Machine Learning Research
2016; 1(1): 19-32
http://www.sciencepublishinggroup.com/j/mlr
doi: 10.11648/j.mlr.20160101.13

A Novel Approach to Detect Text in Various Dynamic-Colour Images

S. Kannadhasan, R. Rajesh Baba

Department of Electronics and Communication Engineering, Anna University, Madurai, Tamilnadu, India

Email address:
kannadhasan.ece@gmail.com (S. Kannadhasan)

To cite this article:
S. Kannadhasan, R. Rajesh Baba. A Novel Approach to Detect Text in Various Dynamic-Colour Images. *Machine Learning Research.* Vol. 1, No. 1, 2016, pp. 19-32. doi: 10.11648/j.mlr.20160101.13

Received: November 24, 2016; **Accepted**: December 24, 2016; **Published**: January 19, 2017

Abstract: Detecting text in multi-colour images is an important prerequisite. The RBG image is converted into YUV image, after that the multidimensional filter is used to reduce the noise in the YUV image. Canny edge detection is used to measure the continuity of the edges in the images. A efficient text detection is proposed using stroke width transformation method based on contours which can effectively remove the interference of non-stroke edges in complex background and the importance of recent feature (inter-frame feature), in the part of caption extraction(detection, localization). The horizontal and vertical histogram basis is used to calculate the luminance and chrominance which defines the background. Moreover the morphological operation which removes non text areas in the boundaries. Since some background pixels can also have the similar colour, some false stroke areas or character pixels are possible to appear in the output image, which will degrade the recognition rate of OCR (optical character recognition). It exploits the temporal homogeneity of colour of text pixels to filter out some background pixels with similar colour. Optical character recognition enables us to successfully extract the text from an image and convert it into an editable text document. Experimental results evaluated on the Neural network classifier which performance training and testing methods. Training dataset show that our accession yields higher precision and performance compared with forefront methods. The experimental results demonstrate the proposed method will provides efficient result than the existing technique.

Keywords: Image Segmentation, Stroke Width Transformation (SWT), Connected Component Analysis (CCA), Histogram of Gradients (HOG), Edge Detection, Neural Network Classifier, Optical Character Recognition

1. Introduction

In recent years the current use of analyzing techniques receives intensive attention for the increasing use of digital image capturing devices, such as mobile phones, digital camera, PDAs, static image and dynamic images etc. Text information has inspired great interests, among all the content in an image, since both human and computer can easily understand the wide applications like sign detection and translation, mobile text recognition, license plate reading, content-based web image search and so on, define an integrated image text detection and information extraction system (TDIE, shown in Fig.1) with six stages: text detection, image segmentation, text localization, text extraction, text enhancement and recognition (ORC).). Among these stages, text detection and localization are complicated to the overall system performance. In multiple methods it have been proposed to address video image and video text detection and localization difficulties, and some of them have concurred impressive results for specific applications. However, speed and precise of the text detection and localization in Multi-colour images is still a competition due to the variations of text font, size, shape, colour and lining orientation, and it is often influenced by complicated background, brighter changes, image distortion and degrading. The existing methods can be roughly categorized into two groups: region-based method and connected component (CC)-based method. Region-based methods define to detect and separate localize text regions by texture analysis. Currently, the vector separated from each local region is fed into a classifier for estimating the likelihood of text then neighbouring text regions are combined to create text blocks. Because text regions from non-text content to detect and localize texts have distinct textural properties accurately even when the images are

noisy. On the other hand, Connected component based methods segment an image are difficulties to extracting texts into a set of Connected component group to successively from small Connected component to larger ones, then the final Connected component can be either text or background by analyzing their mathematical characteristics. Problem description of the existing methods has reported devastating localization performance; there still remain several problems to solve. For region-based methods, the speed is intensively down and the performance is sensitive to text line orientation. On the other hand, Connected component based methods cannot segment the text components preciously without exact knowledge of text position and scale. Moreover, designing speed and authentic connected component analyzer is defiant since there are many non-text regions which are easily perplexed with texts when analyzed individually. The proposed method is a combination of two algorithms. In the literature survey many algorithms were developed for segmentation. But they aren't benefit for various types of the text detection images. So we go for a combination of three algorithms that will segment the text image. The proposed method consists of six modules.

