**การจำแนกประเภทเรือจากภาพด้วยวิธีถ่ายโอนเรียนรู้**

**Ship Classification from Photos based on Transfer Learning**

|  |  |
| --- | --- |
| เรือโทหญิง จิตรา เรืองรุ่งโสม\* | รศ.ดร. อรฉัตร จิตต์โสภักตร์\*\* |
| Lt.JG Chittra Roungroongsom, WRTN | Assoc.Prof.Dr. Orachat Chitsobhuk |

**บทคัดย่อ**

การระบุประเภทเรือในน่านน้ำ ถือเป็นงานแขนงหนึ่งที่มีบทบาทในงานด้านการตรวจการณ์ทางทะเล โดยการระบุเรือจากข้อมูลที่เป็นภาพถ่าย เป็นอีกแหล่งข้อมูลหนึ่งที่ถูกรวบรวมได้ง่าย งานวิจัยนี้ จึงนำเสนอการจำแนกประเภทเรือจากภาพถ่ายสีโดยใช้การเรียนรู้เชิงลึกจากไลบรารี่แบบเปิด โดยจะทำการเปรียบเทียบการจำแนกจากโครงสร้างแบบจำลองการเรียนรู้เชิงลึกที่เป็นที่นิยม 5 แบบ คือ แบบจำลอง CNN อย่างง่ายที่ทำการฝึกข้อมูลตั้งแต่เริ่มต้น และแบบจำลองที่ได้รับการฝึกมาแล้วนำมาทำ Transfer Learning 4 แบบ ได้แก่ VGG16, InceptionV3, ResNet50 และ MobileNet บนชุดข้อมูลภาพถ่ายเรือแบบเปิด ประกอบด้วยข้อมูลสำหรับฝึกจำนวน 2,500 รูป และสำหรับทดสอบจำนวน 500 รูป จากเรือทั้งหมด 5 ประเภท ผลการทดสอบพบว่า แบบจำลอง MobileNet ด้วยวิธี Transfer Learning ให้ผลลัพธ์ที่ดีที่สุด โดยมีค่าความแม่นยำอยู่ที่ 89.2%

**Abstract**

Specifying ship types in waters plays a role in the field of maritime domain awareness, especially to classify ship types from photos, which is data sources available in common. In this paper, we present the comparison of efficiency of ship classification from photos using deep learning. The system is implemented by using open source library. We compare the effectiveness of the classification among 5 outstanding models, including, Simple CNN trained from scratch, and four models that apply transfer learning techniques: VGG16, InceptionV3, ResNet50, and MobileNet. The experiments were performed on an open ship photo datasets which composed of 3,000 images with 5 ship type labeled. The results show that MobileNet outperforms the rest by achieving an accuracy of 89.2%.

|  |  |
| --- | --- |
| \* นักศึกษาปริญญาโท | คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง |
| \*\* อาจารย์ | ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ |

1. บทนำ

ในปัจจุบัน ระบบอัตโนมัติได้เข้ามามีบทบาทในการทำงานด้านต่างๆ เป็นอย่างมาก ไม่ว่าจะเป็นเรื่องเล็กๆ ในชีวิตประจำวัน เช่น การใช้เซนเซอร์ควบคุมการปิดเปิดประตูโดยอัตโนมัติ หรืองานในระดับใหญ่ขึ้นมาอย่างการผลิตสินค้าด้วยเครื่องจักรอัตโนมัติในแวดวงอุตสาหกรรม เป็นต้น ระบบอัตโนมัติยังเป็นระบบการทำงานที่มีแนวโน้มที่จะพัฒนาและเติบโตเพิ่มมากขึ้นเรื่อยๆ มีหน้าที่หลักเพื่ออำนวยความสะดวกสบายในการดำเนินชีวิตแก่มนุษย์ มีการทำงานที่สะดวก สามารถควบคุมและช่วยลดต้นทุนการทำงานโดยแรงงานมนุษย์

องค์ประกอบหนึ่งที่สำคัญสำหรับการทำงานของระบบอัตโนมัติคือความสามารถในการจำแนกข้อมูลที่น่าเชื่อถือ ซึ่งจะช่วยให้ระบบสามารถเข้าใจสภาพแวดล้อมในขณะนั้นและสามารถตัดสินใจดำเนินการอย่างเป็นอัตโนมัติเพื่อให้บรรลุผลสำเร็จตามเป้าหมายต่อไป

ศักยภาพของระบบอัตโนมัติดังกล่าว ได้ถูกนำไปใช้ในงานด้านการตรวจการณ์ทางทะเล เช่น การเฝ้าระวังเรือที่รุกล้ำอาณาเขตน่านน้ำที่มีการจำกัดประเภทเรือผ่านเข้าออก การวิเคราะห์พฤติกรรมเรือที่มีแนวโน้มการกระทำผิดกฎหมาย หากเรามีระบบที่มีความสามารถในการระบุทราบประเภทเรือในขั้นแรกแล้วนั้น ก็จะเป็นข้อมูลที่เป็นประโยชน์ต่อการวิเคราะห์ในขั้นอื่นๆ ต่อไป เป็นต้น

ข้อมูลที่เป็นรูปภาพ เป็นอีกแหล่งข้อมูลที่สำคัญในการหาข่าวข้อมูลที่ช่วยในการสนับสนุนภารกิจดังกล่าว โดยข้อมูลรูปภาพที่ใช้ในการวิเคราะห์เรือในอดีต มีหลายประเภท อาทิเช่น ภาพจากเรดาห์ (Synthetic Aperture Radar: SAR), ภาพกล้องถ่ายภาพความร้อน (Forward Looking Infra-Red: FLIR), และข้อมูลจากโซนาร์ เป็นต้น การที่จะได้มาซึ่งข้อมูลเหล่านั้น เราจำเป็นต้องมีอุปกรณ์ตรวจจับบนเรือ ซึ่งมีขนาดใหญ่ มีราคาสูง ไม่สามารถบรรทุกในเรือขนาดเล็ก หรือไม่เหมาะกับการทำงานของอากาศยานในลักษณะที่ต้องอาศัยความเร็วในการเคลื่อนที่ ทว่าเนื่องด้วยความก้าวหน้าของเทคโนโลยีกล้องในปัจจุบัน ทำให้เราสามารถใช้กล้องขนาดเล็กที่มีศักยภาพสูง ที่ติดพ่วงกับพาหนะขนาดเล็ก อากาศยาน หรือกระทั่งหุ่นยนต์ เพื่อถ่ายภาพและนำมาประมวลผลได้

