Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 115 (2017) 251–257

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on Advances in Computing & Communications, ICACC-2017, 22-24 August 2017, Cochin, India

An Improved Transfer learning Approach for Intrusion Detection

Alwyn Mathew*[a], Jimson Mathew[a], Mahesh Govind[b], Asif Mooppan[b]

[a]*Department of Computer Science and Engineering ,IIT Patna-801103, India*
[b]*Vuelogix Technologies Pvt Ltd, Bangalore-560102, India*

Abstract

Its crucial for financial systems to have sound security measures in place. For security reasons customers are not allowed to wear a helmet while using ATM(Automated Teller Machine). An automated helmet detection using ATM surveillance camera feed can help improve security significantly. Recently deep convolutional neural network (DCNN) have shown state of the art accuracy in object detection and localization. In this work, a pretrained Google's inception model have been used and have achieved an accuracy of 95.3% by training the model on a proprietary ATM surveillance dataset. Transferred information from inception model has been feed to multiple fully connected layers with drop outs to achieve better accuracy.

Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications.

Keywords: Transfer learning; Image classification; Deep learning; Video surveillance; Intrusion detection

1. Introduction

Human brain effortlessly and instantly divides images into simple and complex shapes, discrete objects and regions of interest, and categorizes these entities according to their perceived significance and semantic context[1]. Brain makes instant decisions as to whether a shape is a circle, a corner or a straight line, and within milliseconds we make mental associations as to what type of objects this compilation of shapes describes, and whether those objects themselves are components of a larger, more abstract and perhaps more meaningful discrete object category (such as a man with a Helmet, a man without a Helmet), rather than just the sum of their parts. Furthermore, a judgment is made as to whether those objects are worthy of our focus and attention in the overall image presented before us; the man in this picture (see Fig. 1) clearly the object of interest in this picture is the user, whereas the cash machine and the poster are part of the background. It is often desirable for digital images to be treated by the computer in a similar

* Alwyn Mathew. Tel.: +91-9645239304.
E-mail address: alwyn.pcs16@iitp.ac.in

1877-0509
Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications
10.1016/j.procs.2017.09.132

manner. The process by which we try to classify a digital image into different categories is of particular interest to many application.

Fig. 1: ATM surveillance image

The accurate recognition of intrusion at a point in time is the most challenging problem for systems that provide location-based services. For a better understanding of the situation, consider a situation where a user walks around with a helmet inside an ATM, a helmet detection system can be employed to analyse the video stream from the camera in real-time and determine the current context, the user is at a point in time. This can be useful for identifying users those who violate the rules (intrusion detection) and appropriate warning can be given remotely. There has been continuous research in the field of image recognition due to the challenges. Recognition methods analyse information from digital images. Digital images are obtained through a camera or a video camera and they store information as intensity values at each pixel [17]. A recognition method contains a database of the objects or the class of objects to be recognised. Here a class is a type of objects sharing some common features, for example a same horse at different poses [6, 13] or different brands of motorbikes [15] . The database needs to be obtained by training or learning objects from input images, either real-time or offline. Here real-time learning takes much less processing time than offline ones. Processing time is the time taken from the start of learning to the end of adding image to the database.

2. Related Work

Deep learning and convolution neural network (CNN) have revolutionized the way researchers approach problems such as computer vision, speech recognition and natural language processing that were nearly impossible for computers to understand. Deep learning algorithm powered by computation advances and availability of very large dataset have recently been shown to exceed human performance in visual tasks.

Two major barriers still prevail in deep learning. Firstly, computational requirement increased tremendously as models got deeper and huge amount of data are used to train these models. To overcome this hurdle, we went from CPU(Central Processing Unit) to parallel GPU(Graphics Processing Unit) computing which made training deep networks a breeze. Even though GPU price are dropping, high end GPUs capable of handling deeper models are costly. Secondly, while training and testing these data are pulled from the same feature space and same distribution where most machine learning methods works perfectly. But change in distribution makes the model fail and needs to be revamped for better results. Transfer learning enables us to transfer knowledge between task domains(see Fig. 2,3). Transfer learning permits us to use knowledge acquired from a pretrained model to a new task with minimum computation.

Standard feedforward neural network was one of the most popular primitive building blocks in machine learning. The computation complexity of these model restricted researchers and industries to take advantage of it. Introduction of convolutional neural network made a breakthrough in performance and had more learning capabilities than traditional neural network. Very soon researches figured out stacking more convolutional neural layers can give better models.