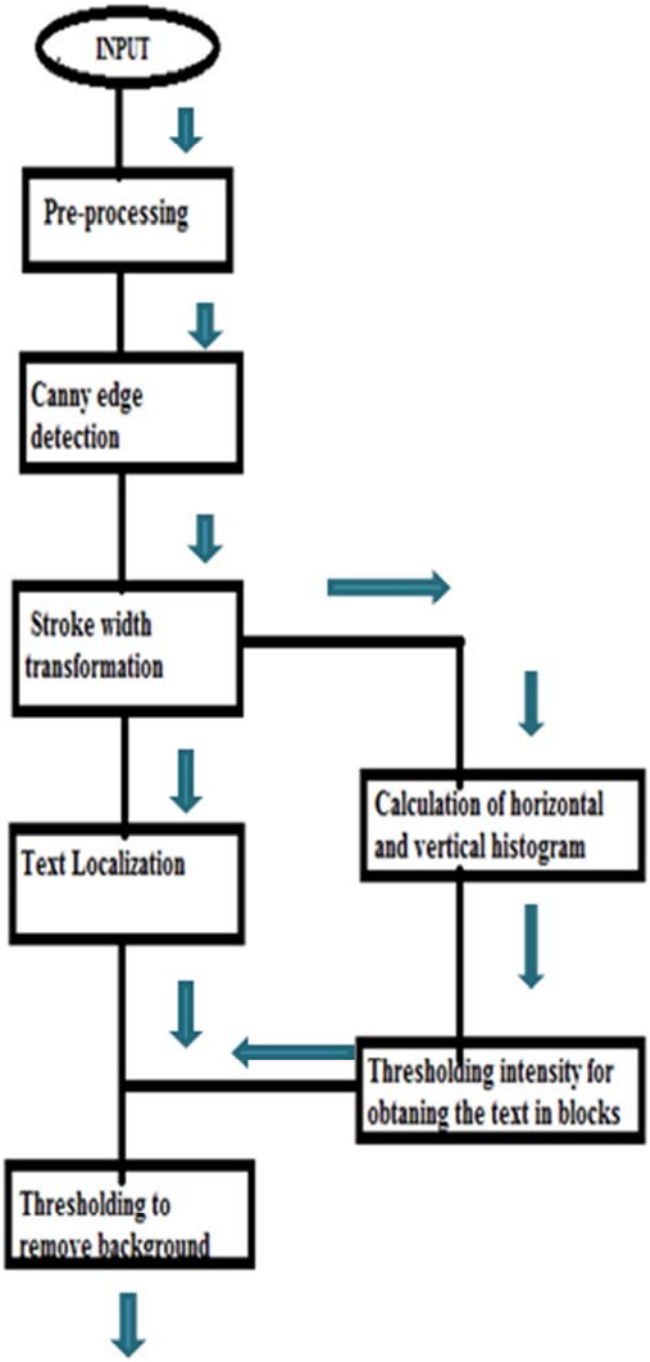

Figure 1a. *Text Detection system.*

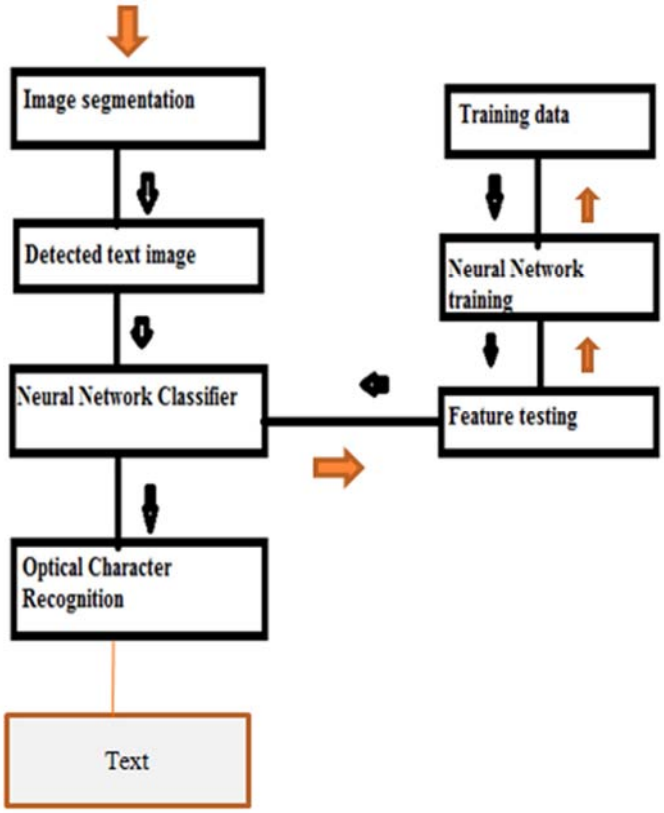

Figure 1b. *Text Extraction system.*

Fig. 1 Architecture of a Text Detection and information extraction system

They are Pre-processing, Edge detection, Stroke width Transformation, Image segmentation, Text localization, Optical character recognition and neural network classifier which consist of both testing and training methods for the future extraction. Multi-Dimensional filter is used to discard the unwanted noise in the YUV transformed image. Edge detection is mentioned by using canny edge detection detector is used to measure edges of the overall boundaries of the horizontal and vertical axis present in the YUV transformed image. Morphological operations are defined by moving a structuring element over a binary image to be modified in such a way that it is centered over an image pixel at some point. Text components are clustered into text lines by words practically with an energy reduction method. Moreover the morphological operation which removes non text areas in the boundaries. Since some background pixels can also have the similar colour, some false stroke areas or character pixels are possible to appear in the output image, which will degrade the recognition rate of OCR (optical character recognition). It exploits the temporal homogeneity of colour of text pixels to filter out some background pixels with similar colour. OCR software enables us to successfully extract the text from an image and convert it into an editable text document. Experimental results evaluated on the Neural network classifier which performance training and testing methods. Training dataset show that our accession yields higher precision and performance compared with forefront methods. The experimental results demonstrate the proposed method will provides efficient result than the existing technique.

2. Related Works

The text regions based on measurement have discrete characteristics from the non-text regions based methods on slope strength and consistency properties. Basically a region-based method consists of two stages: 1) To evaluate the text detection in local image regions using differentiation, and 2) Local text regions into text blocks and text verification to discard non-text regions for future processing using text localization method. An earlier method proposed by Jin Hyung Kim *et al*. [2] Kim support vector machines (SVMs) are used as the texture classifier well even in 2D or 3D dimensional spaces in the absence of a rich set of training and also further extractor is required to decrease the dimensionality of the quality component. SVMs extract their own architecture using kernel function in this paper Lyu *et al.* [3] text edges to detect candidate of various scales with a sobel operator and also using threshold procedure to luminance changes is used to filter out non-text edges and the clustered text regions into text lines by definite profile projection analysis(PPA). In this paper Chen and Yuille [5] To increase up text detection proposed methods using a cascade boost classifier, the selected from a feature pool containing gray-level, gradient and edge features in the weak learners. In this method the detected text regions are

combined into text blocks are alignment by the local binarization. This method performs capitalistic is more than 10 times faster than the other methods. The observations based on the connected component methods have distinct mathematical features, spatial neighbouring components and statistical relationships. These methods usually consist of three stages: 1) Connected component extraction to segment candidate text from images; 2) CCA to filter out non-text components using heuristic rules or classifiers; and 3) post-processing to cluster the text components into text blocks. [6] Stroke width transformation method (SWT) to measure the appropriate height and width of the text in the images. It's used to reduce the unwanted interference in the non-stroke regions in the images. In this paper Ali Mosleh [7] Histogram of gradient basis calculation is used to calculate horizontal and vertical projections in the image. It reduces the background by calculating the threshold intensity of the text alone in the text blocks Experimental results evaluated on the Neural network classifier which performance training and testing methods. Training dataset show that our accession yields higher precision and performance compared with forefront methods.

3. System Overview

It is mentioned earlier that region-based methods or connected component based method can't be able to detect and localize the text, it need some special work based on the local and global information to specify the text detection and recognition task. It's noted that region-based methods and connected component based methods are interrelating. In this the region based method is used to extract the local texture information to segment the image orderly while connected component method can filter out the non-text components and localize the text. The RBG image is converted into YUV image, after that the multidimensional filter is used to reduce the noise in the YUV image. Canny edge detection is used to measure the continuity of the edges in the images. A efficient text detection is proposed using stroke width transformation method based on contours which can effectively remove the interference of non-stroke edges in complex background and the importance of recent feature (inter-frame feature), in the part of caption extraction(detection, localization). The horizontal and vertical histogram basis is used to calculate the luminance and chrominance which defines the background. Moreover the morphological operation which removes non text areas in the boundaries. Since some background pixels can also have the similar colour, some false stroke areas or character pixels are possible to appear in the output image, which will degrade the recognition rate of OCR (optical character recognition). It exploits the temporal homogeneity of colour of text pixels to filter out some background pixels with similar colour. OCR software enables us to successfully extract the text from an image and convert it into an editable text document. Experimental results evaluated on the Neural network classifier which performance training and testing methods. Training dataset show that our accession yields higher precision and performance compared with forefront methods. The experimental results demonstrate the proposed method will provides efficient result than the existing technique and the required output will displayed in text notepad file.

4. Pre-processing

The preprocessing is the process of making an image that is suitable for the next level. It performs filtering of noise and other artefacts in the image and sharpening the edges in the image. RBG to YUV transformation and Sharpening also take place here. After that the preprocessing of data uses some Rule-based filtering (inference) system so the image is enhanced. But the possibilities for the noise in Multi-colour image are very less. Here we are using the Multi-dimensional filter for the noise removal**.**

4.1. RBG Colour Model

A RBG colour image is an image in which each pixel is specified by three values one each for the red component, blue component, and green components of the pixel scalar.

Figure 2. *Pre-processing Image.*

Here each pixel consists of intensity values. The intensity value for single or double arrays, values range from 0 to1 for eight bit integer and values range from zero to two fifty five for sixteen bit integer, the subentries values for sixteen bit integer values range from zero to sixty five thousand five thirty five bits. In RGB colour model, each colour appears in its primary spectral components. The colour of a pixel is made up of three components; red, and blue (RGB), described by their corresponding intensities. Colour component are also known as colour channels or colour planes (components). In the RGB colour model, a colour image can be represented by the intensity function. The intensity of each colour channel is usually stored using eight bits, which indicates that the quantization level is 256. That is, a pixel in a colour's image requires a total storage of 24 bits. A 24 bit memory can express as 224 =256×256×256=16777216 distinct colour's. The number of colour's should adequately meet the display effect of most images. Such images may be called true colour images, where each information of pixel is kept by using a 24-bit memory.

