JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

"IMAGE TEXT DETECTION AND RECOGNITION USING DEEP LEARNING TECHNIQUES"

JANETHA.S — M.TECH, 4TH SEM
JAYASHREE N — ASSISTANT PROFESSOR
SHUBASHINI — ASSISTANT PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
C. BYREGOWDA INSTITUTE OF TECHNOLOGY, KOLAR, INDIA

ABSTRACT

The most challenging problem of nowadays is to find end-end text detection of the natural images. Text detection of natural images has used number of models with carefully hand engineered features and their prior knowledge of the analysis. In the given paper, we consider multilayer neural network with different route and combine its representational power with recent developments in unsupervised feature learning, we train highly accurate text detector and recognizer modules by using a common framework. We integrate these two modules into a full end-to-end, lexicon-driven, scene text recognition module system by using the simple off the shelf methods it makes the state of art on benchmarks like SVT and ICDAR 2003.

I. INTRODUCTION

1.1 Introduction

The text information have become very important in finding text detection in natural scenes, the vital role is played by the text detection prerequisite process in identifying the textual information and extraction. The important feature of the text detection is to remove its background and unwanted details from the image. Recent methods have number of challenges like detection accuracy while algorithms flow easily. Scene test properties are obtained by the features that designed manually but by using the training data the important features of the deep learning methods can be obtained easily.

Thus, The text detection and recognition of the natural scene images have been a wide attention in computer vision and images in their understanding, the number of vital applications used in this method are like banking, traffic name board, image retrieval, understanding the scenes and their visual assistance. Recognizing text from the scanned document in unconstrained images is challenging because of wide variations in many ways like backgrounds, fonts, textures and lighting conditions. As a result, most of the data are retrieved from the image by using the cleverly hand engineered features. To obtain text detection and data recognition output into a complete system it requires sophisticated models such as conditional random fields or pictorial structures. The two fundamental tasks for STR are text detection and recognition. Text detection target to determine its position of text from the given image assigned, the position is represented with the bounding box with different colour for its identification. Text recognition converts the image containing text into machine readable strings.

1.2 Scene Text Detection

We use scene text detection in various fields like identification of labels in books, ID cards, transportation ticket counters, banners, traffic scenarios, road signboards, vehicle name plates, and sign board detection, invalid used of name plate number etc.

1.3 Problem Statement

Text obtained by the scene image has obtained patterns of variety non-character patterns in highly complicated background elimination like small size, low quality, low contrast, irregular shape, blurring effect, low lighting.

II. LITERATURE SURVEY

Literature survey is first basic development process of software. The software need to be developed should contain the necessary time factor, its economy and strength of the software.

2.1 Existing System

It has text identification and detection it takes all components needed for processing using MSER and colour cluster. The text line generation uses word lines is determined to obtain multi-level features. An efficiency of the existing algorithm is limited and in n natural scene has wide challenge in determining patterns light condition, background objects.

2.2 Proposed system

The text detection and recognition uses 32-by-32 pixel window where recognizer decides which of 62 characters which is of 26 uppercase, 26 lowercase letters, and 10 digits in the window.

I. III. METHODOLOGY

3.1 Project Methodology

Semantic segmentation and general object detection has been developed for deep learning which lead to the great success. Number of models is also used in the processing of the data detection. The semantic segmentation work as first it takes the text from the segmentation map by using the FCN. Once it is done the box is created the bounding box for detecting the text in the processing. The text obtained should be of clear size and the orientation should be precise and the bounding box information should be perfect in their processing.

