

การพัฒนาแบบจำลองรู้จำใบหน้าแมวด้วยการเรียนรู้เชิงลึก

A development of cat face recognition model using deep learning

ภัทรพล ศรีรักษ์¹, จิตมนต์ อังสกุล² และ ธรา อังสกุล^{2*}
Pattarapol Srirak¹, Jitimon Angskun² and Thara Angskun^{2*}

Received: 26 April 2024 ; Revised: 24 June 2024 ; Accepted: 25 July 2024

บทคัดย่อ

ปัจจุบันจำนวนสัตว์เลี้ยงในประเทศไทยมีจำนวนสูงมากขึ้นในทุกปี ปัญหาที่เกิดขึ้นกับการเลี้ยงสัตว์คือการที่สัตว์เลี้ยงสูญหายหรือพลัดหลง แมวเป็นหนึ่งในสัตว์เลี้ยงที่มีความนิยม มีอัตราการพลัดหลงแล้วกลับมาเพียงร้อยละ 64 นอกจากนั้นอัตราการเสียชีวิตระหว่างสูญหายของแมวยังสูงกว่าสุนัขถึง 2 เท่า การระบุตัวตนแมวสามารถช่วยทำให้ระบุตำแหน่ง หรือให้ข้อมูลเกี่ยวกับแมวที่พลัดหลงได้ โดยวิธีการระบุตัวตนแมวแบบเดิม เช่น การสักหู การฝังชิป มีข้อจำกัดหลายประการ ไม่ว่าจะเป็นความรุนแรงต่อสัตว์ ความเสี่ยงต่อการติดเชื้อ หรือโอกาสที่อุปกรณ์จะสูญหาย บทความนี้นำเสนอวิธีการระบุตัวตนของแมวด้วย ซึ่งเป็นวิธีการไม่รุนแรงและมีต้นทุนต่ำ โดยเป็นการประยุกต์ใช้วิธีการระบุตัวตนด้วยใบหน้าของมนุษย์ โดยได้นำเสนอวิธีการรู้จำใบหน้าแมวโดยประกอบไปด้วย 3 ขั้นตอน 1) ขั้นตอนการตรวจจับใบหน้าแมวในภาพ ซึ่งสามารถตรวจจับใบหน้าและตำแหน่งของหู ตา และจมูกของแมวได้ 2) ขั้นตอนการตรวจสอบใบหน้า เพื่อเรียนรู้คุณลักษณะจากภาพใบหน้าคู่เหมือนของแมวตัวเดียวกัน และคู่ต่างของแมวคนละตัว 3) ขั้นตอนการระบุตัวตน เป็นกระบวนการที่จะนำคุณลักษณะที่สกัดได้จากใบหน้าแมวมาใช้เพื่อจัดเรียงหรือเปรียบเทียบข้อมูลชุดอื่นๆ โดยใช้วิธีการ K-NN เพื่อหาจำนวนใบหน้าที่คล้ายคลึงและระบุว่าเป็นแมวตัวใดในฐานข้อมูล ผลการทดลองพบว่า แบบจำลองสำหรับการตรวจจับใบหน้าที่พัฒนาขึ้น มีค่า mAP เท่ากับ 0.995 โดยขั้นตอนการระบุตัวตน 5 อันดับแรก โดยมีความแม่นยำอยู่ที่ 89%

คำสำคัญ: การรู้จำใบหน้าแมว, การระบุตัวแมว, การเรียนรู้เชิงลึก

Abstract

Currently, the number of pets in Thailand is increasing every year. Unfortunately, these pets may get lost or lost. Cats are one of the most popular pets. The rate of cats getting lost and returning is only 64 percent. In addition, the death rate of lost cats is twice as high as that of dogs. Cat identification can help to locate them or provide information about lost cats. Traditional cat identification methods, such as ear tattooing and microchipping, have many limitations. Among them are violence against animals, the risk of infection, or the possibility of the device being lost. This article introduces the development of cat face recognition to identify lost cats by their faces. This is a non-violent and low-cost method. It is an application of a method based on human facial identification. The method for recognizing cat faces consists of 3 steps: 1) A step for detecting cat faces in images, which can detect the face and the position of the ears, eyes, and nose. 2) A step to learn features from facial images of identical pairs of the same cat and different pairs of different cats. 3) An identification step is a process in which features extracted from cat faces are used to sort or compare other data sets using the K-NN method to find the number of similar faces and identify which cat is in the database. The experimental results found that the developed model for face detection has a mAP value of 0.995 by the top-5 identification process. The Identification has been shown to be 89% accurate.

Keywords: Cat face recognition, cat identification, deep learning

¹ นักศึกษาปริญญาโท สำนักศาสตร์และศิลปดิจิทัล มหาวิทยาลัยเทคโนโลยีสุรนารี

² รองศาสตราจารย์ สำนักวิชาศาสตร์และศิลปดิจิทัล มหาวิทยาลัยเทคโนโลยีสุรนารี

¹ Master's degree student, Institute of Digital Arts and Science, Suranaree University of Technology

² Associate Professor, Institute of Digital Arts and Science, Suranaree University of Technology

* Corresponding author: Thara Angskun; Email: angskun@sut.ac.th

บทนำ

ในประเทศไทย การเลี้ยงสัตว์เลี้ยงเพิ่มขึ้นเนื่องจากการเปลี่ยนแปลงทางสังคม เช่น จำนวนคนโสด จำนวนคู่ไม่มีบุตร และจำนวนผู้สูงอายุที่เพิ่มขึ้น การแพร่หลายของสื่อโซเชียลมีเดีย อัตราการเกิดลดลงในขณะที่การเลี้ยงสัตว์เลี้ยงเพิ่มขึ้น สะท้อนให้เห็นแนวโน้มการเลี้ยงสัตว์เลี้ยงเป็นสมาชิกในครอบครัว ปัญหาหลักของการเลี้ยงสัตว์เลี้ยง คือ สัตว์เลี้ยงสูญหายหรือพลัดหลง และการเพิ่มขึ้นของจำนวนสัตว์จรจัด จากสถิติของ Kantar Worldpanel (2022) พบว่าแมวเป็นหนึ่งในสัตว์เลี้ยงยอดนิยมและมีอัตราเติบโตเพิ่มสูงขึ้นอย่างต่อเนื่องในประเทศไทย (Figure 1) ซึ่งหากแมวสูญหายหรือพลัดหลงจะพบยาก เนื่องจากความรักอิสระและระมัดระวังตัว

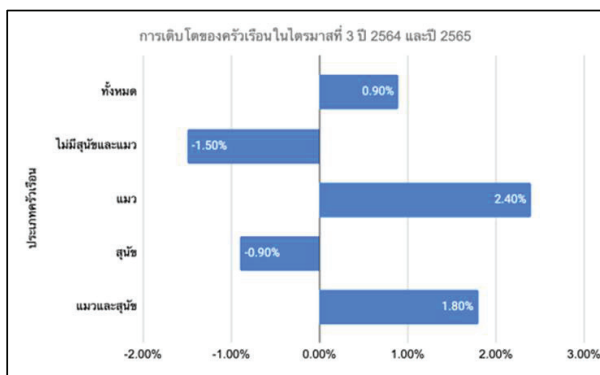


Figure 1 Household growth in the 3rd quarter of 2021 and 2022

วิธีการระบุตัวตนสัตว์เลี้ยงมีหลายแบบ ตั้งแต่ชั่วคราวไปจนถึงถาวร เช่น ปลอกคอ, การประทับเงิน, การสักหู และการฝังไมโครชิป นอกจากนี้ยังมีงานวิจัยที่พยายามใช้ชีวมาตรเพื่อลดค่าใช้จ่ายและความรุนแรงในการระบุตัวตน (Kumar & Singh, 2018) ทั้งนี้การศึกษาวิธีการรู้จำใบหน้าแมวยังมีความท้าทายเช่นการจำแนกแมวที่มีลักษณะเฉพาะที่ผิวหน้ามีลวดลายและพื้นที่ผิวน้อย จะทำให้การระบุตัวตนได้ยาก เพราะต้องใช้คุณลักษณะบนใบหน้าในการระบุตัวตน เช่น ความหนาแน่นของลายบนหน้าผาก ตำแหน่งของตุ่มบนจมูก และเนื้อผิวรอบดวงตา (Lin & Kao, 2018)