4.2. YUV Transformation

YUV transformed color image is an encoded part of dynamic image and static image. Its main purpose is to reduce the bandwidth of the chrominance component and obtain the luminance component alone in the dynamic and static image. In this YUV models identify the color image were the luminance (brightness) component is define as (Y) components and the chrominance (color) component is define as (UV) components. YUV component specifies the analogy encoding and digital encoding of color information in dynamic and static images. YUV signals are significantly developed from Red, Green and Blue color source. Measured values of R, G, and B are defined to produce Y', and measure the overall (brightness) luminance. U and V are computed as scaled differences between Y' and the B and R values. YUV is computed from RGB as follows:

$$Y' = W_R R + W_G G + W_B B$$

$$U = U_{Max} \frac{B - Y'}{1 - W_B} \approx 0.492\left(B - Y'\right)$$

$$V = V_{Max} \frac{R - Y'}{1 - W_R} \approx 0.877\left(R - Y'\right)$$

The value of WR is define as a range of 0.2126 and the value of WB is define as a range of 0.0722. Y values are conventionally shifted and aligned value ranges from sixteen to two thirty five moreover than using complete range of zero to two fifty five.

Figure 3. *YUV transformed image.*

4.3. Multi-dimensional Filter

The main goal of the multi-dimensional filter is to run through the signal process which discards the unwanted noise present in the dynamic and static images. The multi-dimensional filter runs through the process step by step, alternating each step with the conventional of filter with the operated unstable image. The design of multidimensional filters used for digital group forming conventionally requires regular sampling of the unstable data. Regular sampling in the context of this work means for example spatial sampling on a hexagonal grid, although other regular patterns can be used as well.

4.4. Discarding Noise by Multi-dimensional Filtering

Multi-dimensional filtering is similar to using a suitable filter, in that each o/p pixel value is set to an average of the pixel values in the neighbourhood of the responding i/p pixel. Moreover, the multi-dimensional filter, the value of an o/p pixel is expressed by the median of the neighbouring pixels, either by the mean. The median is much less sensitive than the mean compared to intense values (called statistics). Therefore Multi-dimensional filtering is able to discard these statistics without decreasing the sharpness of the image

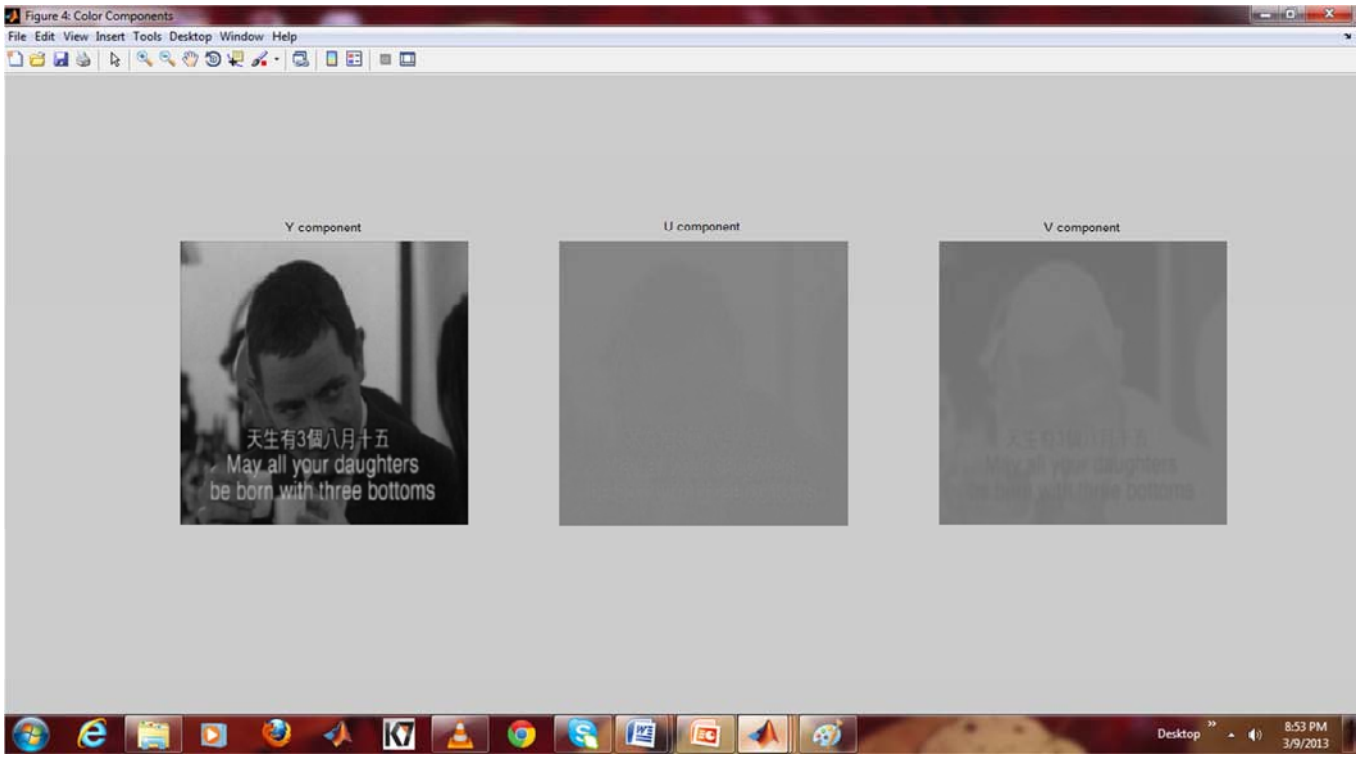

Figure 4. *Removed noise in multi-dimensional Filter.*

5. Edge Detection

Edges are conspicuous local changes of intensity in an image. Edges typically occur on the boundary of two separate regions between images.

5.1. Goal of Edge Detection

Produce a line drawing of a edge from an image of that scene. Efficient features can be separate out from the edges of an image (corners, lines, curves). These features are used by top-level computer vision algorithms (e.g., recognition).

5.2. Causes Intensity Changes

Various physical events cause extremity changes a) Mathematical events1) Object boundary (incoherence in depth and surface colour and texture).2) Bottom layer (incoherence in surface orientation and surface colour and texture). b) Non-mathematical events1) Direct reflection of light, such as a mirror. 2) Shadows. 3) Inter-reflections.

5.3. Steps of Edge Detection

1) Smoothing: Reduce as much noise as possible, without eliminating the true edges.2) Apply a filter to improve the amount of the edges in the image.3) Detection: determine which edge pixels should be removed as noise and which should be detained. 4) Localization: determine the exact location of an edge (*sub-pixel* resolution might be required for some applications, that is, innovate the location of an edge to better than the spacing between two pixels). Edge ticking and lining are usually required in this step. Most edge detection methods work on the appropriative conditions that an edge occurs where there is incoherence in the intensity function. In a continuous image of pixels we can calculate the gradient value by taking the variation of greyscale values between adjacent pixels.

