Trường Đại học Bách Khoa - ĐHQG Tp.HCM
Khoa: Khoa Khoa học và Kỹ thuật Máy tính
Khoa/Bộ môn quản lý MH: Khoa học Máy tính

Tp.HCM, ngày ..... tháng ..... năm .....

Đề cương môn học Sau đại học

HỌC SÂU VÀ ỨNG DỤNG TRONG THỊ GIÁC MÁY TÍNH (DEEP LEARNING AND ITS APPLICATIONS IN COMPUTER VISION)

Mã số MH: CO5255

Số tín chỉ:	**Tc (LT.BT&TH.Tự Học): 3**						**ECTS: 6**				
Số tiết -Tổng:	**75**	LT:	**30**	BT:	**0**	TH:	**0**	ĐA:		BTL/TL:	**45**
Đánh giá:	Thuyết trình:		20%								
	Bài tập lớn:		30%								
	Thi cuối kỳ:		50%								
- Môn tiên quyết:	- Học máy và Ứng dụng									MS: 055241	
- Môn học trước:											
- Môn song hành:											
- CTĐT ngành (Mã ngành):	Khoa Học Máy Tính (8480101)										
- Ghi chú khác:											

1. <u>Mục tiêu môn học:</u>

Môn học này cung cấp những kiến thức nền tảng và những thành tựu của kỹ thuật học sâu trong lĩnh vực thị giác máy tính. Mục tiêu chính của môn học là cung cấp cho học viên nền tảng về mặt lý thuyết và thực tế để có thể sử dụng và phát triển các kiến trúc mạng nơ-ron phức tạp để giải quyết các bài toán khó trong lĩnh vực thị giác máy tính như: phân loại ảnh, phát hiện đối tượng, nhận diện đối tượng, sinh ảnh.

Aims:

This course provides a thorough understanding of the fundamental concepts and recent advances in deep learning. The main objective is to provide students practical and theoretical foundations to use and develop deep neural architectures to solve challenging computer vision tasks such as: image classification, object detection, object recognition, image generation.

2. <u>Nội dung tóm tắt môn học:</u>

Để học môn học này, học viên cần phải có kiến thức nền tảng là xử lý ảnh, thị giác máy tính và học máy. Do đó, nội dung của môn học được chia làm 2 phần. Phần đầu giúp học viên ôn lại các khái niệm căn bản của xử lý ảnh, thị giác máy tính, và học máy, bao gồm: mạng nơ-ron học sâu. Phần thứ hai sẽ giới thiệu các kiến trúc mạng đặc trưng cho từng bài toán trong xử lý ảnh và thị giác máy tính như: phân loại ảnh, phát hiện và nhận diện đối tượng, phân đoạn ảnh, sinh ảnh.

Course outline:

To take this course, students are required to have backgrounds on image processsing, computer vision and machine learning. Therefore, the content of this course is divided into two part. The first part provides the review of concepts on image processing, computer vision and machine learning (e.g: deep neural network). The second part introduces specialized network architectures for each challenging tasks on image processing and computer vision, such as: image classification, object detection and recognition, image segmentation, image generation.

3. Tài liệu học tập:

1. Giáo trình (Textbook)

- Ian Goodfellow, Yoshua Bengio and Aaron Courville, 2016. **Deep Learning**, MIT Press, http://www.deeplearningbook.org/ (**chính**)

- Ahmed Fawzy Gad, 2018. **Practical Computer Vision Applications Using Deep Learning with CNNs**, Apress, (ISBN: 9781-484241677)

2. Tài liệu tham khảo (References)

- Yuxi Liu, 2017. **Python Machine Learning By Example: The easiest way to get into machine learning**, Packt Publishing , (ISBN: 978-1783553112)

- **Huang**, K., **Hussain**, A., **Wang**, Q.-F., **Zhang**, R. (Eds.), 2019. **Deep Learning: Fundamentals, Theory and Applications**, Springer, (ISBN: 978-3030060732)

- Francois Chollet, 2017. **Deep Learning with Python (1st Edition)**, Manning Publications, (ISBN**:** 978-1617294433)