งานวิจัยนี้มุ่งเน้นการรู้จำใบหน้าแมว ซึ่งเป็นเทคนิคใหม่ในการระบุตัวแมว โดยการพัฒนาขั้นตอนวิธีและประยุกต์ใช้ระบบการเรียนรู้เครื่องเพื่อระบุตัวตนแมวที่หายไป โดยงานวิจัยนี้เสนอโมเดลรู้จำใบหน้าแมว 3 ขั้นตอน โดยใช้สถาปัตยกรรม YoLoV5-Face และ Triplet Neural Network ซึ่งมีศักยภาพในการลดค่าใช้จ่ายจากการใช้อุปกรณ์เพื่อระบุตัวตนและเพิ่มความแม่นยำในการระบุตัวตนแมว เพื่อช่วยให้อาณาเขตที่หายไปสามารถกลับคืนสู่เจ้าของได้

งานวิจัยที่เกี่ยวข้อง

การระบุตัวตนสัตว์เลี้ยง

หลายงานวิจัยได้กล่าวถึงความเป็นเมืองอัจฉริยะหรือการเปลี่ยนผ่านเข้าสู่ยุคดิจิทัล (Kumar & Singh, 2018; Moreira *et al.*, 2017) สิ่งที่เกิดขึ้นกับผู้เลี้ยงสัตว์ คือ การพลัดหลงกับสัตว์เลี้ยงหรือหลุดหายไป ทำให้มีการศึกษาเกี่ยวกับวิธีการที่จะระบุตัวตนของสัตว์ว่าจะใช้วิธีการใด จากการศึกษางานวิจัยด้านวิธีการระบุตัวตนสัตว์เลี้ยงในปัจจุบันพบว่าวิธีการระบุตัวตนสัตว์เลี้ยงแบบเดิมแบ่งออกเป็น 3 ประเภท ได้แก่ การระบุตัวตนแบบถาวร การระบุตัวตนแบบกึ่งถาวร และการระบุตัวตนแบบชั่วคราว ซึ่งตัวอย่างของการระบุตัวตนแต่ละแบบได้แก่ การระบุตัวตนแบบถาวร เช่น การสักใบหู การฝังไมโครชิป การเจาะหูหรือการตัดหู การประทับเงิน การประทับเงิน ตัวอย่างการระบุตัวตนแบบกึ่งถาวร เช่น การติดป้ายที่หู การใส่ปลอกคอ ซึ่งจะพบว่าวิธีการแบบถาวรและกึ่งถาวรนั้นเป็นวิธีการที่รุนแรงกับร่างกายของสัตว์ ต้องใช้ผู้เชี่ยวชาญหรือผู้ที่มีประสบการณ์ และมีความเสี่ยงในการติดเชื้อมากขึ้นกับสัตว์ ตัวอย่างการระบุตัวตนแบบชั่วคราว เช่น การย้อมสี การแขวนป้าย RFID วิธีการแบบชั่วคราวจะมีการปฏิสัมพันธ์กับร่างกายของสัตว์น้อยกว่าเนื่องจากการสวมใส่อุปกรณ์ทำให้ไม่มีการใช้ความรุนแรง แต่ก็เป็นวิธีที่สามารถสลาย หรือสูญหายได้ ปัญหาเรื่องการระบุตัวตนจึงเป็นสิ่งที่ท้าทายในงานวิจัยด้านนี้ ซึ่งมีงานวิจัยได้นำเสนอวิธีการแก้ปัญหาโดยใช้ ชีวมาตรสัตว์ โดยแบ่งออกเป็น 4 วิธีดังนี้ 1) การระบุจากจุดบนจมูก 2) การระบุจากรูปแบบดวงตา 3) การระบุจากใบหน้า 4) การระบุจากหลอดเลือดจอประสาทตา ดิจิทัล (Kumar & Singh, 2018) ซึ่งวิธีการระบุตัวตนโดยใช้ชีวมาตรสัตว์จะสามารถลด ความรุนแรง ความเสี่ยงที่จะเกิดโรค และโอกาสสูญหายของอุปกรณ์ได้ ทั้งนี้วิธีการระบุตัวตนสัตว์ด้วยใบหน้านั้น สามารถประยุกต์ใช้กับการค้นหาแมวที่สูญหายได้ดี เนื่องจากสามารถใช้ภาพถ่ายในมุมกว้าง เช่น ภาพจากกล้องโทรทัศน์วงจรปิดได้ งานวิจัยนี้จึงเลือกใช้วิธีการดังกล่าว

การระบุตัวตนสัตว์ด้วยใบหน้า

กระบวนการระบุตัวตนสัตว์เลี้ยงด้วยใบหน้าจะถูกแบ่งออกเป็น 3 ขั้นตอนคือ 1) การระบุใบหน้า 2) การระบุคุณลักษณะ 3) การระบุตัวตน ซึ่งเมื่อเปรียบเทียบในแต่ละขั้นตอนจะมีรายละเอียดดังนี้ ขั้นตอนการระบุใบหน้า เป็นขั้นตอนแรกเริ่มของการรู้จำใบหน้า เพื่อที่จะค้นหาใบหน้าของสิ่งที่สนใจก่อนจะนำมาสกัดคุณลักษณะ ซึ่งการระบุใบหน้าจะแบ่งเป็น 2 ประเภท คือ Feature-Based และ Deep Learning-Based ประเภท Feature-Based อาทิ งานของมอเรราห์และคณะ (Moreira *et al.*, 2017) ได้ใช้ EigenFace และ Fisher Face ในการตรวจจับใบหน้าสุนัข งานของหลวนและคณะ (Ruan *et al.*, 2014) เลือกใช้ Deformable Parts Model (DPM) ซึ่ง

มีจุดเด่นคือความรวดเร็วและเข้าใจได้ง่าย แต่มีข้อจำกัด เรื่องแบบจำลองสำหรับการสกัดคุณลักษณะไม่มีความยืดหยุ่น ทำให้ใช้งานต่อได้ยาก ต้องใช้มนุษย์เป็นผู้กำหนดคุณลักษณะด้วยตัวเองทำให้ต้องใช้ทรัพยากรเช่นผู้ที่มีความรู้ในการกำหนดคุณลักษณะและเวลาที่ใช้ในการกำหนดคุณลักษณะ ทั้งนี้ในระยะหลัง วิธีการเรียนรู้ด้วยเทคนิคการเรียนรู้เชิงลึก (Montenegro *et al.*, 2022) นั้นได้รับความนิยมมากขึ้น เนื่องจากปัจจัยทางด้านฮาร์ดแวร์ที่มีประสิทธิภาพมากขึ้นและซอฟต์แวร์ที่มีการพัฒนาเครื่องมือใหม่ๆ ให้สามารถใช้งานได้ง่ายขึ้น งานวิจัยของยูนและคณะ (Yoon *et al.*, 2021) ได้เลือกใช้ ResNet-like ลินและเคา (Lin & Kao, 2018) ได้ใช้ Faster R-CNN แบบสองขั้นตอนในการค้นหาระนาบที่หน้าสนใจและใบหน้า อตัม ไคลน์ (Klein, 2019) ได้ใช้ YOLOv3 ในการตรวจจับใบหน้า ซึ่งวิธีการเรียนรู้เชิงลึก มีข้อดี คือ ความสามารถในการให้คอมพิวเตอร์เป็นผู้ค้นหาคุณลักษณะข้อมูลเองได้ มีความยืดหยุ่นและสะดวกสามารถนำไปใช้งานต่อในงานใกล้เคียงได้ แต่ก็ต้องใช้พลังในการประมวลผลค่อนข้างสูง ซึ่งโดยสรุปแล้วแต่ละวิธีนั้นก็จะมีข้อดีและข้อเสียที่แตกต่างกัน

ดังนั้นงานวิจัยนี้จึงเลือกใช้ YOLOv5-Face ที่พัฒนาโดยชีและคณะ (Qi *et al.*, 2022) เนื่องจากเป็นวิธีการการเรียนรู้เชิงลึก ซึ่งมีความสามารถในการตรวจจับใบหน้าได้ในหลากหลายสภาพแวดล้อม มีการทำงานที่รวดเร็วและสามารถระบุจุดสำคัญบนใบหน้าได้

