

การตรวจสอบการใช้โครงข่ายประสาทเทียมแบบสยามสำหรับระบุจำใบหน้าภาพสเก็ตซ์และภาพถ่าย

Investigating the Use of the Siamese Network for Face Sketch - Photo Recognition

^{1#}สามารถ หมุดและ, ^{2*}เทียนสิริ เหลืองวิลไล, ^{3#}สิริอร วิทยากร, ^{4#}ปรัชญา พลายเวช และ ^{5#}ชัยวัฒน์ บุระณะ

^{1,3} คณะเทคโนโลยีสารสนเทศ, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

²ภาควิชาคณิตศาสตร์, โรงเรียนนายเรืออากาศนวมินทกษัตริยาธิราช

⁴กรมสื่อสารอิเล็กทรอนิกส์ทหารอากาศ

⁵ศูนย์พิสูจน์หลักฐาน 1, สำนักงานตำรวจแห่งชาติ

^{1#} Samart Moodleah, ^{2*} Thiansiri Luangwilai, ^{3#} Sirion Vittayakorn,

^{4#} Prachaya Plaivech and ^{5#} Chaiwat Burana

^{1,3} Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang

² Mathematics Department, Navaminda Kasatriyadhiraj Royal Air Force Academy

⁴ Directorate of Communications and Electronics, Royal Thai Air Force

⁵ Scientific Crime Detection Center 1, Royal Thai Police

[#]samart@it.kmitl.ac.th, ^{*}Thiansiri.L@gmail.com, [#]sirion@it.kmitl.ac.th, [#]prachaya_p@rtaf.mi.th, [#]chaiwat.bm@hotmail.com.au

Received : March 20, 2023

Revised : November 20, 2023

Accepted : November 21, 2023

บทคัดย่อ

ในปัจจุบันมีการเพิ่มขึ้นของจำนวนกล้องวงจรปิดในทุกสถานที่ แต่อย่างไรก็ตามยังมีอยู่หลายเหตุการณ์ที่กล้องวงจรปิดไม่สามารถบันทึกเหตุอาชญากรรมได้ ทำให้เจ้าหน้าที่ต้องสืบสวนสอบสวนจากผู้เห็นเหตุการณ์ คำบอกเล่าจากผู้เห็นเหตุการณ์เหล่านี้จะถูกใช้ในการเป็นหลักฐานและสร้างภาพสเก็ตซ์ใบหน้าสำหรับติดตามผู้กระทำความผิด แต่อย่างไรก็ตามความถูกต้องของภาพสเก็ตซ์หน้าผู้กระทำความผิดนั้นปกติจะมีความถูกต้องที่น้อยอยู่ ขึ้นอยู่กับความจำและความสามารถในการอธิบายลักษณะหน้าตาผู้กระทำความผิดของผู้เห็นเหตุการณ์ ทำให้การบ่งชี้ตัวบุคคลจากภาพสเก็ตซ์ถือเป็นเรื่องท้าทาย อีกทั้งการจะหาภาพถ่ายบุคคลนั้น ในฐานข้อมูลขนาดใหญ่โดยใช้ภาพสเก็ตซ์เป็นภาพเริ่มต้นนั้นแทบจะเป็นไปไม่ได้

ดังนั้นงานวิจัยชิ้นนี้ จึงได้ดำเนินการตรวจสอบการใช้อัลกอริทึมในการจับคู่ภาพสเก็ตซ์กับภาพถ่ายผู้ต้องสงสัย ในฐานข้อมูลขนาดใหญ่ได้ โดยนำเสนอโครงข่ายประสาทเทียมแบบสยามที่มีโครงข่ายแฝดเป็นโครงข่ายคอนโวลูชัน และทดสอบประสิทธิภาพของโครงข่ายที่นำเสนอด้วย ชุดข้อมูลจาก The Chinese University of Hong Kong (CUHK) จากการทดสอบพบว่าอัลกอริทึมที่นำเสนอมีความสามารถในการจับคู่ภาพสเก็ตซ์กับภาพถ่ายผู้ต้องสงสัย ได้ความแม่นยำมากกว่า 98 เปอร์เซ็นต์

คำสำคัญ: ระบุจำใบหน้าจากภาพสเก็ตซ์, โครงข่ายประสาทเทียมแบบสยาม, โครงข่ายคอนโวลูชัน, ภาพสเก็ตซ์ใบหน้าทางนิติวิทยาศาสตร์

Abstract

Nowadays, there is an increase in the number of CCTV everywhere. However, there are still many cases when CCTV cannot capture criminal activities and officers have to rely heavily on the eyewitness report. The eyewitness report is used as a piece of evidence and creates forensic face sketches of the criminals. The accuracy of a forensic face sketch is normally low depending on the memory of the eyewitness and the ability of the eyewitness to explain. Therefore, identifying persons from forensic face sketches is very challenging. Moreover, manually searching for the suspected photo from a large database by using a forensic face sketch as input is rather impossible.

As a result, in this research, we investigate an algorithm to match a forensic face sketch to a suspected photo in a database. The model used here is the Siamese network with a twin convolution neural network. The model's performance is investigated using the dataset from the Chinese University of Hong Kong (CUHK). It was found that the proposed algorithm is able to match the forensic face sketch to the mugshot photo with an accuracy greater than 98 per cent.

