

การประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 43

The 43rd Electrical Engineering Conference (EECON-43)



28 - 30 ตุลาคม 2563

ณ โรงแรมท็อบแลนด์
อำเภอเมือง จังหวัดพิษณุโลก



ไฟฟ้ากำลัง (PW)

การประมวลผลสัญญาณดิจิทัล (DS)

อิเล็กทรอนิกส์กำลัง (PE)

โฟโตนิกส์ (PH)

ไฟฟ้าสื่อสาร (CM)

วิศวกรรมชีวการแพทย์ (BE)

ระบบควบคุมและการวัดคุม (CT)

คอมพิวเตอร์และเทคโนโลยีสารสนเทศ (CP)

อิเล็กทรอนิกส์ (EL)

พลังงานหมุนเวียน (RE)

งานวิจัยที่เกี่ยวข้องกับวิศวกรรมไฟฟ้า (GN)



การตรวจจับถุงขยะนอกถังขยะ โดยใช้เอฟเฟิเชียนต์ดีท

Garbage Bag Detection using EfficientDet

ทัศนัย พลอยสุวรรณ¹, กฤตพงศ์ มงคลวนิช² และ อารีรัตน์ ปฐมชัยวาลย์³¹ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสยาม²ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสยาม³ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสยาม areerat.pathom@gmail.com

บทคัดย่อ

ปัญหาของการจัดการขยะมีแนวโน้มสูงขึ้นทุกปี งานวิจัยนี้มีวัตถุประสงค์เพื่อตรวจสอบการตรวจจับขยะด้วยเทคนิคเอฟเฟิเชียนต์ดีทที่นำเสนอฐานข้อมูลภาพถุงขยะจาก Garbage object detection Dataset ทั้งหมด 1,050 ภาพ แบ่งออกเป็น ภาพสร้าง Annotation Bounding Box จำนวน 735 ภาพ และภาพ Pseudo-Labeling จำนวน 315 ภาพ โดยใช้เทคนิคการเพิ่มภาพ อาทิ การหมุน การย่อ-ขยายภาพ การกลับด้าน การปรับสีของภาพ จากผลการศึกษา พบว่า เอฟเฟิเชียนต์ดีทที่นำเสนอมีประสิทธิภาพมากยิ่งขึ้น ค่า AP เพิ่มขึ้นที่ 6.9 % และมีค่าความผิดพลาดลดลงที่ 4.3%

คำสำคัญ: ถุงขยะ, การตรวจจับวัตถุ, เอฟเฟิเชียนต์ดีท

Abstract

The problem of waste management is trending higher every year. The research aims to detect garbage bags with the proposed EfficientDet technique. A total of 1,050 images of garbage object detection dataset images are divided into 735 Annotation Bounding Box images and 315 Pseudo-Labeling images by the image augmentation technique such as rotation, zoom, or flip, color adjustment image to the image. Finally, the proposed EfficientDet technique is more effective. AP increased by 6.9%, and samples with miscounted garbage bags decreased by 4.3%.

Keywords: Garbage Bag, Object Detection, EfficientDet

1. บทนำ

ในยุค Social Distancing ประชาชนส่วนใหญ่ในกลุ่มช่วงอายุ 19-38 ปีมักจะ ไม่กังวลในการกักตุนสินค้าอุปโภคบริโภคและการส่งสินค้าหรืออาหารออนไลน์เป็นทางเลือกหนึ่งที่เพิ่มขึ้น อีกทั้งประชาชนส่วนใหญ่มีความกังวลต่อสถานการณ์การแพร่ระบาดของโควิด-19 และมาตรการการปิดเมือง (lock down) ซึ่งเป็นปัจจัยทำให้เกิดพฤติกรรมการตัดสินใจเลือกใช้บริการสั่งสินค้าหรืออาหารออนไลน์มากขึ้น ส่งผลให้มีปริมาณขยะมูลฝอย ขยะคั่งค้าง ขยะล้นถัง และ รถขยะมาเก็บช้า มีการเพิ่มขึ้นของปริมาณของขยะตามจำนวนของประชากรและสภาพทาง

