Review on Convolution Neural Network Based Image Recognition

Mrigana Walia[1], Sofia Hamid[2]

[1]Assistant Professor, [2]M.Tech Scholar

[1,2]Universal Group of Institutions, Mohali, Punjab, India

[1]mrigana9@gmail.com,[2]sofiahamid166@gmail.com

ABSTARCT

Image classification has become, and has a long history with it, a major challenge in machine vision. The challenge includes a wide range of images intra-class caused by color, size, conditions and shape. Big data of the labeled training images are needed and to prepare this big data, it consumes a lot of time and cost as for the purpose of training only. Classification of images arose to reduce the gap between computer vision and human vision by training the computer with the data. The grading of the image is achieved by differentiating the image into the prescribed category based on the vision content. The classification of images is a classic problem in the fields of image processing, computer vision, and machine learning. In this paper we use deep learning to study the classification of the images.

Keywords: Image classification, Deep Learning, Tensorflow, Convolutional Neural Networks, Auto-encoders, Machine Learning.

INTRODUCTION

Image classification has recently grown and is becoming a trend among technology developers, especially with the growth of data in various sectors of the industry, such as e-commerce, automotive, healthcare and gaming. Facebook uses the most obvious example of this technology. Facebook now can detect up to 98 percent accuracy in order to identify your face with only a few tagged images and classified it into your Facebook's album [1]. The technology itself almost beats the capacity of person in image detection or recognition. One of the main methods for this technology is deep learning. Deep learning falls under the Artificial Intelligence category, where it can act or think like a human being. The system

itself will normally be set with hundreds or perhaps thousands of input data to make the' training' session more efficient and rapid. Machine learning is also the commonly used systems for the classification of pictures [2]. Nevertheless, within machine learning there are still pieces that can be changed. The classification of images will therefore be occupied with the deep learning system. When it comes to Image Classification Machine Vision has its own meaning. This technology is capable of recognizing people in pictures, objects, locations, action and writing. The combination of artificial intelligence software and machine vision technology will achieve the image classification outstanding result. Image classification has become, and has a long history with it, a major challenge in machine vision. The challenge involves a wide range of photos intra-class due to color, scale, environmental conditions and form. Big data of labeled training images are needed and in order to prepare this big data, it consumes a lot of time and cost as only for the purpose of training.

DEEP LEARNING

Deep learning is a subfield of machine learning which attempts to use hierarchical architectures to learn high-level abstractions in data. It is a new technique and has been widely applied in conventional fields of artificial intelligence, such as semantic sorting, transfer learning, natural language processing, computer vision, and much more. There are three main reasons for the deep learning boom today: the dramatically increased chip processing capabilities (e.g. GPU units), the significantly reduced computing hardware costs, and the significant advances in machine learning algorithms. Deep Neural Networks (DNNs) have emerged as powerful models of machine learning which exhibit major differences from conventional image classification approaches. DNNs with deep architectures have the ability to learn complex models and allow the learning of powerful representations of objects without handling the built features. Deep networks have been shown to be effective for computer vision tasks, as they are able to extract correct features when discriminating together. Deep learning approaches were widely adopted by numerous researchers in recent ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competitions and achieved top accuracy levels[3]. Deep learning algorithms are classified into four categories:

(1) Convolutionary Neural Networks

(2) Restricted Boltzmann Machines

3) Auto-encoding and Sparse Coding.

Convolutionary Neural Networks (CNNs)

The Convolutionary Neural Networks (CNN) is one of the most popular approaches to deep learning, where multiple layers are robustly equipped. It has been found to be highly effective, and is also the most widely used in various applications for computer vision. In general, a CNN consists of three central neural layers, which are convolutionary layers, pooling layers, and layers that are completely linked. Different types of layers perform different functions. The network is educated in two stages: a forward stage and a backward stage. Firstly, the main objective of the forward stage is to represent the input image in each layer with the current parameters (weights and bias). The predictive output is then used with the ground truth labels to calculate the cost of the loss. Second, based on the cost of loss, the backward stage with chain rules measures the gradients of each parameter. All parameters are modified and prepared for the next forward calculation based on the gradients [4]. The network learning can be stopped after appropriate iterations of the forward and backward phases. Next, we'll first introduce the functions along with each layer's recent developments, then summarize the networks ' widely used training strategies. Finally, we present several well-known CNN models, derived models, and conclude with the current tendency in real applications to use these models.

