جامعة محمد بن زايد للذكاء الاصطناعي
MOHAMED BIN ZAYED UNIVERSITY OF ARTIFICIAL INTELLIGENCE

Core Courses Syllabi

CV706 - Neural Networks for Object Recognition and Detection

Title	Neural Networks for Object Recognition and Detection
Code	CV706
Loading	4 Credit-hours
Prerequisites	• A thorough understanding of the material covered in the following Masters course: 1. CV 703 Visual Object Recognition and Detection • Hands-on experience with Python and Pytorch or equivalent language/ library
Catalog Description	This course provides focused coverage of special topics on object recognition and detection. The students will develop skills to critique the state-of-the-art works on visual object recognition and detection. Moreover, students will be required to implement papers with the following aims: (1) reproduce results reported in the seminal research papers, and (2) improve the performance of the published works. This course builds upon concepts from Human and Computer Vision (CV701), Visual Object Recognition and Detection (CV702) and assumes familiarity with fundamental concepts in machine learning and optimization.
Goal	This PhD course aims to inculcate a deeper understanding of algorithms for image classification, object detection, instance segmentation, semantic segmentation and panoptic segmentation, so the students can become capable of doing research in these specialized topics and can implement these methods for building systems for real-world scene understanding. A significant goal of this course is to enhance students' teamwork skills by requiring them to participate in group projects and develop their communication and analytical skills by engaging them in reading group activities.
Contents	The course covers three modules: **(I)** Image Classification, **(II)** Object Detection, and **(III)** Image Segmentation.
Recommended Textbooks	I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
Recommended References & Supplemental Material	Relevant research papers, tech reports, and surveys for each topic, where needed, are identified in the teaching plan ahead. In addition, the following textbooks will be useful: 1. D. A. Forsyth, and J Ponce, *Computer vision: A Modern Approach, 2nd edition*, Pearson, 2003. 2. R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer Verlag, 2011.

Teaching Week	Topics
1	**Image Classification (Module I)** This module covers the following topics: large-scale image classification and fine-grained image classification. **Lecture** • Overview of image classification • Relevant papers and assigned reading: - A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks", *Advances in neural information processing systems*, 2012. - K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *International conference on learning representations*, 2016. - K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", *IEEE conference on computer vision and pattern recognition*, 2016. - J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks", *IEEE conference on computer vision and pattern recognition*, 2018. **Lab** • Discussion on choosing a relevant paper to implement for the project • Start Project-1 work
2	**Image Classification (contd.)** **Lecture** • Introduction to large-scale image classification • Reading group activity on selected papers related to "large-scale image classification": - A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks", *Advances in neural information processing systems*, 2012. - K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *International conference on learning representations*, 2016. **Lab** • Continue Project-1 work
3	**Image Classification (contd.)** **Lecture** • Continue reading group activity on selected papers related to "large-scale image classification": - K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", *IEEE conference on computer vision and pattern recognition*, 2016. - J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks", *IEEE conference on computer vision and pattern recognition*, 2018. **Lab** • Continue Project-1 work

Teaching Week	Topics
4	**Image Classification (contd.)** **Lecture** • Introduction to fine-grained image classification • Reading group activity on selected papers related to "fine-grained image classification": - T. Y. Lin, A. R. Chowdhury, and S. Maji, "Bilinear cnn models for fine-grained visual recognition", *IEEE international conference on computer vision*, 2015. - Z. Akata, S. Reed, D. Walter, H. Lee, and B. Schiele, "Evaluation of output embeddings for fine-grained image classification", *IEEE conference on computer vision and pattern recognition*, 2015. - H. Zheng, J. Fu, T. Mei, and J. Luo, "Learning multi-attention convolutional neural network for fine-grained image recognition", *IEEE international conference on computer vision*, 2017. **Lab** • Preparation of presentation on Project-1 work • Continue Project-1 work
5	**Image Classification (contd.)** **Assessment 1.1** • Presentation of the projects by different groups **Lab** • Peer review of project reports **Assessment 1.2** • In-class exam covering Module I – Image Classification
6	**Object Detection (Module II)** This module covers single-stage and two-stage object detectors for conventional imagery and aerial imagery. **Lecture** • Overview of object detection • Relevant papers and assigned reading: - R. Girshick, "Fast R-CNN", *IEEE international conference on computer vision*, 2015. - S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", *Advances in neural information processing systems*, 2015. - J. Dai, Y. Li, K. He, and J. Sun, R-FCN: Object detection via region-based fully convolutional networks", *Advances in neural information processing systems*, 2016. - W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector", *European conference on computer vision*, 2016. - J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", *IEEE conference on computer vision and pattern recognition*, 2016. **Lab** • Discussion on choosing a relevant paper to implement for the project • Start Project-2 work

