Neural Network (CNN) ที่มีความสามารถพิเศษในการสกัดคุณลักษณะเด่นของภาพ

(Feature extraction) ที่มีความซับซ้อนที่เหมาะสมกับข้อมูลที่ไม่มีโครงสร้างมาใช้ในการสร้างแบบจำลอง 1 แบบจำลอง และอีก 3 จำลองจากบทความอื่น ๆ มาสร้างแบบจำลองในการวิเคราะห์การมีส่วนร่วมของผู้เรียนและเปรียบเทียบผลของแบบจำลองจากใบหน้าของผู้เรียนแบ่งระดับการมีส่วนร่วมออกเป็น 3 ระดับ ได้แก่ ไม่มีส่วนร่วม มีส่วนร่วมระดับปกติ และส่วนร่วมระดับมาก แบบจำลองที่พัฒนาขึ้นได้ค่าความแม่นยำร้อยละ 92.33 มากกว่า 3 แบบจำลองที่นำมาจากบทความอื่น ๆ

การพัฒนาแบบจำลองในปัจจุบันได้มีเฟรมเวิร์กและไลบรารีที่น่าสนใจมากมาย เช่น TensorFlow Torch, Keras, SciKit-Learn ในบทความนี้ได้เลือกใช้ไลบรารี Keras ในการพัฒนาแบบจำลองซึ่งเป็น API ที่ทำงานอยู่บน TensorFlow มีความเรียบง่าย ยืดหยุ่น และทรงพลังเหมาะสำหรับใช้ในการนำมา ทดลองหรือวิจัยเพื่อนำเสนอแนวคิดต่างๆ ถูกนำมาใช้ในหน่วยงานต่าง เช่น CERN, NASA, NIH และองค์กรทางวิทยาศาสตร์อีกมากมายทั่วโลก [5]

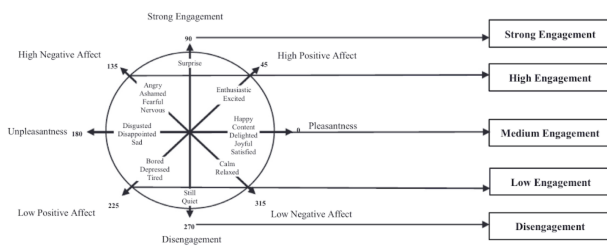
ในการพัฒนาแบบจำลองให้มีประสิทธิภาพจำเป็นต้องอาศัยการปรับพารามิเตอร์ (Hyper Parameter) ให้มีความเหมาะสมและยังมีพารามิเตอร์ที่น่าสนใจในการปรับแต่งเพื่อให้แบบจำลองมีประสิทธิภาพมากยิ่งขึ้น เช่น อัตราการเรียนรู้ (Learning Rate) ซึ่งถูกขนานนามว่าเป็นเจ้าพ่อของพารามิเตอร์ [6] ได้เปรียบเทียบการใช้งาน Learning Rate กับแบบจำลองที่ได้พัฒนาขึ้นและแบบจำลองสำเร็จรูปที่เป็นที่นิยมในการจำแนกภาพของโรคมะเร็งในระยะลุกลาม รวมทั้ง [7] มาใช้ในการจำแนกภาพของโรคมะเร็ง และ [8] ได้นำเสนอผลการเปรียบเทียบการใช้ Learning Rate ของ Machine Learning ซึ่งมีค่าตั้งแต่ 0.10 - 0.80 พารามิเตอร์ต่อมาคือ Optimizer [9] ได้นำเสนอการเปรียบเทียบ optimizer 9 ตัว ได้แก่ SGD vanilla, SGD momentum + nesterov, SGD momentum, AdaGrad, RMSProp, Adam, AdaDelta, Adamax, Nadam [10] ได้นำเสนอการใช้งาน Optimizer อย่าง RAdam ซึ่งนำมาทดสอบกับแบบจำลอง DenseNet, Resnet, VGG16 ในการจำแนกภาพที่บ่งบอกว่าเป็นมะเร็งหรือไม่ [11] ได้นำเสนอการใช้งาน optimizer อย่าง Rectified Adam ในการพยากรณ์อารมณ์ของมนุษย์ และ [12] ได้นำเสนอแบบจำลองในการการตรวจหาโรคอัลไซเมอร์ซึ่งใช้ optimizer อย่าง Rectified Adam และ

Batch Size [13] ได้นำเสนอผลการเปรียบเทียบการใช้งาน Batch Size ได้แก่ 16, 32, 64, 128, 256 ในแบบจำลองของ CNN

บทความนี้เสนอการพัฒนาแบบจำลองที่ใช้ในการวิเคราะห์พฤติกรรมการมีส่วนร่วมของผู้เรียน โดยได้เลือกใช้ อัลกอริทึม CNN ที่มีการปรับแต่งพารามิเตอร์ได้แก่ Learning Rate, Batch Size, Optimizer รวมทั้งการปรับแต่งในชั้นของการสกัดคุณลักษณะ (Feature extraction) และชั้นของการทำ Fully connected และแบ่งระดับการมีส่วนร่วมออกเป็น 3 ระดับ คือ ไม่มีส่วนร่วม มีส่วนร่วมระดับปกติ มีส่วนร่วมระดับมาก ชุดข้อมูลที่นำมาใช้สำหรับการเรียนรู้ของแบบจำลองคือ AffectNet เป็นชุดข้อมูลใบหน้าของมนุษย์ซึ่งเป็นเด็กจนถึงผู้สูงอายุที่ระบุอารมณ์ของใบหน้าจำนวน 8 อารมณ์ อีก 3 กลุ่มที่ไม่ได้ระบุอารมณ์ และมีจำนวนมากกว่า 1 ล้านภาพที่มาจากอินเทอร์เน็ต โดยใช้ 6 ภาษาในค้นหาใบหน้าทำให้ข้อมูลตัวอย่างของใบหน้าที่มีความหลากหลายในแต่ละประเทศประกอบด้วย อารมณ์เป็นกลาง ความสุข ความโศกเศร้า ตื่นเต้น ความกลัว รังเกียจ ความโกรธ และดูถูก [38] เป็นชุดข้อมูลที่มีการนำมาใช้ในการวิจัยอย่างแพร่หลายและมีประสิทธิภาพมากที่สุดในการนำมาใช้งานที่เกี่ยวข้องกับงานวิจัยด้านนี้

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การเชื่อมโยงระหว่างอารมณ์พื้นฐานและการมีส่วนร่วม โดย [14] ได้มีการแบ่งโดย การเจียบ ความสงบ คือ ไม่มีส่วนร่วม เบื่อหน่าย หดหู่ เหิน้อย ง่วงนอน ผ่อนคลาย คือ มีส่วนร่วมระดับน้อย มีความสุข ดีใจ สนุกสนาน พอใจ รังเกียจ เศร้า ผิดหวัง คือ มีส่วนร่วมระดับปกติ กระตือรือร้น ตื่นเต้น โกรธ หวาดกลัว ประหม่า ละอวยใจ คือ มีส่วนร่วมระดับมาก แปลกใจ ประหลาดใจ คือ มีส่วนร่วมระดับมากที่สุด แสดงดังรูปที่ 1 และ [15] ได้มีการแบ่งโดย เบื่อ สับสน หงุดหงิด คือ ไม่มีส่วนร่วม และ [16] เชื่อมโยงกับพฤติกรรมโดยที่พฤติกรรมมีความสนใจกับอารมณ์พึงพอใจและอารมณ์สับสน คือการมีส่วนร่วม ส่วนพฤติกรรมไม่มีความสนใจกับอารมณ์พึงพอใจ สับสน และเบื่อ คือการไม่มีส่วนร่วม



