International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017),e-ISSN:2455-2585
International Conference on
Recent Explorations in Science, Engineering And Technology (ICRESET'19)
Volume-5, Special Issue-March, 2019.

Image Recognitions with Deep Learning and Comparison between Different Convolutional Neural Network Structures using Tensor flow

[1]Nandhini.P, [2]Dr.G.Vijaya

[1]Department of InformationTechnology, InformationTechnology, KGiSL InstituteofTechnology
[2]Headf of Department, KGiSL Institute ofTechnology

I. ABSTRACT

ABSTRACT: Over the past decade, human movements and behaviors' can be monitored by using big data, including Global Positioning System (GPS) data, social media login data and mobile phone tracking data. Deep Learning is the part of an artificial intelligence function that imitates the workings of the human brain in processing data and creating patterns for use in decision making. Deep learning is a subset of machine learning in Artificial Intelligence (AI) that it capable of learning unsupervised from unstructured or unlabeled data. This is also called as Deep Neural Learning or Deep Neural Network. The process of machine learning can be carried out with the help of hierarchical level of artificial neural networks. Automatic number-plate recognition (ANPR) is a technology that uses optical character recognition on images to create vehicle location data from vehicle number plate image. It is used to check if a vehicle is registered or licensed. Automatic number plate recognition can be used to store the captured images by the cameras and also the text from the license plate, with some configurable to store a photograph of the driver. However, traditional methods may not be suitable for extracting comprehensive vehicle information due to the complexity and diversity of human behaviors. Studies have shown that deep neural networks have out paced the abilities of human beings in various fields and that deep neural networks can be explained in a unique manner. Thus, deep neural network methods can potentially be used to understand human behaviors. In this project, a deep learning neural network constructed in Tensor Flow is applied to detect and classify vehicle information in toll gates available across country and the models of these vehicles are analyzed to verify the classification results. The vehicles doesn't need to spend lots of time on the toll gate, as the Tensor Flow system will recognize the number plate and based on the government identification, the amount will auto debit from the respective bank. For the social science classification problem investigated in this study, the deep neural network classifier in Tensor Flow provides better accuracy and more lucid visualization than do traditional neural network methods, even for erratic classification rules. Furthermore, the results of this study reveal that Tensor Flow has considerable potential for application in the human geographyfield.

KEYWORDS- IMAGE PROCESSING, VEHICLE NUMBER PLATE RECOGNITION, DEEP LEARNING, FEATURE EXTRACTION, PATTERN RECOGNITION

II. INTRODUCTION

A. Deeplearning

Deep Learning solves this central problem in representation learning by introducing representations that are expressed in terms of other, simpler representations. Complex concepts have been enabled by Deep learning. Figure 1.1 shows how a deep learning system can represent the concept of an image of a person by combining simpler concepts, such as corners and contours, which are denied in terms of edges. The quintessential example of a deep learning model is the feed forward deep network, or multilayer perceptron (MLP). A multilayer perceptron is a mathematical function mapping in which some set of input values to produce their respective output values. The function is formed by composing many simpler functions. New representation of the input such as different mathematical function can be implemented. The idea of learning the right representation for the data provides one perspective on deep learning. It is very simple to learn a multistep computer program can uses deep learning as another perspective. Computer memory can occupy the different layer of inputs.

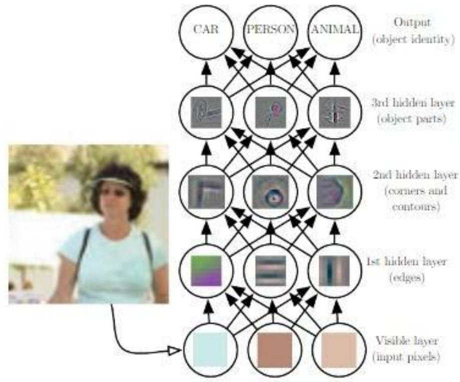

Fig:1.1 Multilayer Perceptron (MLP)

B. Convolutional neural networks(CNNs)

Convolutional neural networks (CNNs) are a specialized kind of neural network for processing input data that has an inherent grid-like topology [9]. Generally, the input data to a CNN will have natural structure to it such that nearby entries are correlated. Examples of this type of data are 1-D audio time series data and 2-D images. The more formal definition of a CNN is a neural network that uses convolution in place of general matrix multiplication in at least one of its layers.