การเรียนรู้เครื่องจักรโดยใช้โครงข่ายประสาทเทียม ถูกนำมาใช้ในการจำแนกประเภทเรือจากภาพในช่วงทศวรรษที่ผ่านมา ตัวอย่างงานวิจัย เช่น [1] ใช้การจำแนกเรือจากภาพ SAR ด้วย เกณฑ์การติดสินใจแบบ Fuzzy Logic จากคุณลักษณะเวคเตอร์พาราเมตริกซ์จากการคำนวณแบริ่งเรือ [2] นำเสนอนำอัลกอริธึมในการแยกคุณลักษณะจากการมอดูเลตคลื่นความถี่เวฟเล็ตแบบไม่ต่อเนื่อง เป็นต้น โดยงานวิจัยเหล่านั้นมักอาศัยการสกัดคุณลักษณะแบบการทำมือด้วยตนเองสำหรับงานที่เฉพาะเจาะจง ซึ่งวิธีดังกล่าวนั้นบางวิธีมีความซับซ้อนในการจัดทำ ทั้งยังสิ้นเปลืองเวลาและอาจเกิดข้อผิดพลาดจากมนุษย์ได้ง่าย รวมถึงปัญหาในเรื่องของการต้องใช้หน่วยประมวลที่สูงในการประมวลผลอีกด้วย แต่ด้วยความก้าวหน้าของเทคโนโลยีอีกเช่นกัน ที่ช่วยให้เราสามารถประมวลผลแบบอัตโนมัติโดยไม่ต้องสกัดคุณลักษณะทำมือ ด้วยวิธีที่เรียกว่า การเรียนรู้เชิงลึก (Deep Learning)

โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) เป็นหนึ่งในวิธีการเรียนรู้เชิงลึกที่นิยมนำมาใช้ในงานด้านการประมวลผลภาพ ซึ่งในปัจจุบัน มีงานวิจัยที่นำเสนอแบบจำลองอย่างหลากหลาย ที่เราสามารถนำมาเลือกใช้และปรับแต่งได้ตามความเหมาะสม ตัวอย่างงานวิจัย เช่น [3] นำเสนอการจำแนกประเภทเรือจากภาพโดยเปรียบเทียบประสิทธิภาพความแม่นยำ 3 วิธี ได้แก่ 1) วิธีดั้งเดิมจำแนกโดยมีการทำ preprocess กับรูปภาพก่อน คือ การตรวจจับขอบเรือในรูป แยกคุณลักษณะด้วย Power spectrum และ Moment generating function และใช้ Support Vector Machine (SVM) เป็นตัวตัดสินใจ 2) วิธี Bag of features โดยใช้คุณลักษณะ SURF และตัดสินใจด้วย Multi-class SVM 3) ใช้ CNN โดยแยกเป็น CNN ที่ออกแบบเองและออกแบบตามแบบจำลอง AlexNet [4] ผลพบว่า AlexNet ได้ผลลัพธ์ดีที่สุดที่ค่าความแม่นยำ 88.22% [5] นำเสนอการใช้ CNN ด้วยวิธี Transfer learning ซึ่งเป็นวิธีการนำแบบจำลองที่ได้รับการฝึกข้อมูลและปรับน้ำหนักมาแล้วมาใช้ปรับใช้กับข้อมูลเรือ โดยเปรียบแบบจำลอง 3 แบบ คือ AlexNet, Inception [6] และ ResNet [7] เพื่อหาค่าพารามิเตอร์ในการปรับน้ำหนักที่เหมาะสมที่สุด ผลลัพธ์พบว่า แบบจำลอง Inception V3 ที่จำนวนเลเยอร์ 48 ชั้น และค่า L2 Regularization เท่ากับ 0.0005 ให้ผลลัพธ์ดีที่สุดอยู่ที่ 78.73%

นอกจากนี้ CNN ยังได้ถูกนำมาใช้จำแนกประเภทจากรูปภาพ SAR และภาพถ่ายทางอากาศอีกด้วย เช่น [8] ได้ออกแบบแบบจำลอง CNN จำนวน 4 แบบ เพื่อทดสอบการตรวจพบและจำแนกเรือ 5 ประเภท ในภาพถ่ายจากดาวเทียม TerraSAR-X ที่ความละเอียดสูง เปรียบเทียบกับวิธีดั้งเดิมที่ใช้ SVM ผลลัพธ์พบว่า CNN ให้ค่าผลลัพธ์ที่ดีกว่า โดยค่าความแม่นยำที่ดีที่สุดอยู่ที่ 94% [9] นำเสนอการนำ CNN มาใช้ในการแยกเรือ 3 ประเภท ด้วยวิธีถ่ายโอนเรียนรู้ โดยทำการเปรียบการทำ Transfer learning และ Fine tuning กับชุดข้อมูล COSMO-SkyMed โดยใช้ VGG16 [10] เป็นแบบจำลองฐาน ผลลัพธ์พบว่า วิธี Fine tuning ให้ผลลัพธ์ที่ดีกว่าที่ค่าความแม่นยำ 97.62% [11] นำเสนอการทำ Data augmentation สำหรับชุดข้อมูลเรือที่มีจำนวนน้อย เพื่อเพิ่มจำนวนรูปภาพในการฝึกจากข้อมูลภาพ TerraSAR-X และทำการเปรียบเทียบผลลัพธ์ความแม่นยำจากการนำไปใช้ในแต่ละแบบจำลอง ได้แก่ CNN แบบดั้งเดิมที่ออกแบบเอง การทำ Transfer learning กับแบบจำลอง ResNet 50 และแบบจำลองอื่นๆ ได้แก่ AlextNet, VGG16, Densenet-121 และ Resnet-34 ผลลัพธ์พบว่าการทำเพิ่มจำนวนรูปและใช้ร่วมกับการทำ Transfer Learning ให้ผลลัพธ์ที่ดีเพียงพอ Densenet-121 ให้ค่าความแม่นยำสูงสุดที่ 98.96% แต่แลกกับการใช้เวลาฝึกที่นาน แบบจำลองอื่นๆ เช่น VGG16 ก็ให้ผลลัพธ์ที่ใกล้เคียงที่ 98.46 แต่ใช้เวลาน้อยกว่า จึงเป็นทางเลือกในการใช้งานตามความเหมาะสม