รูปที่ 1 การเชื่อมโยงระหว่างอารมณ์และการมีส่วนร่วม [14]

การจัดกลุ่มอารมณ์เพื่อให้สอดคล้องกับการวิเคราะห์การมีส่วนร่วมของผู้เรียน มีตั้งแต่ 2 – 5 ระดับ 2 ระดับ ได้แก่ การมีส่วนร่วม และ ไม่มีส่วนร่วม [3, 16, 17-22] 3 ระดับ ได้แก่ ไม่มีส่วนร่วม มีส่วนร่วมระดับปกติ และมีส่วนร่วมระดับมาก [3, 4, 21-26] 4 ระดับ ได้แก่ มีส่วนร่วมน้อยมาก มีส่วนร่วมระดับน้อย มีส่วนร่วมระดับปกติ มีส่วนร่วมระดับมาก [21, 27 – 35] และ 5 ระดับ ได้แก่ ไม่มีส่วนร่วม มีส่วนร่วมระดับน้อย มีส่วนร่วมระดับปกติ มีส่วนร่วมระดับมาก มีส่วนร่วมระดับมากที่สุด [14] ดังแสดงในตารางที่ 1

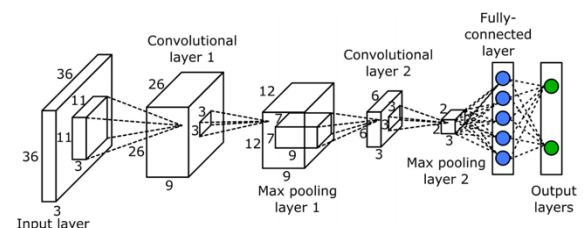
ตารางที่ 1 การแบ่งระดับการมีส่วนร่วม

บทความ	ระดับการมีส่วนร่วม	จำนวน
[3,16,17-22]	การมีส่วนร่วม และ ไม่มีส่วนร่วม	2
[3, 4, 21-26]	ไม่มีส่วนร่วม มีส่วนร่วมระดับปกติ และมีส่วนร่วมระดับมาก	3
[21, 27 – 35]	มีส่วนร่วมน้อยมาก มีส่วนร่วมระดับน้อย มีส่วนร่วมระดับปกติ มีส่วนร่วมระดับมาก	4
[14]	ไม่มีส่วนร่วม มีส่วนร่วมระดับน้อย มีส่วนร่วมระดับปกติ มีส่วนร่วมระดับมาก มีส่วนร่วมระดับมากที่สุด	5

จากการแบ่งการมีระดับการมีส่วนร่วมโดยเชื่อมโยงกับอารมณ์พื้นฐานและการจัดกลุ่มอารมณ์ตั้งแต่ 2 – 5 ระดับส่งผลต่อค่าความแม่นยำของแบบจำลองที่นำมาใช้ในการตรวจสอบ

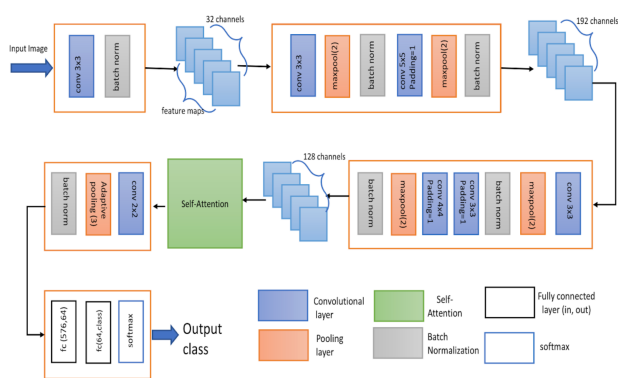
และการพยากรณ์โดยที่ 2 ระดับ เช่น [15] ได้ทำการแบ่งเป็น 2 ระดับได้ค่าความแม่นยำร้อยละ 90.87 [18] แบบจำลองที่พัฒนาขึ้นได้ค่าความแม่นยำร้อยละ 72.38 3 ระดับ เช่น [7] ได้ทำการแบ่งออกเป็น 3 ระดับได้ค่าความแม่นยำร้อยละ 92.33 [24] ได้ค่าความแม่นยำร้อยละ 80.00 และ [3] ได้ค่าความแม่นยำร้อยละ 87.25 4 ระดับ เช่น [39] ได้ทำการแบ่งออกเป็น 4 ระดับได้ค่าความแม่นยำร้อยละ 57.90 [21] ได้ค่าความแม่นยำร้อยละ 60.00 และ [31] ได้ค่าความแม่นยำร้อยละ 58.84 ในการวิจัยนี้ได้แบ่งการมีส่วนร่วมออกเป็น 3 ระดับได้แก่ ไม่มีส่วนร่วม (Dis-engagement) มีส่วนร่วมระดับปกติ (Normally-engagement) มีส่วนร่วมระดับสูง(High-engagement) ซึ่งเป็น 3 ระดับที่ให้ประสิทธิภาพและความแม่นยำมากที่สุดในการพยากรณ์จากวิจัยและบทความต่าง ๆ

Convolutional neural network (CNN) เป็นโครงข่ายประสาทเทียมที่มีความสามารถพิเศษในการสกัดคุณลักษณะเด่นของภาพ (Feature extraction) ที่มีความซับซ้อนเหมาะสมกับข้อมูลที่ไม่ใช่โครงสร้างสถาปัตยกรรมของ CNN ประกอบด้วย 3 ชั้น ดังรูปที่ 2 ได้แก่ 1) ชั้นคอนโวลูชัน (Convolutional Layer) เป็นชั้นที่มีความสำคัญกับ CNN ซึ่งเป็นชั้นที่เป็นการสกัดคุณลักษณะเด่นของข้อมูลประกอบไปด้วยคอร์เนลฟิลเตอร์ (Kernel Filter) และมีขนาดของคอร์เนล (Kernel Size) มีขนาดตั้งแต่ 3X3 , 5X5 , 7X7 เป็นต้น 2) ชั้นพูลลิง (Pooling Layer) จะสกัดส่วนที่สำคัญของข้อมูลและลดขนาดของข้อมูล ได้แก่ ค่ามากที่สุด (max pooling) น้อยที่สุด (min pooling) ผลรวม (sum pooling) และค่าเฉลี่ย (Average Pooling) 3) ชั้นการเชื่อมโยงเต็มรูปแบบ (Fully Connected Layer) ทำให้ข้อมูลจากชั้นก่อนหน้าที่เป็นสองมิติเป็นหนึ่งมิติก่อนที่จะมีการเชื่อมโยงกันทุก ๆ โหนด



รูปที่ 2 โครงสร้างของ CNN [36]

Md Asif Jalal และคณะ [37] ได้นำเสนอการสร้างแบบจำลองในการพยากรณ์อารมณ์โดยใช้ชุดข้อมูล AffectNet และใช้แบบจำลอง AlexNet เป็นพื้นฐานในการพัฒนาแบบจำลองเพิ่มเติมที่นำเสนอในงานวิจัยนี้ซึ่งโครงสร้างแสดงดังรูปที่ 3 ผลทดลองคือ AlexNet ได้ค่าความถูกต้องร้อยละ 58.00 และแบบจำลองที่พัฒนาเพิ่มเติมที่นำเสนอได้ค่าความแม่นยำร้อยละ 93.80