Keywords: Face sketch recognition, Siamese network, Convolution network, Forensic face sketch

1. บทนำ

ปัจจุบันเทคโนโลยีการรู้จำใบหน้า (Face Recognition) โดยอาศัยภาพถ่ายหรือวิดีโอ ถือเป็นเทคโนโลยีที่มีอัตราความสำเร็จสูง และถูกนำไปใช้อย่างแพร่หลายในหลากหลายด้าน เช่น การตรวจจับคนร้ายโดยอาศัยภาพจากกล้องวงจรปิด การเข้าถึงข้อมูลการเงินโดยอาศัยการรู้จำใบหน้า รวมทั้งการรู้จำใบหน้าเพื่อปลดล็อกโทรศัพท์มือถือสมาร์ทโฟน อย่างไรก็ตาม การรู้จำใบหน้าโดยอาศัยข้อมูลจากโดเมนที่แตกต่างกัน (Heterogeneous Face Recognition) เช่น การรู้จำใบหน้าคนร้ายจากภาพสเก็ตช์ก็ยังเป็นปัญหาที่มีความท้าทายเป็นอย่างมาก เนื่องจากปัจจัยต่าง ๆ เช่น ภาพสเก็ตช์ที่ความคลาดเคลื่อนจากกระบวนการวาดเอง หรือความคลาดเคลื่อนในการสื่อสารระหว่างพยานและผู้สเก็ตช์ภาพ [1, 2, 3] ซึ่งความคลาดเคลื่อนต่าง ๆ เหล่านี้ ทำให้อัลกอริทึมสำหรับการรู้จำใบหน้าจากภาพถ่ายสเก็ตช์นั้นไม่ประสบความสำเร็จเท่าที่ควร

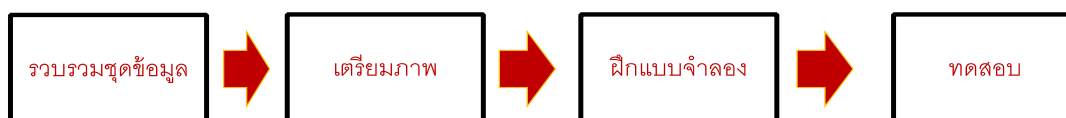
งานวิจัยที่เกี่ยวกับระบบรู้จำใบหน้าจากภาพสเก็ตช์จะมีปัญหาหลัก คือ การที่ต้องดำเนินการวิเคราะห์ภาพที่อยู่ในรูปแบบหรือโดเมนที่แตกต่างกัน [4] ตัวอย่างเช่น ภาพชนิดหนึ่งเป็นภาพถ่ายสี อีกชนิดเป็นภาพลายเส้น หรือเป็นภาพสเก็ตช์ที่สร้างขึ้นมาจากโปรแกรมชิ้นส่วนร่างประกอบ (Composite Face Sketch) การจะฝึกโครงข่ายประสาทเทียมเพื่อให้สามารถจับคู่ภาพสเก็ตช์กับภาพถ่ายได้นั้น จึงมีความซับซ้อนและมีความยุ่งยากกว่าปกติ อีกทั้งในกระบวนการใช้ภาพสเก็ตช์เพื่อค้นหาภาพถ่ายประวัติดอาชญากรในกระบวนการนิติวิทยาศาสตร์เพื่อติดตามผู้กระทำความผิดมาลงโทษก็จะมีอุปสรรคในกระบวนการสร้างภาพเพิ่มเข้ามาอีก เพราะภาพสเก็ตช์ที่สร้างจากคำบอกเล่าของผู้เห็นเหตุการณ์จะมีความผิดพลาดจากความเป็นจริงอยู่มาก [5] ทำให้ความซับซ้อนในกระบวนการฝึกของโครงข่ายประสาทเทียมมีความยากและความถูกต้องลดน้อยลงไปมาก

กระบวนการพัฒนาระบบรู้จำใบหน้าจากภาพสเก็ตช์สามารถแบ่งออกเป็น 2 วิธี [6, 7] โดยวิธีที่หนึ่ง คือการแปลงภาพทั้งสองให้เป็นรูปแบบหรือโดเมนเดียวกัน โดยสามารถแปลงภาพสเก็ตช์เป็นภาพถ่าย หรือแปลงภาพถ่ายเป็นภาพสเก็ตช์ จากนั้นนำภาพทั้งคู่ไปดำเนินการฝึกโครงข่ายประสาทเทียมเพื่อให้มีขีดความสามารถในการจับคู่ภาพสเก็ตช์และภาพถ่ายจริงได้ในขั้นตอนกระบวนการแปลงรูปแบบภาพนั้นสามารถกระทำได้หลากหลายวิธี [4, 8, 9] ในกรณีวิธีที่สองคือ การนำภาพสเก็ตช์และภาพถ่ายดั้งเดิมมาใช้ในการฝึกแบบโครงข่ายประสาทเทียมโดยตรงเลย ให้สามารถพยากรณ์หรือจับคู่ภาพได้ถูกต้อง

กระบวนการวิธีนี้แม้ว่าจะมีความยากในการฝึกโครงข่ายประสาทเทียมในเรื่องความแตกต่างของชุดข้อมูลทั้งสองชุด แต่มีข้อดีคือจะไม่เพิ่มตัวแปรความคลาดเคลื่อน ในขั้นตอนของการแปลงภาพอย่างในวิธีแรก [10, 11, 12]

ในงานวิจัยชิ้นนี้จะเป็นนำเสนอการประยุกต์ใช้งาน โครงสร้างโครงข่ายประสาทเทียมแบบสยาม จากนั้นจะดำเนินการวิเคราะห์ศักยภาพของโครงข่ายในการจับคู่ภาพเหมือนระหว่างภาพถ่ายกับภาพสเก็ชลายเส้น ว่าความสามารถมีศักยภาพ และมีความเป็นไปได้ในการไปประยุกต์ใช้ทางนิติวิทยาศาสตร์เพื่อสามารถค้นกรองชุดภาพจากฐานข้อมูลขนาดใหญ่ให้มีจำนวนน้อยลงเพื่อให้เจ้าหน้าที่สามารถดำเนินการตรวจสอบต่อไปได้หรือไม่