5.4. Canny Edge Detector

Canny edge detection is used to measure the horizontal and vertical axis present around the text region in gray scale image. Most of the changed edges in the greyscale intensity are finding using basically canny algorithm Canny edge detector is the optimal and most widely used algorithm for edge detection. Canny edge detector provides robust edge detection, localization and linking. Compared to other edge detector canny edge detection has better continuity and low contrast lined edges when compared to sobel edge detector and prewit edge detector. Due to the low contrast line edges the text can be easily obtained from the background images. This can have more flexibility and better management which involves the kernel function in canny edge detection algorithm. At each and every edge of the lines is calculated for the output pixels at a particular row. We need input pixels calculate at the row's below and above the text region boundaries. Thus the output at the first and the last rows are undefined and the same happens in case of columns. To incorporate this, the output width and height and the output buffer position changes after each steps.

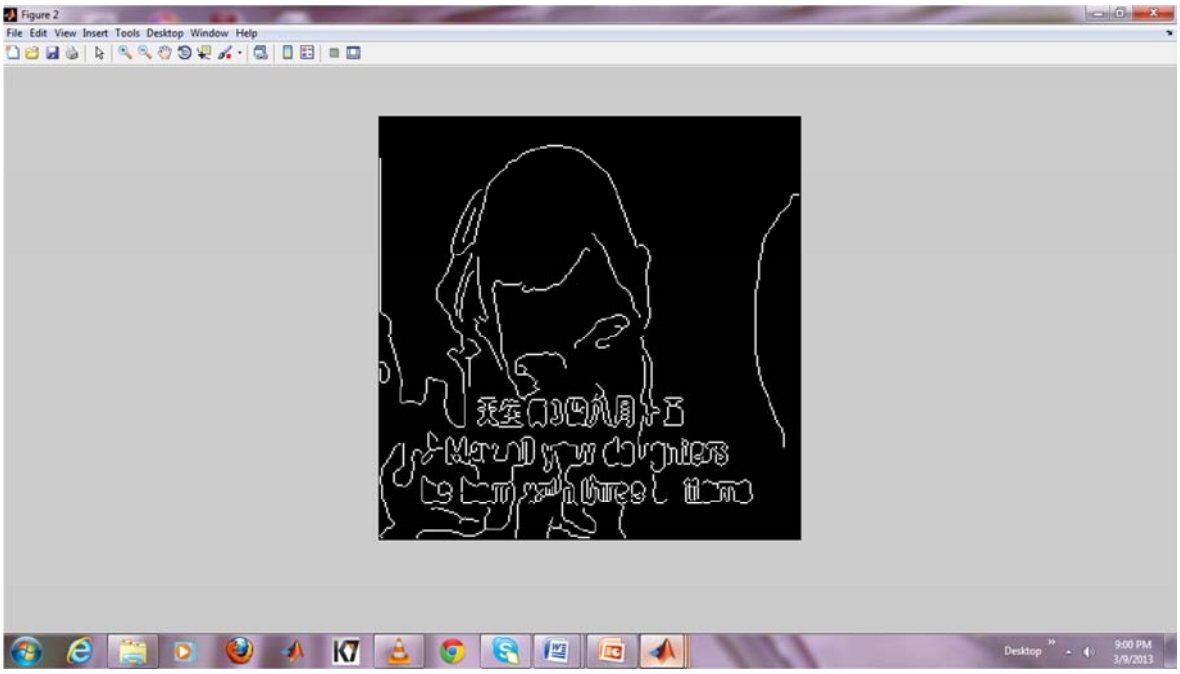

Figure 5. *Canny edge detector.*

5.5. Stroke Width Transformation

In order to find the text from the image, first the image is transferred into edges using canny edge detection method. The threshold value is chosen, that edges of the letters are not diminished while avoiding or generating many edges. Stroke width transformation (SWT) is performed to calculate the approximate width and height of possible stroke where each pixel belongs to starting edge points and tracing along the gradient. In this the letter can have brighter (or) darker color compared to its background. Stroke width transformation (SWT) has to perform by both the forward edge and backward edge along with the gradient edges. The stroke width transformation is a local image which computes with the pixel width that containing the stroke of the likely pixel. Stroke width transformation (SWT) values are more complex in the corners; it's calculated by measuring the height and width of the text in the image. The Stroke width transformation is used to measure the approximate width and height of the text aligned in each rows and column directions in the image.

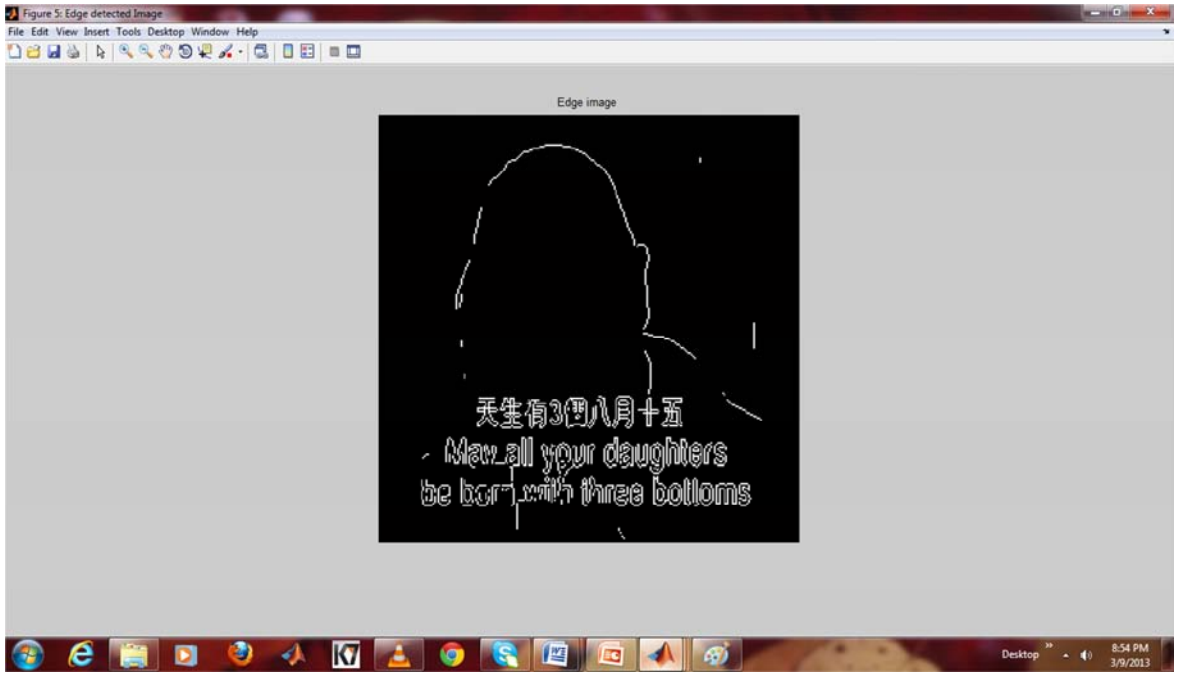

Figure 6. *Applying stroke width tranformation.*

Figure 7. *Measuring the width and height of the image.*

6. Image Segmentation

To segment candidate CCs from the YUV transformed image, Niblack's local binarization algorithm is adopted due to its high efficiency and non-sensitivity to image degrading. The formula to binaries each pixel x is defined as

$$b(x) = \{0\ if\ gray(x) < \mu r(x) - k.\sigma r(x)$$
$$\{255\ if\ gray(x) > \mu r(x) + k.\sigma r(x)$$
$$\{100\ otherwise$$

Binarization is the process of converting a YUV transformed image to a black and white image. A binary image contains a pixel intensity range of 0 to 255 levels. Binarization is done using Thresholding. Thresholding sets all pixels above a defined value to white, and the rest of pixels to black in the image. It is very important to decide the appropriate threshold value to binaries the image, though it is difficult to decide a global value which is suitable for all images. For instance, if threshold value is chosen as X for an image then the pixels of that image having the intensity range equaling X will be binaries to a value 0 and other pixels will be given a value 1. Thus grayscale image will be converted as black & white image. In a black & white image, each pixel has a value of either 0 or 1. The black pixel (digit 1) refers to foreground of the image and white pixel (digit 0) refers to background of the image.