รูปที่ 3 โครงสร้างของแบบจำลองที่นำเสนอ [37]

Mahbub Murshed และคณะ [4] ได้ศึกษาและวิจัยเกี่ยวกับการตรวจสอบการมีส่วนร่วมของ ผู้เรียนในการเรียนผ่านระบบอีเลิร์นนิง (E-learning) โดยการใช้อัลกอริทึมที่มีพื้นฐานมาจาก Convolutional neural network (CNN) 4 แบบจำลอง ได้แก่ 1) All Convolutional Network (All-CNN) 2) Network in Network (NiN-CNN) 3) Very Deep Convolutional (VD-CNN) และ 4) เป็นส่วนที่พัฒนาเพิ่มเติมจากการทำวิจัยนี้ โครงสร้างและพารามิเตอร์ของแต่ละแบบจำลองที่ใช้ในการวิเคราะห์ดังรูปที่ 4 ซึ่งในการวิเคราะห์ที่มีการแบ่งการมีส่วนร่วมออกเป็น 3 ระดับ คือ 1) ไม่มีส่วนร่วม (not-engaged) 2) มีส่วนร่วมปกติ (normally-engaged) 3) มีส่วนร่วมมาก (highly-engaged) และในการวิจัยนี้ใช้วิเคราะห์จากวิดีโอของผู้เรียนที่ถูกบันทึกผ่านทางกล้องเว็บแคมการวิเคราะห์ในงานวิจัยนี้ใช้ชุดข้อมูลที่นำมาสร้างโมเดลในการวิเคราะห์คือ DAISSE ผลที่ได้จากการวิเคราะห์คือ All-CNN ได้ค่าความแม่นยำร้อยละ 75.97 NiN-CNN ได้ค่าความแม่นยำร้อยละ 83.22 VD-CNN ได้ค่าความแม่นยำร้อยละ 86.45 และ 4) เป็นส่วนที่พัฒนาเพิ่มเติมจากการทำวิจัยนี้ได้ค่าความแม่นยำร้อยละ 92.33

AB-CNN [14]	VD-CNN [16]	NIN-CNN [15]	Proposed model
Input 32×32 Grayscale Image			
3×3 conv. 96 ReLU	3×3 conv. 96 ReLU	5×5 conv. 192 ReLU	3×3 conv. 192 BatchNorm ReLU
3×3 conv. 96 ReLU	3×3 conv. 96 ReLU	1×1 conv. 192 ReLU	1×1 conv. 192 ReLU
3×3 conv. 96 ReLU	3×3 conv. 96 ReLU	1×1 conv. 192 ReLU	
3×3 conv. 96 ReLU with stride 2	3×3 max-pooling with stride 2	3×3 max-pooling dropout (0.5)	3×3 conv. 192 BatchNorm ReLU with stride 2 dropout (0.5)
3×3 conv. 192 ReLU	3×3 conv. 192 ReLU	3×3 conv. 96 ReLU	3×3 conv. 96 BatchNorm ReLU
3×3 conv. 192 ReLU	3×3 conv. 192 ReLU	1×1 conv. 96 ReLU	1×1 conv. 96 ReLU
3×3 conv. 192 ReLU	3×3 conv. 192 ReLU	1×1 conv. 96 ReLU	1×1 conv. 96 ReLU
3×3 conv. 192 ReLU with stride 2	3×3 max-pooling with stride 2	3×3 max-pooling dropout (0.5)	3×3 max-pooling with stride 2 dropout (0.5)
3×3 conv. 192 ReLU	3×3 conv. 192 ReLU	3×3 conv. 32 ReLU	3×3 conv. 32 BatchNorm ReLU
1×1 conv. 192 ReLU	1×1 conv. 192 ReLU	1×1 conv. 32 ReLU	3×3 conv. 32 BatchNorm ReLU
1×1 conv. 3	1×1 conv. 3	1×1 conv. 3	1×1 conv. 3
global average pooling	global average pooling	global average pooling	global average pooling
softmax	softmax	softmax	softmax

รูปที่ 4 โครงสร้างและพารามิเตอร์ของแต่ละแบบจำลอง [4]

Saswat Dash และคณะ [16] ได้ศึกษาและวิจัยเกี่ยวกับ การวิเคราะห์การมีส่วนร่วมของ ผู้เรียนในการเรียนในชั้นเรียน ออนไลน์ ในการวิเคราะห์การมีส่วนร่วมของผู้เรียนวิจัยนี้ได้ใช้ อัลกอริทึม ที่มีพื้นฐานมาจาก CNN 5 แบบจำลอง ได้แก่ 1) AllConv 2) VDCov 3) NiNConv 4) Con-Pool Convolutional (CPCov) 5) Cluster Munitions Convolutional (CMConv) โดยมีโครงสร้างและ พารามิเตอร์ ของแบบจำลองที่ใช้ในการวิเคราะห์ดังรูปที่ 5 และใช้ชุดข้อมูล ที่ นำมาสร้างโมเดลคือ DAISEE ในการวิเคราะห์แบ่งกลุ่มการมีส่วน ร่วมออกเป็น 2 ระดับ คือ 1) ไม่มีส่วนร่วม (not-engaged) 2) มีส่วนร่วม (engaged) ผลที่ได้จากการวิเคราะห์ แบบจำลอง AllConv ได้ความถูกต้องที่ร้อยละ 81.80 แบบจำลอง VDCov ได้ความถูกต้องที่ร้อยละ 79.20 แบบจำลอง NiNConv ได้ความ ถูกต้องที่ร้อยละ 81.20 แบบจำลอง CPCov ได้ความถูกต้อง ที่ร้อยละ 83.60 แบบจำลอง CMConv ได้ความถูกต้องที่ร้อยละ 91.10



รูปที่ 5 โครงสร้างและพารามิเตอร์ของแบบจำลอง [16]

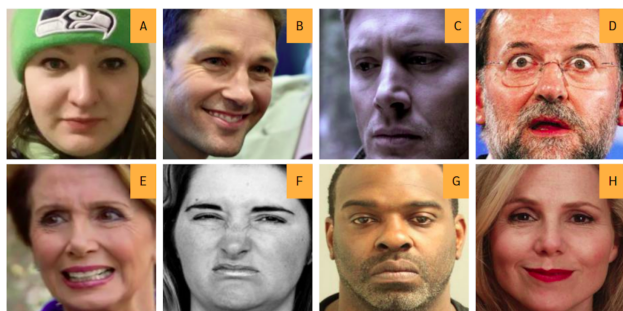
3. วิธีการดำเนินการวิจัย

3.1 เครื่องมือ

งานวิจัยนี้ใช้ไลบรารี Keras [5] และ Tensorflow ในการพัฒนาแบบจำลองด้วยภาษา Python ซึ่งคอมพิวเตอร์สเปค CPU Intel Core i5-10400F หน่วยความจำหลัก 32 GB หน่วยประมวลผลกราฟฟิก Nvidia RTX 3060 (12GB) ระบบปฏิบัติการ Ubuntu เวอร์ชัน 20.04 LTS

