LIBL Bill Automation

Muneebah Mohyiddeen[1], Athira A S[2], Ramseena V S[3], Silpa K R[4], Sruthy E S[5]

Assistant Professor, Department of Computer Science And Engineering,Universal Engineering College ,Vallivattom ,

Thrissur, India[1]

B.Tech Student, Department of Computer Science And Engineering,Universal Engineering College ,Vallivattom ,

Thrissur, India[2,3,4]

Abstract: The project deals with day-to-day account settlement. It is a simple application used to make expense management quicker and easier. It uses OCR methodology to read and process receipts which can be converted to spread sheets, documents, or CSV. The user can take/upload a receipt image for recognition of the images. The bill images are captured through camera by image processing and then the data are extracted and also analysed by recognizing the text. All these actions are performed by using an OCR tool, which performs text detection and recognition on the pre-processed image using the Deep Learning (Two-Step CNN Framework) model. Once the image is clicked, it is then encoded and sent to the server. The extracted text data is used to extract the relevant information (such as date, company name, items, total, etc.). This data is then sent back in the desired format

Keywords: OCR, CNN, Text Recognition, Character Segmentation

I. INTRODUCTION

The financial transactions are maintained traditionally with paper bills to store information. It is very time-consuming to search the information through them. These types of tasks are done easily by a computer, but the paper bills cannot be stored digitally without manually typing their contents. So from these bills we can extract and store this information automatically. This can be achieved by using optical character recognition. In this day and age, very accurate outputs from images of printed text are obtained by using an optical character recognition. So an optical character recognition can be used to access and store this type of printed information as machine-process-able text. To store these bills in an organized manner, the user can use this app to convert images of the bills into text format by any android phone user. This further reduces the time consumed to retrieve text from an image of a bill. It is simple to use because the user need only to click a picture of the bill from a reasonable distance and the information in the bill would be stored. According to their date, this information on the bills can be easily required.

II. THEORY

[A]. Optical character recognition

Optical character recognition or optical character reader (OCR) is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text. It can be from a scanned document, a photo of a document, a scene-photo or from described text covered on an image. This can be electronically edited, searched and stored. During OCR processing word processors like Microsoft Word or Google Docs can be used to edit the documents .The development process usually encircles 6 steps with the help of optical character recognition. The first step is to acquire images of paper documents. The paper documents are mostly black and white, and an OCR scanner can threshold the images. The goal of pre-processing is to make basic data usable by computers. The noise level is reformed and areas outside the text are expelled.

Pre-processing is specifically crucial for recognizing handwritten documents. Pre-processing provides a clean character image to output better results of image recognition. To group the characters into meaningful blocks segmentation is performed. To slice the input data into a set of features and to find essential characteristics for pattern recognizable feature extraction is done. Once all the features are extracted, these are fetched to a neural network to train it to recognize characters.

[B]. Image processing

Image processing is a method to perform some operations on an image, to get an enhanced image or to extract some useful information from it. It is a signal processing where an input is an image and output may be image or features associated with that image. Image processing basically includes the following three steps, firstly importing the image by

image acquisition tools then analysing and manipulating the image. The output will be an corrected image which is based on image analysis. The two types of methods used for image processing are analogue and digital image processing. The hard copies like printouts and photographs uses analogue image processing. Digital image processing techniques help in manipulation of the digital images by using computers. The three general aspects while using digital technique are pre-processing, enhancement, and display, information extraction.

[C].   Text recognition

Text recognition stage converts images of text into string of characters or words. It is important to convert images of text into words. Different approaches of recognition are character recognition and word recognition. The text image is divided into multiple cut-outs of single characters in character recognition. Character recognition approach using Optical Character Recognition module (OCR) is used where originally images are divided into k classes. Before getting fed to OCR, the binary text image is generated to pass through connected components analysis and grey scale consistency constraint module. Support Vector Machine (SVM) based classifier is used for character recognition. Word recognition uses character recognition outputs along with language models to observe words from text image.

[D].   Character Segmentation

Character segmentation involves converting text into multiple sets of single characters. It is suitable while in the case of degraded text or connected characters. It originally determines candidate cut pixels from the characters and then two pass path finding process is used. Helps to find out potential cuts in forward pass and prove true cuts and expel false cuts in backward pass. It is an operation that seeks to dissolve an image of a sequence of characters into sub images of individual symbols. It is one of the decision processes in a system for optical character recognition (OCR).