ขั้นตอนการสกัดคุณลักษณะ เป็นขั้นตอนที่คอมพิวเตอร์จะเรียนรู้ลักษณะของภาพ เช่นเดียวกับขั้นตอนระบุใบหน้าที่การสกัดคุณลักษณะสามารถทำได้ 2 วิธี คือ Feature-Based และ Deep Learning-Based ซึ่งจากการทบทวนวรรณกรรมแต่ละงานวิจัย จะมีการเลือกใช้วิธีการที่แตกต่างกัน อาทิ กลุ่มงานวิจัยที่เลือกใช้วิธีการแบบ Feature-Based ได้แก่ कुमारและซิง (Kumar & Singh, 2018) ได้นำเสนอวิธีการสกัดคุณลักษณะแบบ Fisher Linear Projection and Preservation (FLPP) ที่เป็นกระบวนการสกัดคุณลักษณะของใบหน้าสุนัขพร้อมกระจายชุดข้อมูลระหว่างคลาส งานวิจัยของ ลินและเคา (Lin & Kao, 2018) เลือกใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันในการระบุตัวตน โดยเลือกใช้ EigenFace เพื่อสกัดคุณลักษณะบนใบหน้าแมวจากชิ้นส่วนใบหน้า กลุ่มงานวิจัยที่เลือกใช้วิธีการแบบ Deep Learning Based ได้แก่ ไลและคณะ (Lai *et al.*, 2019) ได้ใช้โครงข่ายประสาทเทียมแบบคอนโวลูชันสถาปัตยกรรม Faster R-CNN ในการเรียนรู้คุณลักษณะของภาพโดยนำกระบวนการแบบ Coarse-to-fine โดยใช้แบบจำลองสองตัว แบบจำลองตัวแรกเพื่อทำนายพื้นที่ของสุนัขและ แบบจำลองตัวที่สองเพื่อสกัดคุณลักษณะของใบหน้าเพื่อระบุตัวตน งานวิจัยของมูเกต์และคณะ (Mougeot *et al.*, 2019) ไม่มีขั้นตอนของการระบุใบหน้า แต่ใช้การสกัดคุณลักษณะจาก

ภาพใบหน้าสุนัขที่ผ่านการประมวลผลมาแล้วด้วยการใช้โครงข่ายประสาทเทียมแบบแฝดสามคู่กับฟังก์ชัน Triplet Loss เพื่อเรียนรู้คุณลักษณะของใบหน้า เช่นเดียวกับงานของ ไคลน์ (Klein, 2019) และยูนและคณะ (Yoon *et al.*, 2021) ซึ่งงานวิจัยนี้เลือกใช้วิธีการแบบ Deep Learning Based โดยเลือกใช้โครงข่ายประสาทเทียมแบบแฝดสามคู่กับฟังก์ชัน Triplet Loss มาใช้ในการพัฒนา เนื่องจากเป็นกระบวนการที่ได้รับความนิยมและได้รับการพัฒนาวิธีการอย่างต่อเนื่อง จึงง่ายต่อการพัฒนาและต่อยอดทั้งในปัจจุบันและอนาคต

ขั้นตอนการระบุตัวตนเป็นขั้นตอนที่จะนำข้อมูลที่ถูกลักษณะคุณลักษณะมารวบรวมไว้ เพื่อเปรียบเทียบกับภาพใบหน้าเป้าหมายที่ถูกนำเข้ามาเพื่อระบุตัวตน ซึ่งกระบวนการที่นิยมนั้นมีหลากหลายเทคนิคดังนี้ เทคนิค Support Vector Machine (SVM) ถูกใช้ในงานวิจัยของ ลินและเคา (Lin & Kao, 2018) เพื่อเปรียบเทียบค่าคุณลักษณะโดยใช้การลากผ่านของค่าข้อมูล เทคนิค K-Nearest Neighbor (K-NN) ถูกใช้ในงานของยูนและคณะ ไคลน์ และมูเกต์และคณะ (Yoon *et al.*, 2021; Klein, 2019; Mougeot *et al.*, 2019) โดยใช้ในการเปรียบเทียบค่าจากข้อมูลในตำแหน่งโดยรอบ ซึ่งจุดเด่นของเทคนิคนี้คือ ความสามารถในการปรับค่าข้อมูลนำเข้า และจำนวนเพื่อนบ้านใกล้เคียงได้ ทำให้มักถูกนำมาวัดผลการทำนาย 1 อันดับจนถึง 5 อันดับได้

การศึกษางานวิจัยด้านวิธีการระบุตัวตนสัตว์เลี้ยงในแง่ผลภาพรวมของวิธีการมีรายละเอียดดังนี้ เฉินและคณะ (Chen *et al.*, 2016) ได้นำเสนอวิธีการระบุตัวตนแมวโดยใช้ลักษณะของจมูก เนื่องจากลักษณะจมูกของสัตว์ส่วนใหญ่จะมีลักษณะเหมือนลายนิ้วของคน เช่นเดียวกับปศุสัตว์ โดยได้มีการทดลองใช้รูปจมูกแมวกว่า 700 รูป จากแมว 70 ตัว เพื่อระบุตัวตนของแมว แต่อย่างไรก็ตาม งานวิจัยดังกล่าวไม่ได้ระบุผลการทดลองเชิงปริมาณอย่างชัดเจน หลิวและคณะ (Liu *et al.*, 2012) ได้ใช้การตรวจจับใบหน้าสุนัขเพื่อคัดแยกสายพันธุ์โดยใช้สไลด์ิงวินโดวส์ (Sliding window) เพื่อระบุจมูกและตาเพื่ออ้างอิงไปยังจุดที่เป็นหูและส่วนประกอบที่เหลือ จากนั้นจึงวิเคราะห์ฮิสโตแกรม (Histogram) และใช้ เอสไอเอฟที (SIFT: Scale Invariant Feature Transform) ในการวิเคราะห์เพื่อหาจุดเด่นของรูปภาพที่ได้รับเข้ามา ซึ่งมีความแม่นยำ 67% จากสุนัข 133 สายพันธุ์ จำนวน 8,351 รูปภาพ งานวิจัยของ ลินและเคา (Lin & Kao, 2018) พัฒนาระบบรู้จำใบหน้าแมวโดยใช้การเรียนรู้เชิงลึกแบบฟาสต์เตอร์อาร์-ซีเอ็นเอ็น (Faster R-CNN) โดยศึกษาจากรูปภาพของแมวจำนวน 150 ตัว ซึ่งในการพัฒนา ผู้วิจัยได้แบ่งแบบจำลองออกเป็นสองส่วน คือ ส่วนสำหรับการตรวจจับใบหน้า และส่วนสำหรับการระบุส่วนประกอบบนใบหน้า ได้แก่ หน้าผาก ตา จมูก และปาก จากนั้นจึงใช้เอสวีเอ็ม (SVM) ในการวิเคราะห์

ลักษณะของส่วนประกอบเพื่อคัดแยกแมวแต่ละตัว ซึ่งในการทดลองสามารถระบุตัวตน โดยมีความแม่นยำภาพ 94.1% แต่อย่างไรก็ตามยังมีข้อจำกัดด้านสภาพแสงและแมวบางลักษณะ

เช่น แมวดำ ที่ไม่สามารถแยกคุณลักษณะเฉพาะออกมาเพื่อคัดแยกได้ งานวิจัยที่เกี่ยวข้องสามารถสรุปแสดงดัง Table 1

Table 1 Research process comparison

Author	Method ¹	Face Detection Method	Face Detection Feature			Feature Extraction Method	Identification Method
			Multiple Face	Face Landmark	Face Alignment		
Adam Klein	D	YOLO	✓			Triplet Neural Network+GOR+Triplet Mining	KNN
Lin, Kao	H	Faster R-CNN	✓	✓		Eigen Face	SVM
Lai <i>et al.</i>	D	Faster R-CNN	✓	✓	✓	CNN	CNN(Prediction)
Mougeot <i>et al.</i>	D					Triplet Neural Network +Triplet Mining	RANSAC
Moreira <i>et al.</i>	H					Eigen Face+LBP +CNN	SVM
Kumar, Singh	M					LBP	SVM+LDA
Chen <i>et al.</i>	M					HOG	SVM
Ruan <i>et al.</i>	M	DPM		✓		HOG	
This Research	D	YOLOv5-Face	✓	✓	✓	Triplet Neural Network+GOR+Triplet Mining	KNN+FAISS