3.2 ชุดข้อมูลและการเตรียมข้อมูล

ในการพัฒนาแบบจำลองวิจัยนี้ได้ใช้ชุดข้อมูล AffectNet [38] ในการเรียนรู้ของแบบจำลอง รูปภาพมีขนาด $224 * 224 * 3$ ซึ่งมีภาพทั้งหมด 959,907 ภาพ 11 ป้ายกำกับ 8 อารมณ์ ได้แก่ เป็นกลาง ความสุข โศกเศร้า ตื่นเต้น ความกลัว รังเกียจ ความโกรธ ดูถูก และไม่มีใบหน้า ไม่มีอารมณ์ ไม่แน่ใจ [38] แสดงในรูปที่ 6 และได้จัดกลุ่มเพื่อให้สอดคล้องกับการวิจัยที่ให้ประสิทธิภาพและความแม่นยำมากที่สุดในการพยากรณ์จากวิจัยและบทความต่าง ๆ โดยแบ่งออกเป็น 3 ระดับได้แก่ ไม่มีส่วนร่วม (Dis-engagement) มีส่วนร่วมระดับปกติ (Normally-engagement) มีส่วนร่วมระดับสูง (High-engagement) [7]



รูปที่ 6 รูปภาพใบหน้าจากชุดข้อมูล AffectNet A) เป็นกลาง B) ความสุข C) ความโศกเศร้า D) ตื่นเต้น E) ความกลัว F) รังเกียจ G) ความโกรธ H) ดูถูก [38]

การแบ่งข้อมูลใหม่จาก 8 อารมณ์เป็น 3 ระดับการมีส่วนร่วม โดยใช้อารมณ์พื้นฐานเชื่อมโยงกับระดับการมีส่วนร่วม [14] และการเชื่อมโยงอารมณ์พื้นฐานในรูปที่ 1 ได้แบ่งข้อมูลใหม่เพื่อใช้ในการวิจัยดังตารางที่ 2 ในแต่ละระดับอารมณ์ถูกเลือกนำมาใช้ทั้งหมด 96,111 ภาพเป็นจำนวนที่น้อยที่สุดของทั้ง 3

ระดับ เพื่อให้ข้อมูลทั้ง 3 ระดับมีปริมาณที่เท่ากัน รวมทั้ง 3 ระดับจะได้จำนวนภาพทั้งหมด 288,333 ภาพ ป้องกันการลำเอียงของข้อมูลที่มีจำนวนไม่เท่ากัน (Bias) และมีการใช้เทคนิค K-fold cross validation มาแบ่งข้อมูลสำหรับการเรียนรู้และการทดสอบแบบจำลอง โดยที่แบ่งข้อมูลออกเป็น 5 กลุ่ม (K=5) จำนวน 3 ชุด สัดส่วน 75:15:10 ตามลำดับ [39] ได้แก่ ชุดสำหรับการฝึกสอน (Training set) ร้อยละ 75 จำนวน 216,149 ภาพ ชุดข้อมูลสำหรับการตรวจสอบ (Validation set) ร้อยละ 15 จำนวน 43,251 ภาพ และชุดข้อมูลสำหรับการทดสอบ (Test set) ร้อยละ 10 จำนวน 28,833 ภาพ ดังแสดงในตารางที่ 3 การแบ่งข้อมูลสำหรับการเรียนรู้ของแบบจำลอง

ตารางที่ 2 การแบ่งกลุ่มข้อมูลใหม่สำหรับใช้ในการวิจัย

ป้ายกำกับใหม่	ป้ายกำกับจากชุดข้อมูล	จำนวน	รวม
ไม่มีส่วนร่วม	ดูถูก เป็นกลาง	4,252 218,516	222,768
มีส่วนร่วม ระดับปกติ	ความสุข ความเศร้า ความเบื่อหน่าย	381,150 46,813 5,193	433,156
มีส่วนร่วม ระดับสูง	ตื่นเต้น ความโกรธ ความกลัว	32,052 53,382 10,677	96,111

ตารางที่ 3 การแบ่งข้อมูลสำหรับการเรียนรู้ของแบบจำลอง

ป้ายกำกับ	Training Set (75%)	Validation Set (15%)	Test Set (10%)	All (100%)
ไม่มีส่วนร่วม	72,083	14,417	9,611	96,111
มีส่วนร่วม ระดับปกติ	72,083	14,417	9,611	96,111
มีส่วนร่วม ระดับสูง	72,083	14,417	9,611	96,111
รวม	216,149	43,251	28,833	288,333

3.3 การวัดประสิทธิภาพ

การวัดประสิทธิภาพของการเรียนรู้แบบจำลองได้ใช้การวัดค่าความแม่นยำ (Accuracy) ความถูกต้อง ระหว่างสิ่งที่ทำนายกับสิ่งที่เกิดขึ้นจริง โดยมีสมการในการคำนวณดังนี้

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

โดย

TP คือ ค่าที่พยากรณ์ถูกต้องเชิงบวก

TN คือ ค่าที่พยากรณ์ถูกต้องเชิงลบ

FP คือ ค่าที่พยากรณ์ผิดพลาดเชิงบวก

FN คือ ค่าที่พยากรณ์ผิดพลาดเชิงลบ

โดยในบทความนี้ Accuracy หมายถึงค่าความแม่นยำในการพยากรณ์ระดับการมีส่วนร่วมจากใบหน้าทั้ง 3 ระดับ ได้แก่ ไม่มีส่วนร่วม มีส่วนร่วมระดับปกติ และการมีส่วนร่วมระดับสูง

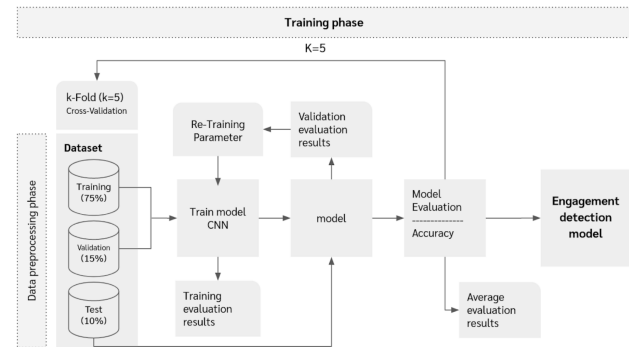
การวัดความสูญเสียหรือค่าการสูญเสีย (Loss Function) คือ ฟังก์ชันที่ใช้วัดความผิดพลาดระหว่างค่าที่คาดการณ์ไว้กับค่าที่แท้จริงในแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) ค่าการสูญเสียจะทำหน้าที่เป็นตัววัดประสิทธิภาพของแบบจำลอง โดยค่าการสูญเสียต่ำจะหมายความว่าแบบจำลองทำงานได้แม่นยำมากขึ้น

3.4 กระบวนการฝึกสอนแบบจำลอง

เมื่อทำการแบ่งข้อมูลทั้ง 3 กลุ่มแล้ว ใช้ฟังก์ชัน ImageDataGenerator ในการทำ One-hot Encoding และแปลงข้อมูลให้เป็นแบบ Numpy แล้วใช้ flow_from_directory ในการเรียกข้อมูลทั้ง 3 กลุ่มที่เตรียมไว้ มีการกำหนดขนาดของภาพ (input_size) เป็น 224 x 224 กำหนดกลุ่มของข้อมูล (class_mode) แบบ categorical และกำหนดปริมาณของข้อมูลในแต่ละรอบ (batch_size) ตามจำนวนที่กำหนดในแต่ละรอบของการทดลอง

จากนั้นเข้าสู่กระบวนการฝึกสอนแบบจำลองกำหนดรอบการเรียนรู้ 15 รอบ (epoch) ใช้ชุดสำหรับการฝึกสอน (Training set) และ ชุดข้อมูลสำหรับการตรวจสอบ (Validation set) ในการฝึกสอน เมื่อครบทั้ง 15 รอบ ใช้ชุดข้อมูลสำหรับการทดสอบ