2. วิธีดำเนินการวิจัย



รูปที่ 1 ขั้นตอนการใช้โครงข่ายประสาทเทียมแบบสยามในการพัฒนาระบบรู้จำใบหน้าที่จากภาพถ่ายและภาพสเก็ช

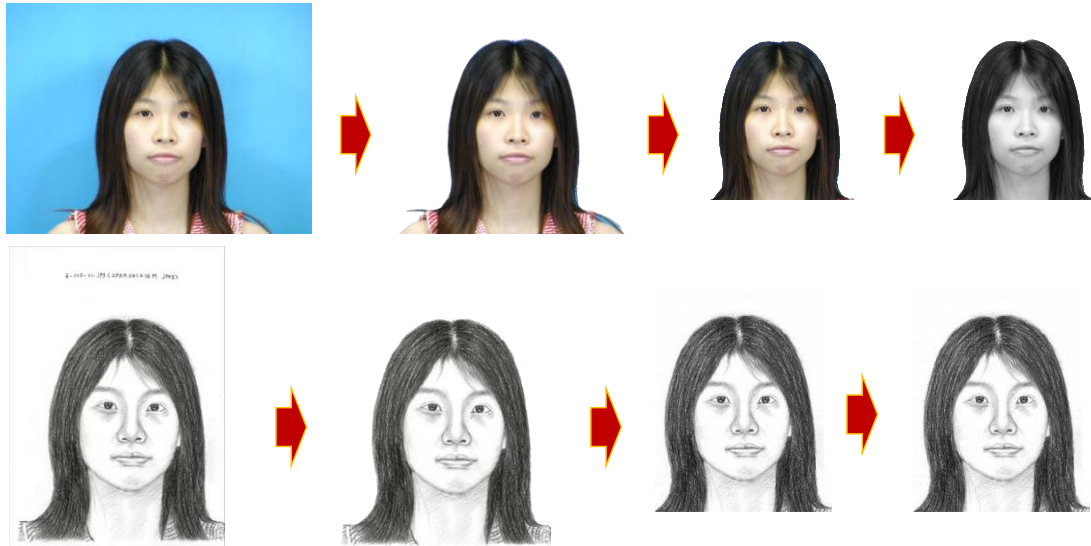
จากรูปที่ 1 แสดงขั้นตอนการฝึกและประยุกต์ใช้แบบจำลองโครงข่ายประสาทเทียมแบบสยามสำหรับระบบรู้จำใบหน้าที่ภาพถ่ายและภาพสเก็ช โดยในงานวิจัยชิ้นนี้เป็นกระบวนการออกแบบและทดสอบประสิทธิภาพพื้นฐานของโครงข่ายประสาทเทียมแบบสยามก่อน ดังนั้นขั้นตอนแรกจะเริ่มต้นด้วยชุดข้อมูลภาพที่ศิลปินวาดให้มีความใกล้เคียงกับต้นฉบับก่อน โดยใช้ชุดข้อมูลภาพ CUHK Face Sketch Database (CUFS) ซึ่งเป็นชุดข้อมูลภาพคู่ระหว่างภาพถ่ายหน้าตรงและภาพสเก็ชลายเส้นของ Chinese University of Hong Kong (CUHK) [13, 14] (แสดงตัวอย่างภาพถ่ายหน้าตรงและภาพสเก็ช โดยใช้ศิลปินวาดลายเส้นในชุดข้อมูล ดังรูปที่ 2) ซึ่งจุดประสงค์ของชุดข้อมูลนี้คือ การสร้างฐานข้อมูลเพื่อฝึกแบบจำลองโครงข่ายประสาทเทียมให้มีขีดความสามารถในการสังเคราะห์สร้างภาพสเก็ชให้เหมือนที่สุดจากภาพถ่ายจริง โดยงานวิจัยนี้เป็นการประยุกต์ใช้ในระบบรู้จำใบหน้าที่จากภาพสเก็ชแทน



รูปที่ 2 ภาพแสดงตัวอย่างภาพถ่ายหน้าตรงและภาพสเก็ช โดยใช้ศิลปินวาดลายเส้นในชุดข้อมูล CUFS [13]

ขั้นตอนที่สองเป็นกระบวนการเตรียมภาพเพื่อให้เหมาะแก่การนำไปใช้กับโครงข่ายประสาทเทียมต่อไป โดยชุดข้อมูลภาพ CUFS นั้นจะถูกแบ่งออกเป็น 100 คู่สำหรับการฝึกโครงข่ายประสาทเทียมและอีก 88 คู่ จะใช้ในการตรวจสอบความถูกต้องและวิเคราะห์ประสิทธิภาพ ซึ่งแต่ละคู่นั้นประกอบไปด้วยภาพสเก็ชลายเส้นหนึ่งภาพและภาพถ่ายหน้าตรงอีกหนึ่งภาพ โดยภาพทั้งคู่นั้นจะถูกกำหนดด้วยหมายเลขเดียวกัน และจะถูกนำมาเข้ากระบวนการเตรียมภาพดังนี้ 1) ดำเนินการตัดพื้นหลัง