สังคมและเศรษฐกิจ โดยเฉพาะอย่างยิ่งปัญหา ขยะติดเชื้จากอุปกรณ์ทางการแพทย์และหน้ากากอนามัยแบบใช้แล้วทิ้งที่มีปริมาณเพิ่มมากขึ้น

ในการป้องกันการแพร่ระบาดของโควิด-19 มีผลทำให้การกำจัดขยะติดเชื้ต้องได้รับการฆ่าเชื้ก่อน และ ต้องกำจัดด้วยการเผาด้วยเตาเฉพาะของขยะติดเชื้ นอกจากนี้เมื่อประชาชนอยู่บ้านมากขึ้นทำให้ต้องมีการเพิ่มจุดสำหรับติดถังขยะติดเชื้ชนิดที่ปิดสนิท และ ติดตั้งระบบการคัดแยกขยะติดเชื้ ผลการปิดเมืองยังทำให้ปริมาณขยะโดยเฉพาะบรรจุภัณฑ์พลาสติกที่ใช้ในการห่อหุ้มสินค้าอุปโภคและบริโภคเพื่อความสะดวกในการขนส่งและความปลอดภัยต่อการบริโภคเพิ่มมากขึ้น เนื่องจากพลาสติกมีคุณสมบัติเบา แข็งแรง ทนทาน และราคาถูก สามารถผลิตได้หลากหลายรูปแบบตามความต้องการ ทำให้มีปัญหามลพิษพลาสติกที่กำจัดไม่ถูกวิธีและก่อให้เกิดปัญหา เช่น ปัญหามลพิษทางอากาศ ปัญหาเศรษฐกิจ และ ปัญหาสิ่งแวดล้อมทั้งทางตรงและทางอ้อมที่ส่งผลต่อประชาชนและสิ่งแวดล้อม เช่น การเผาขยะพลาสติกเป็นส่วนหนึ่งของปัญหาฝุ่น PM 2.5 ทำให้เสี่ยงต่อสุขภาพหรือก่อให้เกิดผลกระทบต่อสุขภาพ ทำให้เกิดโรคทางเดินหายใจ โรคหัวใจและหลอดเลือด รวมทั้งมะเร็งปอด โดยเฉพาะประชาชนกลุ่มเสี่ยง เด็ก หญิงตั้งครรภ์ ผู้สูงอายุ และผู้มีโรคประจำตัว เป็นต้น

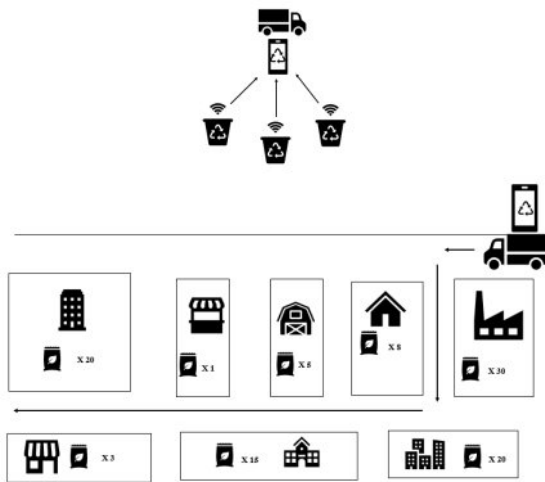
อย่างไรก็ตาม ปัจจัยที่ส่งผลต่อการเพิ่มของปริมาณขยะมีหลายประการรวมถึงการเจริญเติบโตของธุรกิจอี-คอมเมิร์ซ และการสั่งอาหารออนไลน์ที่ขยายตัวอย่างรวดเร็วก่อให้เกิดปัญหาขยะล้น และขยะนอกถังขยะ ดังรูปที่ 1 จนกลายเป็นปัญหาขยะคั่งค้างอยู่ในชุมชน สำหรับปัญหานี้สืบเนื่องมาจากการมีอัตราการบริโภคเพิ่มขึ้น การขยายชุมชนเพิ่มขึ้น และการปรับเปลี่ยนวิถีชีวิตใหม่ (new normal) แต่การจัดเก็บขยะมูลฝอยในแต่ละพื้นที่ความรับผิดชอบและการขนส่งขยะไปยังสถานที่กำจัดขยะซึ่งในแต่ละวันจะมีการเก็บขยะวันละ 1 รอบ หรือบางพื้นที่มีการเก็บขยะ 1 ครั้งต่อสัปดาห์ หรือ บางพื้นที่อาจมีการเก็บขยะที่ต่างกัน ทำให้เกิดปัญหาขยะล้นถังอันนำไปสู่ปัญหาการทิ้งขยะนอกถังขยะ