It can interpret convolution as fixing one vector in place, striding the other vector along it, and for each stride a dot product is computed. Each dot product produces one number that is an entry in the output vector. In the case of CNNs, a convolutional layer is generally composed of three stages. The first stage involves parametrized, learnable filters each performing a convolution in parallel. This convolution operation can be modified from the above definition in different ways, such as how much to stride before computing another dot product. The second stage involves an element-wise non-linearity similar to a fully- connected layer. Finally, the third stage is called pooling. Pooling is a method of down sampling the output vector of the second stage. One way to do this is called max-pooling in which the maximal element in a defined section of the output is taken to represent the entire section. To summarize, a convolutional layer tries to find local patterns in the input. Each filter in the first stage is learned during training in such a way that is task specific. In other words, the CNN attempts to find the most relevant patterns that help determine how to accomplish the given task. An equivalent way of thinking about CNNs is by imagining a convolutional layer as being a fully connected layer with an infinitely strong prior that says weights are shared across input data entries and a majority of them are zero [9]. In addition, it is important to note that the ideas discussed here also generalize to arbitrarily sized, finite tensors.

III. EXISTINGSYSTEM

Type of Number Plate Recognition System

1) Super-resolutiontechnique

Comparisons of super-resolution techniques have been mainly concerned with what assumptions are made in the modeling of the super- resolution problem. The blurring process to be known or those regions of interest among multiple frames are related through global parametric transformations, these are the assumptions one has to make. Other models take into account arbitrary sampling lattices, physical dimension of sensor, a non-zero aperture time, focus blurring, and more advanced additive noise models. To simplify a model many times these assumptions are chosen and are usually used in a specific method. In addition, methods that do not make these assumptions have not demonstrated objectively that removing these assumptions gains better super-resolution reconstruction performance. Signal-to-noise ratio (SNR), Peak signal-to noise ratio (PSNR), Root mean square error (RMSE), mean absolute error (MAE), and mean square error (MSE) have all been used as objective measures of super-resolution accuracy; however, the outstanding method of presenting results is clearly subjective to visual quality.

CHALLENGE ISSUES FOR SUPER RESOLUTION

A. Imageregistration

Image registration is critical for the success of multi-frame SR reconstruction, where spatial samplings of the HR image are fused. The image registration is a basic image processing problem that is well known as ill-posed. The problem is more difficult in the SR setting, where the observations are low-resolution images with heavy aliasing artifacts. The performance of the standard image registration algorithms decreases as the resolution of the observations goes down, resulting in more registration errors. Degradations caused by these registration errors are visually more annoying than the blurring effect resulting from interpolation of a singleimage.

B. ComputationEfficiency

Another difficulty limiting practical application of SR reconstruction is its intensive computation due to large number of unknown samples, which require expensive matrix manipulations Real applications always demand efficiency of the SR reconstruction to be of practical utility.

C. Robustness

Prospects Traditional SR techniques are vulnerable to the presence of deviation due to motion errors, inaccurate blur models, noise, moving objects, motion blur, moving scene etc. Robustness of SR is of interest because the image degradation model parameters cannot be estimated perfectly, and sensitivity to deviations may result in visual degradations, which are unacceptable in many applications, e.g., video standard conversion

2) Modified Stroke WidthTransform

This is a common task performed on unstructured scenes. Unstructured scenes are images that contain undetermined or random scenarios. For example, you can detect and recognize text automatically from captured video to alert a driver about a road sign. This is different than structured scenes, which contain known scenarios where the position of text is known beforehand.

Segmenting text from an unstructured scene greatly helps with additional tasks such as optical character recognition (OCR). The automated text detection algorithm in this example detects a large number of text region candidates and progressively removes those less likely to contain text.