Teaching Week	Topics
7	**Object Detection (contd.)** **Lecture** • Introduction to two-stage object detectors • Reading group activity on selected papers related to "two-stage object detectors": - R. Girshick, "Fast R-CNN", *IEEE international conference on computer vision*, 2015. - S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks", *Advances in neural information processing systems*, 2015. - J. Dai, Y. Li, K. He, and J. Sun, R-FCN: Object detection via region-based fully convolutional networks", *Advances in neural information processing systems*, 2016. **Lab** • Continue Project-2 work
8	**Object Detection (contd.)** **Lecture** • Introduction to single-stage object detectors • Reading group activity on selected papers related to "single-stage object detectors": - W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector", *European conference on computer vision*, 2016. - J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection", *IEEE conference on computer vision and pattern recognition*, 2016. **Lab** • Continue Project-2 work
9	**Object Detection (contd.)** **Lecture** • Applying object detectors on aerial images • Reading group activity on selected papers related to "object detection in aerial images": - G. S. Xia, X. Bai, J. Ding, Z. Zhu, S. Belongie, J. Luo, M. Datcu, M. Pelillo, and L. Zhang, "DOTA: A large-scale dataset for object detection in aerial images", *IEEE conference on computer vision and pattern recognition*, 2018. - J. Ding, N. Xue, Y. Long, G. S. Xia, and Q. Lu, "Learning RoI transformer for detecting oriented objects in aerial images", *IEEE conference on computer vision and pattern recognition*, 2019. **Lab** • Preparation of presentation on Project-2 work • Continue Project-2 work

Teaching Week	Topics
10	**Object Detection (contd.)** **Assessment 2.1** • Presentation of the projects by different groups **Lab** • Peer review of project reports **Assessment 2.2** • In-class exam covering Module II – Object Detection
11	**Image Segmentation (Module III)** This module covers: 1. Instance segmentation, 2. Semantic segmentation, and 3. Panoptic segmentation. **Lecture** • Overview of image segmentation • Relevant papers: - K. He, G. Gkioxari, P. Dollár, R. and Girshick, "Mask R-CNN", *IEEE international conference on computer vision*, 2017. - S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation", *IEEE conference on computer vision and pattern recognition*, 2018. - J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", *IEEE conference on computer vision and pattern recognition*, 2015. - L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, A.L., "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs", *IEEE transactions on pattern analysis and machine intelligence*, 2016. - H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network", *IEEE conference on computer vision and pattern recognition*, 2017. - A. Kirillov, K. He, R. Girshick, and P. Dollár, "Panoptic segmentation", *IEEE conference on computer vision and pattern recognition*, 2019. - A. Kirillov, R. Girshick, K. He, and P. Dollár, "Panoptic feature pyramid networks", *IEEE conference on computer vision and pattern recognition*, 2019. **Lab** • Discussion on choosing a relevant paper to implement for the project • Start Project-3 work

Teaching Week	Topics
12	**Image Segmentation** **Lecture** • Instance segmentation • Reading group activity on selected papers related to "instance segmentation": - K. He, G. Gkioxari, P. Dollár, R. and Girshick, "Mask R-CNN", *IEEE international conference on computer vision*, 2017. - S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation", *IEEE conference on computer vision and pattern recognition*, 2018. **Lab** • Continue Project-3 work
13	**Image Segmentation** **Lecture** • Semantic segmentation • Reading group activity on selected papers related to "semantic segmentation": - J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", *IEEE conference on computer vision and pattern recognition*, 2015. - L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, A.L., "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs", *IEEE transactions on pattern analysis and machine intelligence*, 2016. - H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network", *IEEE conference on computer vision and pattern recognition*, 2017. **Lab** • Continue Project-3 work
14	**Image Segmentation** **Lecture** • Panoptic segmentation • Reading group activity on selected papers related to "panoptic segmentation" - A. Kirillov, K. He, R. Girshick, and P. Dollár, "Panoptic segmentation", *IEEE conference on computer vision and pattern recognition*, 2019. - A. Kirillov, R. Girshick, K. He, and P. Dollár, "Panoptic feature pyramid networks", *IEEE conference on computer vision and pattern recognition*, 2019. **Lab** • Preparation of presentation on Project-3 work • Continue Project-3 work
15	**Image Segmentation** **Assessment 3.1** • Presentation of the projects by different groups **Lab** • Peer review of project reports **Assessment 3.2** • In-class exam covering Module III – Image Segmentation