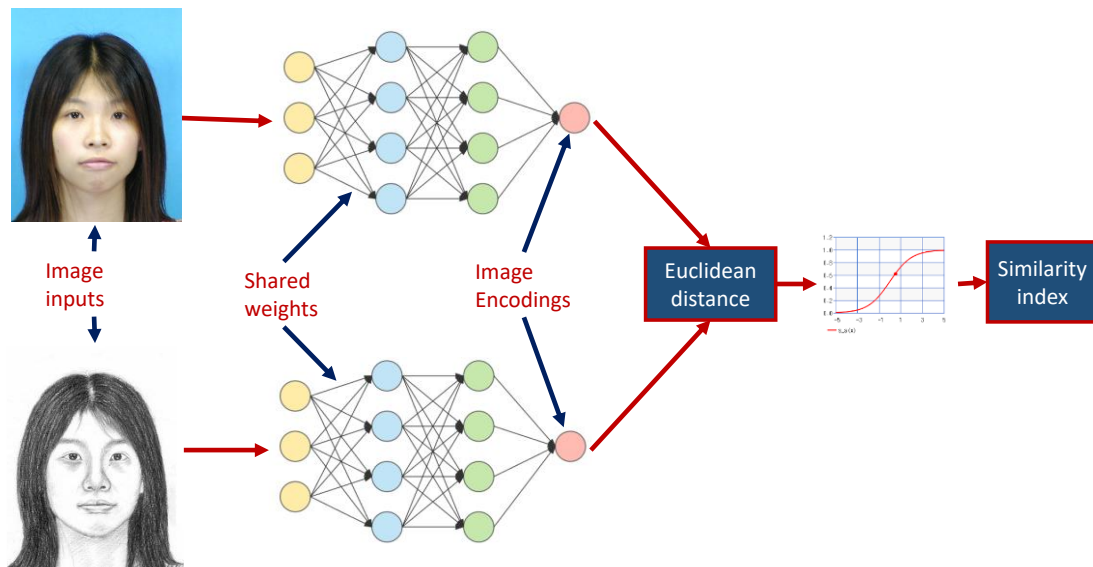
ของภาพออกทั้งหมด 2) เลือกตัดเก็บไว้เฉพาะส่วนหน้า และ 3) แปลงภาพเป็นภาพสีโทนเทา (Greyscale Image) ขนาด 200x250 พิกเซล ดังแสดงตัวอย่างในรูปที่ 3



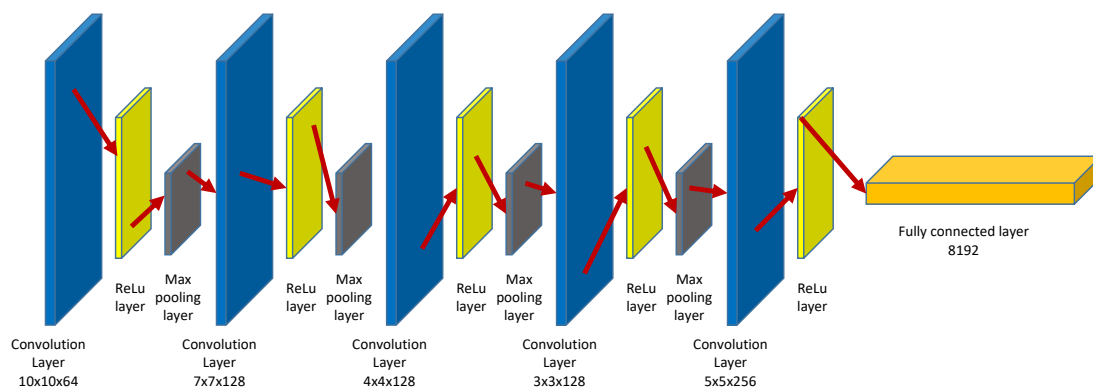
รูปที่ 3 ภาพแสดงตัวอย่างการเตรียมชุดข้อมูลภาพถ่ายและภาพสเก็ทซ์เพื่อจะใช้ในการฝึกโครงข่ายประสาทเทียมแบบสยาม

ขั้นตอนที่สามคือ การดำเนินการฝึกโครงข่ายประสาทเทียมแบบสยาม ซึ่งโครงข่ายชนิดนี้เป็นโครงข่ายแบบแผ่ที่สามารถรับชุดข้อมูลได้สองชุดพร้อมกันได้ ซึ่งโครงข่ายทั้งสองจะใช้น้ำหนักเท่ากัน ในขณะที่ทำงานควบคู่ไปกับเวกเตอร์อินพุตที่แตกต่างกันสองแบบเพื่อคำนวณเวกเตอร์เอาต์พุตที่เปรียบเทียบกันได้ โดยจากพฤติกรรมการทำงานแบบแผ่นี้ ทำให้สามารถแก้ปัญหาเรื่องความแตกต่างของชุดข้อมูลทั้งสองชุดที่ต้องการจะเปรียบเทียบกันได้ อีกทั้งโครงข่ายชนิดนี้ยังได้มีการนำไปใช้ในงานวิจัยและทดสอบในกับรูปภาพมาแล้วหลายงาน เช่น การจับคู่ภาพสิ่งของกับภาพสเก็ทซ์แบบหยาบ [10] การจับคู่ภาพหน้ากับภาพสเก็ทซ์โดยใช้ข้อมูลจากกราฟอิสโตแกรมของภาพ [11] การจับคู่ภาพหน้ากับภาพสเก็ทซ์โดยใช้ข้อมูลจากกราฟของส่วนประกอบของหน้า [12]