Steps to implementation:

Step 1: Detect Candidate Text Regions Using Maximally Stable Extremely Regions [5]
Step 2: Remove Non-Text Regions Based On Basic Geometric Properties
There are several geometric properties that are good for discriminating between text and non-text regions [6,7] including:

- Aspectratio
- Eccentricity
- Eulernumber
- Extent
- Solidity

Step 3: Remove Non-Text Regions Based On Stroke Width Variation
Step 4: Merge Text Regions For Final Detection Result
Step 5: Recognize Detected Text Using OCR

Disadvantages

- Not suited for long and narrow components because it should support their ratio between 0.1 to 10[8]
- Component size should be too long and too short should berejected
- Another common problem that may surround text, such as is connected components signframes.

3) Effect of Character Spacing on the Performance of Automatic Number Plate Recognition (ANPR) Systems throughSimulation.

(i) The extent of the improvement likely to be gained bytime-sharing,
(ii) The factors which will limit the degree ofimprovement,
(iii) The effects on the degree of improvement of any modifications which are seen to befeasible

Disadvantages

The threshold value should be set manually at fixed range.
The following three automatic thresholding methods are affected by the pixel intensity of the objects in the ROI.

- If the objects are dark on a lightbackground
- The automatic methods calculate the high thresholdvalue
- Set the low threshold value to the lower value of the thresholdlimits.

4) Optical CharacterRecognition

While number plate recognition has special type of OCR technology, today optical character recognition (OCR) technology is considered strictly a type of technology - mainly software - that lets you scan paper documents and turn them into electronic, editable files [1]. OCR (optical character recognition) is the recognition of printed or written text characters by a computer. This involves photo scanning of the text character-by-character, analysis of the scanned-in image, and then translation of the character image into character codes, such as ASCII, commonly used in data processing. In OCR processing, the filtered in picture or bitmap is broke down for light and dim zones keeping in mind the end goal to recognize each alphabetic letter or numeric digit. At the point when a character is remembered, it is changed over into an ASCII code. Special circuit boards and computer chips designed expressly for OCR are used to speed up the recognition process [2].Optical character recognition is the electronic conversion of optically processed characters. Character recognition can be offline or online, in online character recognition computer recognitions the character when it isdetected.

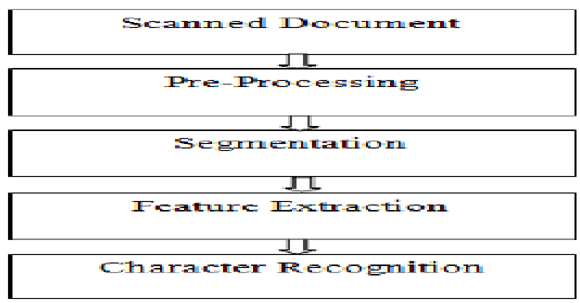

Fig4.1: Traditional working steps of character recognition

In this traditional method of OCR, there are three important steps which are segmentation, Feature Extraction, and Classification. In segmentation we determine the elements of an image. The second important step is to seizure the important characteristics of each character which distinguish each symbol; feature extraction is performed mainly through analyzing the distribution of points, through transformation, series expansion and structural analysis. The third important step of recognition is the classification, identifying each element characters and assigns it to the correct character class [3].Among these main steps, the most significant part is usually the image preprocessing step which enhances/improves the input image to a level that characters can be segmented in a correct method. Therefore, the reliability and accuracy of the ANPR systems rely on the methods that are used in preprocessing. Based on the importance of the pre- processing steps used in approaches to ANPR, we can compare various edge detection filters involved in the process of plate recognition. Edge detection can be identified as a sub-process of the pre-processing techniques that can be applied to an input image tasks such as gray scaling, binary conversion, noise removal, performing morphological functions to recreate or develop the images acquired- can be considered as some of the other tasks which can be performed before an image is passed through edge detection in the preprocessing phase. By applying edge detection filters, the image is converted into an image with boundaries of the objects which exist in input image. This narrows down the process of identifying characters so that the objects identified through processing boundaries can be used to segment the characters, which are then used in character recognition. Sobel, Canny, Gabor and Log-Gabor edge detection filters are the four candidate methods which will be analyzed. These are the filters/algorithms normally used in Automatic Number Plate Recognition[6].