Figure 8. *Image segmentation using stroke width transformation.*

7. Morpological Operation

Morphological operations are defined by moving a structuring element over a binary image to be modified in such a way that it is centred over an image pixel at some point. The process of removing certain details in an image which is smaller than certain reference shape is called dilation of image processing and the reference shape is called structuring element. The structuring element in a morphological operation plays an important role with its different shape and size. It uses the dilation process to reduce the distortion image in the background and detect the text region in the image. When a structuring element of any size

containing complement of 0's & 1's is placed on a binary image, a specified logical operation will be performed between structuring element and underlying binary image at each pixel position. The binary result of that logical operation is stored in the output image at that pixel position. Binarization is the process of converting a YUV transformed image to a black and white image. A binary image contains a pixel intensity range of 0 to 255 levels. Binarization is done using thresholding. Thresholding sets all pixels above a defined value to white, and the rest of pixels to black in the image. It is very important to decide the appropriate threshold value to binaries the image, though it is difficult to decide a gray value which is suitable for all images. For instance, if threshold value is chosen as X for an image then the pixels of that image having the intensity range equaling X will be binaries to a value 0 and other pixels will be given a value 1. Thus YUV transformed image will be converted as black & white image. In a black & white image, each pixel has a value of either 0 or 1. The black pixel (digit 1) refers to foreground of the image and white pixel (digit 0) refers to background of the image.

Figure 9. *Morphological operation based image segmentation.*

Histogram transformation method, many applications are used to measure the desirable shape of the text in the image. In this the histogram is used to measure the desirable text region by calculating the horizontal axis of the text and vertical axis of the text in order to reduce the background region. It uses the laplacian transformation technique to calculate the rows and columns in the background region and detected the text alone in the multi-color images. Histogram of gradient (HOG) method is used to smoothen the background noise in the horizontal and vertical region and identify the text clearly from the background images. Histogram of gradient (HOG) method is used to calculate the threshold intensity of the text in the text region and separate the text alone from the background image.

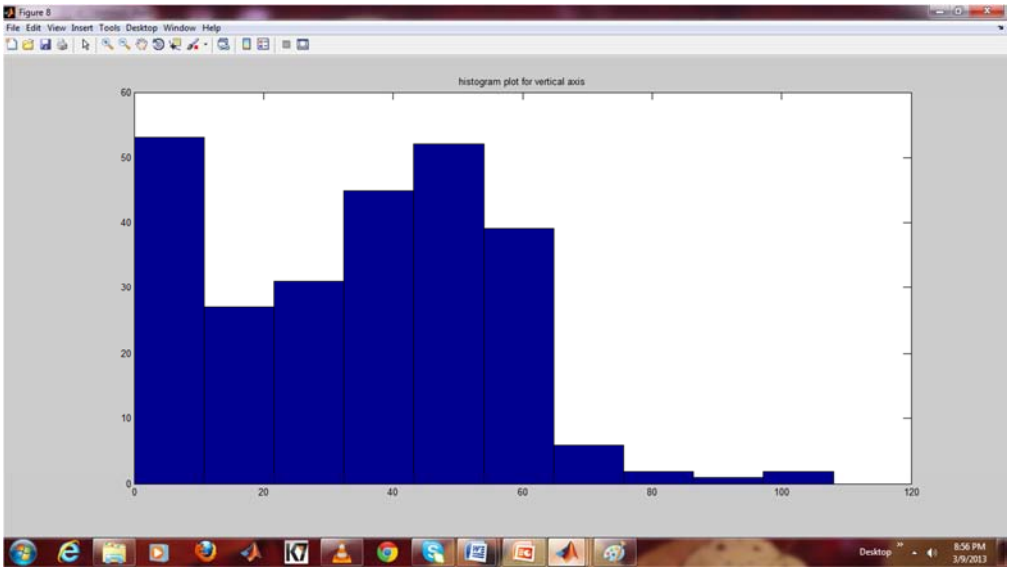

Figure 10. *Horizontal background removed using histogram.*

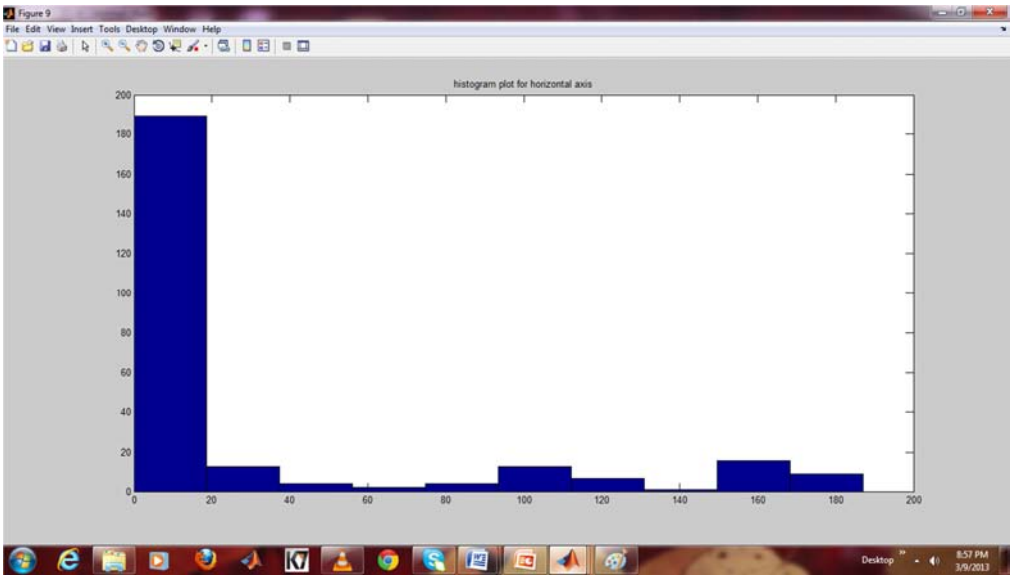

Figure 11. *Image Vertical background using histogram Image.*

8. Localization

The combined horizontal and vertical projection method is a resourceful way to localize text string along the horizontal orientation assumption. The horizontal and vertical projections cannot separate complex text layouts frequently appeared in the multi-colours image with passing through the exponentially. The performs using multiple passes of horizontal and vertical projections are engaged the region growing technique with initial bounding boxes to the segmented text regions. The localize text may project various shapes. The bilinear transformation method modifies the various horizontal and vertical axes into a rectangular format. It's used to align the major axis and minor axis of the horizontal and vertical text regions.

The output of the change detection module is the binary image that contains only two labels, i.e., '0' and '255', representing as 'background' and 'foreground' pixels respectively, with some noise. The goal of the connected component analysis is to detect the large sized connected foreground region or object. This is one of the important operations in motion detection. The pixels that are collectively connected can be clustered into changing or moving objects by analysing their connectivity. In binary image analysis, the object is extracted using the connected component analysis operation, which consist of assigning an individual label to each maximally connected foreground region of pixels. When a foreground pixel with two or more than two foreground neighbouring pixels carrying the same label is found, the labels associated with those pixels are registered as being equivalent. That means these regions are from the same object. The handling of equivalent labels and merging thereafter is the most complex task.