ในงานวิจัยนี้ เลือกทดสอบการใช้โครงข่ายประสาทเทียมแบบสยาม ที่ประกอบไปด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) สองข้าง โดยโครงข่ายแผ่ด้านนี้สามารถรับภาพนำเข้า (Input Image) สำหรับการฝึกได้พร้อมกันสองภาพและนำผลที่ได้มาเปรียบเทียบกันเพื่อหาค่าความเหมือน (Similarity Index) ดังตัวอย่างในรูปที่ 4 และแผนภาพแสดงโครงสร้างของประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network) ที่นำเสนอในงานวิจัยนี้ แสดงในรูปที่ 5



รูปที่ 4 ภาพแสดงโครงสร้างของโครงข่ายประสาทเทียมแบบสยาม (Siamese Neural Network) ของงานวิจัยนี้



รูปที่ 5 ภาพแสดงโครงสร้างของโครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolutional Neural Network)

ที่ใช้เป็นโครงข่ายแฝดในโครงข่ายประสาทเทียมแบบสยาม

เนื่องจากงานวิจัยนี้ต้องการทดสอบความเป็นไปได้และประสิทธิภาพในการใช้โครงข่ายประสาทเทียมแบบสยามที่มีโครงข่ายคอนโวลูชันอยู่ภายในสองชุด ดังนั้นงานวิจัยนี้จึงเลือกรูปแบบมาตรฐานของโครงข่ายคอนโวลูชันจำนวน 5 ชั้น ซึ่งแสดงด้วยแผ่นสีน้ำเงิน และระหว่างนั้นจะมีชั้นของ Rectifier Layer (ReLU Layer) และชั้นของ Max Pooling Layer ซึ่งแสดงด้วยแผ่นสีเหลืองและสีเทาตามลำดับ กันอยู่ระหว่างกลางของแต่ละชั้นของชั้นคอนโวลูชัน โดยในชั้นสุดท้ายจะเป็นชั้นของ Fully Connected Layer ซึ่งแสดงด้วยแผ่นสีส้ม

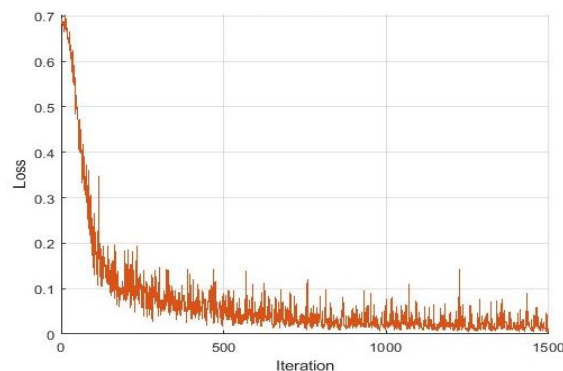
ในขั้นตอนการฝึกนั้น ในแต่ละรอบของการฝึกนั้นจะมีการกำหนดเลือกภาพมาจำนวน 40 ภาพ จากนั้นอัลกอริทึมจะดำเนินการสุ่มตัวเลขระหว่าง 0 และ 1 หากได้ค่ามากกว่า 0.5 จะเลือกภาพถ่ายหน้าตรงและภาพสเก็ชซ์ของบุคคลคนกันเดียวมา ในทางตรงกันข้ามหากได้ค่าน้อยกว่า 0.5 จะดำเนินการสุ่มภาพถ่ายหน้าตรงและภาพสเก็ชซ์ของบุคคลคนละคนกัน จากนั้นจะดำเนินการส่งผ่านภาพแต่ละคู่เข้าสู่โครงข่ายประสาทเทียมแบบสยาม เมื่อเสร็จหนึ่งคู่จะดำเนินการทำซ้ำจนครบทั้งหมดทุกภาพ

ถือว่าครบหนึ่งรอบของการฝึก จากนั้นจะดำเนินการคำนวณค่าฟังก์ชันเป้าหมาย (Loss Function) เฉลี่ยรวมที่ได้จากทุกคู่ภาพ จากนั้นจะดำเนินการปรับปรุงค่าของตัวแปรต่าง ๆ ให้ค่าฟังก์ชันเป้าหมายมีค่าความคลาดเคลื่อนลดน้อยลง

ในขั้นตอนสุดท้ายนั้นคือ การทดสอบ โดยการนำชุดภาพที่เหลืออีก 88 ภาพมาใช้ในการทดสอบ โดยจะสุ่มภาพมารอบละ 50 คู่ ซึ่งใน 50 คู่นี้ จะมีทั้งคู่ที่เป็นภาพของบุคคลเดียวกันและคู่ที่เป็นภาพของบุคคลคนละคนกัน จากนั้นจะนำภาพแต่ละคู่เข้าสู่โครงข่ายประสาทเทียมแบบสยาม และตรวจสอบว่าโครงข่ายประสาทเทียมสามารถพยากรณ์ภาพคู่เหมือนและคู่ที่ต่างกันได้เท่าใด และดำเนินการทำซ้ำจนครบ 10 รอบ เพื่อทำการคำนวณอัตราความถูกต้องของการพยากรณ์ว่ามี ความถูกต้องทั้งหมดเป็นกี่เปอร์เซ็นต์จากการทำนายทุกคู่ภาพ ซึ่งผลที่ได้นั้นจะเป็นการชี้วัดประสิทธิภาพของ โครงข่ายการเรียนรู้เชิงลึกที่ออกแบบและพัฒนาในงานวิจัยนี้