- Các bài báo khoa học tại các hội nghị và tạp chí đầu ngành (the state-of-the-art research papers)

4. Các hiểu biết, các kỹ năng cần đạt được sau khi học môn học:

STT	Chuẩn đầu ra môn học (CĐRMH)	Công cụ đánh giá CĐRMH	Đóng góp CĐR Chương trình (CĐRCT)		
			Ứng dụng	Nghiên cứu	
CĐRMH.1	Hiểu được các kiến thức nền tảng của học sâu sử dụng trong xử lý ảnh và thị giác máy tính	Thuyết trình, Bài tập lớn, Thi cuối kỳ	a, f		1.2, 1.3, 3.1
CĐRMH.2	Có khả năng thiết kế và hiện thực các kiến trúc mạng khác nhau cho bài toán về xử lý ảnh và thị giác máy tính.	Bài tập lớn, Thi cuối kỳ	a, b, e, f		1.3, 3.1

Learning outcomes:

No.	Course learning outcomes (CLO)	CLO assessment	Matching with PLO		
			Coursework	Research	
L.O.1	An understanding of the fundamental concepts of deep learning used in digital image processing and computer vision	Presentation, Assignments, Final Examination	a, f		1.2, 1.3, 3.1
L.O.2	Able to design and implement various deep neural network architectures for digital image processing and computer vision problems.	Assignments, Final Examination	a, b, e, f		1.3, 3.1

Bảng ánh xạ chuẩn đầu ra môn học và chuẩn đầu ra chương trình ứng dụng:

	Chuẩn đầu ra của chương trình (CĐRCT)										
Chuẩn đầu ra môn học (CĐRMH)	a	b	c	d	e	f	g	h	i	j	k
CĐRMH.1	✓					✓					
CĐRMH.2	✓	✓			✓	✓					

Bảng ánh xạ chuẩn đầu ra môn học và chuẩn đầu ra chương trình nghiên cứu:

	Chuẩn đầu ra của chương trình (CĐRCT)										
Chuẩn đầu ra môn học (CĐRMH)	a	b	c	d	e	f	g	h	i	j	k
CĐRMH.1											
CĐRMH.2											

5. <u>Hướng dẫn cách học - chi tiết cách đánh giá môn học:</u>

1. Đây là môn học hướng đến cả lý thuyết và thực hành. Học viên cần chuẩn bị các tư liệu sau trong đầu cho môn học:

a. Về lý thuyết:

- Học viên cần trang bị tài liệu tham khảo đầy đủ cho mỗi chương.

- Đọc trước tài liệu cho mỗi buổi lên lớp. Thời gian trên lớp chủ yếu dành cho mở rộng khái niệm, liên hệ thực tiễn, và thảo luận.

b. Về thuyết trình:

- Học viên cần tìm hiểu và đọc kỹ các bài báo khoa học liên quan đến chủ đề đã chọn.

- Chuẩn bị slide và thuyết trình thử nhiều lần trước khi thuyết trình trên lớp.

c. Về bài tập lớn:

- Học viên cần có khả năng tổng hợp tài liệu liên quan đến đề tài bài tập lớn.

- Học viên cần trang bị các phương tiện lập trình cho môn học, cụ thể là Python và OpenCV.

- Hiện thực nội dung bài tập lớn.

Đánh giá sẽ dựa trên:

- Thuyết trình nội dung bài báo khoa học: 20%

- Bài tập lớn: 30%

- Thi cuối kỳ: 50%

Đánh giá sẽ dựa trên:

- Thuyết trình nội dung bài báo khoa học: 20%

- Bài tập lớn: 30%

- Thi cuối kỳ: 50%

Learning strategies & Assessment Scheme:

1. This course is designed to help students to strongly acquire both of the theoritical and the practical aspect of image and video processing. Student attending the course should prepare the following materials at the beginning of the course.

a. Theoretical parts:

- Student should prepare reference materials for each chapter before the lecture.