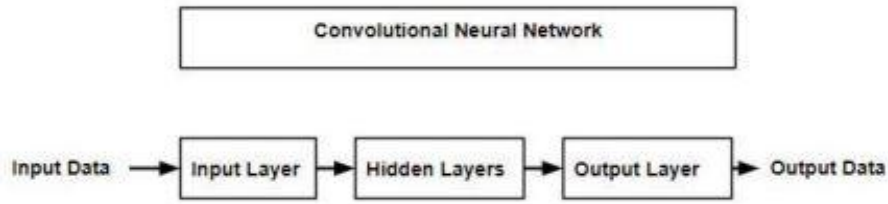

Fig1. Basic block diagram of CNN[4]

Restricted Boltzmann Machines (RBMs)

An RBM is a variant of the Boltzmann Machine, with the limitation that the visible units and hidden units have to form a bipartite graph. This limitation allows for more efficient training algorithms, in particular the contrastive divergence algorithm based on gradients. Deep Belief Networks (DBNs), Deep Boltzmann Machines (DBMs) and Deep Energy Models (DEMs) can be composed using RBMs as learning modules.

Autoencoder

The autoencoder is a particular type of artificial neural network used to learn efficient encoding. Instead of training the network to predict some target value Y given inputs X, an autoencoder is trained to recreate its own inputs X, hence the output vectors have the same dimensionality as the input vector. A deep autoencoder is often equipped with a back-propagation model, for example the method of conjugating gradient. Though often relatively efficient, if errors are present in the first few layers this model could become quite ineffective. This may lead the network to learn how to reconstruct the training data on average. A proper approach to removing this problem is pre-training the network with initial weights approximating the final solution [5]. There are also proposed variants of autoencoder to make the representation as "constant" as possible with regard to input changes.

Sparse coding

The purpose of sparse coding is to learn to define the input data by over-complete set of basic functions. Sparse coding has many advantages: (1) It can better recreate the descriptor using several bases and capturing the similarities between similar descriptors that share bases; (2)Sparsity helps the representation to capture salient image properties; (3) it is compatible with the biological visual system, which suggests that sparse signals are useful for learning; (4)the analysis of image statistics shows that image patches are sparse signals; (5)the patterns with sparse characteristics are more linearly separable.

THE BASIC PRINCIPLES OF DEEP LEARNING

Depth learning is the creation of artificial neural networks, which aim to mimic the human visual system in order to process visual information and thus make classification smarter. Depth learning is a method of machine learning by integrating low-level characteristics to construct more abstract high-level representations (attributes, groups, or functions) to discover distributed data features [6]. Unlike conventional shallow learning and deep learning in principle, two things are emphasized: 1) the structural layers of the learning model are emphasized, the model should usually have more than four layers of depth; 2) the characteristics of the learning process are clearly illustrated, which is more appropriate for the classification of high-level features created by integrating low-level features[8].

a) Differentiated depth structure

The goal is to provide a discriminative capacity to identify patterns, generally describing the subsequent distribution of data. Typical systems with a discriminative range include convolutionary neural networks. The convolutionary neural network (CNNs) is the first truly successful learning algorithm for training models in multilayer networks Inspired by the visual system structure, a transformation invariance feature can be obtained when a neuron with the same parameters is applied to different positions in the previous layer.

b) Generative depth structure

The structure defines the higher-order correlation properties of the data or the mutual probability distribution of the observed data and the classes concerned. A standard depth structure for the generative has a DBN network. The DBN consists of a series of Boltzmann Restricted Machine (RBM) modules consisting of a cascade. RBM is characteristic of a neural artificial network. The visible layer unit and secret layer unit are interconnected, but there is no connection within the layer. The secret units will obtain the higher order correlation of the visual units data, thus extracting the higher the number of DBN layers and the more abstract features are.

c) Mixed structure

This is in fact a combination of generative thinking and differentiated structure of thinking, that is, by means of a certain strategy to make the output of the generative structure with the ability to distinguish patterns. Yet sadly, so far there is no clear model, yet mixed structure will be very well established with the academic attention to the learning depth.