3. ผลการวิจัย

ขั้นตอนการดำเนินการฝึกโครงข่ายประสาทเทียมนั้นได้ดำเนินการตามที่อธิบายไว้ดังกล่าวข้างต้น โดยผลลัพธ์ของกราฟ แสดงค่าฟังก์ชันเป้าหมาย (Loss Function) ที่ได้ในแต่ละรอบของการฝึกแบบจำลองแสดงในรูปที่ 6 โดยในกราฟนี้เป็นการฝึก โครงข่ายประสาทเทียมจำนวน 1500 รอบ ซึ่งค่าของฟังก์ชันเป้าหมาย (Loss Function) ได้มีการลดต่ำลงอย่างรวดเร็วจาก 0.7 จนต่ำกว่า 0.1 ภายใน 200 รอบของการฝึก จนมีค่าใกล้กับ 0 ที่การฝึก 1500 รอบ แสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบสยาม ตอบสนองได้ดีต่อชุดข้อมูลภาพ CUFS



รูปที่ 6 กราฟแสดงค่าฟังก์ชันเป้าหมาย (Loss Function) ที่ได้ในแต่ละรอบของการฝึกโครงข่ายประสาทสยาม (Siamese Neural Network)

เมื่อโครงข่ายประสาทเทียมแบบสยามได้ถูกฝึกเป็นที่เรียบร้อยแล้ว โครงข่ายประสาทเทียมนี้จะถูกทดสอบ ประสิทธิภาพในการพยากรณ์ และคำนวณหาค่าดัชนีความคล้ายของคู่ภาพสเก็ชกับภาพถ่ายหน้าตรง ซึ่งจากกระบวนการ ทดสอบโครงข่าย (Model Validation) พบว่า โครงข่ายประสาทเทียมแบบสยามที่ถูกพัฒนาขึ้นในงานวิจัยนี้ มีค่าความถูกต้อง (Accuracy) ในการพยากรณ์หรือทำนายว่าภาพสเก็ชกับภาพถ่ายหน้าตรงเป็นบุคคลเดียวกันหรือไม่ มีความถูกต้อง ถึง 98.6 เปอร์เซ็นต์ แสดงให้เห็นขีดความสามารถของโครงข่ายประสาทเทียมแบบสยามนี้ว่าเหมาะสมกับชุดภาพสเก็ช ลายเส้นชนิดนี้ แต่อย่างไรก็ตามยังมีความเป็นไปได้ในการต่อยอดกับภาพสเก็ชชนิดอื่น ๆ อีก

รูปที่ 7 แสดงตัวอย่างการคำนวณหาดัชนีความคล้ายของภาพจำนวน 10 คู่ที่สุ่มขึ้นมา เพื่อทดสอบโครงข่ายประสาทเทียม แบบสยามที่ผ่านการฝึกมาแล้ว ตัวอย่างในภาพนี้จะเห็นว่าภาพคู่ใดที่เป็นคู่ภาพของบุคคลเดียวกันทั้งภาพสเก็ชและภาพถ่ายหน้าตรง ค่าคะแนนดัชนีความคล้ายจะมีค่าสูงและใกล้เคียงกับหนึ่งมาก ซึ่งโครงข่ายประสาทเทียมแบบสยามก็จะสามารถพยากรณ์ได้ ว่าเป็นภาพบุคคลเดียวกันได้อย่างแม่นยำ ในทางตรงกันข้ามหากคู่ภาพของภาพสเก็ชและภาพถ่ายหน้าตรงมาจาก บุคคลคนละคนกันแล้ว ค่าดัชนีความคล้ายจะมีค่าต่ำและใกล้เคียงศูนย์ ทำให้โครงข่ายประสาทเทียมแบบสยามก็จะทำนาย

ออกมาว่าเป็นภาพของบุคคลคนละคนกัน ได้อย่างถูกต้อง โดยขอบเขตค่าดัชนีความคล้ายที่กำหนดไว้ในกระบวนการฝึก และการพยากรณ์ว่าเป็นบุคคลเดียวกันคือ ค่าดัชนีความคล้ายที่มากกว่า 0.5



รูปที่ 7 ตัวอย่างการคำนวณค่าดัชนีความคล้ายของภาพที่สุ่มมาเพื่อใช้ในการพยากรณ์คู่ภาพเหมือนระหว่างภาพถ่ายจริงและภาพสเก็ตซ์

นอกจากนี้แล้ว จากการปรับแต่งไฮเปอร์พารามิเตอร์พื้นฐาน พบว่า

- 1) การเพิ่มจำนวนชั้นของโครงข่ายคอนโวลูชันมากกว่า 5 ชั้น ทำให้มีประสิทธิภาพเพิ่มขึ้นเพียงเล็กน้อย แต่การฝึกและการพยากรณ์ต้องใช้เวลาอย่างมากขึ้นอย่างสูง อีกทั้งโครงข่ายปัจจุบันสามารถพยากรณ์ได้ถูกต้องถึง 98.6 เปอร์เซ็นต์ แต่ในทางตรงกันข้าม หากลดจำนวนชั้นของโครงข่ายคอนโวลูชันน้อยกว่า 5 ชั้น พบว่า ประสิทธิภาพลดลงมาก
- 2) ขนาดของรูปอินพุตที่ใหญ่ขึ้น ทำให้ต้องลดจำนวนภาพในแต่ละรอบการฝึกลง ทำให้ได้ประสิทธิภาพลดลงไปด้วย
- 3) โครงสร้างโครงข่ายปัจจุบันสามารถให้ค่าความถูกต้องที่มากกว่า 90 เปอร์เซ็นต์ จากการฝึกเพียง 500 รอบ

4. แนวทางการนำผลงานวิจัยไปใช้ประโยชน์

งานวิจัยนี้ได้แสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบสยัมที่ประกอบไปด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชัน 4 ชั้น มีขีดความสามารถในการพยากรณ์คู่เหมือนระหว่างภาพถ่ายจริงและภาพสเก็ตซ์ ซึ่งจากขีดความสามารถนี้สามารถนำไปประยุกต์ในกระบวนการนิติวิทยาศาสตร์ เพื่อสืบค้นหาภาพถ่ายประวัติอาชญากรจากฐานข้อมูลขนาดใหญ่ได้จากการใช้ภาพสเก็ตซ์จากคำบอกเล่าเป็นภาพเริ่มต้นในกระบวนการสืบค้น