จากปัญหาดังกล่าวเพื่อให้เกิดประสิทธิภาพในการจัดการขยะ ดังรูปที่ 2 ได้มีการนำเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence: AI) และอินเทอร์เน็ตของสรรพสิ่ง (Internet of Things: IoT) โดยการตรวจจับวัตถุ (Object Detection) และระบุตำแหน่ง พร้อมทั้งนับจำนวนของถุงขยะนอกถังขยะ ในการเก็บรวบรวมข้อมูลและประมวลผลแบบทันที

(real-time processing) ผ่านกล้องไร้สายอัจฉริยะที่ติดตั้งในบริเวณจุดทิ้งขยะ โดยที่จำนวนขยะนอกถังนี้เป็นดัชนีชี้วัดปริมาณขยะส่วนเกินของชุมชนได้ และจัดเก็บบันทึกเป็น big data เพื่อนำข้อมูลไปวิเคราะห์ใน ระบบการบริหารการจัดเก็บขยะ



รูปที่ 1 ตัวอย่างขยะล้นและขยะนอกถังขยะ

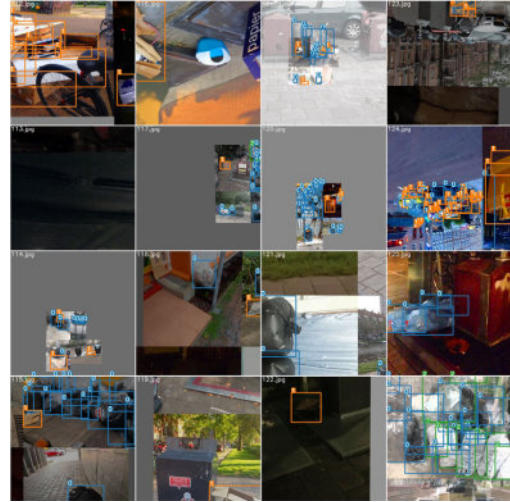


รูปที่ 2 การจัดการขยะ

ในปัจจุบันงานวิจัยที่เกี่ยวกับการตรวจจับและระบุตำแหน่งวัตถุในภาพ ได้แก่ การหาตำแหน่งพิกัดของรถ มนุษย์ หรือ Custom objects ต่าง ๆ ในรูปภาพได้แก่เทคนิค Faster R-CNN [1], RetinaNet [2], YOLOv3 [3], YOLOv4 [4] และ EfficientDet (D0-D7) [5] โดยทั่วไปในการทดสอบเปรียบเทียบประสิทธิภาพของอัลกอริทึมจะทำการทดสอบบนฐานข้อมูล COCO [6] ซึ่งเป็นฐานข้อมูลขนาดใหญ่สำหรับใช้ทำงานวิจัยด้าน object detection และ segmentation

งานวิจัยนี้ได้เสนอการนำ EfficientDet ซึ่งเป็นการตรวจจับและตรวจนับจำนวนของถุงขยะนอกถังขยะ โดยการนับถุงขยะจากจำนวน Bounding box และการใช้เทคนิค Pseudo-Labeling เพื่อช่วยลดจำนวน

การสร้าง labeled dataset จาก unlabeled dataset นอกจากนี้ทำการ data augmentation นำเทคนิค Mosaic [4] ในการเพิ่มประสิทธิภาพในการ train ของระบบ ดังรูปที่ 3 ตัวอย่าง Mosaic