(Test set) ใช้ในการทดสอบแบบ จำลองที่ได้จากการเรียนรู้และใช้เทคนิคการแบ่งข้อมูลแบบ K-fold cross validation คือ K=5 ในการแบ่งข้อมูลทั้ง 5 ชุดเพื่อทดสอบความถูกต้องของแบบจำลอง ดังแสดงในรูปที่ 7



รูปที่ 7 กระบวนการฝึกสอนแบบจำลอง

4. ผลการวิจัย

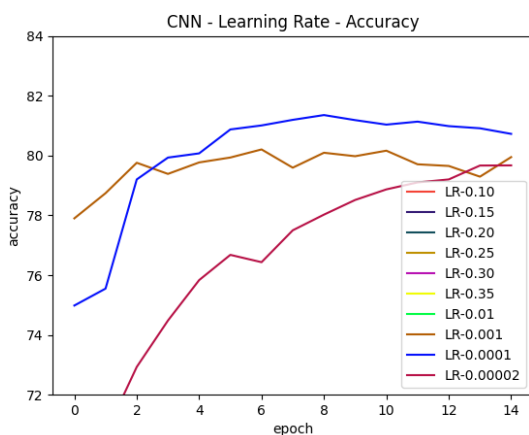
ในการพัฒนาแบบจำลองในการตรวจสอบการมีส่วนร่วมในชั้นเรียนออนไลน์ได้นำโครงสร้างแบบจำลองจากเว็บไซต์ Keras - Simple MNIST convnet [39] เป็นโครงสร้างแบบ จำลองตั้งต้นในการพัฒนาแบบจำลอง แบ่งการทดลองออกเป็น 3 กลุ่ม ซึ่งในการทดลองได้ทำตามลำดับกลุ่มการทดลองที่ 1-3 ตามลำดับและเป็นการนำแบบจำลองจากกลุ่มการทดลองก่อนหน้านี้ที่ได้ผลดีที่สุดมาทดลองต่อในกลุ่มการทดลองถัดไป

กลุ่มที่ 1 คือ การปรับแต่งโครงสร้างชั้นการสกัดคุณลักษณะเด่นของภาพ (Feature extraction) ได้กำหนดรูปแบบในการทดลองออกเป็น 6 รูปแบบคือ 1) การเพิ่มขึ้นของ จำนวน NN สม้าเสมอ 2) การเพิ่มขึ้นของจำนวน NN สม้าเสมอและลดลงอย่างสม้าเสมอเท่ากับการ เพิ่มขึ้น 3) การลดจำนวนของ NN สม้าเสมอ 4) จำนวนของ NN เท่าเดิม 5) จำนวนของ NN สลับกัน (สองค่า) 6) จำนวนของ NN เพิ่มและลดไม่เท่ากัน ดังรูปที่ 12 ผลการทดลองที่ดีที่สุดคือ Feature extraction Filter : 128-256-512 ค่าความแม่นยำร้อยละ 79.62 และค่าความสูญเสียร้อยละ 0.52 ซึ่งได้ผลค่าความแม่นยำใกล้เคียงกับ Filter : 128-256-512-1024 แต่ค่าความสูญเสียสูงกว่า

กลุ่มที่ 2 คือ การทดลองแต่ละพารามิเตอร์ ได้แก่ Learning Rate (LR), Batch Size (BZ), Optimizer (OTM), Fully connected (FC)

● พารามิเตอร์ Learning Rate (LR)

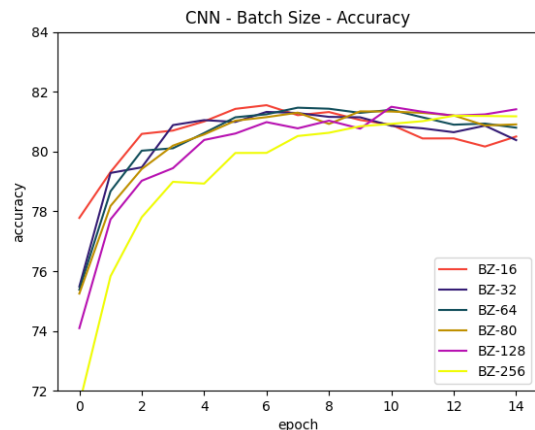
ผลการทดลองค่า Learning Rate ได้แก่ 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.01 ได้ ค่าความแม่นยำ 33.33 และค่าความสูญเสีย 1.10 เนื่องจากมีค่าที่สูงเกินไปจึงทำให้ไม่สามารถคำนวณต่อไป ส่วน Learning Rate 0.0001 ให้ผลการทดลองที่ดีที่สุด ผลการทดลองคือค่าความแม่นยำ 80.75 และค่าความสูญเสีย 0.51 และเป็นค่าตั้งต้นของโครงสร้างแบบจำลองที่นำมาใช้ ดังรูปที่ 8



รูปที่ 8 กราฟเปรียบเทียบผลการทดลองค่าความแม่นยำของ Learning Rate

● พารามิเตอร์ Batch Size (BZ)

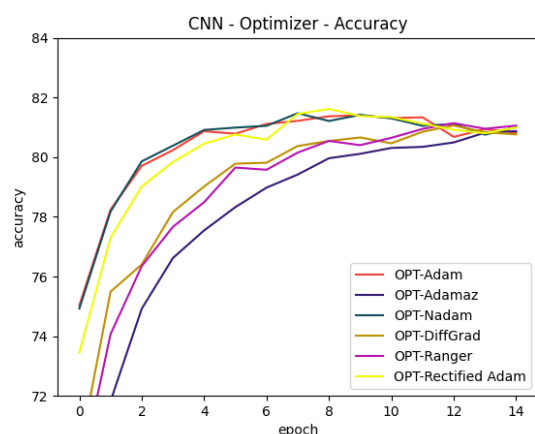
กำหนดค่า Batch Size ได้แก่ 16, 32, 64, 80, 128, 256 ซึ่งค่าที่ให้ผลการทดลองที่ดีที่สุดคือ Batch Size 128 ผลการทดลองคือค่าความแม่นยำ 81.27 และค่าความสูญเสีย 0.50 จากกราฟจะเห็นได้ว่าค่าความแม่นยำของโมเดลจะเพิ่มขึ้นอย่างรวดเร็วในช่วงแรก เมื่อขนาด Batch Size เพิ่มขึ้นจาก 16 เป็น 32 จากนั้นความแม่นยำจะเพิ่มขึ้นอย่างช้าลง เมื่อขนาด Batch Size เพิ่มขึ้นจาก 32 เป็น 64 และ 80 เมื่อเพิ่มขึ้นเป็น 128 ค่าความแม่นยำของโมเดลจะลดลงเล็กน้อย จากนั้นความแม่นยำจะลดลงอย่างรวดเร็วเมื่อขนาดชุดข้อมูลเพิ่มขึ้นเป็น 256 ดังรูปที่ 9



รูปที่ 9 กราฟเปรียบเทียบผลการทดลองค่าความแม่นยำของ Batch Size

● พารามิเตอร์ Optimizer (OTM)