โดยในหัวข้อนี้จะแสดงตัวอย่างแนวทางในการนำไปใช้ของโครงข่ายประสาทเทียมแบบสยัมในกระบวนการสืบค้นหาภาพถ่ายผู้ต้องสงสัยในฐานข้อมูล โดยจะใช้ภาพสเก็ตซ์ลายเส้นเป็นข้อมูลตั้งต้นในการค้นหา ซึ่งในตัวอย่างนี้ได้เลือกภาพสเก็ตซ์หมายเลขที่ 8 ดังแสดงในรูปที่ 8 (ก) จากนั้นภาพนี้จะถูกสมมุติเป็นภาพสเก็ตซ์ของคนบุคคลผู้ต้องสงสัยที่ได้จากคำบอกเล่าของผู้เห็นเหตุการณ์ ขั้นตอนต่อไปภาพนี้จะถูกส่งเข้าไปในโครงข่ายประสาทเทียมแบบสยัมเพื่อเปรียบเทียบกับภาพถ่ายหน้าตรงภาพอื่น ๆ ในฐานข้อมูล โดยผลลัพธ์ที่ได้ในแต่ละการเปรียบเทียบ เป็นค่าดัชนีความคล้ายซึ่งแสดงอยู่ในรูปที่ 9 จากนั้นเมื่อได้ดัชนีความคล้ายจากการเปรียบเทียบภาพสเก็ตซ์กับภาพถ่ายหน้าตรงในฐานข้อมูลมาครบแล้ว จะดำเนินการเรียงลำดับค่าดัชนีจากน้อยไปมาก เพื่อหาภาพถ่ายหน้าตรงที่มีความคล้ายหรือมีคุณลักษณะเหมือนกับภาพสเก็ตซ์ที่เป็นภาพตั้งต้นมากที่สุด

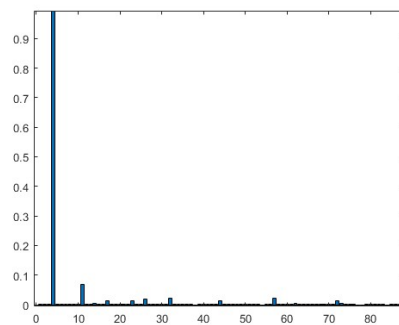


(ก)

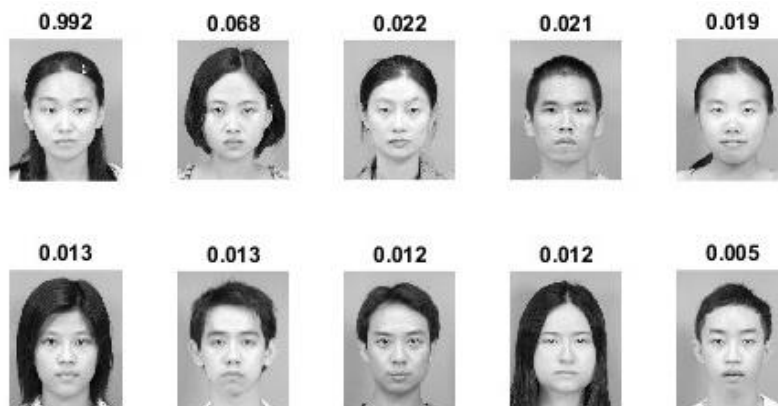


(ข)

รูปที่ 8 ตัวอย่างภาพสเก็ตช์ (ก) และภาพถ่ายหน้าตรงบุคคล (ข) หมายเลข 8 ในชุดข้อมูล CUFS [13]



รูปที่ 9 ค่าดัชนีความคล้ายของภาพสเก็ตช์ในรูปที่ 8 (ก) เทียบกับภาพถ่ายหน้าตรงในฐานข้อมูลภาพที่ใช้ทดสอบ



รูปที่ 10 แสดงภาพหน้าตรงที่มีดัชนีความคล้ายกับภาพสเก็ตช์ในภาพที่ 8 (ก) จำนวน 10 อันดับแรก

ในรูปที่ 10 แสดงภาพถ่ายหน้าตรง 10 อันดับแรกที่มีดัชนีความคล้ายมากที่สุด เมื่อเปรียบเทียบกับภาพสเก็ตช์หมายเลข 8 ในภาพที่ 8 (ก) จากผลที่ได้ในภาพนี้จะเห็นได้ว่า โครงข่ายประสาทเทียมแบบสยามมีประสิทธิภาพในการค้นหาภาพถ่ายหน้าตรงจากภาพสเก็ตช์อย่างสูง ดังเห็นได้จากภาพของบุคคลคนเดียวกัน จะมีค่าดัชนีความคล้ายอยู่ในระดับที่สูงมาก อยู่ที่ 0.992 ซึ่งเป็นภาพของบุคคลคนเดียวกันและแสดงอยู่ในรูปที่ 8(ข) ในขณะที่อันดับสองมีดัชนีความคล้ายอยู่เพียง 0.068 และอันดับที่ 10 จะมีค่าดัชนีความคล้ายน้อยลงไปอีกมาก อยู่ที่ 0.005 ซึ่งค่าดัชนีความคล้ายของภาพคู่เหมือนและภาพคู่ต่างนั้น มีความแตกต่างกันอย่างชัดเจน ซึ่งแสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบสยามนี้ สามารถแบ่งแยกภาพถ่ายบุคคล และภาพสเก็ตช์ได้อย่างมีประสิทธิภาพ

