Ngô Đức Tuấn – 18520186

Lê Hoàng Long – 18521027

Trần Lê Huy – 18520757

**Nhóm: N013**

**Bài tập 7/10/2020**

*Câu 1: 30 bài báo cáo khoa học (paper) về topic: Object detection trong giai đoạn 2016 - 2020*

**1. Feature pyramid networks for object detection**

- Author: Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie

- Year: 2017

- Cited: 4744

- Link: <https://openaccess.thecvf.com/content_cvpr_2017/html/Lin_Feature_Pyramid_Networks_CVPR_2017_paper.html>

**2. Focal loss for dense object detection**

- Author: Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollar

- Year: 2017

- Cited by: 4535

- Link: <https://openaccess.thecvf.com/content_iccv_2017/html/Lin_Focal_Loss_for_ICCV_2017_paper.html>

**3. Relation networks for object detection**

- Author: Han Hu, Jiayuan Gu, Zheng Zhang, Jifeng Dai, Yichen Wei

- Year: 2018

- Cited by: 359

- Link: <https://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Relation_Networks_for_CVPR_2018_paper.html>

**4. Object detection with deep learning: A review**

- Author: Zhong-Qiu Zhao, Shou-Tao Xu, Peng Zheng, Xindong Wu

- Year: 2019

- Cited by: 590

- Link: <https://ieeexplore.ieee.org/abstract/document/8627998/>

**5. Detecting People in Artwork with CNNS**

- Author: Nicholas Westlake, Hongping Cai, Peter Hall

- Year: 2016

- Cited by: 22

- Link: <https://paperswithcode.com/paper/detecting-people-in-artwork-with-cnns>

**6. Cascade r-cnn: Delving into high quality object detection**

- Author: Zhaowei Cai, Nuno Vasconcelos

- Year: 2018

- Cited by: 709

- Link: <https://openaccess.thecvf.com/content_cvpr_2018/html/Cai_Cascade_R-CNN_Delving_CVPR_2018_paper.html>

**7. R-fcn: Object detection via region-based fully convolutional**

- Author: Jifeng Dai, Yi Li, Kaiming He, Jian Sun

- Year: 2016

- Cited by: 3020

- Link: <http://papers.nips.cc/paper/6465-r-fcn-object-detection-via-region-based-fully-convolutional-networks>

**8. Nas-fpn: Learning scalable feature pyramid architecture for object detection**

- Author: Golnaz Ghiasi, Tsung-Yi Lin, Quoc V. Le

- Year: 2019

- Cited by: 170

- Link: <https://openaccess.thecvf.com/content_CVPR_2019/html/Ghiasi_NAS-FPN_Learning_Scalable_Feature_Pyramid_Architecture_for_Object_Detection_CVPR_2019_paper.html>

**9. Deep learning for generic object detection: A survey**

- Author: Li Liu, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu, Matti Pietkainen

- Year: 2020

- Cited by: 351

- Link: <https://link.springer.com/article/10.1007/s11263-019-01247-4>

**10. A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection**

- Author: Zhaowei Cai, Quanfu Fan, Rogerio S.Feris, Nuno Vasconcelos

- Year: 2016

- Cited by: 916

- Link: <https://link.springer.com/chapter/10.1007/978-3-319-46493-0_22>

**11. Camouflaged object detection**

- Author: Deng-Ping Fan, Ge-Peng Ji, Guolei Sun, Ming-Ming Cheng, Jianbing Shen, Ling Shao

- Year: 2020

- Cited by: 18

- Link: <http://openaccess.thecvf.com/content_CVPR_2020/html/Fan_Camouflaged_Object_Detection_CVPR_2020_paper.html>

**12. Beyound Anchor-Bases object detection**

- Author: Fuchun Sun, Huaping Liu, Yuning Jiang, Lei Li, Jianbo Shi

- Year: 2020

- Cited by: 74

- Link: <https://ieeexplore.ieee.org/abstract/document/9123553/>

**13. Scalable and efficient object detection**

- Author: Mingxing Tan, Ruoming Pang, Quoc V. Le

- Year: 2020

- Cited by: 95

- Link: <http://openaccess.thecvf.com/content_CVPR_2020/html/Tan_EfficientDet_Scalable_and_Efficient_Object_Detection_CVPR_2020_paper.html>