Figure 12. *Text Separation using localization method.*

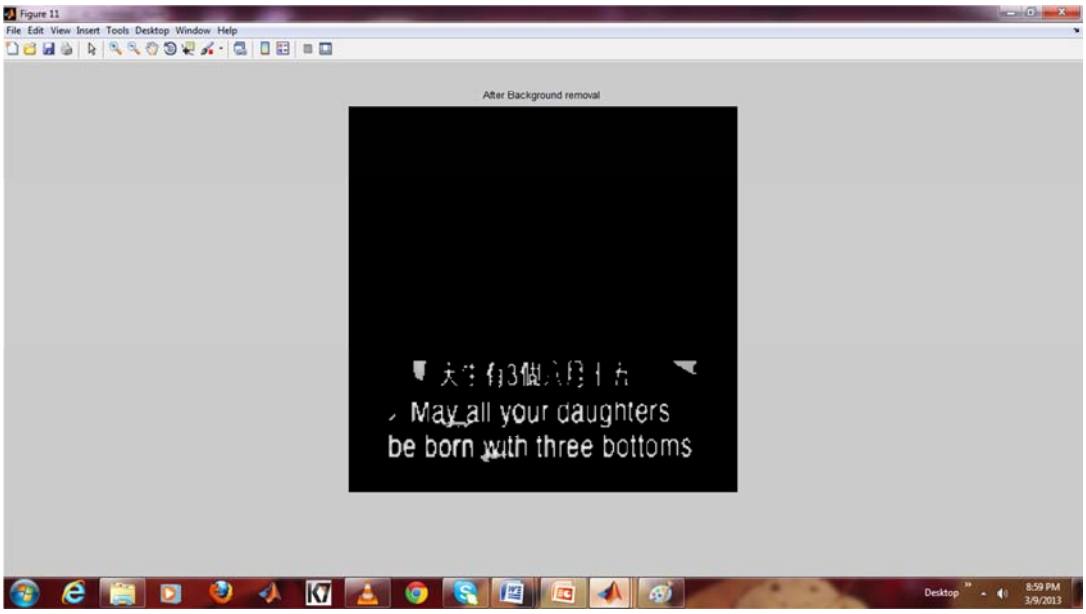

Figure 13 *Connected component analysis using text detection.*

9. Text Extraction

The text extraction method is used to extract the text from various dynamic images after the conversion of localization process and image segmentation. The detected image is clustered and grouped using connected component analysis. In text extraction process the text is extracted using different types of classifiers and optical character recognition method is used to extract the text and display the text in text note pad file.

9.1. Connected Component Classification

Connected component classification was performed on training and validation for train set and evaluation for test set. In connected component analysis both the text component and non-text component were clustered for the training classifiers from the training images. There are three types of classifiers used for the connected component classifications which are SLP (Single layer perceptron's), MLP (Multi-layer perceptron's), and NNC (Neural network classifiers), were unitedly trained using back propagation algorithm. The parameters of SLP and MLP were jointly trained with the combining coefficient, after initialization by back propagation. There is insufficient data to rely on this approach which causes lack of Convergence and low guarantee of global minimization are further drawbacks of SLP and MLP training so NNC method is performed for text extraction. The accuracy and precision of the MLP classifier value range is 60-65% and the SLP classifier value range is 70-75% accuracy. But still there is no sufficient data available to extract the text exactly. The schematic graph of recall rate and false positive rate is mentioned below.