5. สรุปผลการวิจัย

งานวิจัยชิ้นนี้ได้แสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบสยวมที่ประกอบไปด้วยโครงข่ายประสาทเทียมแบบคอนโวลูชันแฝดนั้น มีศักยภาพเพียงพอในการพยากรณ์ความเหมือนระหว่างภาพจริงและภาพสเก็ตช์ โดยในงานวิจัยนี้ได้แสดงให้เห็นถึงความเป็นไปได้ในการนำไปใช้ในกระบวนการนิติวิทยาศาสตร์เพื่อสืบค้นหาภาพถ่ายประวัติอาชญากรจากฐานข้อมูลขนาดใหญ่ โดยใช้ภาพสเก็ตช์จากคำบอกเล่าของผู้เห็นเหตุการณ์เป็นภาพเริ่มต้นในกระบวนการสืบค้น

โดยแนวทางในการพัฒนาต่อยอดงานวิจัยชิ้นนี้ คือการประยุกต์ใช้โครงข่ายประสาทเทียมแบบสยวมนี้กับภาพสเก็ตช์ในรูปแบบอื่น อย่างเช่นชุดข้อมูลภาพสเก็ตช์ที่สร้างขึ้นจากซอฟต์แวร์ ที่มีวัตถุประสงค์เพื่อลดเวลาทำงานของศิลปินและได้ภาพจำนวนมาก โดยจะเลือกชิ้นส่วนต่าง ๆ ของหน้ามาประกอบขึ้นเป็นหน้าบุคคล โดย Galea & Farrugia [15] ได้อธิบายและรวบรวมชุดข้อมูลภาพสเก็ตช์ชนิดนี้ไว้ในงานวิจัยของเขา และในกรณีสุดท้าย คือการต่อยอดไปใช้กับภาพสเก็ตช์บุคคลต้องสงสัยที่สร้างจากกระบวนการเฉพาะทางของสำนักงานตำรวจแห่งชาติ เพื่อให้อัลกอริทึมและโครงข่ายประสาทเทียมนี้มีศักยภาพที่สูงขึ้นและสามารถนำไปใช้ได้จริงในทางนิติวิทยาศาสตร์

6. เอกสารอ้างอิง

- [1] Klare, B., Li, Z., & Jain, A. K. (2010). Matching forensic sketches to mug-shot photos. *IEEE transactions on pattern analysis and machine intelligence*. 33(3): 639-646.
- [2] Klare, B. F., & Jain, A. K. (2012). Heterogeneous face recognition using kernel prototype similarities. *IEEE transactions on pattern analysis and machine intelligence*. 35(6): 1410-1422.
- [3] Ouyang, S., Hospedales, T. M., Song, Y. Z., & Li, X. (2016). Forgetmenot: Memory-aware forensic facial sketch matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp.5571-5579).
- [4] Ounachad, K., Oualla, M., Souhar, A., & Sadiq, A. (2020, March). Face sketch recognition-an overview. In *Proceedings of the 3rd International Conference on Networking, Information Systems & Security* (pp.1-8).
- [5] Bushra, S. N., & Maheswari, K. U. (2021). Crime Investigation using DCGAN by Forensic Sketch-to-Face Transformation (STF)-A Review. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1343-1348). IEEE.
- [6] Lahlali, S. E., Sadiq, A., & Mbarki, S. (2015). A REVIEW OF FACE SKETCH RECOGNITION SYSTEMS. *Journal of Theoretical & Applied Information Technology*. 81(2).
- [7] Ashwini, B., & Prajakta, K. (2016). A survey of face recognition from sketches. *International Journal of Latest Trends in Engineering and Technology*. 6(3): 150-158.
- [8] Liu, Q., Tang, X., Jin, H., Lu, H., & Ma, S. (2005). A nonlinear approach for face sketch synthesis and recognition. In *2005 IEEE Computer Society conference on computer vision and pattern recognition*. 5 (1): 1005-1010). IEEE.
- [9] Wang, N., Tao, D., Gao, X., Li, X., & Li, J. (2014). A comprehensive survey to face hallucination. *International journal of computer vision*. 106(1): 9-30.
- [10] Qi, Y., Song, Y. Z., Zhang, H., & Liu, J. (2016). Sketch-based image retrieval via siamese convolutional neural network. In *2016 IEEE international conference on image processing (ICIP)* (pp. 2460-2464). IEEE.
- [11] Fan, L., Liu, H., & Hou, Y. (2019). An improved siamese network for face sketch recognition. In *2019 International Conference on Machine Learning and Cybernetics (ICMLC)* (pp. 1-7). IEEE.

- [12] Fan, L., Sun, X., & Rosin, P. L. (2021). Siamese graph convolution network for face sketch recognition: an application using graph structure for face photo-sketch recognition. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 8008-8014). IEEE.
- [13] Wang, X., & Tang, X. (2008). Face photo-sketch synthesis and recognition. *IEEE transactions on pattern analysis and machine intelligence*. 31(11): 1955-1967.
- [14] Zhang, W., Wang, X., & Tang, X. (2011). Coupled information-theoretic encoding for face photo-sketch recognition. In CVPR 2011 (pp. 513-520). IEEE.
- [15] Galea, C., & Farrugia, R. A. (2016). A large-scale software-generated face composite sketch database. In 2016 International Conference of the Biometrics Special Interest Group (BIOSIG) (pp. 1-5). IEEE.