Disadvantage:

1. It can use SVM for training the datasets. The biggest disadvantages are choosing appropriately hyper parameters of the SVM that will allow for sufficient generalization performance. Also, choosing the appropriate kernel function can betricky.
2. In a way the SVM moves the problem of over-fitting from optimizing the parameters to model selection. Sadly kernel models can be quite sensitive to over-fitting the model selectioncriterion

5) Artificial Neural Network (ANN)

Artificial Neural Network (ANN) sometimes known as neural network is a mathematical term, which contains interconnected artificial neurons. Several algorithms such as are based on ANN. In multi layered perceptron (MLP) ANN model is used for classification of characters.

In contains input layer for decision making, hidden layer to compute more complicated associations and output layer for resulting decision. Feed forward back-propagation (BP) algorithm was used to train ANN.

IV. PROPOSEDSYSTEM

Tensor Flow handles all the details for user by providing a function to compare the model predictions contained in logits with labels_placeholder, the correct category labels. The output of sparse_softmax_cross_entropy_with_logits() is the loss value for each image. TensorFlow shines here, using a technique known as *auto-differentiation*; it calculates the gradient of the loss—with respect to the parameter values. It calculates the influence of each parameter on the overall loss and the extent to which decreasing or increasing it by small amounts would serve to reduce the loss. It seeks to improve accuracy by recursively adjusting all of parameter values. After this step, the process restarts with the next image group. The next two lines in our code (below) take measures of the accuracy. It indices of the category with the highest score, which are category label predictions. These labels are compared to the correct category class labels, which returns a vector of boolean values—which are cast into float values (either to 0 or 1), whose average is the fraction of correctly predicted images. TensorFlow contains various optimization techniques for translating gradient information into updates for the parameters. For our purpose in this Paper to choose the simple *gradient descent* option, which only examine the current state of the model for determining how to update the parameters and doesn't consider previous parameter values. The process for categorizing input images, comparing predictions with the correct categories, calculating loss, and adjusting parameter values is repeated many, many times. Computing duration and cost would quickly escalate with larger and more complex models, but our simple model here doesn't require much patience or high-performance equipment to see meaningfulresults.

V. CONCLUSION

Number plates with different format can be identified by splitting the higher level of their network into different sub-networks, each one assuming a different number of digits in the output. A parallel sub-network then decides how many digits are present. I suspect this approach would work here, however it's not implemented it for this project. To detect number plates which have much more varied fonts. One possible solution would be to make my training data more varied by drawing from a selection of fonts. The slowness is a killer for many applications: A modestly sized input image takes a few seconds to process on a reasonably powerful GPU. I don't think it's possible to get away from this without introducing a (cascade of) detection stages .It would be an interesting exercise to see how other ML techniques compare, in particular pose regression (with the pose being an affine transformation corresponding with 3 corners of the plate) looks promising. A much more basic classification stage could then be tacked on the end. This solution should be similarly terse if an Machine Learning library such as scikit-learn is used. In conclusion, this project will shown that a single CNN (with some filtering) can be used as a passable number plate detector / recognizer, however it does not yet compete with the traditional hand-crafted (but more verbose)pipelines in terms ofperformance.

References:

1. Chirag Patel ,Dipti Shah,Atul Patel, PhD. "Automatic License PlateRecognition"
2. Harish D. Kendre,Gaurav V. Talokar, Mohan Girhe, Tejas Pidkalwar, " The Automatic Number Plate Recognition System (ANPR)", INTERNATIONAL JOURNAL OF MATHEMATICS AND COMPUTER RESEARCH, Volume 1 issue 3 April 2013 ISSN :2320-7167, pp. 99-102.
3. Swapnil Desai, Ashima Singh," Optical character recognition using template matching and back propagationalgorithm".
4. Lubna, M. F. Khan and N. Mufti, "Comparison of Various Edge Detection Filters for ANPR",The sixth international conference on innovative computing technology(INTECH2016).
5. Chen, Huizhong, et al. "Robust Text Detection in Natural Images with Edge-Enhanced Maximally Stable Extremal Regions." Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE,2011.
6. Gonzalez, Alvaro, et al. "Text location in complex images." Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE, 2012.
7. Li, Yao, and Huchuan Lu. "Scene text detection via stroke width." Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE,2012.
8. Boris Epshtein, Eyal Ofek ,Yonatan Wexler " Detecting Text in Natural Scenes with Stroke Width Transform"
9. I.Goodfellow, et al., Deep LearningDeep Learning. MIT Press, 2016
10. YN. Chen, CC. Han, CT. Wang, BS. Jeng, and KC. Fan, "The application of a convolution neural network on face and license plate detection," in Proceedings of 18th International Conference on Pattern Recognition (ICPR 2006), Hong Kong, 2006, pp.552-555.
11. J. Li, C. Niu, and M. Fan, "Multi-scale convolutional neural networks for natural scene license plate detection," in Advances in Neural Networks–ISNN 2012, Heidelberg: Springer, 2012, pp.110-119.
12. IJ. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet, Multi-digit number recognition from street view imagery using deep convolutional neural networks; Apr, 2014, 13. LM. Belue, and KW. Bauer, "Determining input features for