LITERATURE REVIEW

Rocha et al. [2010] Presented an automated method for the classification of fruit and vegetables by extracting the statistical characteristics. This research was done on the basis of super market data, and using k-means algorithm subtracts the history of images. Although a supermarket dataset of 2633 images was used to enforce the method, their work did not mention the accuracy of the classification result.

Zhang and Wu [2012] In their research work, a multi-class kernel support vector supporting machine (kSVM) classification system was proposed and 18 separate groups of fruit were

categorized. We applied the fruit images using the split-and-merge segmentation algorithm for the purpose of extraction of features.

Zawbaa et al. [2014] Derived an automated fruit identification scheme for the classification of several fruit varieties, which extracts two features by shape and color characteristics and transforms invariant features (SIFT). They have used support vector machine (SVM) and neighbouring k-nearest (k-NN) algorithms to classify the images. Their introduced model is developed using the 178 fruit images dataset and classifiers have achieved several accuracies for different fruit types, and so on.

Sang et al. [2014] The methodology for identifying and figuring fruits from images in cluttered greenhouses was presented and a two-stage approach for determining and counting fruits was created. Although the method is developed using a huge dataset of 28000 color images of fruits and plants using SVM as a classifier, the model performs an external correlation of 74.2 per cent of any linear improvement for a comprehensive dataset.

Shivaji et al. [2016] Classify fruit images by scale, consistency, color and health through the creation of an embedded based system. They have been followed to classify two steps which are color and edge detection but no classification algorithm and the success rate is stated in the method being proposed.

Naik and Patel [2017] Classification and grading of fruits are discussed in these research works over a small size of data set. Random forest classifier is used here and is then compared to k-NN and SVM only.

Jana et al. [9] Developed a technique for classifying different fruit forms using texture and color characteristics based on statistical color characteristics and co-occurrence matrix (GLCM) at the gray-level. SVMs are used in the approach to classify pictures of the fruits. Their work is done with two types, namely texture and colour, features but no data size has been mentioned, as well as the accuracy result with the classifier.

Hossain et al. [2018] Classify fruits for industrial applications use deep learning techniques. The proposed model is based on a deep learning model based on convolutionary neural network and pretrained VGG-16 (also called OxfordNet). Here clear images of the fruits and general images of the fruits using two datasets. A strong degree of precision is reached in this research work but they have not contrasted with other classifiers.

Astuti et al. [2018] suggested an artificial neural network comparison technique for classifying fruits using SVM based on their shapes. Fast Fourier Transform (FFT) is extracted in this research work and subsequently applied to the SVM-based identifier as input. No segmentation method is mentioned here and ANN provides 66.7 percent classification in training frames which means that all fruits are not properly classified by the ANN.

Sekar et al. [2018] Discussed a Fruit Identification approach to computer vision. Different fruit disease detection methods as well as their advantages and disadvantages were explored in their research work. Several models and features of the classifiers are discussed in their work but they did not propose a computer vision approach to classify or detect fruit or fruit diseases etc.

Andrej Jokie, Nikola Vukovic (2018) This literature paper mainly focuses on the extraction based on features. Their main goal is to reduce the computing power used for the ALPR system, and less should be the data used for the training. It also incorporates pre-processing segmentation, and they proposed a compressive sensing-based dimensionality reduction system.

Mohamed Yousef, Khaled F. Hussain, and Usama S. Mohammed(2018) Their emphasis in this literature review was on preventing loss of accuracy. They took advantage of the CTC loss feature to train their model. And introduced Highway networks for image classification. They also successfully implemented it on the recognition of the captcha, and on street images. They used lots of different datasets for various implementations.

Rayson Laroca , Evair Severo , Luiz A. Zanlorensi, Luiz S. Oliveira, Gabriel Resende Gonc ¸alves, William Robson Schwartz and David Menotti(2018) Their proposed model also performed some pre-processing works, such as segmentation They implemented Convolutionary Neural Networks to classify the image. We also developed two techniques such as inverted license plates and flipped characters for augmentation of approach. They used a dataset called UFPR-ALPR which contains more than 150 videos with 4500 frames of vehicles captured.