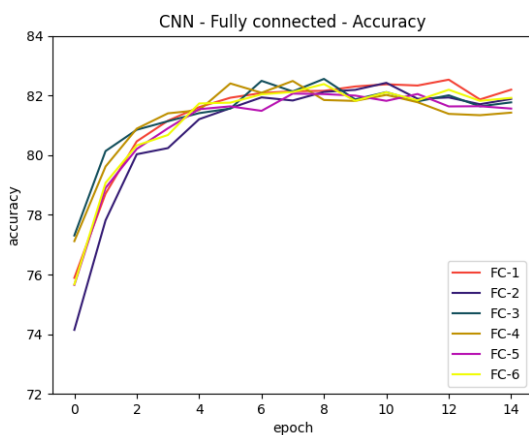
ผลการทดลองพารามิเตอร์ Optimizer (OTM) ได้แก่ Adam และ Adamax , Nadam, DiffGrad , Ranger, Rectified Adam ซึ่งทั้งหมดถูกพัฒนาต่อมาจาก Adam จากกราฟจะเห็นว่า Adam มีประสิทธิภาพสูงสุด โดยมีความแม่นยำสูงกว่าตัวอื่น ๆ ในช่วงเริ่มต้นจากนั้น Adamax จะสูงกว่าตัว Adam ในช่วงกลางและ Nadam จะสูงกว่า Adamax ในช่วงท้ายของและผลการทดลองที่ดีที่สุดคือ Ranger ผลการทดลองคือค่าความแม่นยำ 81.04 และค่าความสูญเสีย 0.49 แต่ค่าความแม่นยำของแต่ละตัวมีจำนวนที่ใกล้เคียงกัน ดังรูปที่ 10



รูปที่ 10 กราฟเปรียบเทียบผลการทดลองค่าความแม่นยำของ Optimizer

● พารามิเตอร์ Fully connected (FC)

ผลการทดลองซึ่งได้กำหนดรูปแบบในการทดลองออกเป็น 6 รูปแบบดังรูปที่ 12 ซึ่งเป็นจำนวนจำนวนของ นิวรัลเน็ตเวิร์ก (neural networks) ผลการทดลองที่ดีที่สุดคือ FC : 50-100 และ FC : 50-50 ให้ผลการทดลองที่เท่ากันคือ ผลการทดลองคือค่าความแม่นยำ 82.20 และค่าความสูญเสีย 0.49 และเมื่อนำผลการทดลองในแต่ละรอบมาเปรียบเทียบพบว่า FC : 50-50 มีการโอเวอร์ฟิตมากกว่าการทดลอง FC : 50-100 ดังรูปที่ 11 โดยที่ FC-1 คือ FC: 32, FC-2 คือ FC: 64-32, FC-3 คือ FC: 64-64, FC-4 คือ FC: 50-100, FC-5 คือ FC: 50-50, FC-6 คือ FC: 50-100-50

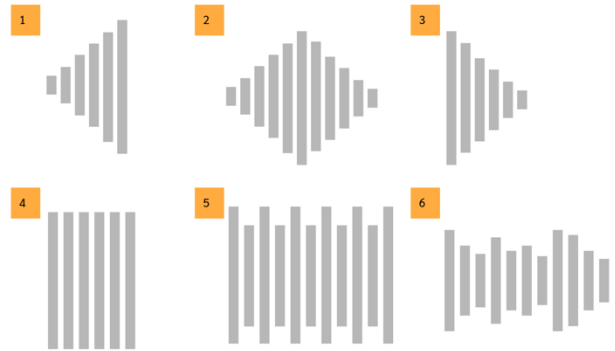


รูปที่ 11 กราฟเปรียบเทียบผลการทดลองค่าความแม่นยำของชั้น Fully connected

กลุ่มที่ 3 คือ การทดลองโดยการจับคู่ 2 พารามิเตอร์

- Optimizer + Batch Size (No. 1)
- Optimizer + Learning Rate (No. 2)
- Batch Size + Learning Rate (No. 3)
- Optimizer + Fully connected (No. 4)
- Batch Size + Fully connected (No. 5)
- Learning Rate + Fully connected (No. 6)

ผลการทดลองที่ดีที่สุดคือ Optimizer + Fully connected ผลการทดลองคือค่าความแม่นยำ 82.30 และค่าความสูญเสีย 0.46 ดังสรุปผลในตารางที่ 5



รูปที่ 12 รูปแบบการทดลองเพิ่ม/ลดจำนวน Node ในการทดลอง

4.1 การเปรียบเทียบผลการทดลองกลุ่มที่ 1 - 3 ที่ได้ผลดีที่สุดกับการทดลองตั้งต้น (Base Model)

ดังแสดงในตารางที่ 6 พร้อมทั้งแสดงค่าความแม่นยำ และค่าความสูญเสียที่เพิ่มขึ้น/ลดลง การทดลองที่ 1 เป็นของแบบจำลองตั้งต้น (Base Model) การทดลองที่ 2 เป็นผลการทดลองที่มีผลดีที่สุดในกลุ่มการทดลองที่ 1 การทดลองที่ 3-6 เป็นการทดลองที่มีผลดีที่สุดในกลุ่มการทดลองที่ 2 การทดลองที่ 7-12 เป็นผลการทดลองในกลุ่มการทดลองที่ 3 พบว่าการทดลองที่ 10 Optimizer + Fully connected สามารถทำให้ค่าความแม่นยำเพิ่มขึ้นได้มากถึง 16.51 และสามารถลดค่าความสูญเสียได้ถึง 0.31 ซึ่งใกล้เคียงกับการทดลองที่ 5 อยู่ในกลุ่มการทดลองที่ 2 FC : 50-100 สามารถทำให้ค่าความแม่นยำเพิ่มขึ้นได้มากถึง 16.41 และสามารถลดค่าความสูญเสียได้ถึง 0.28 และเมื่อนำผลมาเปรียบเทียบจะเห็นความสัมพันธ์ของผลการทดลองเมื่อเปรียบเทียบกับการทดลองที่ 1 คือการทดลองที่ 2 ที่ไม่มีชั้น Fully connected สามารถเพิ่มค่าความแม่นยำได้ถึง 14.31 และลดค่าความสูญเสียได้ถึง 0.22 การทดลองที่ 6 (การทดลองที่ 1 เพิ่มขึ้น Fully connected) สามารถเพิ่มค่าความแม่นยำได้ถึงร้อยละ 16.41 และลดค่าความสูญเสียได้ถึงร้อยละ 0.28 และการทดลองที่ 10 (การทดลองที่ 6 และเปลี่ยน Optimizer เป็น Ranger) สามารถเพิ่มค่าความแม่นยำได้ถึงร้อยละ 16.51 และลดค่าความสูญเสียได้ถึงร้อยละ 0.31 กล่าวคือ เมื่อมีการเพิ่มขึ้น Fully connected สามารถทำให้ค่าความแม่นยำเพิ่มขึ้นและ

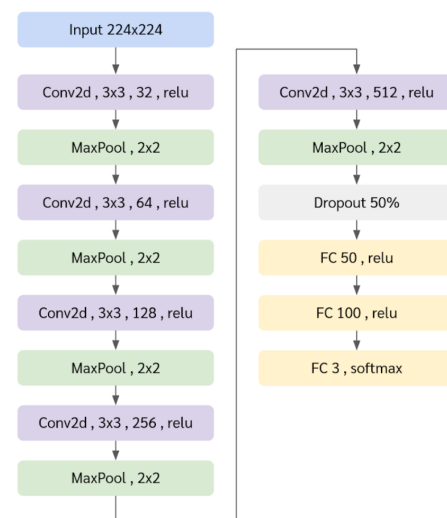
เมื่อมีการเปลี่ยน Optimizer จากเดิมคือ Adam เป็น Ranger สามารถทำให้ค่าความสูญเสียลดลงได้

