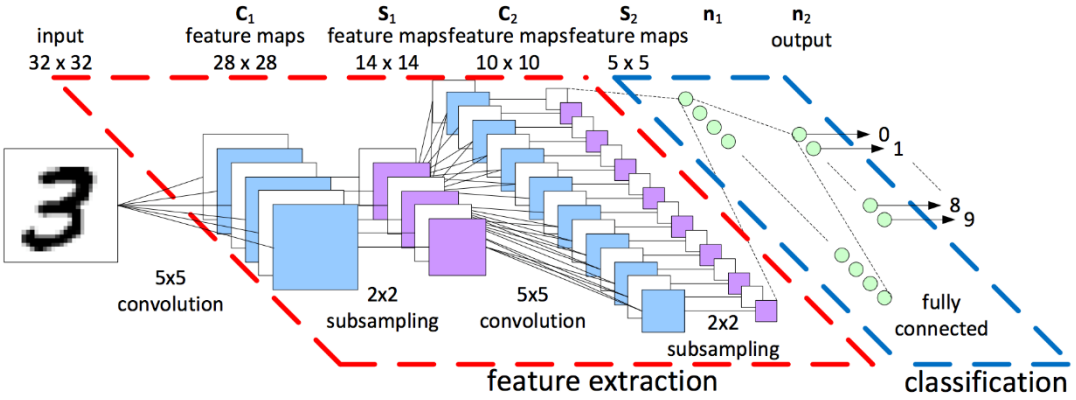
New Machine Learning Program Recognizes Handguns in even Low-Quality Video

Jingyue Shen UID:704797256

A team of computer scientists at the University of Granada in Spain has developed an automatic handgun detection alert system that can reliably detect handguns based on visual recognition and classification. The system is capable of catching guns from even low-quality YouTube footage in just under a quarter second. They plan to equip surveillance or control cameras with this system to lower the crime rate caused by guns.

In the last five years, deep learning in general and Convolutional Neural Networks (CNNs)**[1]** in particular have achieved superior results to all the classical machine learning methods in image classification, detection and segmentation in several applications. Instead of manually selecting features, deep learning CNNs automatically discover increasingly higher level features from data**[2].**With the success of those application, it might seem quite easy to develop the system. Experts can use convolutional neural networks(CNNs) **[1]** to train a model to recognize handguns, which is the mainstream in this area. One example of convolutional neural network is shown in the image below.

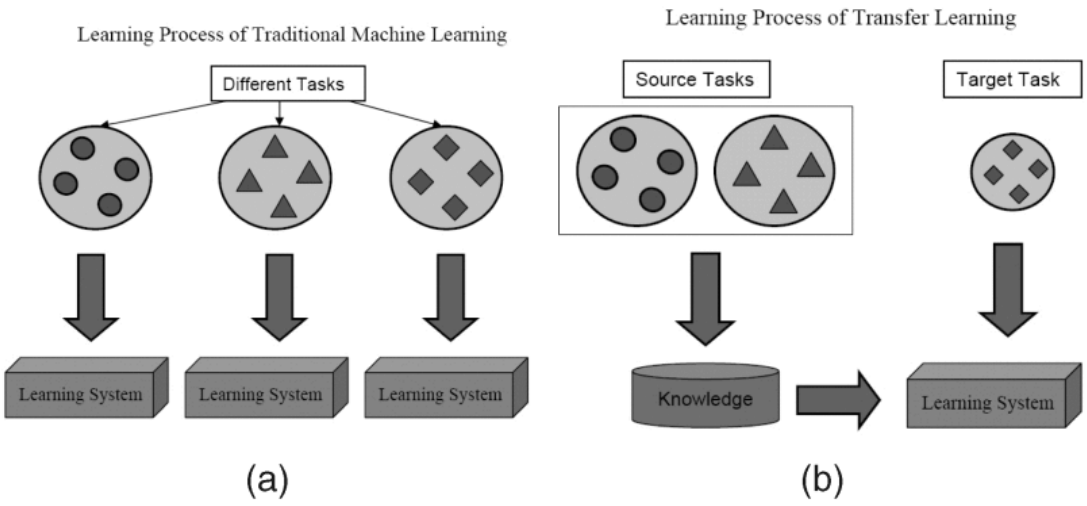


It has many complicated features like 3-D volumes of neurons, partial connectivity, different types of layers, etc**.[3]**

However, a proper training of deep CNNs, which contain millions of parameters, requires very large datasets, in the order of millions of samples, as well as High Performance Computing (HPC) resources, e.g., multi-processor systems accelerated with GPUs. And there are not so many available images of handguns for researchers. To solve this problem, researchers used a technique called transfer learning**[4]**.

Currently, a major assumption in many machine learning and data mining algorithms is that the training and future data must be in the same feature space and have the same distribution. However, in many real-world applications, this assumption may not hold: the training set might be too small to learn a useful model**[4]**. That is where transfer learning comes into play. Basically,

the idea of transfer learning is to “fine-tune” a pre-trained accurate model from a related task domain to obtain an accurate model for the task domain that we are interested in. Below is a picture explaining about the difference between traditional learning and transfer learning**[4]:**



For example, if we already have a good model trained for recognizing cars. It's possible to take that car recognition model and "fine tune" it using new images of trucks. The model for cars already contains a lot of semantic information about trucks because trucks are pretty similar to cars. So, through the fine tuning process, we can retrain the existing car model to be a truck model by teaching it the difference between a car and a truck.

In the researchers work, they used a VGG-16 based classification model pre-trained on the ImageNet dataset that can classify input images with respect to 1000 different object classes and fine-tuned on their own dataset of 3000 images of guns taken in a variety of contexts**[2].** Using transfer learning, they got a model that can reach accuracy between 90%-95% and eliminate almost all the false positive. Furthermore, the model can detect pistols even in low-quality videos under a quarter second**[2]**, which is quite good for real-world application.

I think their work is very impressive and has great potential to be applied to real life, especially in countries where the possession of gun is legal, such as U.S. As we all know, the crime rate caused by gun is quite high in U.S.. If we can apply the system to most of the surveillance cameras, numerous crimes related to gun might be prevented or under quick control due to the system’s fast reaction time. However, there is still one problem need to be solved. The system developed so far still issues significant set of false negatives(i.e. the system fails to detect the presence of pistols) due to low contrast or luminosity of the scene. Researchers need to figure out how to tackle this problem in their future work to obtain a more perfect model.

[1] ImageNet Classification with Deep Convolutional Neural Networks
http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[2] Automatic Handgun Detection Alarm in Videos Using Deep Learning
https://arxiv.org/pdf/1702.05147.pdf

[3] Convolutional Neural Networks Overview
ttp://cs231n.github.io/convolutional-networks/#overview

[4] A Survey on Transfer Learning
http://ieeexplore.ieee.org/document/5288526/