Keras [12] เป็นไลบรารี่ภาษา Python ที่ใช้ในการเขียนโปรแกรมสำหรับงานทางด้านการเรียนรู้เครื่องจักร มีโมดูลต่างๆ ให้สามารถเรียกใช้งานได้ และสามารถดาวน์โหลดเพื่อใช้งานได้ ซึ่งหลายงานวิจัยเองก็ใช้ Keras ในการจำแนกวัตถุจากภาพ ทว่างานวิจัยเกือบทั้งหมดที่นำเสนอการจำแนกรูป มักทดลองการฝึกข้อมูลด้วยคอมพิวเตอร์ที่มีหน่วยประมวลผลขนาดใหญ่ (GPU) ซึ่งแม้จะหาได้ทั่วไป แต่มีราคาแพงสำหรับการใช้งานทั่วไป ยังไม่ค่อยมีงานวิจัยใดที่เปรียบเทียบการฝึกข้อมูลภายใต้ข้อจำกัดของหน่วยประมวล CPU แบบทั่วไป

งานวิจัยนี้ จึงนำเสนอผลการเปรียบเทียบประสิทธิภาพด้านความแม่นยำของการจำแนกประเภทเรือ 5 ประเภทจากภาพถ่ายโดยใช้แบบจำลองการเรียนรู้เชิงลึกบน Keras ด้วยวิธี Transfer Learning โดยจะทำการเปรียบเทียบค่าความแม่นยำจากการจำแนกในแบบจำลอง CNN จำนวน 5 แบบ ได้แก่ 1) แบบจำลอง CNN อย่างง่าย 2) VGG16 3) InceptionV3 4) ResNet50 5) MobileNet [13] ทำการฝึกและทดสอบบนภาพเรือด้วยชุดข้อมูลเปิดของ Analytics Vidhya [14] จำนวน 3,000 ภาพ เพื่อหาแบบจำลองที่ให้ผลลัพธ์ความแม่นยำในการทำนายดีที่สุด ภายใต้ขอบเขตการฝึกข้อมูลบน CPU

ลำดับเนื้อหาที่เหลือในงานวิจัย มีดังนี้ ส่วนที่ 2 กล่าวถึงหลักการพื้นฐานของ CNN ที่งานวิจัยนี้นำมาใช้ ส่วนที่ 3 อธิบายข้อมูลและวิธีที่นำเสนอ ส่วนที่ 4 แสดงผลการทดลอง และสุดท้าย ส่วนที่ 5 สรุปผลจากการวิจัยครั้งนี้

1. โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN)

CNN เป็นโครงข่ายประสาทเทียมที่จำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นที่ย่อยๆ และนำกลุ่มของพื้นที่ย่อยๆมาผสานกัน เพื่อดูว่าสิ่งที่เห็นอยู่เป็นอะไรกันแน่ การมองพื้นที่ย่อยของมนุษย์จะมีการแยกคุณลักษณะ (feature) ของพื้นที่ย่อยนั้น เช่น ลายเส้น และการตัดกันของสี ซึ่งการที่มนุษย์รู้ว่าพื้นที่ตรงนี้เป็นเส้นตรงหรือสีตัดกัน เพราะมนุษย์ดูทั้งจุดที่สนใจและบริเวณรอบๆ ประกอบกัน

การแยกคุณลักษณะจากภาพของ CNN ใช้หลักการเดียวกันกับคอนโวลูชันเชิงพื้นที่ (Spatial Convolution) กล่าวคือเป็นการคำนวณจากการนำตัวกรอง (filter หรือ kernel) ของคุณลักษณะหนึ่งๆ ที่ต้องการไปทาบบนพิกเซลของภาพแล้วทำการดอทเมตริกซ์เพื่อดึงค่าคุณลักษณะที่สนใจ เลื่อนตัวกรองและทำซ้ำไปเรื่อยๆ จนครบทุกพิกเซล ก็จะได้ผังคุณลักษณะ (feature map) ออกมา แสดงสมการ

เมื่อ f คือเมตริกซ์รูปภาพนำเข้า h คือเมตริกซ์ตัวกรอง m และ n คือดัชนีของแถวและคอลัมน์ของเมตริกซ์ผลลัพธ์จากการดอทเมตริกซ์

โครงสร้างของ CNN มีลักษณะเป็นชั้นโครงข่ายซ้อนกัน โดยมีองค์ประกอบหลักๆ 5 อย่าง คือ

1. ชั้น Convolution ทำการกรองโดยการดอทให้เป็นผังคุณลักษณะ
2. ชั้น Activation เป็นฟังก์ชันการตัดสินใจในการส่งต่อผลลัพธ์ของแต่ละนิวรอน
3. ชั้น Pooling เช่น Max Pooling คือตัวกรองที่ใช้ในการเลือกค่าที่มากที่สุดในบริเวณเมทริกซ์เดียวกันของแต่ละเมทริกซ์เพื่อดึงคุณลักษณะเด่นออกมา
4. ชั้น Flatten เป็นการทำ feature map ที่เป็นเมตริกซ์หลายมิติ ให้เหลือมิติเดียว เพื่อสะดวกในการวิเคราะห์ข้อมูล
5. ชั้น Fully-Connected เป็นตัวตัดสินใจความน่าจะเป็นของ feature จากชั้นก่อนว่าตรงกับคลาสใดและส่งต่อค่าน้ำหนักไปยังชั้นถัดไปด้วย จนได้ผลลัพธ์เป็นคลาสในชั้นสุดท้าย



รูปที่ ภาพรวมโครงสร้างของ CNN [15]