1 D = Deep Learning (การเรียนรู้เชิงลึก),
H = Hybrid (วิธีการแบบผสมผสาน),
M = Manual (วิธีการสกัดคุณลักษณะด้วยมือ)

แบบจำลองการรู้จำใบหน้าแมว

สถาปัตยกรรมของแบบจำลอง

การออกแบบระบบรู้จำใบหน้าสัตว์เลี้ยงด้วยการเรียนรู้เชิงลึกประกอบด้วยสามขั้นตอนหลัก ได้แก่ 1) การตรวจจับใบหน้าแมวภายในภาพ ซึ่งมีความท้าทายจากสภาพแวดล้อมที่หลากหลาย อาทิ ระยะห่าง แสง และเงา 2) การเปรียบเทียบใบหน้า ซึ่งใช้กระบวนการสกัดคุณลักษณะและฝึกฝนแบบจำลองเพื่อแยกคุณลักษณะของแมวแต่ละตัว และ 3) การระบุตัวตน, โดยการสร้างชุดเวกเตอร์ของข้อมูลและใช้พื้นที่การฝังตัวข้อมูล (Embedding Space) เพื่อเปรียบเทียบกับข้อมูลในฐานข้อมูลโดยใช้วิธีการ KNN เพื่อระบุตัวตนของแมว (Figure 2)

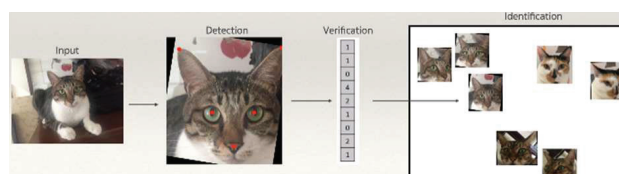


Figure 2 Architecture of the cat face recognition model

2. ชุดข้อมูลนำเข้า

ชุดข้อมูลนำเข้าในงานวิจัยชิ้นนี้ ประกอบด้วย 2 ชุด ได้แก่ ชุดข้อมูลจากแค็กเกิล (Kaggle) ซึ่งประกอบไปด้วยข้อมูลใบหน้าของแมว 9,993 รูปภาพ โดยแบ่งเป็น 7,284 ภาพสำหรับการฝึกฝนและ 2,709 ภาพสำหรับการทดสอบ (Figure 3) ชุดข้อมูลดังกล่าว เป็นชุดข้อมูลเพื่อใช้สำหรับการฝึกแบบจำลองให้สามารถตรวจจับใบหน้าของแมวได้ ชุดข้อมูลชุดที่สองเป็นชุดข้อมูลใบหน้าของแมวจากในงานวิจัยหลินและเคา (Lin & Kao, 2018) (Figure 4) ซึ่งได้รวบรวมใบหน้าแมว 519 ตัวจากสถานรับเลี้ยงสัตว์ ซึ่งแมวแต่ละตัวจะมีภาพไม่ต่ำกว่า 10 ภาพมุมที่แตกต่างกัน โดยจากข้อมูลดิบจำนวน 13,536 ภาพ เมื่อคัดแยกภาพที่มีความละเอียดต่ำ มีภาพน้อยกว่า 5 รูป ภาพซ้ำ ภาพเบลอ และไม่สมบูรณ์อื่นๆ มีภาพแมวคงเหลือจำนวน 396 ตัว โดยแบ่งเป็นชุดสำหรับฝึกฝนจำนวน 328 ตัว 8,125 ภาพ ชุดสำหรับการตรวจสอบ 68 ตัว 2,258 ภาพ เพื่อนำมาใช้ฝึกและทดสอบแบบจำลอง เพื่อให้สามารถรู้จำใบหน้าแมวแต่ละตัวได้ และใช้ชุดข้อมูลดังกล่าวสำหรับการทดสอบความแม่นยำในการทำนายผล



Figure 3 Example of cat image in dataset from Kaggle.

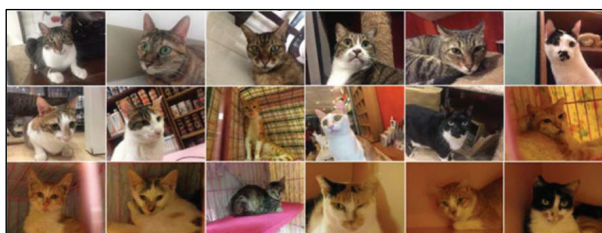


Figure 4 Examples of cat faces presented in Lin and Kou's research

กระบวนการตรวจสอบใบหน้า

กระบวนการตรวจสอบใบหน้าในงานวิจัยนี้ ได้ทดลองใช้ pre-trained โมเดล โดยวิธีการแรกใช้ Haar Cascade Classifier และ FHOG-based SVM ซึ่งพบว่าผลลัพธ์การตรวจจับใบหน้าแมวของทั้งสองวิธีนั้นมีความผิดพลาดสูง จึงได้เปลี่ยนเป็นวิธีการเรียนรู้เชิงลึก (Deep Learning) โดยใช้ CNN ซึ่งมีความแม่นยำกว่าสองวิธีการแรก แต่มีข้อจำกัดคือใช้ระยะเวลาประมวลผลนาน สามารถค้นหาได้เพียง 1 ใบหน้าต่อภาพ งานวิจัยนี้จึงได้ปรับเปลี่ยนโดยใช้เทคนิค Faster R-CNN ซึ่งเป็นเทคนิคเดียวกันกับงานวิจัยของ Lin และ Kou ซึ่งมีข้อจำกัดเรื่องสถาปัตยกรรมถูกออกแบบมาเพื่อให้ระบุ Bounding Box ของใบหน้าเท่านั้น หากต้องการระบุจุดสำคัญบนใบหน้าจะต้องพัฒนาแบบจำลองเพิ่มสำหรับการประมวลผลซึ่งจะต้องใช้เวลาและทรัพยากรในการประมวลผลเพิ่มขึ้น

ดังนั้นงานวิจัยนี้จึงเลือกใช้สถาปัตยกรรมตระกูล YOLO ที่มีจุดเด่นในเรื่องความเร็วและความแม่นยำเช่นเดียวกับ Faster R-CNN โดยแบบจำลองได้ถูกออกแบบมาเพื่อใช้สำหรับการค้นหาวัตถุที่หน้าจะมีวัตถุและสร้าง Bounding Box ขึ้นมารอบวัตถุเท่านั้น โดยงานวิจัยนี้จึงเลือกใช้ YOLOv5-Face ในเวอร์ชันย่อยที่ได้รับการปรับปรุงให้สามารถใช้การระบุจุดสำคัญบนใบหน้าได้ (Figure 5)



Figure 5 Sample image of the results of cat face landmark and alignment

กระบวนการตรวจสอบใบหน้า

แบบจำลองที่ใช้ในขั้นตอนตรวจสอบใบหน้าแมวนี้ ได้พัฒนาในรูปแบบโครงข่ายประสาทเทียมแบบแฝดสาม (Triplet Neural Network) ซึ่งเป็นแบบจำลองสามารถรับข้อมูลนำเข้าทั้งหมด 3 ชุด และมีการแชร์น้ำหนัก (Weight) กันภายในแบบจำลอง เพื่อให้ภาพเดียวกันได้ผลลัพธ์ออกมาเหมือนกัน (Figure 6)

งานวิจัยนี้ใช้การถ่ายโอนความรู้ (Transfer Learning) เข้ามาช่วยเพื่อให้ประหยัดเวลาและทรัพยากรในการพัฒนา โดยแบบจำลองที่เลือกใช้จะเป็นแบบจำลอง EfficientNetV2S ซึ่งโมเดลที่ถูกฝึกฝนไว้ล่วงหน้า (Pre-trained Model) โดยจะไม่ใช้เลเยอร์ชั้นสุดท้าย แต่จะใช้เพียงเลเยอร์รองสุดท้ายเพราะต้องการผลลัพธ์ข้อมูลในลักษณะของเวกเตอร์ โดยจะเชื่อมเลเยอร์ต่อท้ายเพื่อปรับผลลัพธ์ให้เป็นเวกเตอร์ขนาด 128 จากนั้นผลลัพธ์จะถูกใช้ฟังก์ชัน L2 เพื่อ Normalize ข้อมูล