III.   RELATED WORK

Here we introduce each papers based on the technologies used in the toilet and this are arranged in technologies bases
The aim of this paper [1] OCR is a classical topic in pattern recognition and machine learning research. It is still a summons problem for handwritten Chinese character recognition. Transfer learning provides a solution to this; the training samples are making used in the source domain to benefit the training process in the target domain. The CNN-based transfer learning can be done in three steps firstly, CNN L is well trained in the source domain to learn feature extractors and classification boundaries, and then weights of CNN L are transferred to CNN T for initialization. The advantages include better final performance and feasibility to combine CNN based transfer learning with traditional transfer learning. The drawback of this paper is lacking tolerably labelled training samples and also low quality character image samples.

The paper [2] examines a technique for classification of the font style based on character image which is further proposed by employing the distance profile quality with respect to left, right and diagonal directions of a character image.The system for font style recognition and classification system mainly involves three main stages. In the stage 1, for processing an input image is obtained and predate for pre-processing and in stage two further quality estimations are done. Finally, a SVM classifier is used for feature computation. To separate knowledge base for each type of font, the classification of font style develops the complexity of OCR.The advantage is that it optimizes the complexity of OCR and provides satisfactory accuracy.This method provides satisfactory accuracy of an average accuracy of 80%. The drawback of this work is that it can be further improved by normalization, so as to achieve more precision in the results.

The system [3] a pre-processing method is presented for improving Tesseract Optical Character Recognition performance on images with colourful background. The proposed method consists of two steps. At first, to extract the text from the colourful background text segmentation is performed and input image is clustered into k images. In the second step, a classifier is used to identify the image containing text resulting from the previous step. OCR is then performed on the identified images which took from STB devices and are cropped. After image clustering into k clusters, identification of the image that contains text is performed in which for that an appropriate input variables are choose. So image features are extracted. The advantages include Tesseract OCR performance improvements and give an increased accuracy. The drawback of this paper is that for larger images and analysed images with one line manual cropping is needed to be done.

The main intent of this paper [4] is to perceive poor quality English characters, thus an improved OCR with tolerable training data is needed. The proposed OCR system has been described with the help of various modules. The process is split into 4 major blocks. The first block is input acquisition and pre-processing of receipts and old newspapers. The inclined documents are scanned into images and these images will be polished and segmented using Maximally Stable External Regions algorithm. The deep neural network used as a pre-trained Inception V3 model with transfer learning.The advantages include reduction in training time and also improve the recognition accuracy for poor quality

images. The drawback of this paper is that there will be more confusion in case of character recognition for the quality of image.

In this paper [5] is to detect text in natural scenes by using a camera connected to a Raspberry Pi. The detected text is undoubtedly input into a database. So the first step is image pre-processing, select an image. So the advertising image is the input. It extracts all the text from image through pre-processing which involves three parts; Edge detection, Erosion and dilation, Binarization. The next step is smoothing which is the process of removing noise. After image preprocessing has been completed, the contours of the text are determined. The advantages include the processing speed of OCR information is fast and also reduces the labour required for the job and detect faces, eyes and pedestrians. The drawback of this paper is that it is only used for destination objects.

The paper [6] is to take in a number of images of documents. The documents like identity proofs of individuals and then classify them into classes, such as passport and license data will be taken as the basic input. The proposed system has Tesseract OCR which uses a pytesseract package. The implementation of image classification is done by using Convolutional Neural Network.The extracted details from the images are then stored in the database.Then, the features will be extracted from the images. The blocks which are identified as text are next given as input to an OCR. The advantages include better accuracy and performance, reduces the problem of over fitting. But the drawback of this paper clarifies that there is no worth of doing small amount of data and also difficulty of poor quality handwriting.

This paper [7] presents image classification of the proposed system using CNN. The proposed system uses a pytesseract package of Tesseract OCR. The aim of pre-processing is to improve the image quality by removing undesirable distortion and the input image for further processing is raised. The camera sensor detects the noise in food images. The portable cameras or cameras connected to a wearable glass, cap or hat with high resolution are used to capture the food images. The system can also support the visually impaired perceive the food on the platter with its features like colour, texture and shape. The advantage is that it provides higher efficiency and accuracy. The drawback of this paper is that it is computationally expensive and also poor quality of images.

The image classification task is [8] one of the important topics in various computer view tasks. The Naive Bayes classifier is based on Bayes' theorem of probability. The proposed classifier handles the Naive Bayes classifier for minimizing the training time. The aim is to extract a kind of format from a sample of objects. The classification algorithm gets the training set and builds a model to determine new objects. Decision tree is used where a test on an attribute represents an internal node, and a class label is holed by a leaf node. The advantages include good accuracy and results within minimum training time. The drawback is that when the size of the dataset is small then the precision will decrease.

Machine learning is a field of study [9] that gives computers the ability to study without being especially programmed .The machine gets classifying the given data and obtains a decision boundary. Logistic regression and neural network methods for a face recognition technique are used. These methods transfer a colour image to gray image and then denoised using a low pass filter. So the local intensity deviations around eyebrows, eyelids, nose and mouth are captured. The size of image is normalized and decrement to its original size using nearest neighbour interpolation method. The advantages include minimization of computational space and training time. The drawback is that sometimes it will be difficult to identify and recognize text characters.