**14. Object detection in optical remote sensing images**

- Author: K Li, G Wan, G Cheng, L Meng, J Han

- Year: 2020

- Cited by: 54

- Link: <https://www.sciencedirect.com/science/article/pii/S0924271619302825>

**15. Recent advances in deep learning for object detection**

- Author: Xiongwei Wu, Doyen Sahoo, Steven C.H. Hoi

-Year: 2020

- Cited by: 30

- Link: <https://www.sciencedirect.com/science/article/pii/S0925231220301430>

**16. Prime sample attention in object detection**

- Author: Yuhang Cao, Kai Chen, Chen Change Loy, Dahua Lin

- Year: 2020

- Cited by: 20

- Link: <http://openaccess.thecvf.com/content_CVPR_2020/html/Cao_Prime_Sample_Attention_in_Object_Detection_CVPR_2020_paper.html>

**17. Rethinking RGB-D Salient object detection: Models, Data sets and Large=Scale Benchmarks**

- Author: Deng-Ping Fan, Zheng Lin, Zhao Zhang, Menglong Zhu, Ming-Ming Cheng

- Year: 2020

- Cited by: 57

- Link: <https://ieeexplore.ieee.org/abstract/document/9107477/>

**18. Probabilistic object detection: definition end evaluation**

- Author: David Hall, Feras Dayoub, John Skinner, Haoyang Zhang, Dimity Miller, Peter Corke, Gustavo Carneiro, Anelia Angelova, Niko Suenderhauf

- Year: 2020

- Cited by: 17

- Link: <http://openaccess.thecvf.com/content_WACV_2020/html/Hall_Probabilistic_Object_Detection_Definition_and_Evaluation_WACV_2020_paper.html>

**19. Imbalance problems in object detection**

- Author: Kemal Oksuz, Baris Can Cam, Sinan Kalkan, Emre Akbas

- Year: 2020

- Cited by: 22

- Link: <https://ieeexplore.ieee.org/abstract/document/9042296/>

**20. Context R-CNN: Long Term Temporal Context for Per-Camera Object Detection**

- Author: Sara Beery, Guanhang Wu, Vivek Rathod, Ronny Votel, Jonathan Huang

- Year: 2020

- Cited by: 2

- Link: <https://paperswithcode.com/paper/long-term-temporal-context-for-per-camera>

**21. Looking Fast and Slow: Memory-Guided Mobile Video Object Detection**

- Author: Mason Liu, Menglong Zhu, Marie White, Yinxiao Li, Dmitry Kalenichenko

- Year: 2019

- Cited by: 8

- Link: <https://paperswithcode.com/paper/looking-fast-and-slow-memory-guided-mobile>

**22. Optimizing Video Object Detection via a Scale-Time Lattice**

- Author: Kai Chen, Jiaqi Wang, Shuo Yang, Xingcheng Zhang, Yuanjun Xiong, Chen Change Loy, Dahua Lin

- Year: 2018

- Cited by: 37

- Link: <https://paperswithcode.com/paper/optimizing-video-object-detection-via-a-scale>

**23. Learning Where to Focus for Efficient Video Object Detection**

- Author: Zhengkai Jiang, Yu Liu, Ceyuan Yang, Jihao Liu, Peng Gao, Qian Zhang, Shiming Xiang, Chunhong Pan

- Year: 2020

- Cited by: 2

- Link: <https://paperswithcode.com/paper/learning-motion-priors-for-efficient-video>

**24. Do Image Classifiers Generalize Across Time?**

-Author: [Vaishaal Shankar](https://arxiv.org/search/cs?searchtype=author&query=Shankar%2C+V), [Achal Dave](https://arxiv.org/search/cs?searchtype=author&query=Dave%2C+A), [Rebecca Roelofs](https://arxiv.org/search/cs?searchtype=author&query=Roelofs%2C+R), [Deva Ramanan](https://arxiv.org/search/cs?searchtype=author&query=Ramanan%2C+D), [Benjamin Recht](https://arxiv.org/search/cs?searchtype=author&query=Recht%2C+B), [Ludwig Schmidt](https://arxiv.org/search/cs?searchtype=author&query=Schmidt%2C+L)