รูปที่ 3 ตัวอย่าง Mosaic

2. EfficientDet

EfficientDet (D0-D7) [5] เป็นอัลกอริทึมที่มีแนวคิดของการทำนายตำแหน่งและขนาดของวัตถุที่มีประสิทธิภาพสูงสุดเทียบกับอัลกอริทึมอื่น ได้แก่ YOLOv4, YOLOv3, Faster R-CNN และ RetinaNet พบว่า EfficientDet (D0) มีค่าความถูกต้อง (33.8 AP) ซึ่งมากกว่า YOLOv3 ที่ (33 AP) และ มีค่า Floating Point Operations (FLOPs) น้อยกว่าถึง 28 เท่า เมื่อเทียบกับ RetinaNet และ Mask R-CNN. พบว่า EfficientDet (D1) มีจำนวนพารามิเตอร์น้อยกว่า 6x - 8x เท่า และค่า FLOPs น้อยกว่า 16x - 25x เท่า

EfficientDet (D0-D7) มี 3 ส่วน ได้แก่

1. ส่วน backbone ใช้โครงสร้าง EfficientNet (D0-D7) [7] โดยขนาดภาพนำเข้า (D0 = 512x512 พิกเซล, D1 = 640x640 พิกเซล, D2 = 768x768 พิกเซล, D3 = 896x896 พิกเซล, D4 = 1,024x1,024 พิกเซล, D5 = 1,280x1,280 พิกเซล, D6 = 1,280x1,280 พิกเซล และ D7 = 1,536x1,536 พิกเซล)

2. ส่วน Bi-directional feature pyramid network (BiFPN) ดังรูปที่ 4 เป็นการนำ features ในลักษณะ multi scale resolution ในลำดับชั้น $\{P_3, P_4, \dots, P_7\}$ จาก backbone โดยที่ความละเอียด ในลำดับชั้น $P_i = P_0 / 2^i$ ตัวอย่างเช่น EfficientDet (D1) มีภาพนำเข้าขนาด 640x640 พิกเซล จะมีขนาด $P_3 = 640/2^3 = 80$ (80x80 พิกเซล), $P_4 = 640/2^4 = 40$ (40x40 พิกเซล), $P_5 = 640/2^5 = 20$ (20x20 พิกเซล) $P_6 = 640/2^6 = 10$ (10x10 พิกเซล) และ $P_7 = 640/2^7 = 5$ (5x5 พิกเซล) ตามลำดับ ดังรูปที่ 5 แต่ละ Layer ของ BiFPN ดังสมการที่ (1)

$$\begin{aligned}
 P_7^{out} &= \text{conv}(P_7^{in}) \\
 P_6^{td} &= \text{conv}\left(\frac{w_1 P_6^{in} + w_2 \text{Resize}(P_7^{in})}{w_1 + w_2 + \mathcal{E}}\right) \\
 P_6^{out} &= \text{conv}\left(\frac{w_1' P_6^{in} + w_2' P_6^{td} + w_3' \text{Resize}(P_5^{out})}{w_1' + w_2' + w_3' + \mathcal{E}}\right) \\
 P_5^{td} &= \text{conv}\left(\frac{w_1 P_5^{in} + w_2 \text{Resize}(P_6^{td})}{w_1 + w_2 + \mathcal{E}}\right) \\
 P_5^{out} &= \text{conv}\left(\frac{w_1' P_5^{in} + w_2' P_5^{td} + w_3' \text{Resize}(P_4^{out})}{w_1' + w_2' + w_3' + \mathcal{E}}\right) \\
 P_4^{td} &= \text{conv}\left(\frac{w_1 P_4^{in} + w_2 \text{Resize}(P_5^{td})}{w_1 + w_2 + \mathcal{E}}\right) \\
 P_4^{out} &= \text{conv}\left(\frac{w_1' P_4^{in} + w_2' P_4^{td} + w_3' \text{Resize}(P_3^{out})}{w_1' + w_2' + w_3' + \mathcal{E}}\right) \\
 P_3^{out} &= \text{conv}\left(\frac{w_1' P_3^{in} + w_2' P_4^{td}}{w_1' + w_2' + \mathcal{E}}\right)
 \end{aligned} \quad (1)$$

โดยให้ Resize คือการทำ upsampling หรือ downsampling ขึ้นกับทิศทางของลำดับชั้นในพีระมิด