Handwritten character recognition [10] is always an area of research in the field of pattern recognition. Digital image processing techniques such as Feature Extraction, Image Restoration and Image Enhancement are used. Optical recognition is an offline process i.e. after writing or printing has been completed the recognition will be started. Because of high sufficiency artificial neural networks are mostly used. The advantages include image quality enhancement, highlighting background and foreground. It is actually very time consuming and costly which is a drawback.

This paper [11] proposes that after the coming of digital computers, mixing human functions to computers. Optical Character Recognition is used. Flat bed scanners or digital cameras are used for digitizing. In case of input device, there are two classes of character recognition: on-line and off-line recognition. The first group of systems uses devices like digitizer tablets for data acquisition and while writing the recognition will be done. On-line recognition system is for concurrent data collection format whereas off-line recognition system needed definite procedures for making the image. Low cost and huge processing power and its advantages. But the accuracy rate is low which becomes its disadvantage.

Neural network is playing an important role in handwritten character recognition [12], a space of pattern recognition. Handwritten characters are scanned and translated into 1024 binary pixels. The skeletonization process will be used to binary pixel image and the extra pixels will be deleted. Characters are normalized into 30X30 pixels and shifted to the top left corner of pixel window. It uses neural Feed Forward Multi-player Perceptron network for English handwritten character recognition. For training, back-propagation algorithm can be used. The inputs of MLP are skeletonized and normalized binary pixels of English characters. Good accuracy is provided which is an advantage. But for that a lot of effort is required which becomes its drawback.

The paper presents a deep learning match to no-reference quality assessment of document images [13]. The degraded document image quality is precisely related to the performance of optical character recognition. Firstly, the document image is split into patches and non-informative patches are sifted out using Otsu's binarization technique. Secondly,

using a Convolutional Neural Network (CNN) quality scores are attained and the patch scores are balanced to obtain the document score. With local normalization a gray scale document is pre-processed then crops the image into patches. Dividing the image into patches brings heavy samples for training CNN. The advantage is that it can achieve art performance. But there will be large variations in one valuation which is a disadvantage.

The aim of this paper [14] is to achieve important error reduction for various classification problems networks. A DNN is a feed-forward neural network which has more than one hidden nonlinear layer is proposed using trainable activation functions. First of all crop the patches from the pre-processed images and check patches on binary map. If the patch on the binary map is regular, then patch is neglected. By cropping images into patches get enough training samples. Since the training patches are all labelled. The Convergence rate is much faster so that simultaneous training can be done which is an advantage but the performance is not that much satisfying.

Convolutional Neural Networks (CNN) is a popular deep learning method [15] .Deep learning implement mechanization of feature extraction function. To extract features CNN is used. The character image is processed for removing all the unwanted entities from an image which is done in the preprocessing stage. The input images are resized to a suitable format. The numbers of convolution layers, max pooling layers, ReLu layers and fully connected layers are chosen. The four major affine transformations are translation, scaling, sheering and rotation. First it pre-processes the input image and then analyse the unlabelled test data. In the post processing stage, the integer labels will be obtained as output for the classifier. The advantages are it help to remove undesired qualities of image and also reduces overfitting.But handcrafted features are required for finding features in the text which is a drawback.

Many document images are fixed with background images [16] like checks, deposit books, drive licenses, passports, certificates, etc. Background images can be removed directly without losing the quality of text characters. In order to remove the colour background without changing the text in the foreground, the first step is to assure every pixel has R, G and B three values. In a RGB image, the pixel value is composed of R, G and B three values. After enlargement, pixels from colour background standout and thus can be differentiated from text pixels effectively. It provides improved performance and also effective in removing background images. But still there is a problem for black and white images which is a drawback.

The main purpose of the paper [17] is to separate text information from complex background in colour document images. Thresholding method is used for extracting text from background. For the separation of text from the complex background hybrid approach is used. The threshold value for each detected text region is derived automatically from the data. Canny edge detector is used to detect edges.The image segments containing text are threshold locally to separate the foreground pixels from the background pixels. It provides good accuracy and better performance but also failed to detect single letter word which doesn't contain a hole which is actually a drawback.

The aim of the paper [18] is that an enlarged version of the method, which applies the local color segmentation and the region growing. Methods based on the color clustering are efficient of separating character patterns from images with complex backgrounds. It can extract very small characters with thin strokes as well as large characters with thick strokes from images with complex backgrounds. The k-nearest neighbour method is used. A grayscale image is created from the color image temporarily, and the k- nearest neighbour pixels are found in the grayscale image. In order to find local model colors in each sub image, a histogram-based method is used. The advantage is that it can handle even complex layout but the disadvantage is that it lacks the ability of discriminating colour.