- Year: 2019

- Cited by: 1

- Link: <https://paperswithcode.com/paper/a-systematic-framework-for-natural>

**25. Relation Distillation Networks for Video Object Detection**

- Author: Jiajun Deng, Yingwei Pan, Ting Yao, Wengang Zhou, Houqiang Li, Tao Mei

- Year: 2019

- Cited by: 15

- Link: <https://paperswithcode.com/paper/relation-distillation-networks-for-video>

**26. Deep Residual Learning for Image Recognition**

- Author: Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun

- Year: 2016

- Cited by: 57502

- Link: <https://paperswithcode.com/paper/deep-residual-learning-for-image-recognition>

**27. Objects as Points**

- Author: [Xingyi Zhou](https://arxiv.org/search/cs?searchtype=author&query=Zhou%2C+X), [Dequan Wang](https://arxiv.org/search/cs?searchtype=author&query=Wang%2C+D), [Philipp Krähenbühl](https://arxiv.org/search/cs?searchtype=author&query=Kr%C3%A4henb%C3%BChl%2C+P)

- Year: 2019

- Cited by: 251

- Link: <https://paperswithcode.com/paper/objects-as-points>

**28. Searching for MobileNetV3**

- Author: Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, Quoc V. Le, Hartwig Adam

- Year: 2019

- Cited by: 339

- Link: <https://paperswithcode.com/paper/searching-for-mobilenetv3>

**29. Single-shot refinement neural network for object detection**

- Author: Shifeng Zhang, Longyin Wen, Xiao Bian, Zhen Lei, Stan Z. Li

- Year: 2018

- Cited by: 488

- Link: <https://openaccess.thecvf.com/content_cvpr_2018/html/Zhang_Single-Shot_Refinement_Neural_CVPR_2018_paper.html>

**30. Multi-view 3d object detection network for autonomous driving**

- Author: Xiaozhi Chen, Huimin Ma, Ji Wan, Bo Li, Tian Xia

- Year: 2017

- Cited by: 837

- Link: <https://openaccess.thecvf.com/content_cvpr_2017/html/Chen_Multi-View_3D_Object_CVPR_2017_paper.html>

*Câu 2: Lọc ra 10 bài báo về Object Detection ấn tượng. Lí do?*

**1. Beyound Anchor-Bases object detection**

- Author: Fuchun Sun, Huaping Liu, Yuning Jiang, Lei Li, Jianbo Shi

- Year: 2020

- Cited by: 74

- Link: <https://ieeexplore.ieee.org/abstract/document/9123553/>

- Lí do:

* Do có nhiều trích dẫn.
* Một khuôn khổ chính xác, linh hoạt và hoàn toàn không có mỏ neo để phát hiện đối tượng. Trong khi hầu hết tất cả các máy dò đối tượng hiện đại đều sử dụng các mỏ neo được xác định trước để liệt kê các vị trí, quy mô và tỷ lệ khung hình có thể có cho việc tìm kiếm các đối tượng, hiệu suất và khả năng tổng quát của chúng cũng bị giới hạn trong việc thiết kế các mỏ neo

**2. Object detection in optical remote sensing images**

- Author: K Li, G Wan, G Cheng, L Meng, J Han

- Year: 2020

- Cited by: 54

- Link: <https://www.sciencedirect.com/science/article/pii/S0924271619302825>

- Lí do: Tập dữ liệu chứa 23.463 hình ảnh và 192.472 phiên bản, bao gồm 20 lớp đối tượng. Bộ dữ liệu DIOR được đề xuất (1) có quy mô lớn trên các danh mục đối tượng, trên số cá thể đối tượng và trên tổng số hình ảnh; (2) có một loạt các biến thể về kích thước đối tượng, không chỉ về độ phân giải không gian, mà còn về khía cạnh thay đổi kích thước giữa các lớp và giữa các đối tượng; (3) có sự thay đổi lớn khi hình ảnh thu được với các điều kiện hình ảnh khác nhau, thời tiết, mùa và chất lượng hình ảnh; và (4) có độ tương đồng giữa các lớp cao và tính đa dạng trong lớp.