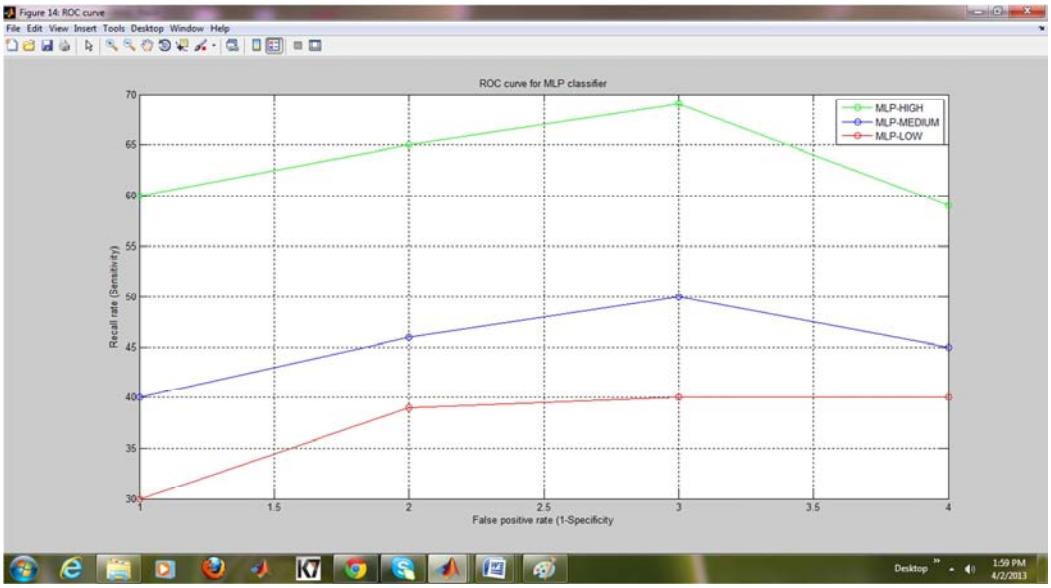

Figure 14. *Roc curve for Multi-layer perceptron (MLP).*

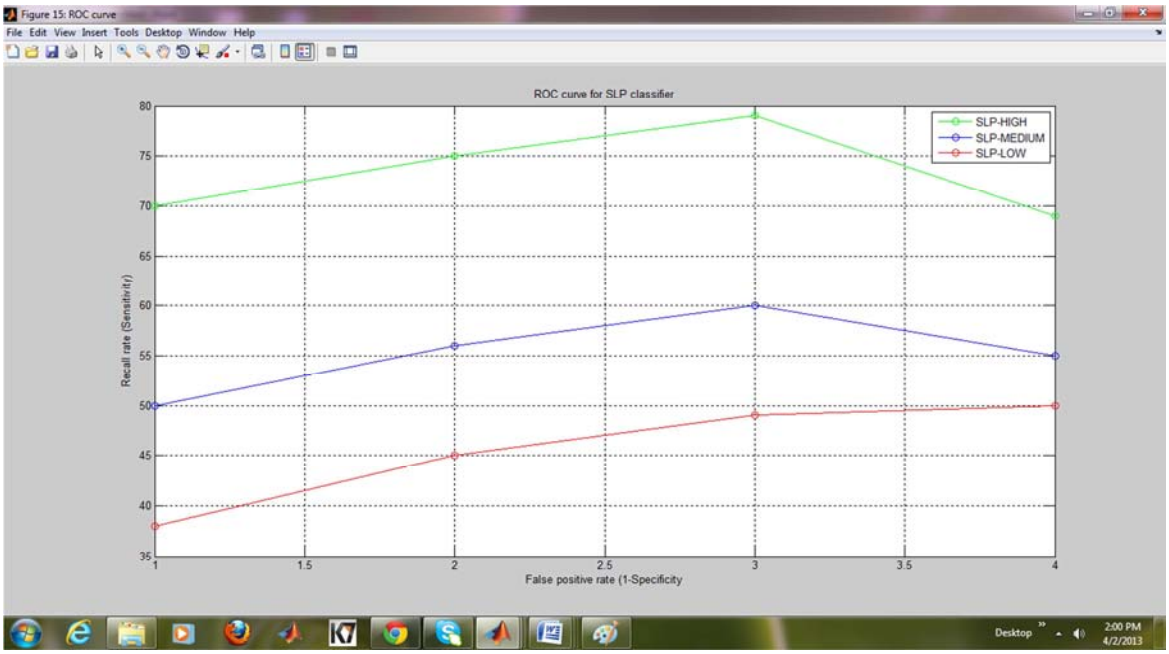

Figure 15. *Roc curve For Single layer perceptron(SLP).*

10. Optical Character Recognition

The handwritten recognition printed text by computer are identified the process is also called as the optical character recognition. The input device are transmitted the signal in a real time with the help of digitizer tablet (pen-based computers and personal digital assistants) or pen position includes the timing information (signature capture) in the valuable recognition. In this paper for calculating and extracting the text Neural network classifier is used, in NNC method both testing and training method is performed. The NNC method gives more accuracy when extracting the text in dynamic images.

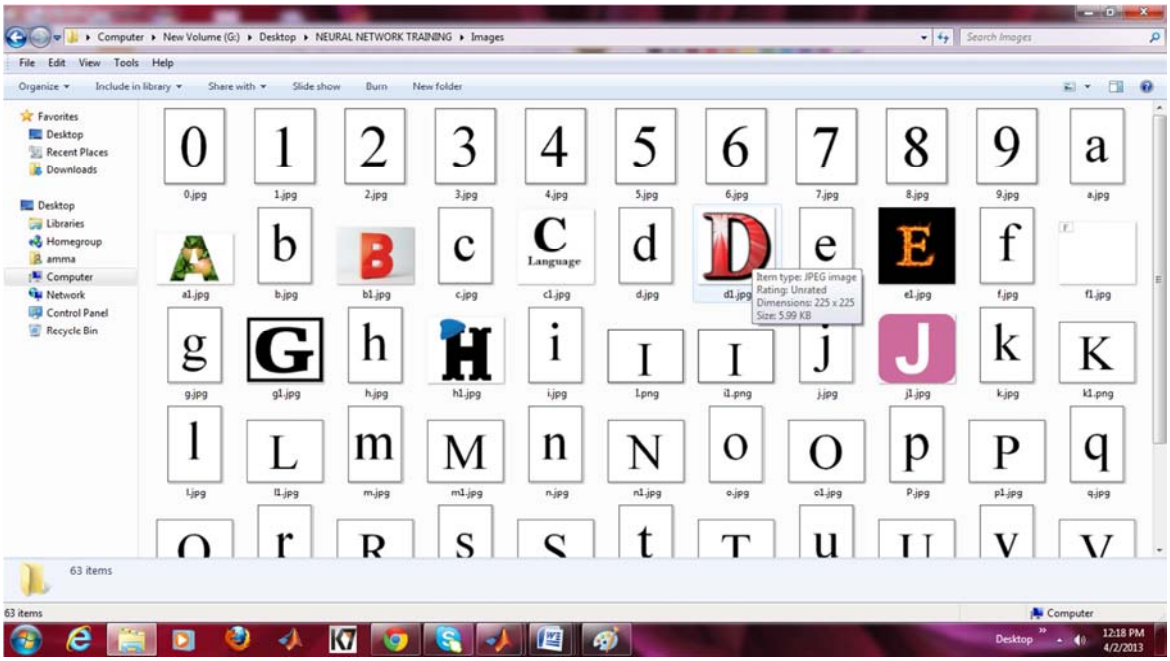

Figure 16. *Character stored in database.*

The accuracy and precision vales ranges from 85-90% when compared with MLP and SLP classifiers. For testing and training the data are compared with non-text component and text component, each and every text are analyzed by comparing the text with the character stored in the database with the help of optical character recognition. In optical character recognition (OCR) it is frequently used as data format or character format which stores the character in the database for translation text mining.

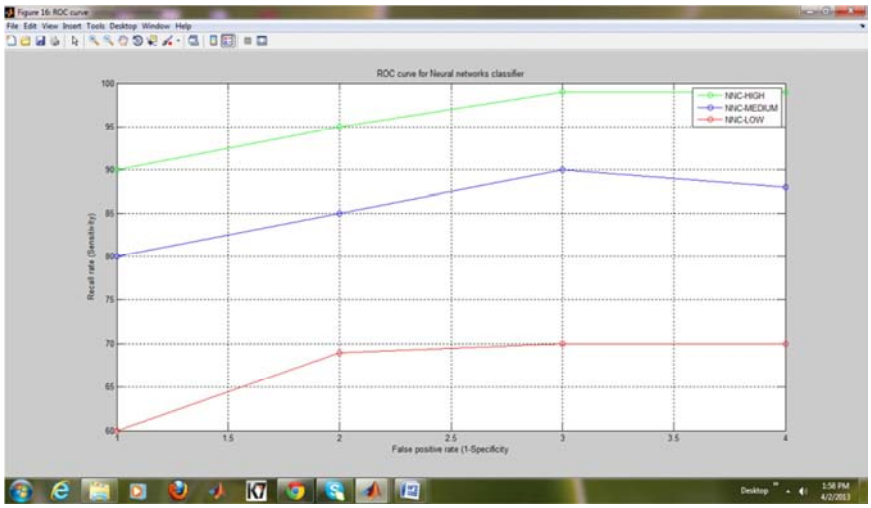

Figure 17. *Roc curve for Neural network classifier (NNC).*

In this the noise removed and extracted text image is compared with characters stored in data base. In this each and every single text component and non-text component are tested and compared with the character stored in the database and the text alone is separated and mentioned in text notepad file.