แม้ว่า CNN จะสามารถทำงานได้ดีกว่าการสกัดคุณลักษณะจากรูปภาพแบบดั้งเดิม แต่ยังมีข้อจำกัดบางประการ คือ การฝึกแบบจำลองให้ได้ความแม่นยำสูง ต้องใช้ข้อมูลฝึกขนาดใหญ่ ยิ่งมีจำนวนมาก เครื่องก็จะสามารถเรียนรู้ได้มากและทำนายได้แม่นยำขึ้น และชั้นโครงข่ายที่ซับซ้อนก็จะใช้เวลาในการฝึกนานยิ่งขึ้นด้วย ซึ่งได้มีการนำเสนอวิธีรับมือกับปัญหานี้ 2 แนวทาง คือ 1) การทำ Data Augmentation เพื่อเพิ่มจำนวนรูปภาพ กล่าวคือ เป็นการทำซ้ำรูปโดยปรับให้แตกต่างจากเดิม เช่น การหมุนรูปในองศาอื่น การตัดรูป การปรับแสงสว่าง เป็นต้น 2) การทำ Transfer Learning คือการใช้แบบจำลองที่มีการปรับค่าน้ำหนักจากการฝึกกับชุดข้อมูลขนาดใหญ่มาแล้ว มีค่าน้ำหนักที่คลอบคลุมวัตถุทั่วไปหลากหลายชนิด แล้วสามารถนำมาปรับใช้กับงานเฉพาะทางโดยไม่ต้องทำฝึกข้อมูลใหม่ทั้งหมด ซึ่งปัจจุบัน มีงานวิจัยที่พัฒนาโครงสร้างแบบจำลอง CNN ที่สามารถนำมาใช้เป็นเป็นแบบจำลองฐานในการทำ transfer learning มากมาย ตัวอย่างแบบจำลองที่นิยมใช้ เช่น

1. AlexNet (2012) เป็นแบบจำลองแรกที่เป็นจุดเริ่มต้นให้ CNN ได้รับความนิยม โครงสร้างมีชั้นโครงข่าย 8 ชั้น ประกอบด้วย Conv 5 ชั้นและ FC 3 ชั้น ประมวลผลคู่ขนานโดยใช้โครงสร้างเดียวกัน

2. VGGNet (2014) แบบจำลองที่มีจำนวนชั้นมากขึ้น โดยเป็นชั้น Conv 2 ชั้นคู่กับชั้น Max Pooling ทบไปเรื่อยๆ และมี FC 3 ชั้น เลขด้านหลังชื่อแบบจำลองระบุจำนวนชั้นที่ใช้ เช่น VGG16 เป็นต้น

3. GoogleNet หรือ Inception (2014) ใช้โมดูลที่เรียกว่า inception บนโครงสร้าง 22 ชั้น โดยการเพิ่มตัวกรองขนาด 1x1 เพื่อปรับลดมิติข้อมูล เพื่อปรับปรุงประสิทธิภาพการคำนวณให้เร็วขึ้น

4) Residual Network หรือ ResNet (2015) นำเสนอแนวคิดการเพิ่มจุด skip connection ไว้ ข้ามชั้นไว้ด้วย ทำให้สามารถผ่านข้อมูลในชั้นที่ลึกขึ้นแต่ไม่ได้เพิ่มภาระทางด้าน performance การเทรนนัก

5) MobileNet (2017) เป็นแบบจำลองขนาดเล็กที่ทำงานได้เร็ว Latency ต่ำ ใช้พลังงานในการประมวลผลไม่มาก ถูกออกแบบมาสำหรับงานที่มีทรัพยากรจำกัด มีชั้นที่เรียกว่า Depthwise separable convolution แยกฝึกเป็นส่วนๆ โดยใช้ 1 ตัวกรองต่อหนึ่ง channel และใช้ Conv ขนาด 1x1 รวมส่วนย่อยแบบเส้นตรงในชั้น Pointwise convolution

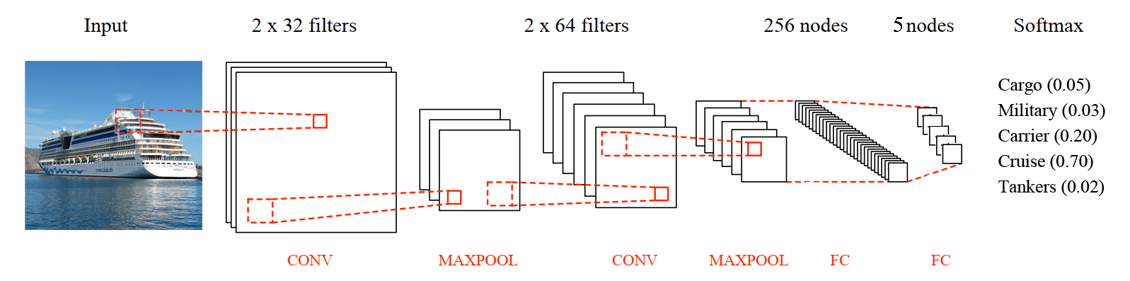
1. วิธีที่นำเสนอ

**3.1 ประเด็นพิจารณาในงานวิจัย** มีดังนี้

1. จำแนกประเภทเรือ 5 ประเภทโดยใช้ 2 ปัจจัยเป็นเกณฑ์จำแนกคือ รูปร่างเรือและสีเรือ
2. สร้างโปรแกรมจำแนกประเภทเรือจากภาพถ่าย โดยใช้ Keras ซึ่งเป็นไลบรารี่เปิด
3. ใช้วิธีสร้าง CNN อย่างง่าย และวิธี Transfer Learning โดยใช้ Pre-trained Model จำนวน 4 แบบเป็นแบบจำลองฐาน ได้แก่ VGG16, InceptionV3, ResNet50, และ MobileNet โดยเฉพาะ MobileNet ซึ่งเป็นแบบจำลองที่ยังไม่มีงานวิจัยที่นำเสนอแบบจำลองนี้มาใช้ในการจำแนกภาพเรือ
4. ทำ Data Augmentation สำหรับข้อจำกัดที่ข้อมูลการฝึกมีน้อย
5. ทดสอบฝึกข้อมูลบน CPU ซึ่งงานวิจัยส่วนมากมักนำเสนอการใช้บน GPU
6. เปรียบเทียบประสิทธิภาพของ 5 แบบจำลอง โดยใช้ค่าความแม่นยำในการทำนายเป็นเกณฑ์