**3. Recent advances in deep learning for object detection**

- Author: Xiongwei Wu, Doyen Sahoo, Steven C.H. Hoi

-Year: 2020

- Cited by: 30

- Link: <https://www.sciencedirect.com/science/article/pii/S0925231220301430>

- Lí do: Cuộc khảo sát thành ba phần chính: (i) các thành phần phát hiện, (ii) chiến lược học tập và (iii) ứng dụng & điểm chuẩn. Trong cuộc khảo sát, chúng tôi đề cập chi tiết đến nhiều yếu tố ảnh hưởng đến hiệu suất phát hiện, chẳng hạn như kiến ​​trúc máy dò, học tập tính năng, tạo đề xuất, chiến lược lấy mẫu, v.v. Cuối cùng, chúng tôi thảo luận về một số hướng trong tương lai để tạo điều kiện và thúc đẩy nghiên cứu trong tương lai để phát hiện đối tượng trực quan với học sâu.

**4. A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection**

- Author: Zhaowei Cai, Quanfu Fan, Rogerio S.Feris, Nuno Vasconcelos

- Year: 2016

- Cited by: 916

- Link: <https://link.springer.com/chapter/10.1007/978-3-319-46493-0_22>.

- Lí do: Use A unified deep neural network, is called CNN (MS-CNN), is proposed for fast multi-scale object detection. The unified network is learned end-to-end, by optimizing a multi-task loss. Feature upsampling by deconvolution is also explored, as an alternative to input upsampling, to reduce the memory and computation costs. State-of-the-art object detection performance, at up to 15 fps.

**5. Cascade r-cnn: Delving into high quality object detection**

- Author: Zhaowei Cai, Nuno Vasconcelos

- Year: 2018

- Cited by: 709

- Link: <https://openaccess.thecvf.com/content_cvpr_2018/html/Cai_Cascade_R-CNN_Delving_CVPR_2018_paper.html>

- Lí do: A multi-stage object detection architecture, the Cascade R-CNN, is proposed to address problems with detection performance tends to degrade with increasing the IoU (intersection over union) because overfitting during training and inference-time mismatch between the IoUs. Experiments also show that the Cascade R-CNN is widely applicable across detector architectures, achieving consistent gains independently of the baseline detector strength.

**6. Feature pyramid networks for object detection**

- Author: Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie

- Year: 2017

- Cited: 4744

- Link: <https://openaccess.thecvf.com/content_cvpr_2017/html/Lin_Feature_Pyramid_Networks_CVPR_2017_paper.html>

- Lí do:

* Get many citations.
* Feature pyramids are a basic component in recognition systems for detecting objects at different scales. This paper exploits the inherent multi-scale, pyramidal hierarchy of deep convolutional networks to construct feature pyramids with marginal extra cost. A top-down architecture with lateral connections is developed for building high-level semantic feature maps at all scales, is called a Feature Pyramid Network (FPN), shows significant improvement as a generic feature extractor in several applications. Using a basic Faster R-CNN system, our method achieves state-of-the-art single-model results on the COCO detection.

**7. Focal loss for dense object detection**

- Author: Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, Piotr Dollar

- Year: 2017

- Cited by: 4535

- Link: <https://openaccess.thecvf.com/content_iccv_2017/html/Lin_Focal_Loss_for_ICCV_2017_paper.html>

- Lí do:

* Get many citations.
* Use R-CNN, the highest accuracy object to date. This paper proposes to address this class imbalance by reshaping the standard cross entropy loss such that it down-weights the loss assigned to well-classified examples. Focal Loss focuses training on a sparse set of hard examples and prevents the vast number of easy negatives from overwhelming the detector during training. To evaluate the effectiveness of their loss, they design and train a simple dense detector we call RetinaNet. The results show that when trained with the focal loss, RetinaNet is able to match the speed all existing state-of-the-art two-stage detectors.