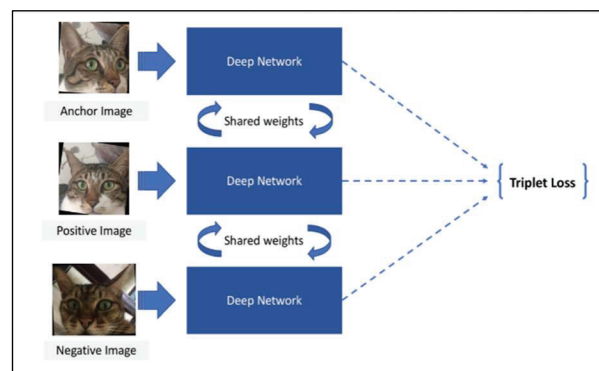


Figure 6 Example image of cat face training with Triplet Network

และเมื่อได้ผลลัพธ์แบบเวกเตอร์แล้ว ผลลัพธ์ดังกล่าวก็จะถูกนำไปใช้ในฟังก์ชัน Triplet Loss เพื่อทำให้ภาพใบหน้าที่แมวตัวเดียวกันอยู่ในกลุ่มเดียวกัน และแมวที่แตกต่างกันจะกระจายตัวจากกัน (Figure 7)

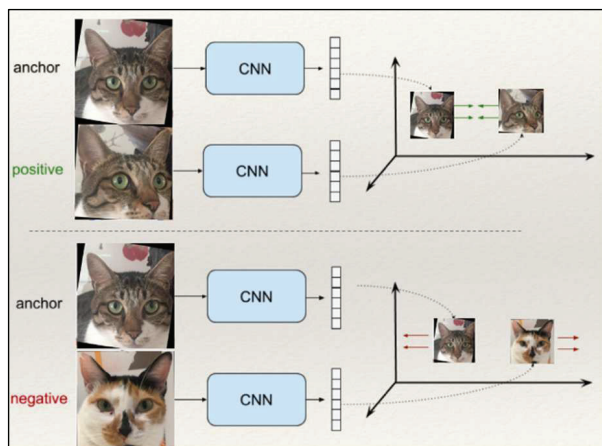


Figure 7 Process of feature extraction and cat face identification

อย่างไรก็ตามการใช้ Triplet Neural Network และ Triplet Loss ยังไม่สามารถสร้างการกระจายตัวของข้อมูลได้ดีเพียงพอ เพราะบางครั้งลักษณะใบหน้าแมวในคลาสอื่น ก็อาจจะมีค่าระยะทางที่ใกล้กว่าใบหน้าแมวในคลาสเป้าหมาย งานวิจัยนี้จึงประยุกต์ใช้วิธีการในการช่วยเลือกข้อมูลและการปรับค่า Loss ของข้อมูลเพื่อให้แบบจำลองเรียนรู้คุณลักษณะที่เหมือนและแตกต่างกันมากขึ้นโดยกระบวนการ Hard Batch ที่ถูกนำเสนอโดย เฮอร์แมนและคณะ (Hermans *et al.*, 2012) คือการทำ Triplet Mining ซึ่งจะช่วยให้แบบจำลองได้เรียนรู้ภาพของคลาสที่มีความใกล้เคียงกัน โดยเมื่อนำ Triplet Mining แบบ Hard Batch มาใช้ ในกระบวนการคัดเลือกรูปภาพแต่ละรอบ โดยสุ่มคลาสจากจำนวน $\text{BatchSize}/4$ คลาส จากนั้นหนึ่งคลาสจากที่สุ่มมาจะถูกเลือกเพื่อเป็นภาพ Anchor และภาพ Positive คลาสที่เหลือจะถือว่าเป็นภาพ Negative ในการเลือกภาพ จะคำนวณระยะทางด้วยวิธีการ Euclidian Distance ระหว่างภาพ Anchor กับทุกภาพที่เหลือในคลาส และจะเลือกภาพที่มีระยะทางสูงสุด (Farthest Positive) เพื่อใช้หาภาพแมวตัวเดียวกันที่มีความแตกต่างกันมากที่สุด จากนั้นคลาสที่เหลือที่เป็น Negative จะถูกสุ่มเลือกมาและหาภาพที่มีความคล้ายกับภาพ Anchor ที่สุดโดยการคำนวณหาระยะทางใกล้ที่สุด (Nearest Negative) จากนั้นค่าจะนำค่าที่ได้ไปคำนวณและปรับค่า Loss ทำให้ข้อมูลมีการกระจายตัวที่ดีขึ้น (Figure 8)



Figure 8 Example of comparison between Hard Triplet

นอกจากนี้ภายใน Triplet Loss Function ได้มีการนำ Generalized Outlier Rejection (GOR) ที่เสนอโดยจางและคณะ (Zhang *et al.*, 2017) มาประยุกต์ใช้เพื่อปรับปรุงประสิทธิภาพของแบบจำลอง โดย GOR จะทำหน้าที่เป็น Regularization ภายใน Triplet Loss Function ซึ่งปกติจะใช้ในการลดค่า Loss ที่เกิดจากการเรียนรู้ของข้อมูล Anchor, Positive และ Negative อย่างไรก็ตาม การเพิ่ม GOR เข้าไปจะช่วยเพิ่มโทษให้กับคู่ข้อมูล Anchor และ Negative ที่มีระยะห่างไม่เพียงพอ ทำให้โมเดลมีความสามารถในการแยกแยะระหว่าง Anchor กับ Negative ได้ดีขึ้น

GOR จะช่วยกระจายตัวข้อมูลให้ดีขึ้นและลดความเสี่ยงของการเกิด overfitting โดยการเพิ่มโทษให้กับค่าที่ไม่มีความแตกต่างอย่างชัดเจนระหว่าง Anchor และ Negative ซึ่งกระบวนการนี้จะส่งผลให้ Triplet Loss Function สามารถแยกแยะข้อมูลได้ดีขึ้น ทำให้โมเดลสามารถเรียนรู้ได้อย่างมีประสิทธิภาพมากขึ้นและสามารถ generalize ได้ดียิ่งขึ้นเมื่อทดสอบกับข้อมูลใหม่ที่ไม่เคยเห็นมาก่อน

การคำนวณ GOR ประกอบด้วยการคำนวณค่า M1 ซึ่งเป็นค่าเฉลี่ยของ dot product ระหว่าง Anchor และ Negative และค่า M2 ซึ่งเป็นค่าเฉลี่ยของ dot product ที่ยกกำลังสอง ระหว่าง Anchor และ Negative จากนั้นกำหนดเงื่อนไข GOR โดยใช้สูตร แสดงดังสมการที่ 1

$$GOR = M1^2 + \max(M2 - \frac{1}{embedding_size}, 0) \quad (1)$$

เพื่อเพิ่มโทษให้กับคู่ข้อมูลที่มีระยะห่างไม่เพียงพอ การใช้ GOR ช่วยให้ข้อมูลมีการกระจายตัวในมิติที่มากขึ้น ส่งผลให้การเรียนรู้มีประสิทธิภาพสูงขึ้นและโมเดลสามารถ generalize ได้ดียิ่งขึ้นเมื่อทดสอบกับข้อมูลใหม่ที่ไม่เคยเห็นมาก่อน

กระบวนการระบุตัวตน

เมื่อได้ผลลัพธ์เป็นคุณลักษณะของใบหน้าในรูปแบบเวกเตอร์แล้ว ในขั้นการระบุตัวตน งานวิจัยนี้จะใช้ Facebook AI Similarity Search (FAISS) เพื่อค้นหาตามความคล้ายคลึงกัน ด้วยจุดเด่นคือด้านความเร็วและความสามารถในการจัดกลุ่มของข้อมูลในรูปแบบ Vector โดยในงานวิจัยนี้จะใช้แบบจำลองในขั้นตอน Verification สร้างเวกเตอร์จากใบหน้าแมวในฐานะข้อมูล และใช้ FAISS เพื่อสร้าง Index ของใบหน้าแมวเก็บไว้ในหน่วยความจำเพื่อการเรียกใช้งาน เมื่อมีการส่งใบหน้าแมวเข้ามาใหม่ จะประมวลผลด้วยแบบจำลอง ก่อนจะนำเวกเตอร์ที่ได้มาเปรียบเทียบกับใบหน้าที่ถูกทำ Index ไว้ในหน่วยความจำ ด้วยวิธีการ KNN ซึ่งจะถูกใช้ในการทำนายใบหน้าจำนวน M รูป ซึ่งค่า K ถูกใช้ในกระบวนการทดสอบ คือ 1,3,4 และ 5