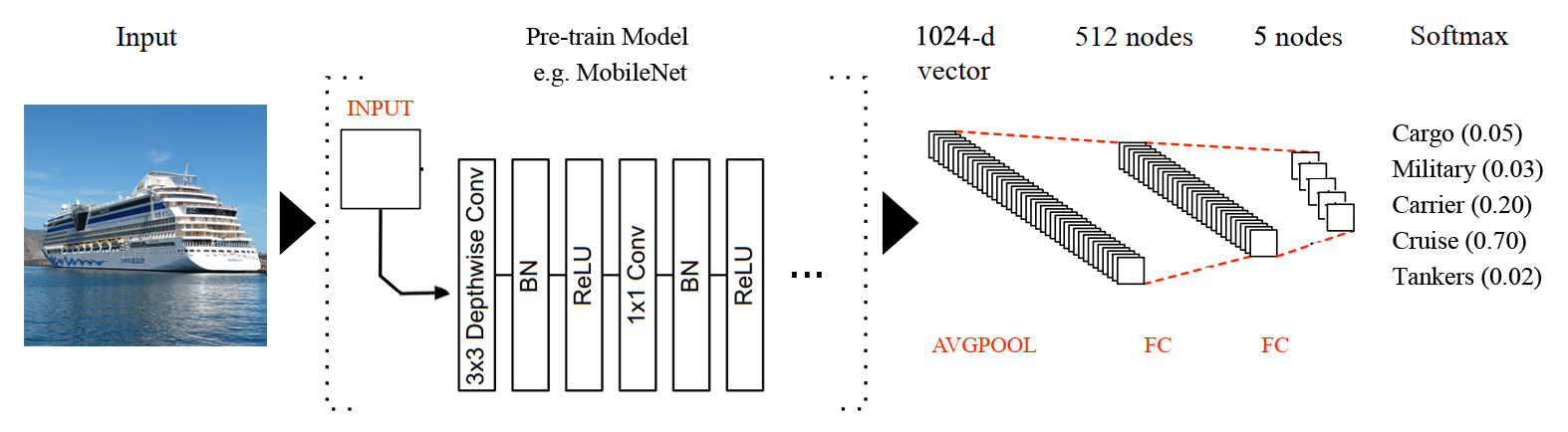
**3.2 โครงสร้างแบบจำลอง**

1) แบบจำลองอย่างง่าย (Simple CNN) ใช้สถาปัตยกรรมจาก [16] นำมาประยุกใช้ในงานจำแนกประเภทเรือ โดยจะเริ่มฝึกข้อมูลตั้งแต่ต้น (train from scratch) มีโครงสร้าง ดังนี้



รูปที่ โครงสร้างแบบจำลองอย่างง่าย [16]

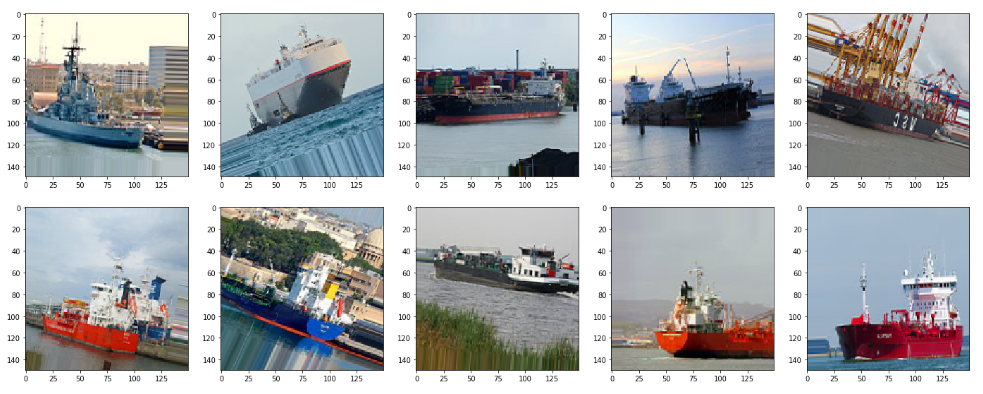
2) แบบจำลองที่ได้รับการฝึกมาแล้ว (Pre-train Model) สำหรับ Transfer Learning โดยตัดส่วน Fully-connected (FC) เดิมออก และกำหนดเลเยอร์ FC ใหม่ 2 ชั้น และชั้น Activation ด้วยฟังก์ชัน Softmax เป็นตัวจำแนก 5 คลาส



รูปที่ โครงสร้างแบบจำลองที่ใช้วิธี Transfer Learning ตัวอย่างการเรียงชั้นโดยคร่าวของ MobileNet

**3.3 การทำ Data Augmentation**

กำหนดการสุ่มค่าการปรับภาพดังนี้ : หมุนไม่เกิน 30 องศา, เลื่อนแนวตั้ง/แนวนอน = 0.15, การขยายภาพ = 0.2, ยืดภาพ = 0.2, พลิกภาพแนวนอน, เติม padding ด้วยพิกเซลใกล้เคียง



รูปที่ ตัวอย่างรูปภาพจากการ Data Augmentation

1. การทดสอบ

**4.1 ชุดข้อมูลและเครื่องมือ**

ใช้ชุดข้อมูลภาพเรือ “Game of Deep Learning: Ship datasets” [14] ที่รวบรวมโดยบริษัท Analytics Vidhya ที่ใช้เป็นเกณฑ์การทดสอบในการแข่งขันออกแบบแบบจำลองการเรียนรู้เชิงลึก ประกอบด้วยคลาสเรือ 5 ประเภท ได้แก่ Cargo (เรือบรรทุกตู้สินค้า), Military (เรือรบ), Carrier (เรือบรรทุกสินค้าล้อเลื่อน), Cruise (เรือสำราญ), และ Tankers (เรือบรรทุกสินค้าเหลว) ข้อมูลที่ใช้มีจำนวน 3,000 รูป แบ่งเป็นข้อมูลสำหรับฝึก 2,500 รูป ข้อมูลสำหรับทดสอบ 500 รูป รูปภาพที่ใช้มีลักษณะดังนี้ 1) เป็นภาพสี 2) ภาพเดี่ยวรูปเรือ ส่วนมากมีพื้นหลังเป็นรูปลอยอยู่ในทะเล 3) เป็นภาพซูมเรือระยะใกล้ 4) มี resolution ภาพประมาณ 210x140 พิกเซล 5) อัตราส่วนพิกเซลรูปต่อขนาดจริง (เมตร) ประมาณ 1 พิกเซล : 2.2 ม.