multilayer perceptrons," Neurocomputing, vol. 7, no. 2, pp. 111-121,1995.

14. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324,1998.
15. Apple Inc, in Part of iOS 6 DeveloperLibrary.
16. C. Gerber, and M. Chung, "Two-step convolutional neural network approach for improved number plate localization on iOS," in Proceedings of 2014 Korea Computer Congress (KCC2014), Busan, Korea, 2014, pp.868-870.
17. California Institute of Technology, Image_Datasets; 2015,http://www.vision.caltech.edu/Image_Datasets/.
18. Sameer Khan and Suet-Peng Yong "A Deep Learning Architecture for Classifying Medical Image of Anatomy Object", Annual Summit and Conference, ISBN 978-1- 5386-1543-0, pp. 1661-1668,2017
19. Rui Wang, Wei Li, Runnan Qin and JinZhong Wu "Blur Image Classification based on Deep Learning", IEEE, ISBN 978-1-5386-1621-5 pp. 1-6, 201746
20. Teny Handhayani, Janson Hendryli, Lely Hiryantyo "Comparison of Shallow and Deep Learning Models for Classification of Lasem Batik Patterns", ICICoS, ISBN 978-1-5386-0904-0, pp. 11-16,2017
21. Laila Ma‟rifatul Azizah, Sitti Fadillah Umayah, Slamet Riyadi, Cahya Damarjati, Nafi Ananda Utama "Deep Learning Implementation using Convolutional Neural Network in Mangosteen Surface Defect Detection", ICCSCE, ISBN 978-1-5386-3898- 9, pp. 242-246, 2017 22.Ye Tao, Ming Zhang, Mark Parsons "Deep Learning in Photovoltaic Penetration Classification", IEEE Power & Energy Society General Meeting, ISBN 978-1-5386- 2213-1, pp. 1-5,2017
23. Sebastian Stabinger, Antonio Rodr´ıguez-S´anchez "Evaluation of Deep Learning on an Abstract Image Classification Dataset", IEEE International Conference on Computer Vision Workshops (ICCVW), ISBN 978-1-5386-1035-0, pp. 2767-2772,2017
24. Hasbi Ash Shiddieqy, Farkhad Ihsan Hariadi, Trio Adiono "Implementation of Deep-Learning based Image Classification on Single Board Computer", International Symposium on Electronics and Smart Devices (ISESD), ISBN 978-1-5386-2779-2, pp. 133-137,2017
25. Hangning Zhou, Fengying Xie, Zhiguo Jiang, Jie Liu, Shiqi Wang, Chenyu Zhu "Multi-classification of skin diseases for dermoscopy images using deep learning", International Conference on Imaging Systems and Techniques (IST), ISBN 978-1- 5386-1621-5, pp. 1-5,2017
26. Sachchidanand Singh, Nirmala Singh "Object Classification to Analyze Medical Imaging Data using Deep Learning", International Conference on Innovations in information Embedded and Communication Systems (ICIIECS), ISBN 978-1-5090- 3295-2, pp. 1-4,2017
27. Rika Sustika, Asri R. Yuliani, Efendi Zaenudin, Hilman F. Pardede "On Comparison of Deep Learning Architectures for Distant Speech Recognition", 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), ISBN 978-1-5386-0659-9, pp. 17-21,2017