**8. Searching for MobileNetV3**

- Author: Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, Quoc V. Le, Hartwig Adam

- Year: 2019

- Cited by: 339

- Link: <https://paperswithcode.com/paper/searching-for-mobilenetv3>

- Lí do:

* MobileNetV3 được điều chỉnh cho phù hợp với các CPU của điện thoại di động thông qua sự kết hợp của tìm kiếm kiến ​​trúc mạng phần cứng (NAS) được bổ sung bởi thuật toán NetAdapt và sau đó được cải thiện thông qua các tiến bộ kiến ​​trúc mới.
* Hai mô hình MobileNet mới để phát hành: MobileNetV3-Large và MobileNetV3-Small được nhắm mục tiêu cho các trường hợp sử dụng tài nguyên cao và thấp. Các mô hình này sau đó được điều chỉnh và áp dụng cho các nhiệm vụ phát hiện đối tượng và phân đoạn ngữ nghĩa.
* MobileNetV3-Large chính xác hơn 3,2% về phân loại ImageNet trong khi giảm độ trễ 20% so với MobileNetV2.
* MobileNetV3-Small chính xác hơn 6,6% so với kiểu MobileNetV2 có độ trễ tương đương.
* Phát hiện MobileNetV3-Large nhanh hơn 25% với độ chính xác gần giống như MobileNetV2 trên COCO detection
* MobileNetV3-Large LRASPP nhanh hơn 34% so với MobileNetV2 R-ASPP với độ chính xác tương tự đối với Cityscapes segmentation

**9. Objects as Points**

- Author: [Xingyi Zhou](https://arxiv.org/search/cs?searchtype=author&query=Zhou%2C+X), [Dequan Wang](https://arxiv.org/search/cs?searchtype=author&query=Wang%2C+D), [Philipp Krähenbühl](https://arxiv.org/search/cs?searchtype=author&query=Kr%C3%A4henb%C3%BChl%2C+P)

- Year: 2019

- Cited by: 251

- Link: <https://paperswithcode.com/paper/objects-as-points>

- Lí do:

* Trong bài báo này, chúng tôi có một cách tiếp cận khác. Chúng tôi mô hình hóa một đối tượng như một điểm duy nhất - điểm chính giữa của hộp giới hạn của nó. Máy dò của chúng tôi sử dụng ước tính điểm chính để tìm điểm trung tâm và hồi quy cho tất cả các thuộc tính đối tượng khác, chẳng hạn như kích thước, vị trí 3D, hướng và thậm chí cả tư thế.
* Phương pháp tiếp cận dựa trên điểm trung tâm của chúng tôi, CenterNet, có thể phân biệt từ đầu đến cuối, đơn giản hơn, nhanh hơn và chính xác hơn so với các công cụ phát hiện dựa trên bounding box.
* CenterNet đạt được sự cân bằng độ chính xác về tốc độ tốt nhất trên tập dữ liệu MS COCO, với 28,1% AP ở 142 FPS, 37,4% AP ở 52 FPS và 45,1% AP với thử nghiệm đa thang ở 1,4 FPS.
* Chúng tôi sử dụng cùng một cách tiếp cận để ước tính hộp giới hạn 3D trong điểm chuẩn KITTI và tư thế con người trên tập dữ liệu COCO keypoint.

**10. Context R-CNN: Long Term Temporal Context for Per-Camera Object Detection**

- Author: Sara Beery, Guanhang Wu, Vivek Rathod, Ronny Votel, Jonathan Huang

- Year: 2020

- Cited by: 2

- Link: <https://paperswithcode.com/paper/long-term-temporal-context-for-per-camera>

- Lí do:

* Giúp cải thiện khả năng nhận dạng trong các camera giám sát thụ động
* Context R-CNN, lập chỉ mục thành ngân hàng bộ nhớ dài hạn được xây dựng trên cơ sở mỗi máy ảnh và tổng hợp các tính năng theo ngữ cảnh từ các khung hình khác để tăng hiệu suất phát hiện đối tượng trên khung hình hiện tại .
* Áp dụng Context R-CNN cho hai cài đặt: (1) phát hiện loài bằng bẫy camera và (2) phát hiện phương tiện trong camera giao thông, cho thấy trong cả hai cài đặt rằng Context R-CNN dẫn đến tăng hiệu suất so với các đường cơ sở mạnh. Hơn nữa, cho thấy rằng việc tăng khoảng thời gian theo ngữ cảnh dẫn đến kết quả được cải thiện.