13 VISION AND AI

PJD's opening remarks

For a very long time, computer scientists have been interested in finding objects in images. Recall my story of visiting the Stanford AI lab in 1968 and learning that the most difficult part of getting the robot arm to stack blocks was to discern the sizes and positions of the blocks from a 2D video image. They had found a way that worked only with 3 identical blocks but could not be generalized to more blocks or blocks of different sizes.

Or recall the story of Luis Van Ahn at Carnegie-Mellon university who, in the early 2000s, resorted to a two-player game to accelerate the process of labeling images in the Google database.

About ten years later, Google had programs that labelled images; humans did not have to be involved. iPhone and Facebook had ways to identify the faces in photos and label them if they matched your contacts list.

What happened? In between a technology called Convolutional Neural Networks was perfected and greatly accelerated with graphics processing chips, or GPUs. This technology started in the 1980s and has progressed rapidly ever since. In fact, image recognition has been the Holy Grail pursuit that pushed the development of neural networks.

Today's speaker is Professor Mathias Kölsch of the Computer Science Department and MOVES. He came to NPS in 2005 and started a vision laboratory in MOVES. His students developed algorithms that found weapons in real time video feeds of moving crowds. He took three years of leave to join an AI company dealing with augmented reality and is now in the process of returning to NPS in time for Winter quarter.