Rayson Laroca, Luiz A. Zanlorensi, Gabriel R. Gonc¸alves, Eduardo Todt, William Robson Schwartz, David Menotti(2019) This literature study introduces an independent layout ALPR system based on state-of - the-art YOLO object detector. They have chosen a single methodology for detecting and classifying layouts in post-processing methods. They have used networks to train models using images from various datasets and with additional augumentation techniques.

Satadal Saha(2019) This study of literature focuses on image extraction of features. Preprocessing approaches such as binarization, localisation, segmentation have been used. Multilayer perceptron (MLP) has been introduced as a classifier, and Quad Tree Based Longest Run (QLTR) is used for network testing. And get the characters it predicts.

PROPOSED WORK

Classification is a systematic arrangement in groups and categories based on its features. Image classification came into existence for decreasing the gap between the computer vision and human vision by training the computer with the data. The image classification is achieved by differentiating the image into the prescribed category based on the content of the vision. In this Research work, we explore the study of image classification using deep learning. Convolutionary neural networks (CNNs) have been commonly utilized in automated picture recognition systems. In most cases, features from the top layer of CNN are used for classification; however, those features may not contain enough useful information to correctly predict the image. In certain instances, characteristics of the lower layer have greater unequal influence than that of the upper layer. Therefore, extending the characteristics of a particular layer only to classification appears to be a mechanism that will not completely exploit the inherent differential influence of CNN studied. This intrinsic property contributes to the need to combine functionality from different levels. To fix this problem, we propose a method of integrating multi-layer features in CNN models.

Objectives:

1) Take a Database of images

2) Pre-processing and cleaning of images dataset

3) Study different image recognition systems based on deep learning.

4) Performance evaluation on the basis of Convolutional neural network technique.