รูปที่ ตัวอย่างรูปภาพเรือที่ใช้ในการฝึกและทดสอบ

ใช้ไลบรารี่ Keras เป็นเครื่องมือในการนำเข้าแบบจำลอง เขียนโปรแกรมด้วย Jupyter Notebook ประมวลผลการฝึกและทดสอบบน Intel(R) Core(TM) i5-7300 CPU @ 2.60GHz 2.71 GHz RAM 8 GB

**4.2 ผลการทดสอบ**

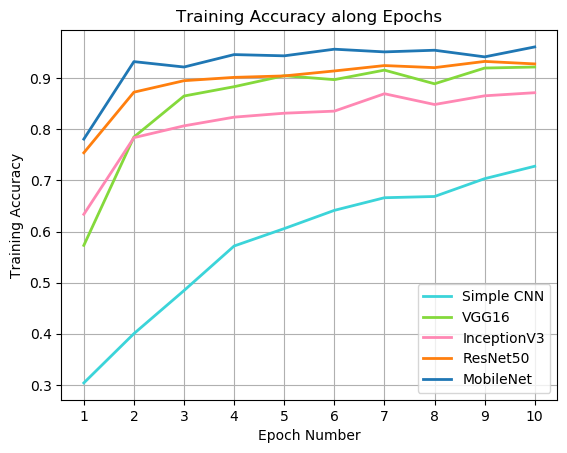
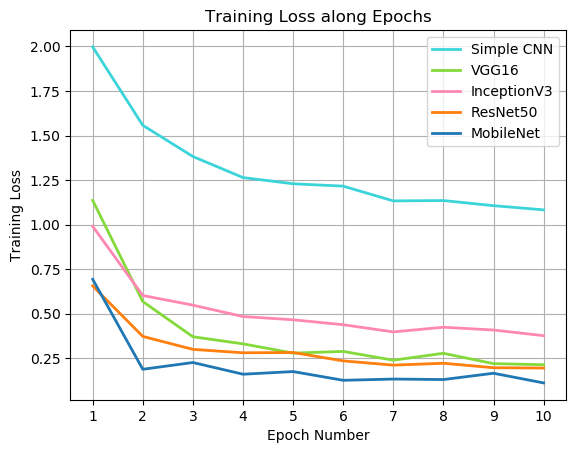
เกณฑ์การคำนวณที่ใช้ค่าความแม่นยำที่ใช้วัดในที่นี้คือ ค่าความแม่นยำในการทำนายประเภทเรือ (Test Accuracy) จากชุดข้อมูลทดสอบ

การทดสอบการฝึกข้อมูลบนหน่วยประมวลผล CPU ความจุ RAM 8 GB Batch size = 32, Epoch = 10 รอบ ได้ผลแสดงดังตาราง 1

ตารางที่ ผลลัพธ์ค่าความแม่นยำและเวลาที่ใช้ฝึกของแต่ละแบบจำลองทั้งห้า

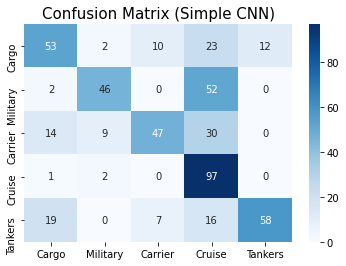
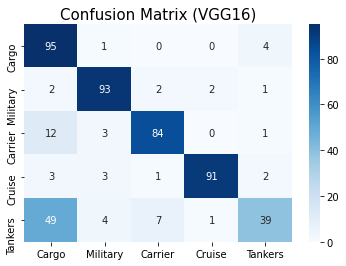
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Total Time Used (sec)** | **Average Time per Epoch (sec)** | **Training Accuracy** | **Training**  **Loss** | **Test Accuracy** |
| Simple CNN | 922 | 92.20 | 0.7277 | 1.0836 | 0.6020 |
| VGG16 | 6744 | 674.44 | 0.9218 | 0.2106 | 0.8040 |
| InceptionV3 | 2315 | 231.50 | 0.8716 | 0.3720 | 0.6180 |
| ResNet50 | 1935 | 193.50 | 0.9278 | 0.1941 | 0.7960 |
| MobileNet | 967 | 96.70 | 0.9611 | 0.1100 | **0.8920** |

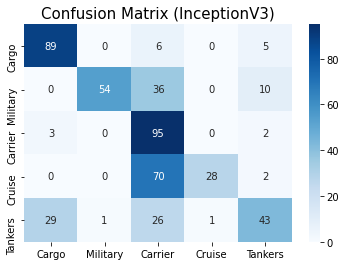
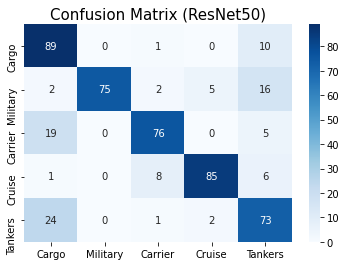
จากผลการทดสอบพบว่า MobileNet ให้ค่าความแม่นยำสูงที่สุดที่ 89.2% เมื่อเทียบกับแบบจำลองประเภทอื่นๆ นอกจากนี้ เมื่อพิจารณาเวลาที่ใช้ในการฝึกข้อมูล ยังใช้เวลาที่น้อยกว่าแบบจำลองอื่นๆ มาก แม้จะใช้เวลามากกว่าการ Simple CNN เล็กน้อย แต่ก็ถือว่าอยู่ในเกณฑ์ที่ดี

รูปที่ กราฟแสดงค่า Training Accuracy และ Loss ระหว่างการฝึกข้อมูลของแต่ละแบบจำลอง

Confusion Matrix คือการประเมินผลลัพธ์การทำนายเปรียบเทียบกับผลลัพธ์จริงๆ ผลลัพธ์ที่ถูกต้องจะแสดงในแนวทแยง confusion matrix ของแต่ละแบบจำลอง มีดังนี้