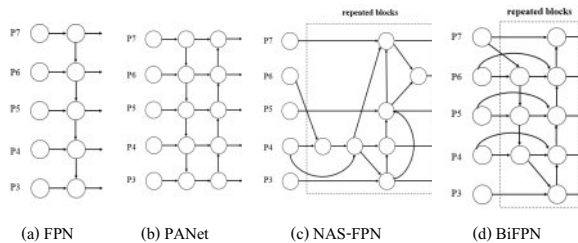
conv คือ Convolutional layer (CNN)

P_i^{td} คือ intermediate ในลำดับชั้น i

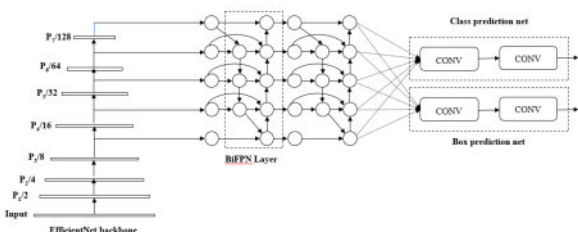
P_i^{out} คือ output ในลำดับชั้น i

3. Box/class prediction network ใน ส่วน ของ การ classification

และ หา Bounding Box ของวัตถุ



รูปที่ 4 Feature network design



รูปที่ 5 EfficientDet

3. Pseudo-Labeling

การเรียนรู้แบบ Semi-Supervised Learning ใช้เทคนิค Pseudo-Labeling [8] โดยการ labeled data แบบอัตโนมัติจากฐานข้อมูล unlabeled dataset ขนาดใหญ่ได้ ขั้นตอนหลัก 5 ขั้นตอน ดังนี้

- 1) Train on batch เพื่อหาค่าพารามิเตอร์ของโมเดล EfficientDet ส่วนที่เป็น label data
- 2) นำ predicted labels และ Bounding Box จาก labeled data มาคำนวณหา Label Loss
- 3) นำโมเดล EfficientDet ที่ผ่านการเทรนจาก labeled data มาทำนาย class prediction และ Bounding Box จาก unlabeled dataset
- 4) นำ predicted labels และ Bounding Box จาก unlabeled dataset มาคำนวณหา Unlabeled Loss
- 5) หา Loss จากสมการที่ (2)

$$\text{Loss} = \text{Label Loss} + \alpha(t) \times \text{Unlabeled Loss} \quad (2)$$

จากสมการที่ (2) เมื่อ $\alpha(t)$ ใช้สำหรับควบคุมความสัมพันธ์ระหว่าง Loss และ Unlabeled Loss โดยให้ t คือ เวลา ซึ่งเป็น ค่า epoch

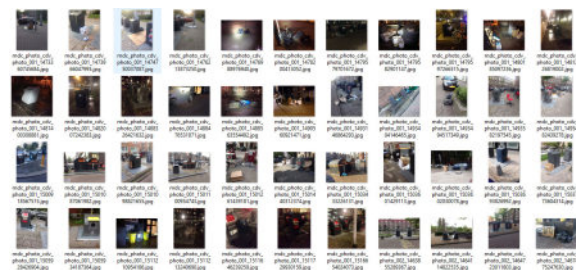
ค่า weight $\alpha(t)$ จะให้ความสำคัญกับ unlabeled dataset เมื่อ epoch มีค่ามากขึ้นและมีค่าเท่ากับ 0 เมื่อ epoch ต่ำ ดังสมการที่ (3)

$$\alpha(t) = \begin{cases} 0 & t < T_1 \\ \frac{t - T_1}{T_2 - T_1} \alpha_r & T_1 \leq t < T_2 \\ \alpha_r & T_2 \leq t \end{cases} \quad (3)$$

โดยให้ $T_1 = 100$, $T_2 = 600$, $\alpha_r = 3$ [8]