กระบวนการวัดผล

งานวิจัยนี้ประเมินประสิทธิภาพของแบบจำลองในการตรวจจับและระบุตัวตนของแมวโดยใช้ค่า Confusion Matrix ดังที่แสดงใน Table 2 โดยจะใช้การวิเคราะห์ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) ค่าความแม่นยำ คำนวณจากอัตราการทำนายถูกต้องต่อการทำนายทั้งหมด ในขณะที่ค่าความระลึกคือจำนวนที่ทำนายถูกต้องต่อจำนวนที่มีจริง การวัดประสิทธิภาพจะพิจารณาจากค่าความแม่นยำเฉลี่ย (AP) และค่าความแม่นยำเฉลี่ยโดยรวม (mAP) สำหรับแต่ละคลาส โดยมีเป้าหมายที่จะทำนายการระบุตัวตนของแมวในสภาพแวดล้อมที่มีวัตถุอื่นๆ อย่างแม่นยำ

Table 2 Confusion Matrix

	Predicted classes	
	Yes	No
Actual class	Yes True Positive	No False Negative
	No False Positive	True Negative

ค่าความแม่นยำ (Precision) คือ อัตราการทำนายถูกต้องต่อการทำนายทั้งหมด โดย n คือจำนวนที่แบบจำลองพยายามทำนายทั้งหมด เช่น การทำนาย 2 ตัว โอกาสที่จะทำนายถูกต้องทั้งหมด คือ ทีพี (TP) = 2 หรือ ทำนายถูกต้อง 1 ตัวทำนาย ผิด 1 ตัว, ทีพี (TP) =1 เอฟพี (FP) =1 หรือ ทำนายผิดหมดทั้งหมด แสดงดังสมการที่ 2

$$precision = \frac{tp}{tp + fp} = \frac{tp}{n} \quad (2)$$

ค่าความระลึก (Recall) คือ จำนวน ที่ทำนายถูกต้องต่อจำนวนที่มีจริง (Ground Truth) ทั้งหมด แสดงดังสมการที่ 3

$$recall = \frac{tp}{tp + fn} \quad (3)$$

ค่าความแม่นยำเฉลี่ย (AP: Average Precision) คำนวณได้จากพื้นที่ใต้กราฟของค่าความแม่นยำ (Precision) กับค่าความระลึก (Recall) โดยจะมีค่าอยู่ระหว่าง 0 ถึง 1 ซึ่งการหาพื้นที่ใต้กราฟจะใช้สมการที่ 4

$$AP = \sum_{k=1}^N p(k) \Delta r(k) \quad (4)$$

โดยที่ N คือ จำนวน รูปทั้งหมดที่ทำนายออกมาทั้งถูกต้องและไม่ถูกต้อง ส่วน P(k) คือ ค่าความแม่นยำ (Precision) ที่ตำแหน่ง ที่เกินขีดแบ่ง (threshold) ที่ค่า k นั้นๆ ส่วน $\Delta r(k)$ คือการเปลี่ยนแปลงของ ค่าความระลึก (recall) ระหว่าง k-1 และ k ซึ่งการคำนวณหาค่าความแม่นยำเฉลี่ย (AP) นี้ จะหาเฉพาะ เพียง 1 คลาส เท่านั้น ซึ่งภายหลังจากการหาค่าความแม่นยำเฉลี่ยในทุกๆ คลาส แล้วจากนั้นจะนำค่าเอพีม่าเฉลี่ย เพื่อคำนวณหาค่าเอ็มเอพี (mean Average Precision : mAP) ของแบบจำลองซึ่งแสดงดังสมการที่ 5

$$mAP = AP/(n) \quad (5)$$

โดย

n คือจำนวนครั้งที่ทำการประเมิน

ผลการทดลองและอภิปรายผล

งานวิจัยนี้ได้ทดสอบประสิทธิภาพของแบบจำลองที่พัฒนาขึ้นโดยแบ่งเป็นสองส่วน คือ การตรวจจับใบหน้าและการระบุตัวตนของแมว ขั้นตอนการตรวจจับใบหน้ามุ่งเน้นที่การหาใบหน้าแมวในภาพ รวมถึงการระบุจุดสำคัญบนใบหน้า เช่น หู, ตา, และจมูก ส่วนการระบุตัวตนจะปรับแต่งค่า Hyper Parameter เพื่อเพิ่มประสิทธิภาพการตรวจจับ ข้อมูลที่ใช้ในการทดสอบประกอบด้วย 2,709 ภาพใบหน้าแมวจาก Kaggle พร้อมเพิ่มข้อมูลคุณลักษณะที่ระบุตำแหน่งจุดสำคัญบนใบหน้า

การทดสอบการตรวจจับใบหน้า

แบบจำลองสำหรับการตรวจจับใบหน้าได้ถูกฝึกฝนเป็นจำนวน 300 epoch เมื่อทดสอบมีค่า mAP อยู่ที่ 0.995 (Figure 9) โดยข้อมูลใบหน้าแมวจากหลากหลายขนาดถูกนำเข้า เพื่อตรวจจับใบหน้าโดยมีค่า Threshold 0.5 พบว่าใบหน้าแมวสามารถตรวจจับได้แม้อยู่ในต่างมุม สภาพแวดล้อมหรือมีการบดบัง นอกจากนี้ยังสามารถระบุจุดสำคัญบนใบหน้าของแมวแต่ละตัวได้ (Figure 10)

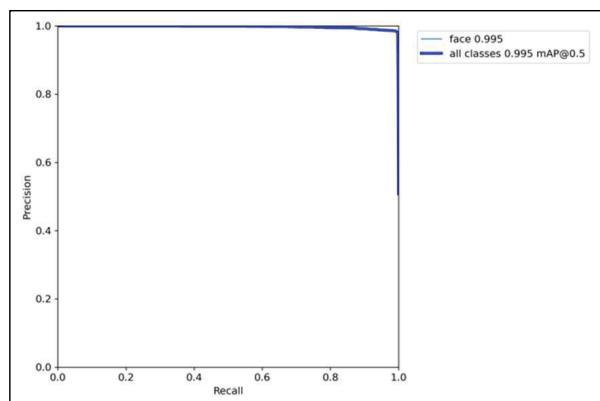


Figure 9 mAP results from testing a model for cat face detection

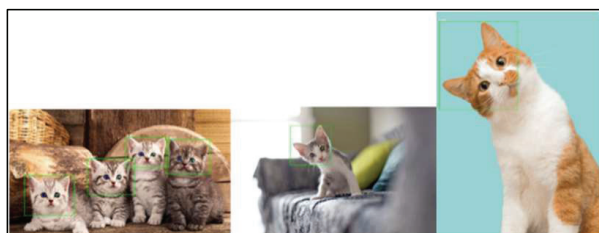


Figure 10 Cat face detection and landmarking

การทดสอบการระบุตัวตน

ในการทดสอบการระบุตัวตน โดยใช้ข้อมูลชุดทดสอบจากงานวิจัยของ Lin และ Kou ซึ่งเป็นภาพใบหน้าของแมว 69 ตัว จำนวน 2,258 ภาพ ประกอบไปด้วยภาพของแมวที่มีทั้งลักษณะที่คล้ายกันและแตกต่างกัน ใบหน้าของแมวถูกนำเข้าในแบบจำลองตรวจจับใบหน้าเพื่อให้ได้เฉพาะส่วนใบหน้าของแมว และปรับตำแหน่งของใบหน้าให้อยู่ในลักษณะหน้าตรง โดยให้ดวงตาดูอยู่ในระนาบเดียวกันเพื่อให้ข้อมูลใบหน้าของแมวอยู่ในลักษณะเดียวกัน ตำแหน่งภาพที่หลุดจากเฟรมใบหน้าแมวจะถูกแทนที่ด้วยสีดำ จากนั้นปรับขนาดภาพมีขนาด 224x224 เพื่อใช้ในการประมวลผลต่อไป (Figure 11) แบบจำลองสำหรับการระบุตัวตนจากใบหน้าได้ถูกฝึกฝนเป็นจำนวน 150 epoch ซึ่งมีค่า Loss จากการฝึกฝนอยู่ที่ 1.86 และการทดสอบอยู่ที่ 2.43 (Figure 12)