รูปที่ Confusion Matrix ของแบบจำลองทั้ง 5 แบบ

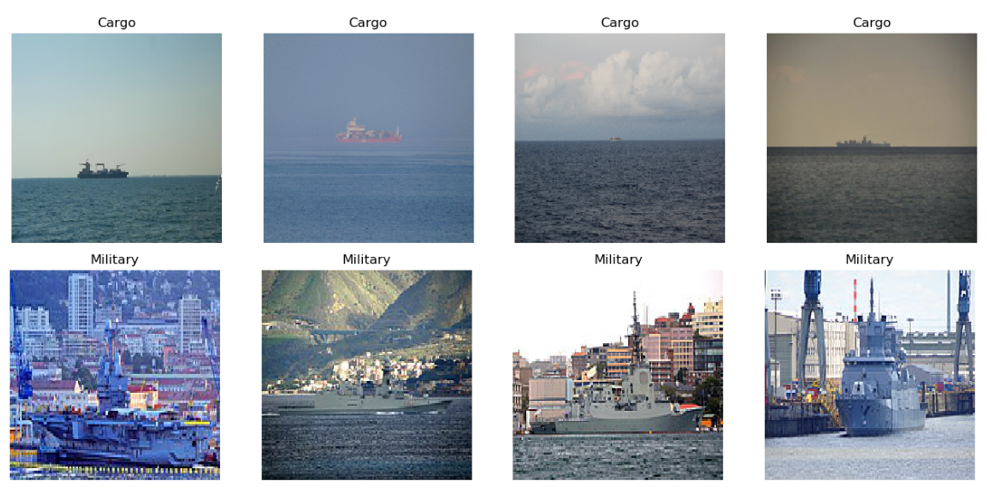
แบบจำลอง MobieNet ให้ผลลัพธ์การทำนายประเภทเรือได้ผิดพลาดน้อยที่สุด เมื่อพิจารณาภาพรวมของทุกแบบแล้วจะเห็นว่า คลาส Cargo เป็นคลาสที่มีโอกาสทายผิดมากกว่าคลาสอื่น รองลงมาคือ Military ดังตัวอย่างการรูปที่ทำนายผิดในรูปที่ 8 จึงสันนิษฐานว่าปัจจัยของข้อมูลที่มีผลต่อความแม่นยำของโมเดล นอกจากในเรื่องของรูปร่างและสีเรือแล้ว ยังอาจมีปัจจัยอื่นร่วมด้วย ได้แก่ พื้นหลังในรูปและสัดส่วนรูปเรือในภาพจากการถ่ายระยะไกล ทำให้รายละเอียดของภาพไม่ชัดเจน จึงทำการทดสอบด้วยรูปเรือทั้งหมด 20 รูป ที่มีองค์ประกอบตามปัจจัยดังกล่าว แบ่งเป็นรูปเรือประเภท Cargo ที่เป็นรูปถ่ายระยะไกล 10 รูป และรูปเรือ Military ที่มีพื้นหลังที่ไม่เกี่ยวข้อง ทดสอบบนแบบจำลอง MobileNet ได้ผลดังตาราง 2



รูปที่ ตัวอย่างภาพทดสอบที่ทำนายผิดจากปัยจัยทางลักษณะข้อมูล

ตารางที่ Confusion Matrix ของรูปทดสอบที่ถ่ายระยะไกลและรูปที่มีพื้นหลังกำกวม

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | ***Predicted*** | | | | | |
| ***Actual*** |  | **Cargo** | **Military** | **Carrier** | **Cruise** | **Tankers** |
| **Cargo** | 5 | 2 | 0 | 0 | 3 |
| **Military** | 0 | 6 | 0 | 1 | 3 |

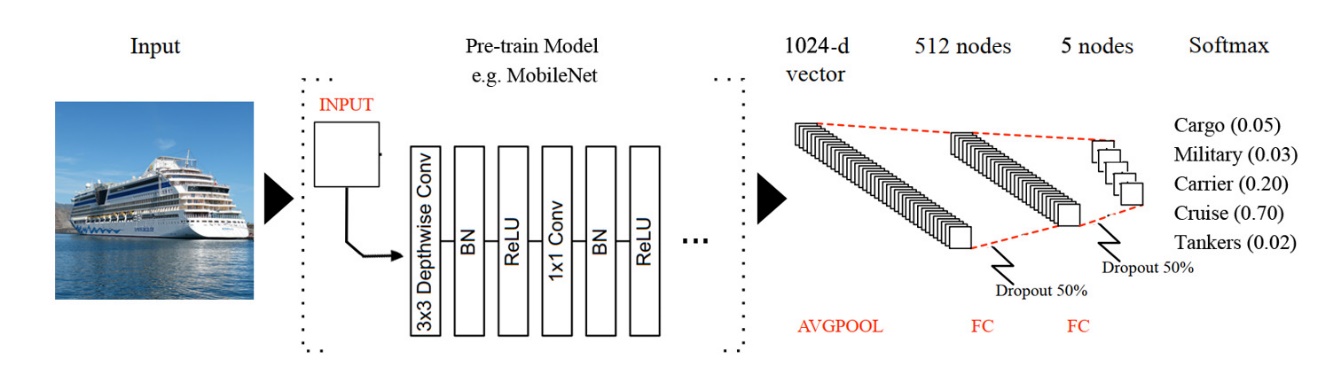


รูปที่ ตัวอย่างรูปทดสอบระยะไกลและรูปมีพื้นหลังกำกวมที่ทำนายผิด

จากผลลัพธ์ Confusion Matrix ตาราง 2 จะเห็นได้ว่า เมื่อใช้แบบจำลอง ยังมีจำนวนทำนายผิดอยู่บ้างเป็นจำนวน 9 จาก 20 ภาพ ตัวอย่างภาพที่ทำนายผิดดังแสดงในรูปที่ 9 ปัจจัยทางด้านรูประยะไกลและรูปที่มีพื้นหลังไม่เกี่ยวข้อง จึงอาจมีส่วนเป็นสาเหตุที่ทำให้แบบจำลองทำนายผิดพลาด รูปที่นำมาควรมีความชัดเจนเพียงพอที่จะเห็นรายละเอียดคุณลักษณะที่เด่นชัดของเรือแต่ละประเภท

นอกจากนี้ ผู้วิจัยได้โดยเลือกแบบจำลอง MobileNet มาทำการทดลองเพิ่มเติมใน 2 ประเด็น ดังนี้

1) ปรับโครงสร้างในชั้น Fully-connected (FC) โดยเพิ่มชั้น Dropout = 0.5 คั่นทั้งสองชั้น FC การใส่ชั้น Dropout เป็นวิธีการแก้ปัญหา Overfitting ที่อาจเกิดขึ้นของแบบบจำลอง กล่าวคือ แบบจำลองเรียนรู้ข้อมูลที่ผูกติดกับรายละเอียดของข้อมูลนั้นๆ มากเกินไป ทำให้จะไม่สามารถไปทำนายผลข้อมูลที่มีรายละเอียดนอกเหนือจากนั้นหรือในอนาคตได้