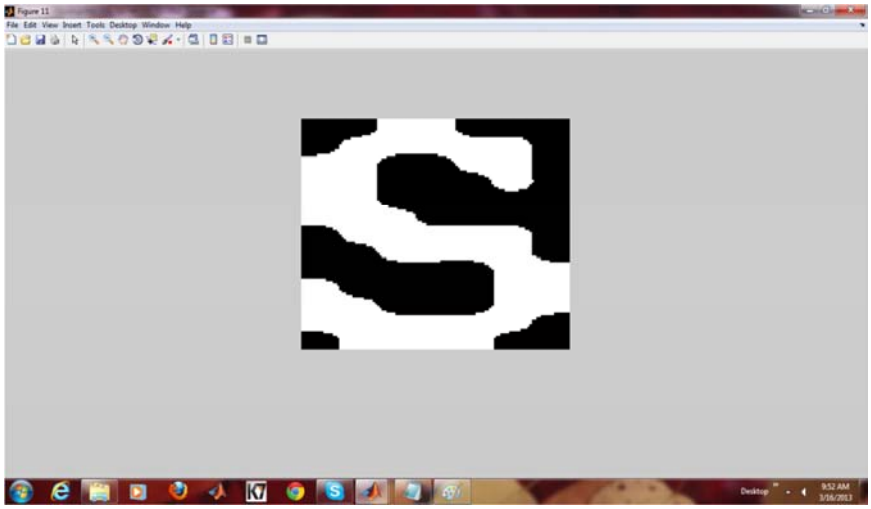

Figure 18. *Compared characters of text component.*

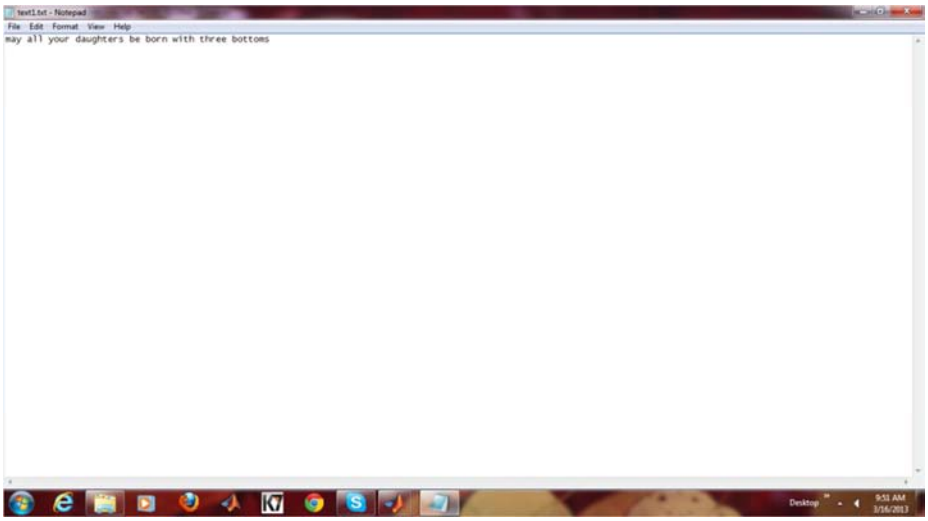

Figure 19. *Extracted text is mentioned in notepad file.*

11. Conclusion

In this paper, we presented application software designed for the location of a various color image. First, we convert the image into YUV transformed image then multi-dimensional filter is used to remove the unwanted noise in the YUV transformed image. Edge detection is mentioned by using canny edge detection. Stroke width transformation technique is used to calculate the approximate width and height of the text. Image segmentation using local binarization to segment the binary image. Morphological operations are defined by moving a structuring element over a binary image to be modified in such a way that it is centered over an image pixel at some point. The histogram of gradients is used to reduce the horizontal and vertical axis in the background image. The goal of the connected component analysis is to detect the large sized connected foreground region or object. Text extraction of the text component using classification process with different types of classifier like SLP classifier, MLP classifier which is performed under connected component classification methods, and finally using the neural network classifier and optical character recognition which will produce the statically output and shows the accuracy among the classifiers. The Extraction of the text image displayed in text notepad file.

References

[1] Yi-Feng Pan, Xinwen Hou, and Cheng-Lin Liu "A Hybrid Approach to Detect and Localize Texts in Natural Scene Images" IEEE transaction on image processing,vol. 20,no. 3,march 2011.

[2] Huiping Li, David Doermann, and Omid Kia "Automatic Text Detection and Tracking in Digital Video" IEEE transaction on image processing, vol. 9, no. 1, january 2000.

[3] Rainer Lienhart, Member, IEEE, and Axel Wernicke "Localizing and Segmenting Text in Images and Videos" IEEE transaction on circuits and systems for video technology, vol. 12, no. 4, april 2002.

[4] Kwang In Kim, Keechul Jung, and Jin Hyung Kim "Texture-Based Approach for Text Detection in Images Using Support Vector Machines and Continuously Adaptive Mean Shift Algorithm" IEEE transaction on pattern analysis and machine intelligence, vol. 25, no. 12, december 2003.

[5] Bingbing Ni, Zheng Song, and Shuicheng Yan, Senior Member, IEEE "Web Image and Video Mining Towards Universal and Robust Age Estimator" IEEE transaction on multimedia, vol. 13, no. 6, december 2011.

[6] Xiaoqian Liu, Member, IEEE, and Weiqiang Wang, Member, IEEE "Robustly Extracting Captions in Videos Based on Stroke-Like Edges and Spatio-Temporal Analysis" IEEE transaction on Multimedia, vol. 14, no.12, April 2012.

[7] Xilin Chen, Member, IEEE, Jie Yang, Member, IEEE, Jing Zhang, and Alex Waibel, Member, IEEE "Automatic Detection and Recognition of Signs From Natural Scenes" IEEE transaction on image processing, vol. 13, no. 1, january 2004.

[8] Yuri Boykov, Member, IEEE, Olga Veksler, Member, IEEE, and Ramin Zabih, Member, IEEE "Fast Approximate Energy Minimization via Graph Cuts" IEEE transaction on parttern analysis and machine intelligence, vol. 23, no. 11, november 2001.

[9] frames Datong Chen∗, Jean-Marc Odobez, Herv/e Bourlard Dalle molle 'Text detection and recognition in images and video frames' molle Institute for Perceptual Articial Intelligence (IDIAP), Rue du Simplon 4, Case postale 592, CH1920 Martigny, Switzerland Received 30 December 2002; accepted 20 June 2003.

[10] Shaolei Feng, R. Manmatha and Andrew McCallum "Exploring the Use of Conditional Random Field Models and HMMs for Historical Handwritten Document Recognition" Center for Intelligent Information Retrieval University of Massachusetts Amherst, MA, 01003 [slfeng, manmatha, mccallum]@cs.umass.edu,IEEE JANUARY 2002.

[11] Niladri B. Puhan, Anthony T. S. Ho, "Localization and text sequence restoration using noise pixels in binary document image watermarking" Journal of Electronic Imaging 18 (2), 023012 (Apr–Jun 2009).

[12] Kohei Arai1 and Herman Tolle2 "Text Extraction from TV Commercial using Blob Extraction Method" International Journal of Research and Reviews in Computer Science (IJRRCS) Vol. 2, No. 3, June 2011.

[13] Shriram Kishanrao Waghmare, A. K. Gulve, Vikas N. Nirgude,Nagraj P. Kamble "Automatic number plate recognition (ANPR) system for Indian condition using support vector machine (SVM)" International Journal of Computer Science and its Applications.

[14] Baoming Shan "Vehicle License Plate Recognition Based on Text-line Construction and Multilevel RBF Neural Network" journal of computers, vol. 6, no. 2, february 2011.

[15] Y C Kiran and Lavanya N C, "Text extraction and verification from video based on SVM" World Journal of Science and Technology 2012.

[16] Sushmita Mitra, Rajat K. De, and Sankar K. Pal, "Knowledge-Based Fuzzy MLP for Classification and Rule Generation" IEEE transaction on neural networks, vol. 8, no. 6, November 1997.

[17] A. Majumdar and B. B. Chaudhuri, "A MLP Classifier for Both Printed and Handwritten Bangla Numeral Recognition" ICVGIP 2006, LNCS 4338, pp. 796 – 804, 2006. © Springer-Verlag Berlin Heidelberg 2006.

[18] T Y Lim, M M Ratnam and M A Khalid, "Automatic classification of weld defects using simulated data and an MLP neural network" Insight Vol 49 No 3 March 2007.

[19] Somkid Amornsamankul, Jairaj Promrak, Pawalai Kraipeerapun, "Solving Multiclass Classification Problems using Combining Complementary Neural Networks and Error-Correcting Output Codes" International journal of mathematics and computers in simulations, Issue 3, Volume 5, 2011.

[20] Guoqiang Peter Zhang, "Neural Networks for Classification: A Survey" IEEE Transactions on systems, man, and cybernetics-part c: applications and reviews, vol. 30, no. 4, November 2000.