Fig 3.1 FCN structure for the detection of text

3.1.1 Pipeline

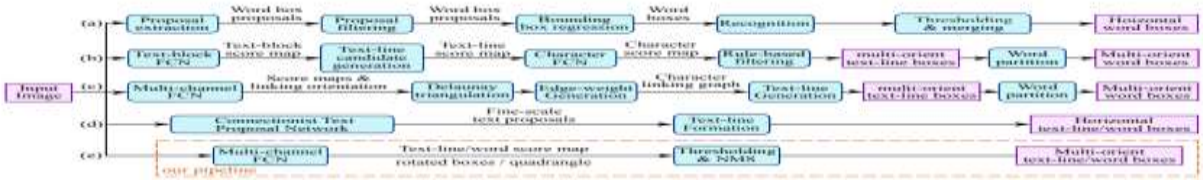

Fig 3.2 Pipeline

The predicted channel here is the score map where pixel is of range [0, 1].rest of the channel represent shape by the pixel view point. If it is of same location its prediction will be precise. The two geometric shapes are (RBOX) and (QUAD) for the processing .Threshold plays a vital role in the predicted region processing. The final output of the processing is mainly concern with the threshold and the pixel generated according to the box obtained in the text for processing thus the result of NMS are considered .

IV.SYSTEM DESIGN

The essential building blocks of our full end-to end system to our text detector and recognition module.

4.1 System Design

The system design uses the system architecture to determine the hypermedia of WebApp. The webApp is obtained for the user who visit and move around the philosophy play a major goal of design. For the presentation and navigation we use structure of the content object.

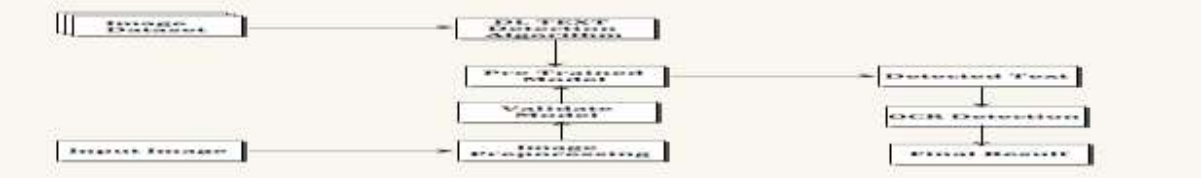

Figure 4.1 The architecture of the system.

V. IMPLEMENTATION

Implementation is mainly uses the two steps data detection and recognition there are further few stages of text detection and recognition.

5.1 Text Detection and Localization

The text detection and localization is the basic step for the processing of the data here we see the data is collected the collected data is sent for the further processing there the data present is the input to the processing and localization is the location or the position of the data by removing all the back ground details of the accessing .the text detection and localization process is the component analysis of the data further these analysis of then find a region based methods.it plot a graph these graph is pointed by the pixels in the analysis considering the edges and it extract cc and then analyses the components. The region based method we used the divides the data into small pixels of data uses 64x32 pixels window and classifies the 16 spatial patches of images.it further had 32x32 pixels of images for future extraction of classifier.

5.2 Classification

The first step we saw is the text detection and recognition after the process the next step is the text classification. Even though the first step reduces the non-text region it may have few non-text regions along with the text regions those text regions should be eliminated the process we use for the elimination of non-text region is the classification step it is also known as verification step. We use classification algorithm in this step it includes and analysis based on the colour, shape, size of the data. We have supervised and unsupervised are the two parts of classification algorithm .the supervised knows properties if the text where as non-supervised don't have idea of the text in knowledge previously.

Fig 5.1Text Detection

5.3 Text Recognition

The next step after the text detection is the text recognition here all the images are converted to the strings or words. The conversion of the word is important only then the word can be easily read

by the humans. The separation of the data is the very important step of the processing of the data. The optical character recognition module is the most important step in the processing of the data. We make segments of the images into k classes them the image is analysed to get the result by the ocr.it is further processed according to the requirement of the analysis of the data.