4. วิธีดำเนินการวิจัย

งานวิจัยนี้นำภาพลงขยะที่ถูกวางทิ้งไว้นอกถังขยะสถานที่ต่าง ๆ จำนวน 1,050 ภาพ แบ่งเป็นภาพถ่ายตามสถานที่สาธารณะจำนวน 450 ภาพ และภาพจากฐานข้อมูล Garbage object detection Dataset [9] จำนวน 600 ภาพ ดังรูปที่ 6 ตัวอย่าง Datasets ในการสร้าง Annotation Bounding Box จำนวน 735 ภาพ (70%) และภาพไม่สร้าง Annotation Bounding Box จำนวน 315 ภาพ (30%) ใช้ Pseudo-Labeling ในการสุ่มเรียก Datasets โดยกำหนดค่า k เท่ากับ 5 สำหรับ EfficientDet (D1) โดยกำหนดภาพขนาด 640x640 พิกเซล ในการใช้เทคนิค Pseudo-Labeling และ Mosaic เพื่อเปรียบเทียบกับ Standard EfficientDet กำหนด Batch size เท่ากับ 64 และใช้ GPU รุ่น NVIDIA Titan RTX



รูปที่ 6 ตัวอย่าง Datasets [9]

5. ผลการวิจัยและสรุป

ตัวอย่างผลการนับจำนวนถุงขยะ และ Bounding Box ดังรูปที่ 7 จากการวิจัย พบว่า EfficientDet มีค่าผิดพลาดที่ 8.7% กับ EfficientDet ที่นำเสนอ มีค่าความผิดพลาดที่ 4.3% และค่า AP ของ EfficientDet (D1) มีค่า 36.2 และ EfficientDet ที่นำเสนอ มีค่า AP ที่ 38.7 EfficientDet ที่นำเสนอมีประสิทธิภาพมากกว่า EfficientDet ตัวอย่างผลการวิจัยการตรวจนับถุงขยะนอกถังขยะ <https://youtu.be/E18iK76u-ZI>



รูปที่ 7 ผลการนับจำนวนถุงขยะ และ Bounding Box

6. เอกสารอ้างอิง

- [1] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," presented at the Neural Information Processing Systems (NIPS), 2015.
- [2] T.-Y. Lin, P. Goyal, R. B. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2999-3007, 2017.
- [3] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement." Available: <https://arxiv.org/abs/1804.02767> [Accessed April 8, 2019].
- [4] A. Bochkovskiy, C. Wang, and H.M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," ArXiv, vol. abs/2004.10934, 2020.
- [5] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," ArXiv, vol. abs/1911.09070, 2020.

- [6] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," in Computer Vision – ECCV 2014, Cham, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., 2014// 2014: Springer International Publishing, pp. 740-755.
- [7] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," ArXiv, vol. abs/1905.11946, 2019.
- [8] E. Arazo, D. Ortego, P. Albert, N. E. O'Connor, and K. McGuinness, "Pseudo-Labeling and Confirmation Bias in Deep Semi-Supervised Learning," ArXiv, vol. abs/1908.02983, 2019.
- [9] <https://medium.com/maarten-sukel/garbage-object-detection-using-pytorch-and-yolov3-d6c4e0424a10> [Accessed November12, 2019].



ทศนัย พลอยสุวรรณ

ปริญญาเอก สาขาวิศวกรรมไฟฟ้า
จุฬาลงกรณ์มหาวิทยาลัย ปัจจุบันดำรง
ตำแหน่ง ผู้ช่วยศาสตราจารย์ ประจำ
ภาควิชาวิศวกรรมไฟฟ้า มหาวิทยาลัย
สยาม งานวิจัยที่สนใจ Deep learning,
Data Science, AI



กฤตพงศ์ มงคลวนิช

ปริญญาโท สาขา Learning, Design, and
Technology, Stanford University
ปริญญาตรี สาขา Computer Science,
University of Wisconsin – Madison
งานวิจัยที่สนใจ Deep Learning,
Computer Science Education, Human-
Computer Interaction, Operating
Systems, Programming Languages



อริรัตน์ ปฐมชัยวาลย์

ปริญญาเอก สาขาเทคโนโลยีทางภาพ
จุฬาลงกรณ์มหาวิทยาลัย
งานวิจัยที่สนใจ Computer Vision, High
Dynamic Range, Color Management
System, Color Reproduction, Color
Vision, Print Media Technology, Digital
Printing Technology