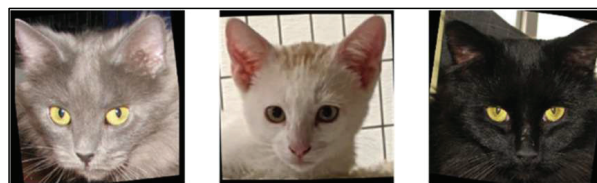


Figure 11 Example of cat face results in a dataset for identification

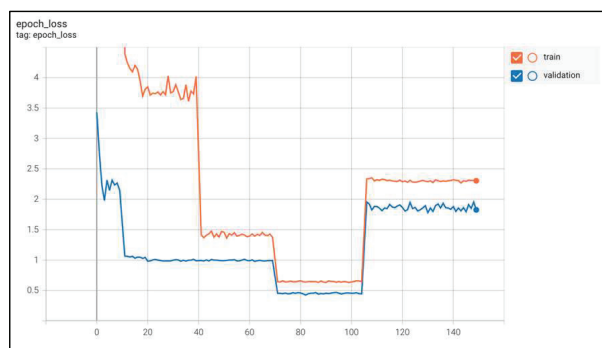


Figure 12 Loss in training dataset for face classification

กระบวนการทดสอบการรู้จำใบหน้า ภาพใบหน้าแมวทั้งหมดจะถูกนำเข้าและประมวลผลด้วย Verification Model และได้ผลลัพธ์ออกเป็นเวกเตอร์ของข้อมูลขนาด 128 มิติ โดยก่อนการประมวลผลภาพใบหน้าแมวที่ต้องการค้นหาจะถูกสุ่มจากชุดข้อมูล และใบหน้าที่เหลือจะถูกนำไปประมวลผลและสร้างฐานข้อมูลด้วย FAISS เพื่อใช้สำหรับการทำนายคลาสที่ต้องการของข้อมูล โดยเมื่อนำมาลดมิติของข้อมูลด้วยเทคนิค UMAP (Uniform Manifold Approximation and Projection) ให้เหลือ 3 มิติและแสดงในรูปแบบกราฟ จะเห็นการกระจายตัวใบหน้าแมวแต่ละตัว (Figure 13)

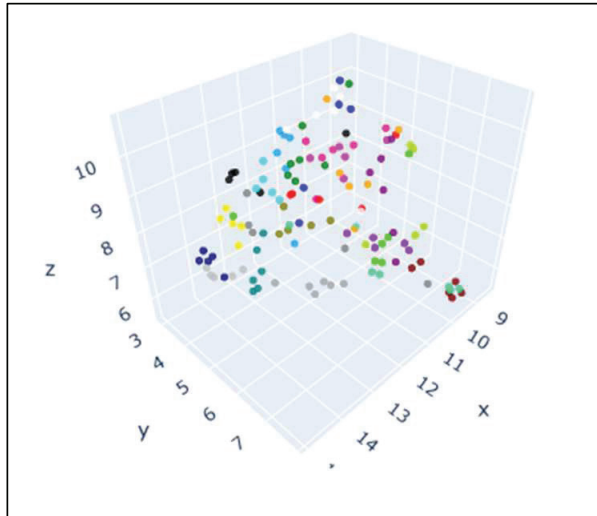


Figure 13 Example distribution of cat face data that has been processed and reduced to 3D

การประเมินผลงานวิจัยชิ้นนี้ใช้วิธีการเดียวกับการบรรณการของมูเกต์และคณะ (Mougeot *et al.*, 2019) โดยการทดสอบในแต่ละรอบ จะสุ่มคลาสของแมวออกมาจำนวน c คลาส และจำนวนรูปของแมวที่ถูกนำมาประมวลผลถูกแทนด้วย N_c โดยให้จำนวน M แทนค่าจำนวนใบหน้าของแมวต้องการค้นหา เช่น $M=2$ จะเป็นการนำใบหน้าของแมวที่ต้องการค้นหา 2 รูป ก็จะรวมเวกเตอร์ของทั้งสองรูปเข้าด้วยกันมาใช้ค้นหาข้อมูลที่เหลือในฐานข้อมูล และ k แทนค่าเพื่อนบ้านใน Nearest Neighbor เป็นจำนวนใบหน้าใกล้เคียงที่ต้องการค้นหา โดยค่า k จะเป็นค่าระหว่าง 1 ถึง 5 เท่านั้นเนื่องจากการตั้งค่า k ให้เป็นค่าที่สูงเกินไปอาจทำให้ประสิทธิภาพของโมเดลลดลงและไม่สามารถจำแนกใบหน้าของแมวได้อย่างมีประสิทธิภาพ นอกจากนี้การจำกัดค่า k อยู่ระหว่าง 1 ถึง 5 ยังช่วยให้การประมวลผลมีความรวดเร็วและแม่นยำมากยิ่งขึ้น ซึ่งในการทดสอบจะทำซ้ำทั้งหมดจำนวน 500 รอบ โดยแต่ละรอบจะสุ่มคลาสจากชุดข้อมูล โดยมีค่า $c=24$ คลาส และวัดค่าความถูกต้องใน 100 ครั้ง ตารางที่ 3 แสดงค่าความแม่นยำเฉลี่ย ค่าสูงสุด ค่าต่ำสุด โดยที่ $M=1$ และ $k=1$ จะเป็นตัวแทนของการเรียนรู้แบบ one-shot และสำหรับค่าอื่นๆ ค่า $k=M+1$ เสมอ โดย Figure 14 แสดงตัวอย่างผลลัพธ์การทำนายด้วยค่า $M=1$ และ $k=5$ และผลลัพธ์ค่า Confusion Matrix จากการประเมิน (Figure 15)

Table 3 Model accuracy when adjusting M and k .

M	k	Identification accuracy		
		Average	Minimum	Maximum
1	1	75.9%	71%	80%
2	3	85.2%	78%	89%
3	4	86.8%	84%	89%
4	5	82.8%	78%	87%

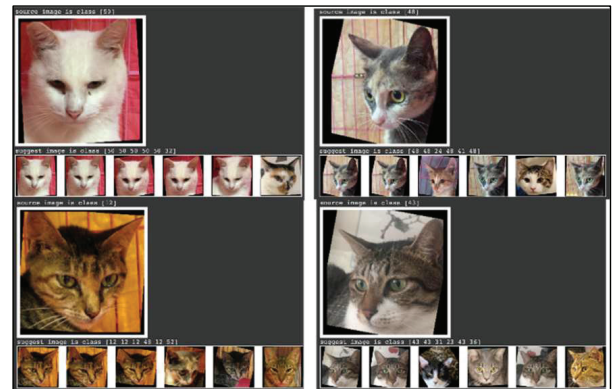


Figure 14 Example of prediction when $M=1$ and $K=5$

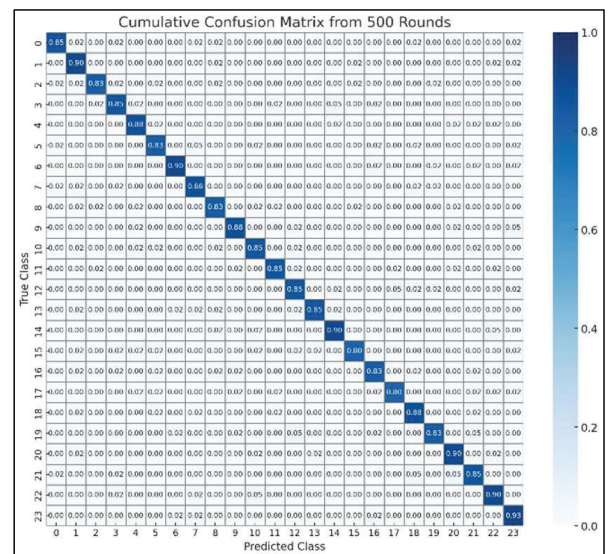


Figure 15 Result of Confusion Matrix size 24x24 of 24 classes of cats

โดยแต่ละแถวของเมทริกซ์แทนคลาสจริง ในขณะที่แต่ละสัณฐานแทนคลาสที่ทำนาย ค่าที่ตำแหน่ง (i, j) ในเมทริกซ์แสดงถึงจำนวนตัวอย่างที่เป็นคลาส i และทำนายว่าเป็นคลาส j หาก $i=1$ และ $j=1$ คือค่าที่ตำแหน่ง $(1,1)$ ในเมทริกซ์ค่านี้แสดงถึงจำนวนของ True Positives (TP) สำหรับคลาสที่ 1 หมายความว่า โมเดลทำนายว่าข้อมูลเป็นคลาสที่ 1 และ

