Technical Talk Report of AI Deep Learning: Blurring the Line between Real and the Fake

by Ir. Dr Mui Kai Yin and Ir. Dr Lee Choo Yong

Ir. Dr Mui Kai Yin is currently the deputy chairman in Electronic Engineering Technical Division (eETD).

Ir. Dr Lee Choo Yong is currently the chairman in Electronic Engineering Technical Division (eETD).

Technical Talk:
AI Deep Learning: Blurring the Line between Real and the Fake

Synopsis
Deep learning enables one to generate high-quality and realistic audio-visual content. The rise of synthetic media will inevitably shift the paradigm in media creation and consumption. In this talk, speaker will share recent advances of generative adversarial networks and its various applications.

First, speaker will present a method that can edit a target portrait footage by taking a sequence of audio as input to synthesize a photo-realistic video. Speaker will then show the possibility of decomposing objects in a single image for creative content creation. Speaker will also talk about how deep generative prior can be exploited to achieve flexible image restoration and manipulation. While synthetic media technologies will significantly accelerate creative expression, it will also bring with them challenges around truth and identity provenance. Speaker will discuss how we may protect ourselves from malicious usages of such technologies.

Technical talk entitled "AI Deep Learning: Blurring the Line between Real and the Fake" was presented by Dr. Loy Chen Change via Zoom virtual platform hosted by IEM eETD and IEM Penang Secretariat on 3rd December 2020. The technical talk was organized by Electronic Engineering Technical Division (eETD). Deputy Chairman of IEM eETD, Ir. Dr. Mui Kai Yin kick start the talk by providing a brief introduction about the talk, follow with introducing and welcome the speaker Dr. Loy to begin his sharing.
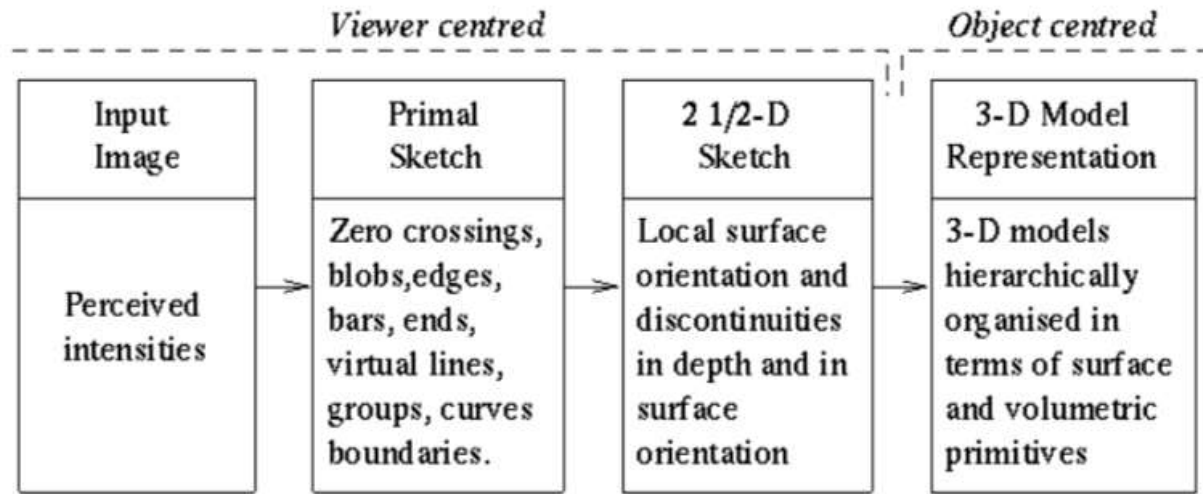
Following are the key contents of Dr. Loy's talk:
- Computer vision background
- A little history
- ImageNet and breakthrough
- Applications and success
- Fundamentals of machine learning
  - Why learning?
  - Statistical learning

Dr. Loy start off with the introduction of "deep learning in computer vision", the history and evolution of this technical domain, and shared the applications and successes. He said Deep learning enables breakthroughs in artificial intelligence (AI) voice recognition, facial recognition, image recognition, game playing, autonomous driving, robo-advisor, and natural language processing.

He presented many practical examples and applications, including identity authentication, entertainment (example: change body shape), image and video editing, and etc.

He emphasized one of David Marr's most important contributions that was in the level of representation and algorithm where Marr proposed a representation framework for vision as following (refer to picture below), with Marr vision, prima sketch, 2.5D sketch, 3D model, that lead to many subsequent development works.