- Read reference materials for each session before attending class. Lecture hours are reserved for extending concepts, for relating to applications, and for attending discussion.

b. Presentation parts

- Students should figure out and read carefully research papers related to the chosen topic.

- Students should prepare slides and rehearsal before presenting in the class.

Assignments:

Student should prepare the ability to obtain/search related materials for assignments.

Students should prepare fundamental library/software for practice, especiallly Python and OpenCV.

Implementing the assignments.

Grading is based on:

- Presentation (state-of-the-art papers): 20%

- Assignments: 30%

- Final Examination: 50%

Grading is based on:

- Presentation (state-of-the-art papers): 20%

- Assignments: 30%

- Final Examination: 50%

6. Nội dung chi tiết:

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
1	Giới thiệu học máy và học sâu	- Giới thiệu về học máy và học sâu- Các khái niệm cơ bản trong học máy và học sâu	CĐRMH.1	Chương 1, 5
1	Huấn luyện mạng nơ-ron học sâu	- Mạng nơ ron học sâu- Các kỹ thuật regularization cho học sâu. - Các kỹ thuật tối ưu cho huấn luyện mô hình học sâu	CĐRMH.1, CĐRMH.2	Chương 6, 7, 8
1	Mạng nơ-ron tích chập	- Toán tử tích chập- Lớp tích chập- Lớp pooling	CĐRMH.1, CĐRMH.2	Chương 9

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
1	Hiểu và trực quan hóa mạng CNN	- Trực quan hóa các bộ lọc- Sinh ảnh tối đa hóa điểm số nhận dạng- Trực quan dữ liệu gradient	CĐRMH.1, CĐRMH.2	Zeiler, M.D. and Fergus, R., 2014, September. Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
1	Các kiến trúc mạng rút trích đặc trưng cơ bản cho ảnh	- AlexNet- GoogLe Net- VGG- ResNet	CĐRMH.1, CĐRMH.2	Tham khảo các bài báo khoa học bên dưới:1. Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).2. Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.3. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*(pp. 1-9).4. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
1	Phát hiện và nhận dạng đối tượng	- Giới thiệu bài toán phát hiện và nhận diện đối tượng- Các kiến trúc mạng đặc thù: SSD, YOLO- Các kỹ thuật hậu xử lý	CĐRMH.1, CĐRMH.2	[1] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C., 2016, October. Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.[2] Redmon , J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
1	Phân đoạn đối tượng trong ảnh	- Giới thiệu bài toán phân đoạn ảnh- Các kiến trúc mạng đặc thù: FCN, U-Net, PSPNet- Các kỹ thuật hậu xử lý	CĐRMH.1, CĐRMH.2	[1] Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).[2] Ronneberger, O., Fischer, P. and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.[3] Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J., 2017. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2881-2890).
1	Mạng nơ-ron hồi quy và ứng dụng trong bài toán theo vết đối tượng	- Mạng nơ-ron hồi quy- Mạng nơ-ron hồi quy hai chiều- Ứng dụng trong bài toán theo vết đối tượng	CĐRMH.1, CĐRMH.2	Chương 10
1	Autoencoders và những biến thể của nó	- Undercomplete Autoencoders- Regularized Autoencoders- Denoising Autoencoders	CĐRMH.1, CĐRMH.2	Chương 14
1	Mô hình sinh ảnh dựa trên học sâu	- Máy Boltzmann- Máy Restricted Boltzmann- Deep Belief Networks- Generative Adversarial Networks	CĐRMH.1, CĐRMH.2	Chương 20

7. <u>Giảng viên tham gia giảng dạy:</u>

CBGD chính:	TS. Nguyễn Hồ Mẫn Rạng
CBGD tham gia:	TS. Lê Thành Sách

XÁC NHẬN CỦA HỘI ĐỒNG XÂY DỰNG CHƯƠNG TRÌNH ĐÀO TẠO VÀ KHOA

Tp. Hồ Chí Minh, ngày ..... tháng ..... năm .....

GIẢNG VIÊN LẬP ĐỀ CƯƠNG

TS. Phạm Hoàng Anh