CSIE 2019 Fall - Computer Vision and Deep Learning (電腦視覺及深度學習)

Instructor: Jenn-Jier James Lien (連 震 杰), Professor, jjlien@csie.ncku.edu.tw, Ext. 62540

Website: http://robotics.csie.ncku.edu.tw/course.html

Class Time/Location: 09:10 ~ 12:00 Thursday / room CSIE65405

Spoken Language: English **Prerequisites:** C, C++ or Python

Syllabus: (W: Week - Ch: Szeliski's book chapter)

Computer Vision	01-01	Introduction to **industry 4.0** - Intelligent robotics and automation: Sensors, machine vision, deep learning, big data and IoT (Internet of Things).
	01-02.1, 03.6	Sensor - **Camera model**: Geometric transformations between 2D and 3D.
	02-06.3	Sensor - **Camera calibration**: Optimization process and **AR** (Augmented reality).
	03-11.1, 11.3 12.1, 12.2	Sensor - **3D**: 1) **Stereo**, 2) ToF (Time-Of-Flight, Kinect 2, SoftKinetic), and 3) Structured light (Kinect 1, DLP projector).
	04-B.	From **AI** (artificial intelligence) to **ML** (machine learning), to **DP** (deep learning): -From **Bayes' Rule** (posterior probability) to Gaussian model, to similarity measure (likelihood probability: Mahalanobis distance, SSD (sum of squared differences) and correlation (or **pattern matching**)), to PCA (linear combination). -From supervised、unsupervised、semi-supervised learning, to DP - Reinforcement
	05-04.1	**SIFT** (and brief HOG): Feature extraction. **OpenCV assignment 1.**
	06-05.3.1, 12.6.4	**Background subtraction/modeling**: Real-time motion detection using **GMM**.
	07-08.1	**Optical flow**: Real-time motion estimation (or feature tracking) for facial expression extraction.
Machine Learning	08-14.2, A.1	**PCA** (principal component analysis, dimensionality reduction, domain knowledge): Real-time face detection (eigenfeature), face recognition (eigenface) and facial expression recognition (eigenflow). And **LDA** (linear discriminant analysis): Linear classification for face recognition.
	09-	**Midterm exam.** (from W01~07)
	10-14.1	**AdaBoost**: Face detection.
	11-14.1	**SVM** (support vector machines): Non-linear classification.
	12-03.7	**VQ** (vector quantization): Clustering and K-means. And **HMM** (discrete-time hidden Markov model): Facial expression recognition in video.
Deep Learning	13-	Deep learning: LeNet, AlexNet, VGG16 and ResNet **OpenCV assignment 2.**
	14-	Deep learning: Faster R-CNN
	15-	Deep learning: RetinaNet
	16-	Deep learning: Reinforcement learning
	17-	Deep learning: Generative Adversarial Network
	18-	**Final exam.** (from W01~17)

Grading: Assignment x 2: 25% x 2 = 50%, Exam x 2: 25% x 2 = 50%.

Textbooks/Reference Books:

1. Textbook: Class lecture notes.
2. Ref1. Computer Vision: Algorithms and Applications by Richard Szeliski, Springer, 2010.
3. Ref2. Learning OpenCV, Computer Vision with the OpenCV Library by Gary Bradski and Adrian Kaebler, O'Reilly, 2008.
4. Ref3. Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville, MIT Press, 2016.

課程概述：

這門課除了讓學生能深入了解電腦視覺、機器學習及人工智慧-深度學習的理論知識，與分析深度學習的原理是如何結合人工智慧及電腦視覺發展而來的相互關係外,技術功能面會以授課老師多年的產學合作經驗來舉實際的範例解釋。課程會先教如何從 2D 影像重建 3D 物體及增擬實境的基本電腦視覺技術開始，接著就會傳授電腦視覺基本但實用的技術，包括即時偵測、追蹤及辨識系統的設計。再來藉由機器學習的連接帶入深度學習領域，教授如何藉由深度學習的原理來開發更好的即時偵測、追蹤及辨識技術來解決實際的問題。本課程期待培養學生於電腦視覺、機器學習及深度學習領域技術設計及整合實作的能力，透過作業實作來建立學生獨立研究、設計及創新的能力，並可把所學的理論基礎應用到工業界的實務面。

教學目標:

培養學生具備以下的基本知識及實作能力：

1）具備基本的電腦視覺、機器學習及深度學習理論基礎；

2）融匯貫通深度學習的原理是如何結合人工智慧及電腦視覺發展而來的；

3）具備以電腦視覺及深度學習的技術來解決即時偵測、追蹤及辨識的生活上實際問題。