	Viewer centred		Object centred
Input Image	Primal Sketch	2 1/2-D Sketch	3-D Model Representation
Perceived intensities	Zero crossings, blobs, edges, bars, ends, virtual lines, groups, curves boundaries.	Local surface orientation and discontinuities in depth and in surface orientation	3-D models hierarchically organised in terms of surface and volumetric primitives

He also quoted the real time face detection is one of the interesting works discovered by Viola & Jones. (Refer to picture below).

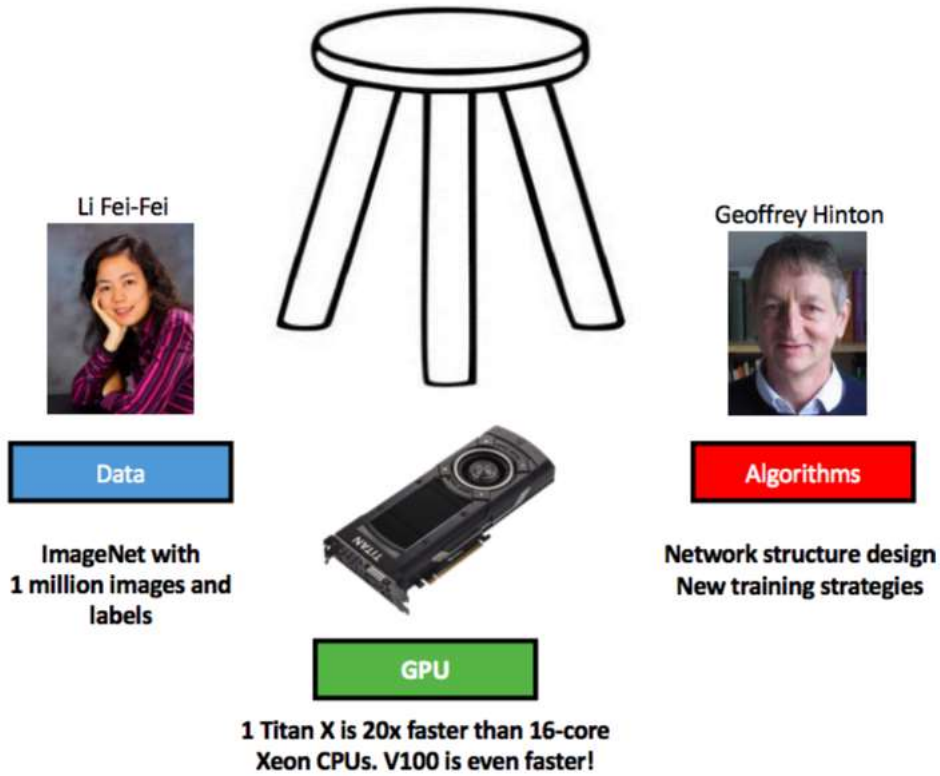Viola-Jones detection - a window of the target size is moved over the input image

- Simple rectangle features (Haar features) in constant time on integral images
- Learning weak classifiers by boosting
- Classifier cascade provides a focus-of-attention mechanism

Haar Feature that looks similar to the bridge of the nose is applied onto the face

Viola and Jones, Rapid Object Detection Using a Boosted Cascade of Simple Features, CVPR 2001

Dr. Loy also talked about the history of deep learning, Neuro network, LeNet, and related developments. What leads to breakthrough? He mentioned that Geoffrey Hinton summarized the findings up to today into the following 4 points:

1. Our labeled database were thousand times too small.
2. Our computers were millions of times too slow.
3. We initialize weights in a stupid way.
4. We used a wrong type of non-linearity.

Li Fei-Fei

Geoffrey Hinton

Data

Algorithms

ImageNet with
1 million images and
labels

Network structure design
New training strategies

GPU

1 Titan X is 20x faster than 16-core
Xeon CPUs. V100 is even faster!

The key application and success of deep learning for AI including face recognition, image recognition, natural language processing, robo-advisor, finance trading bot, autonomous driving, game playing and voice recognition. (Refer to picture below).

Facial Recognition

Voice Recognition

Image Recognition

Game Playing

Deep Learning
Enabling machines to acquire knowledge and skills
inducted from massive data

Autonomous Driving

Robo-advisor, Finance Trading Bot

Natural Language Processing

Dr. Loy also introduced the fundamental of machine learning, supervised learning with various image recognition, speech processing, natural language processing, robotic control and others,

Toward the end of the talk, Dr. Loy also shared some useful resources for deep learning, such as Pycoach (python), colab google, openmmlab.com, coursera for participants to explore.

The talk was attended by 33 participants comprises IEM members, and university students and others. A group photograph was taken toward the end of the evening talk.

AI Deep Learning: Blurring the Line between Real and the Fake
Group Photo