4.2 สรุปผลการวิจัย

ในการทดลองได้ดำเนินการแบ่งการทดลองออกเป็น 3 กลุ่มตามลำดับ ได้แก่

- กลุ่มที่ 1 คือ การปรับแต่งโครงสร้างชั้นการสกัดได้ผลดีที่สุด คือ Feature extraction Filter 128-256-512 ค่าความแม่นยำร้อยละ 79.62 และค่าความสูญเสียร้อยละ 0.52 และนำโครงสร้างไปใช้ในกลุ่มถัดไป
- กลุ่มที่ 2 คือ การทดลองแต่ละพารามิเตอร์โดยใช้โครงสร้างในกลุ่มที่ 1 ประกอบด้วย พารามิเตอร์ Batch Size พบว่า Batch Size 128 ได้ผลดีที่สุดคือค่าความแม่นยำร้อยละ 81.27 และค่าความสูญเสียร้อยละ 0.50 พารามิเตอร์ Optimizer พบผลการทดลองที่ดีที่สุดคือ Ranger คือค่าความแม่นยำร้อยละ 81.04 และค่าความสูญเสียร้อยละ 0.49 พารามิเตอร์ Fully connected พบว่าผลการทดลองที่ดีที่สุดคือ FC : 50-50 คือค่าความแม่นยำร้อยละ 82.20 และค่าความสูญเสียร้อยละ 0.49 โดยโอเวอร์ฟิตน้อยกว่า
- กลุ่มที่ 3 คือ การทดลองโดยการจับคู่ 2 พารามิเตอร์ โดยใช้โครงสร้างแบบจำลองและผลการปรับแต่งพารามิเตอร์จากการทดลองทั้ง 2 กลุ่มก่อนหน้ามาทดลองโดยการจับคู่ 2 พารามิเตอร์ พบว่าผลการทดลองที่ดีที่สุดคือ Optimizer + Fully connected ผลการทดลองคือค่าความแม่นยำร้อยละ 82.30 และค่าความสูญเสียร้อยละ 0.46 ดังผลการเปรียบเทียบในหัวข้อ 4.1 การเปรียบเทียบผลการทดลอง กลุ่มที่ 1 - 3 จะได้โครงสร้างแบบจำลองและพารามิเตอร์ที่ได้ผลดีที่สุด พารามิเตอร์ที่ได้ผลดีที่สุดคือ Optimizer = Range , Batch Size = 80, Learning Rate = 0.0001 และโครงสร้างแบบจำลองที่ได้ผลดีที่สุดดังแสดงในรูปที่ 13
 - Feature Extraction Layers: มีจำนวน 5 ชั้น มีฟิลเตอร์ (Filters) แต่ละชั้นมีดังนี้ 32, 64, 128, 256, 512 ตามลำดับและขนาดคอร์เนล 3x3 ใช้ activation function คือ ReLU

- Maxpool Layer: ในแต่ละชั้นของ Convolutional Layer จะมี Maxpool Layer ขนาด 2x2 และในชั้นสุดท้ายของ Maxpool Layer จะมี Dropout Layer 50% เพื่อลดการโอเวอร์ฟิต
- Fully Connected Layers: จำนวน 3 ชั้นมีจำนวนโหนด 50, 100 ตามลำดับซึ่งใช้ ReLU เป็น activation function
- Output Layer: จำนวน 1 ชั้น จำนวน 3 โหนด และใช้ activation function Softmax



รูปที่ 13 โครงสร้างแบบจำลองที่ได้ผลดีที่สุด

5. ข้อเสนอแนะ

ในการทดลองครั้งถัดไปควรจะมีการทำ Data Augmentation เพื่อเป็นการเพิ่มปริมาณข้อมูลที่ใช้ในการเรียนรู้ของแบบจำลอง โดยการรวมชุดข้อมูลอื่นเช่น Fer2013, EmotioNet, DAiSEE, EmotiW, CK+48 ซึ่งอาจจะส่งผลให้ประสิทธิภาพของแบบจำลองมากยิ่งขึ้น รวมทั้งการทำ Feature Selection โดยนำแบบจำลองอื่น ๆ ที่เป็นที่นิยมมาใช้ เช่น InceptionResNetV2, EfficientNet, NASNetLarge เพื่อหา Feature ที่เหมาะสำหรับการใช้ในการสร้างแบบจำลองใช้เทคนิคอื่น ๆ เข้ามาใช้ในการพัฒนาแบบจำลอง เช่น การเพิ่มปริมาณของข้อมูล (Augmentation) การทำ Ensemble ได้แก่ Bagging และ Boosting

ตารางที่ 4 ตารางผลการปรับแต่งโครงสร้างชั้นการสกัดคุณลักษณะเด่นของภาพ (กลุ่มการทดลองที่ 1)

Feature extraction	Fully connected	Optimizer	Batch size	Learning rate	Accuracy Validation	Loss Validation
64-64-64-64	FC: 0 (output 3)	adam	80	0.0001	70.70	0.66
32-16-32-16	FC: 0 (output 3)	adam	80	0.0001	70.21	0.67
1024-512-256	FC: 0 (output 3)	adam	80	0.0001	73.05	0.63
128-256-512-1024	FC: 0 (output 3)	adam	80	0.0001	78.98	0.87
128-256-512	FC: 0 (output 3)	adam	80	0.0001	79.62	0.52

ตารางที่ 5 ตารางผลการทดลองโดยการจับคู่ 2 พารามิเตอร์ (กลุ่มการทดลองที่ 3)

No.	Fully connected	Optimizer	Batch size	Learning rate	Accuracy Validation	Loss Validation
1	FC: 0 (output 3)	Ranger	128	0.0001	80.66	0.49
2	FC: 0 (output 3)	Ranger	80	0.0001	80.89	0.48
3	FC: 0 (output 3)	adam	128	0.0001	80.91	0.50
4	FC: 50-100	Ranger	80	0.0001	82.30	0.46
5	FC: 50-100	adam	128	0.0001	82.16	0.48
6	FC: 50-100	adam	80	0.0001	81.72	0.50

ตารางที่ 6 ตารางเปรียบเทียบผลการทดลองกลุ่มที่ 1 - 3 ที่ได้ผลดีที่สุดกับการทดลองตั้งต้น (Base Model)

No.	Fully connected	Optimizer	BZ	LR	Acc	+-	Loss	+-
1	FC : 0 (output 3)	adam	80	0.0001	65.79	-	0.77	-
2	FC : 0 (output 3)	adam	80	0.0001	80.10	+14.31	0.55	-0.22
3	FC : 0 (output 3)	adma	80	0.0001	80.75	+14.96	0.51	-0.26
4	FC: 0 (output 3)	adma	128	0.0001	81.27	+15.48	0.50	-0.27
5	FC: 0 (output 3)	Ranger	80	0.0001	81.04	+15.25	0.49	-0.28
6	FC : 50-100	adam	80	0.0001	82.20	+16.41	0.49	-0.28
7	FC : 0 (output 3)	Ranger	128	0.0001	80.66	+14.87	0.49	-0.28
8	FC : 0 (output 3)	Ranger	80	0.0001	80.89	+15.10	0.48	-0.29
9	FC : 0 (output 3)	adam	128	0.0001	80.91	+15.12	0.50	-0.27
10	FC : 50-100	Ranger	80	0.0001	82.30	+16.51	0.46	-0.31
11	FC : 50-100	adam	128	0.0001	82.16	+16.37	0.48	-0.29
12	FC : 50-100	adam	80	0.0001	81.72	+15.93	0.50	-0.27