ความจริงคือตัวอย่างเหล่านั้นอยู่ในคลาสที่ 1 เช่นกัน ยิ่งค่าที่ตำแหน่ง (1,1) มีค่าสูงเท่าไรก็แสดงว่าโมเดลทำนายคลาสที่ 1 อย่างแม่นยำมากขึ้น

การอภิปรายผล

จากการทดสอบรู้จำใบหน้าโดยแยกออกเป็นสองส่วน คือกระบวนการตรวจจับใบหน้า และการบวนการระบุตัวตน ผลลัพธ์ของการตรวจจับใบหน้าที่มีค่า mAP อยู่ที่ 0.995 โดยสามารถระบุใบหน้าได้แม้อยู่ในสภาพแวดล้อมที่แตกต่าง และสามารถตรวจจับใบหน้าที่มีหลากหลายขนาดและหลากหลายใบหน้าที่ได้ภายในภาพเดียวกัน

ในส่วนของการกระบวนการระบุตัวตนผลลัพธ์สามารถระบุตัวตนได้ตั้งแต่ขั้น one-shot recognition โดยเมื่อเพิ่มจำนวนภาพที่ใช้สำหรับการค้นก็ยิ่งเพิ่มความแม่นยำมากขึ้น โดยสูงสุดเมื่อ $M=3$ และ $k=4$ จะมีความแม่นยำเฉลี่ย 86.8%

สรุปและงานวิจัยในอนาคต

บทความนี้ได้เสนอแบบจำลองในการรู้จำใบหน้าแมว โดยกระบวนการรู้จำใบหน้าถูกแบ่งออกเป็น 2 ส่วนคือส่วนของแบบจำลองที่ใช้ในการรู้จำใบหน้าและแบบจำลองสกัดคุณลักษณะเพื่อเปรียบเทียบใบหน้าและนำมาระบุตัวตน โดยในส่วนของการตรวจจับใบหน้าที่จะมุ่งเน้นการตรวจหาใบหน้าแมวในภาพที่มีสภาพแวดล้อมที่ตรวจจับได้ยาก โดยต้องสามารถตรวจจับได้หลายใบหน้าที่พร้อมกัน และสามารถที่จะระบุจุดสำคัญบนใบหน้าได้ ได้แก่ ปลายหูซ้าย ปลายหูขวา ตาซ้าย ตาขวา และจมูก เพื่อให้สามารถที่จะนำภาพไปจัดเตรียมและประมวลผลในการตรวจจับได้ใบหน้าได้ในส่วนของกระบวนการระบุตัวตนนั้น แบบจำลองสามารถสกัดคุณลักษณะและระบุตัวตนได้ตั้งแต่ขั้น one-shot recognition โดยเมื่อเพิ่มจำนวนภาพที่ใช้สำหรับการค้นก็ยิ่งเพิ่มความแม่นยำมากขึ้น แบบจำลองที่สร้างขึ้นมีความแม่นยำเฉลี่ยสูงสุด 86.8% การพัฒนาแบบจำลองสำหรับตรวจจับใบหน้าแมวสามารถใช้การถ่ายโอนความรู้ได้โดยประยุกต์ทรัพยากร มีความแม่นยำสูง และยังสามารถประยุกต์ใช้ร่วมกับงานอื่นๆ เช่น การพัฒนาระบบลงทะเบียนสัตว์เลี้ยง การนำไปประยุกต์ใช้ในเมืองอัจฉริยะสำหรับสัตว์เลี้ยง หรือการประยุกต์ใช้งานวิจัยเช่นด้านอื่นๆ เช่นการรู้จำสัตว์ป่า การตรวจจับพฤติกรรมสัตว์ป่า เป็นต้น อย่างไรก็ตามงานวิจัยนี้ยังไม่ได้นำเสนอผลการเปรียบเทียบของการทำ Face alignment และการเปรียบเทียบประสิทธิภาพระหว่างอัลกอริทึมต่างๆ เช่น FAISS, KNN และ SVM ซึ่งอาจเป็นประเด็นสำหรับการศึกษาเพิ่มเติมในอนาคต เพื่อพัฒนาประสิทธิภาพของแบบจำลองให้ดียิ่งขึ้น

เอกสารอ้างอิง

- ณัฐโชติ พรหมฤทธิ์, & สัจจาภรณ์ ไวจรรยา. (2564). *Fundamental of deep learning in practice*. Hytexts Interactive Limited.
- ปริญา สงวนสัตย์. (2562). *Artificial intelligence with machine learning, AI สร้างได้ด้วยแมชชีนเลิร์นนิง*. Infopress.
- Chen, Y. C., Hidayati, S. C., Cheng, W. H., Hu, M. C., & Hua, K. L. (2016). Locality constrained sparse representation for cat recognition. In *MultiMedia Modeling* (pp. 140-151).
- Ethem, A. (2010). *Introduction to machine learning second edition*. The MIT Press Cambridge.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Hermans, A., Beyer, L., & Leibe, B. (2017). *In defense of the triplet loss for person re-identification*. arXiv preprint arXiv:1703.07737.
- Kantar Worldpanel. (2022, December 5). *Thailand pet food market trends*. <https://www.kantarworldpanel.com/th/news/2022-Thailand-Pet-Food-Market-Trends>
- Klein, A. (2019). *Pet cat face verification and identification*. Stanford University, CS230 Fall 2019.
- Kumar, S., & Singh, S. K. (2018). *Monitoring of pet animal in smart cities using animal biometrics*. Indian Institute of Technology.
- Kumar, S., & Singh, S. (2014). Biometric recognition for pet animal. *Journal of Software Engineering and Applications*, 7, 470-482.
- Lai, K., Tu, X., & Yanushkevich, S. (2019). Dog identification using soft biometrics and neural networks. In *International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8).
- Lin, T. Y., & Kuo, Y. F. (2018). Cat face recognition using deep learning. *ASABE Annual International Meeting*.
- Liu, J., Kanazawa, A., Jacobs, D., & Belhumeur, P. (2012). Dog breed classification using part localization. In *Computer Vision – ECCV 2012* (pp. 172-185).
- Montenegro, L., Abreu, M., Fred, A., & Machado, J. M. (2022). Human-assisted vs. deep learning feature extraction: An evaluation of ECG features extraction methods for arrhythmia classification using machine learning. *Applied Sciences*.

- Moreira, T. P., Perea, M. L., Werneck, R. O., & Valle, E. (2017). Where is my puppy? Retrieving lost dogs by facial features. *Multimedia Tools and Applications*, 76, 15325-15340.
- Mougeot, G., Li, D., & Jia, S. (2019). A deep learning approach for dog face verification and recognition. In *PRICAI 2019: Trends in Artificial Intelligence* (pp. 418-430).
- Qi, D., Tan, W., Yao, Q., & Liu, J. (2022). YOLO5Face: Why reinventing a face detector. arXiv preprint arXiv:2105.12931. Retrieved from <https://arxiv.org/abs/2105.12931>
- Redmon, J., Divvala, S. K., Girshick, R. B., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640. Retrieved from <https://arxiv.org/abs/1506.02640>
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- Ruan, Z., Wang, G., Xue, J., Lin, X., & Jiang, Y. (2014). Detection of user-registered dog faces. *Neurocomputing*, 142, 256-266.
- Yoon, B., So, H., & Rhee, J. (2021). A methodology for utilizing vector space to improve the performance of a dog face identification model. *Applied Sciences*, 11, 2074.
- Zhang, X., Yu, F. X., Kumar, S., & Chang, S. F. (2017). Learning spread-out local feature descriptors. In *IEEE International Conference on Computer Vision (ICCV)* (pp. 4605-4613).