METHODOLOGY

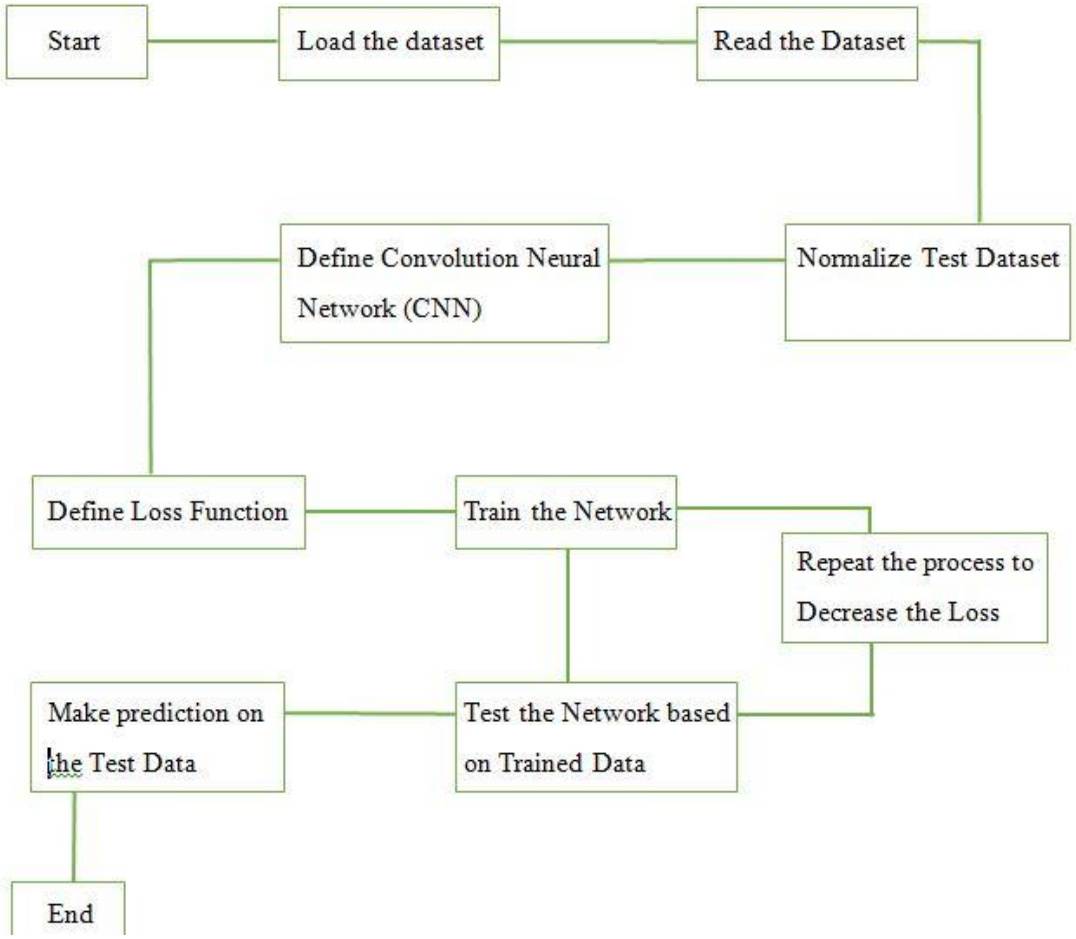

Fig2. Flow chart of the methodology

CONCLUSION AND FUTURE WORK

Recently, image classification is growing and becoming a trend among technology developers, particularly with data growth in various sectors of industry such as e-commerce, automotive, healthcare, and gaming. Deep learning is one of the leading methods for that technology. Deep learning falls under the Artificial Intelligence category, where it can act or think like a human being. The system itself will normally be set with hundreds or perhaps thousands of input data to make the' training' session more efficient and rapid. It begins by giving a kind of' training' with all the input data[7]. Machine learning is also the commonly used systems for the classification of pictures. Nevertheless, within machine learning there are still pieces that can be changed. The classification of images will therefore be dedicated to the deep learning system.

REFERENCES

[1] Rocha, Automatic fruit and vegetable classification from images, Computers and Electronics in Agriculture. 70 (1) (2010) 96–104.

[2] Zhang, (2012). Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine Sensors

[3] Zawbaa, (2014). Automatic Fruit Classification Using Random Forest Algorithm. 2014 14th International Conference on Hybrid Intelligent Systems, HIS 2014.

[4] Md. Abu Bakr Siddique “Study and Observation of the Variations of Accuracies for Handwritten Digits Recognition with Various Hidden Layers and Epochs using Convolutional Neural Network” - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-basic-block-diagram-of-CNN_fig1_330872480 [accessed 8 Jun, 2020]

[5] U. Shivaji R, Automated Fruit Classification System. International Journal of Advanced Research in Computer and Communication Engineering, vol. 5, issue 7, July (2016).

[6] Naik, Sapan & Patel, Bankim. (2017). Machine Vision based Fruit Classification and Grading - A Review. International Journal of Computer Applications.

[7] Jana, 2017: Automatic Fruit Recognition from Natural Images using Color and Texture Features.

[8]Hossain,"Automatic Fruit Classification Using Deep Learning for Industrial Applications." IEEE Transactions on Industrial Informatics 15.2 (2018): 1027-1034.

[9] Astuti, et al. "Automatic fruit classification using support vector machines: a comparison with artificial neural network." IOP Conference Series: Earth and Environmental Science. Vol. 195. No. 1. IOP Publishing, 2018.

[10] Raja Sekar L, "Fruit Classification System Using Computer Vision: A Review" International Journal of Trend in Research and Development (IJTRD), ISSN: 2394-9333, Volume-5 | Issue-1, February 2018.

[11] Andrej Jokic,”License Plate Recognition with Compressive Sensing Based Feature Extraction”.arXiv:1902.05386v1 [cs.CV] 7 Feb 2018.[12] Mohamed Yousef, Khaled F. Hussain, and Usama S. Mohammed”Accrate Vehicle Number Plate Recognition and Realtime Identification Using RASPBERRY PI”.e-ISSN: 2395-0056.p-ISSN: 2395-0072.Volume: 05 Issue: 04 Apr-2018.

[13] Rayson Laroca , Evair Severo , Luiz A. Zanlorensi, Luiz S. Oliveira, Gabriel Resende Gonc ¸alves , William Robson Schwartz and David Menotti "An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO detector". arXiv:1909.01754v2 [cs.CV] 5 Sep 2018.

[14] Rayson Laroca, Luiz A. Zanlorens, Gabriel R. Gonc alves, Eduardo Todt, William Robson Schwartz, David Menotti1. "An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO detector". arXiv:1909.01754v2 [cs.CV] 5 Sep 2019.

[15] Satadal Saha. "A Review on Automatic License Plate Recognition System".Students" Technical Article Competition: PRAYAS-2018,29th april 2019.