6. อ้างอิง

- [1] World Health Organization: WHO. (12 October 2020). Q&A on coronaviruses (COVID-19). [Online] Available:
<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/>
- [2] COVID-19 CORONAVIRUS PANDEMIC (2020). (29 September 2020). [Online] Available:
<https://www.worldometers.info/coronavirus>
- [3] Dewan. M. Lin. F, Wen. D. Murshed. M and Uddin. Z, "A Deep Learning Approach to Detecting Engagement of Online Learners," pp. 1895 - 1902, 2018.
- [4] Murshed. M, Dewan. M, Lin. F, and Wen. D, "Engagement Detection in e-Learning Environments using Convolutional Neural Networks," IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech)., pp. 80–86., 2019.
- [5] Keras. (28 September 2020). [Online] Available:
<https://keras.io/>
- [6] Qiu. G, Yu. X, Sun. B, Wang. Y and Zhang. L, "Metastatic Cancer Image Classification Based on Deep Learning Method," pp. 658-661, 2020.
- [7] Zhong. Z, Zheng. M, Mai. H, Zhao. J and Liu. X, "Cancer image classification based on DenseNet model," pp. 1-6, 2020.
- [8] Igiri. Chinwe, Uzoma. Anyama and Ita. Silas, "Effect of Learning Rate on Artificial Neural Network in Machine Learning," International Journal of Engineering Research., Vol. 4, pp. 359-363, 2021.
- [9] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala and C. O. Aigbavboa, "A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks," International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)., pp. 92-99, 2018.
- [10] Z. Wan, Z. Yuxiang, X. Gong, Zhanghuali and B. Yu, "DenseNet model with RAdam optimization algorithm for cancer image classification," IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)., pp. 771-775, 2021.
- [11] Melinte. Octavian and Vladareanu. Luige, "Facial Expressions Recognition for Human-Robot Interaction Using Deep Convolutional Neural Networks with Rectified Adam Optimizer," Sensors., Vol. 20, pp. 1-21, 2020.
- [12] H. S. Suresha and S. S. Parthasarathy, "Alzheimer Disease Detection Based on Deep Neural Network with Rectified Adam Optimization Technique using MRI Analysis," Third International Conference on Advances in Electronics, Computers and Communications (ICAEECC)., pp. 1-6, 2020.
- [13] Kandel. Ibrahim and Castelli. Mauro, "The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset," ICT Express., Vol. 6, pp. 312-315, 2020.
- [14] Altuwairqi, K., Jarraya, S. K., Allinjawi, A and Hammami, M., "A new emotion-based affective model to detect student's engagement," Vol. 33, pp. 1-11, 2021.

- [15] Dewan. M, Lin. F, Wen. D, Murshed. M and Uddin. Z, "A Deep Learning Approach to Detecting Engagement of Online Learners," pp. 1895-1902, 2020.
- [16] Dash. S, Akber Dewan. M. A, Murshed. M, Lin. F, Abdullah-Al-Wadud. M and Das. A, "A Two-Stage Algorithm for Engagement Detection in Online Learning," 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1-4, 2019.
- [17] Mane. S. S and Surve. A. R, "Engagement Detection using Video-based Estimation of Head Movement," pp. 1745-1749, 2018.
- [18] Nezami. O. M, Dras. M, Hamey. L, Richards. D, Wan. S and Paris. C, "Automatic Recognition of Student Engagement using Deep Learning and Facial Expression," pp. 1-18, 2018.
- [19] Zhang. Z, Li. Z, Liu. H, Cao. T and Liu. S, "Data-driven Online Learning Engagement Detection via Facial Expression and Mouse Behavior Recognition Technology," Vol. 58, pp. 63-86, 2019.
- [20] Chang. C, Zhang. C, Chen. L and Liu. Y, "An Ensemble Model Using Face and Body Tracking for Engagement Detection," pp. 616-622, 2018.
- [21] Huang. T, Mei. Y, Zhang. H, Liu. S and Yang. H, "Fine-grained Engagement Recognition in Online Learning Environment," pp. 338-341, 2019.
- [22] Sümer. Ö, Goldberg. P, Mello. S, Gerjets. P, Trautwein. U and Kasneci. E, "Multimodal Engagement Analysis from Facial Videos in the Classroom," pp. 1-15, 2021.
- [23] Sharma. P, Joshi S, Gautam. S, Maharjan. S, Filipe. V and Reis. M. J. S, "Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning," pp. 1-18, 2019.
- [24] Booth. B. M, Ali. A. M, Narayanan. S. S, Bennett. I, and Farag. A. A, "Toward active and unobtrusive engagement assessment of distance learners," pp. 470-476, 2017.
- [25] Liu. Y, Chen. J, Zhang. M and Rao. C, "Student engagement study based on multi-cue detection and recognition in an intelligent learning environment. Multimedia Tools and Applications," Vol. 77, pp. 28749-28775, 2018.
- [26] T. S. A and Guddeti. R. M. R, "Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks," Vol. 25, pp. 1387-1415, 2021.
- [27] Kaur. A, Mustafa. A, Mehta. L and Dhall. A, "Prediction and Localization of Student Engagement in the Wild," pp. 1-9, 2018.
- [28] Li. Y. Y and Hung. Y. P, "Feature Fusion of Face and Body for Engagement Intensity Detection," pp. 3312-3316, 2019.
- [29] Thiruthuvanathan. M, Krishnan. B and Rangaswamy. M, "Engagement Detection through Facial Emotional Recognition Using a Shallow Residual Convolutional Neural Networks," Vol. 14, pp.236-247, 2021.
- [30] Chang. C, Zhang. C, Chen. L and Liu. Y, "An Ensemble Model Using Face and Body Tracking for Engagement Detection," pp. 616-622, 2018.
- [31] Liao. J, Liang. Y and Pan. J, "Deep facial spatiotemporal network for engagement prediction in online learning," Vol. 51, pp. 6609-6621, 2021.
- [32] Gupta. A, Jaiswal. R, Adhikari. S and Balasubramanian. V, "DAISEE: Dataset for Affective States in E-Learning Environments," Vol. 14, pp.1-12, 2016.

- [33] Alnafjan. A, Hosny. M, Al-Wabil. A and Al-Ohali. Y, "Classification of Human Emotions from Electroencephalogram (EEG) Signal using Deep Neural Network," , Vol. 8, pp. 419-425, 2017.
- [34] Abedi. A and Khan. S. S, "Improving state-of-the-art in Detecting Student Engagement with Resnet and TCN Hybrid Network," , pp. 1-7, 2021.
- [35] Li. J, Ngai. G, Leong. H. V and Chan. S. C. F, "Multimodal Human Attention Detection for Reading from Facial Expression, Eye Gaze, and Mouse Dynamics," , Vol. 16, pp. 37-49, 2016.
- [36] Sengupta. S, Basak. S, Saikia. P, Paul. S, Tsalavoutis. V, Atiah. F, Ravi. V and Peters. A, "A review of deep learning with special emphasis on architectures, applications and recent trends," , pp. 1-29, 2021.
- [37] Jalal. M. A, Mihaylova. L and Moore. R. K, "An End-to-End Deep Neural Network for Facial Emotion Classification," 2019 22th International Conference on Information Fusion (FUSION),, pp. 1-7., 2019.
- [38] Mollahosseini. A, Hasani. B and Mahoor. M. H, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," , pp. 1-18, 2017.
- [39] Simple MNIST convnet. (28 September 2020). [Online] Available : https://keras.io/examples/vision/mnist_convnet