Fig 5.2 Text Recognition

5.3.3 Convolutional Neural Networks

The image obtained after processing is further divided into regions. These regions are allotted with the nodes. The regions are called as the kernel. The filtration process is continued for the accessing of the data .the filtration process has the wide range of analysis of the data. The analysis is done by x-axis and y-axis. The 2d is further analysed and done with 3d array using the layer further. We use the convolution neural network (CNN). The layer uses pooling layers it reduces the spatial dimensions of the array .the spatial dimensions are the dimensions of the height and weight of the images. The array of 3d is further reduced to the most decomposed by the filter window of the kernel of the array then further the processing of the 2d array is done in the processing by removing the spatial dimensions after the further processing of the array the final output the fully convoluted neural network is obtained which is of the 1 dimension as the final image to be displayed.

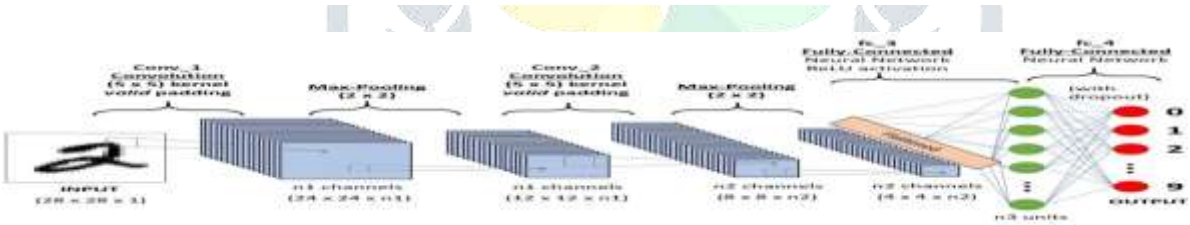

Fig 5.3 Convolutional Neural Networks

5.4 Model training

The model training is the next step of the process after the model is designed the training of the model. Multiple convolutional and pooling layers are used where both are used together in the processing. These layers are connected together to give the output layer there are 6 neurons with the further category of classes.

5.5 Model testing and evaluation

The model testing needed training to carry the data of the model. The location of the data is set to the phase. This data set is checked with the accuracy of the model. The processed model is then further

saved in the real world for the processing. Finally the models are evaluated and display the final outcome of the project.

VI. RESULT

Fig 6.1 Input Image Loading

Fig 6.2 Image PreTrained Model fot Text Detection

Fig 6.3 Text Detection Parameter

Fig 6.4 Text Recognition

Fig 6.5 Final Result

REFERENCES

[1] . Greenhalgh, J.; Mirmehdi, M. Recognizing Text-Based Traffic Signs. IEEE Trans. Intell. Transp. 2015, 16, 1360–1369. [CrossRef]

[2]. Yin, X.-C.; Zuo, Z.-Y.; Tian, S.; Liu, C.-L. Text Detection, Tracking and Recognition in Video: A Comprehensive Survey. IEEE Trans. Image Process. 2016, 25, 2752–2773. [CrossRef] [PubMed]

[3]. Ham, Y.K.; Kang, M.S.; Chung, H.K.; Park, R.-H.; Park, G.T. Recognition of raised characters for automatic classification of rubber tires. Opt. Eng. 1995, 34, 102. [CrossRef]

[4]. Shilkrot, R.; Huber, J.; Liu, C.; Maes, P.; Nanayakkara, S.C. FingerReader: A wearable device to support text reading on the go. In CHI'14 Extended Abstracts on Human Factors in Computing Systems; ACM: New York, NY, USA, 2014; pp. 2359–2364.

[5]. Hedgpeth, T.; Black, J.A., Jr.; Panchanathan, S. A demonstration of the iCARE portable reader. In Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility, New York, NY, USA, 23 October 2006; Volume 279.

[6]. Smith, R. A simple and efficient skew detection algorithm via text row accumulation. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 2, pp. 1145–1148. [CrossRef]

[7]. LEE, C.-M.; Kankanhalli, A. Automatice extraction of characters in complex scene images. Int. J. Pattern Recogn. 1995, 9, 67–82. [CrossRef]