รูปที่ ปรับโครงสร้างแบบจำลองที่ใช้วิธี Transfer Learning ที่ใช้ MobileNet จากรูปที่ 3 โดยเพิ่มชั้น Dropout

2) ทดสอบเพิ่มจำนวนรอบ Epoch เป็น 100 รอบ ทั้งแบบมีและไม่มี dropout ได้ผลดังนี้

ตารางที่ ผลลัพธ์ค่าความแม่นยำและเวลาที่ใช้ฝึกของแบบจำลอง MobileNet 4 รูปแบบ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **รูปแบบ MobileNet** | **Total Time Used (sec)** | **Average Time per Epoch (sec)** | **Training Accuracy** | **Training**  **Loss** | **Test Accuracy** |
| No dropout, epoch = 10 | 967 | 96.70 | 0.9611 | 0.1100 | 0.8920 |
| Dropout, epoch = 10 | 987 | 98.70 | 0.9267 | 0.2299 | 0.9040 |
| No dropout, epoch = 100 | 5853 | 58.53 | 0.9907 | 0.0301 | 0.9200 |
| Dropout, epoch = 100 | 6052 | 60.52 | 0.9785 | 0.0700 | **0.9300** |

จากผลการทดสอบแสดงดังตารางที่ 3 พบว่า แบบจำลอง MobileNet ที่เพิ่มชั้น Dropout และฝึกข้อมูลที่ epoch = 100 ให้ค่าความแม่นยำสูงที่สุดที่ 93% กล่าวได้ว่า การเพิ่มชั้น Dropout และจำนวนรอบ epoch ที่เพิ่มขึ้น มีผลให้แบบจำลองมีความแม่นยำขึ้น ซึ่งต้องแลกกับการใช้เวลาเทรนที่นานขึ้นนั่นเอง

1. สรุปผล

งานวิจัยนี้ นำเสนอการการจำแนกประเภทเรือจากภาพถ่ายโดยใช้แบบจำลองวิธีการเรียนรู้เชิงลึก โดยแสดงการเปรียบเทียบประสิทธิภาพความแม่นยำในการการจำแนก ทดลองกับแบบจำลองที่แตกต่างกัน 5 แบบ คือ แบบจำลอง CNN อย่างง่ายที่ทำการฝึกข้อมูลตั้งแต่เริ่มต้น และแบบจำลองที่ได้รับการฝึกมาแล้วนำมาทำ Transfer Learning 4 แบบ ได้แก่ VGG16, InceptionV3, ResNet50 และ MobileNet มีเป้าหมายเพื่อหาแบบจำลองที่เหมาะสมที่สุดสำหรับชุดข้อมูลเรือ มีการทำ Data Augmentation เพิ่มเติมเพื่อจัดการข้อจำกัดในด้านข้อมูลที่มีปริมาณไม่มากนัก และทำการทดสอบการฝึกฝนโดยใช้เครื่องคอมพิวเตอร์หน่วยประมวลผล CPU ผลลัพธ์แสดงให้เห็นว่า MobileNet เป็นแบบจำลองที่เหมาะสมที่สุด โดยมีค่าความแม่นยำจากการทำนายบนชุดข้อมูลทดสอบ 89.2% นอกจากนี้ เมื่อทำการฝึกฝนแบบจำลอง โดยการเพิ่มชั้น Dropout ในแบบจำลอง MobileNet รวมทั้งทำการฝึกข้อมูลด้วยจำนวนรอบมากขึ้น พบว่าให้ผลลัพธ์ค่าความแม่นยำที่มากขึ้นเท่ากับ 93% อีกด้วย

**เอกสารอ้างอิง**

1. Margarit, G. & Tabasco, A. (2011). “Ship Classification in Single-Pol SAR Images Based on Fuzzy Logic,” in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 8, pp. 3129-3138, Aug. 2011.
2. Y. Liu, X. Zhang and Y. Yu, "Classification of vessel targets using wavelet statistical features," in *2012 5th International Congress on Image and Signal Processing*, Chongqing, 2012, pp. 1551-1555.
3. Bartan, B. (2017). Ship Classification Using an Image Dataset.
4. A. Krizhevsky, I. Sutskever, & G. Hinton, (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems.*
5. M. Leclerc, R. Tharmarasa, M. Florea, A. Boury-Brisset, T. Kirubarajan & N. Duclos-Hindie, (2018). “Ship Classification Using Deep Learning Techniques for Maritime Target Tracking,” *Information Fusion.* 750-757.
6. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, & Z. Wojna. (2016). “Rethinking the inception architecture for computer vision” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.
7. K. He, X. Zhang, S. Ren, and J. Sun, (2016). “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
8. C. Bentes, D. Velotto & B. Tings, (2018). “Ship Classification in TerraSAR-X Images With Convolutional Neural Networks,” in IEEE Journal of Oceanic Engineering, Jan. 2018, vol. 43, no. 1, pp. 258-266.
9. Y. Wang, C. Wang, H. Zhang. (2018). “Ship Classification in High-Resolution SAR Images Using Deep Learning of Small Datasets,” in Sensors 2018, 18, 2929
10. K. Simonyan & A. Zisserman. (2014). “Very Deep Convolutional Networks for Large-Scale Image Recognition”.
11. C. Lu, & W. Li. (2018). “Ship Classification in High-Resolution SAR Images via Transfer Learning with Small Training Dataset,” In *Sensors (Basel, Switzerland)*, 19(1), 63.
12. “Keras: The Python Deep Learning library” (2015) https://keras.io/.
13. A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, & H. Adam (2017). “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”.
14. “Game of Deep Learning: Computer Vision Hackathon” (2019). Retrieved March 2, 2020. https://datahack.analyticsvidhya.com/contest/game-of-deep-learning/
15. PrabHu. (2018) “Understanding of Convolutional Neural Network (CNN) - Deep Learning,” Retrieved March 2, 2020. https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148
16. T. Hiippala. (2017). “Recognizing military vehicles in social media images using deep learning”.