The paper proposes an efficient and effective background determination method for colour document images [19] to solve the problem in color documents. The RGB colour space is transferred to grey level is done. The background and the foreground distributions entered are analysed to determine the background ranges. Finally, the document image is converted into binary image by the background range. The method works in both gray level and colour document images. It is more efficient because it scan only the border pixels in the document image which is an advantage but since it contain some broken characters too.

Efficient Optical Character Recognition [20] in images captured from Set-Top Boxes plays an important role in STB testing.OCR is performed with Tesseract 3.5 and Tesseract 4.0. The first preprocessing method practiced for improving OCR performance for low quality images is image resizing by bucolic interpolation. The second preprocessing method is an image sharpening method. Gaussian low-pass filter is used to obtain smoothed image. The idea is to isolate the text from the background. K-means clustering algorithm is used. It removes noise and makes OCR performance more efficient. But in case of more complex backgrounds causes some problems and furthermore difficulties in reading text which becomes a drawback.

IV. CONCLUSION

Automated billing systems have an effective solution to a wide range of problems that are caused by sudden increases in the number of clients, errors made during manual billing, etc. This is the best solution for reducing human intervention in receipt recognition. Here uses deep learning and OCR methodologies to read and process receipts. We can organize our accountings to plan for future expenses. It will help in business analysis; can make the billing process in the future

much faster and easier and also in more convenient way. Thus, this project proposes an innovation which will serve a wide range of users and their laborious tasks into quite easy ones.

REFERENCES

[1]. YejunTang, L. Peng, Q. Xu, Y. Wang and A. Furuhata, "CNN Based Transfer Learning for Historical Chinese Character Recognition," 2016 12th IAPR Workshop on Document Analysis Systems (DAS), IEEE

[2]. V Bharath and N. S. Rani, "A font style classification system for English OCR," 2017 International Conference on Intelligent Computing and Control (I2C2).

[3]. Matteo Brisinello, Ratko Grbi, Dejan Stefanovic, Robert Peckai-Kovac "Optical character recognition on images with colourful background", 2018 IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin).

[4]. Tan Chiang Wei, U. U. Sheikh, Ab Al-Hadi Ab Rahman, "Improved optical character recognition with deep neural network", Signal Processing & Its Applications (CSPA) 2018 IEEE 14th International Colloquium.

[5]. A. Pise and S. D. Ruikar, "Text detection and recognition in natural scene images," 2014 International Conference on Communication and Signal Processing.

[6]. R. Deepa, Associate Professor Department of Information Technology Loyola-ICAM College of Engineering and Technology Chennai, India [ICECA 2019] "Image lassification and Text Extraction using Machine Learning"IEEE Conference Record # 45616; IEEE

[7]. Pathanjali C, Vimuktha E Salis, Jalaja G, Latha A "A Comparative Study of Indian Food Image Classification Using KNearest-Neighbour and Support-Vector-Machines" - International Journal of Engineering & Technology, 2018

[8]. Shih-chung Hsu1, I-chieh Chen And Chung-lin Huang "Image Classification Using Naive Bayes Classifier with Pairwise Local Observations" - Journal of information science and engineering, 2017

[9]. Debachandra Singh. N "Binary Face Image Recognition using Logistic Regression and Neural Network", 2017- International Conference on Energy, Communication, Data Analytics and Soft Computing(ICECDS).

[10]. Dr Jmgala Sasi Kiran, N Vijaya kumar, N Sashiprabha M Kavya "Image preprocessing 2015" -Bonfring International Journal of Image Preprocessing, Vol. 2, No. 1.

[11]. Berchmans, D., & Kumar, S. S,Optical character recognition: an overview and an insight, 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE.

[12]. A. Pal, D. Singh, Handwritten English character recognition using neural network, 2010, IJCSC(International Journal of Computer science and communication), vol 1.

[13]. L. Kang, P. Ye, Y. Li and D. Doermann, "A deep learning approach to document image quality assessment," 2014 IEEE International Conference on Image Processing (ICIP).

[14]. Hoon Chung, Sung Joo Lee and Jeon Gue Park, "Deep neural network using trainable activation functions," 2016 International Joint Conference on Neural Networks (IJCNN).

[15]. P. P. Nair, A. James and C. Saravanan, "Malayalam handwritten character recognition using convolutional neural network," 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT).

[16]. M. Shen and H. Lei, "Improving OCR performance with background image elimination," in 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD).

[17]. N. Shivananda and P. Nagabhushan, "Separation of Foreground Text from Complex Background in Color Document Images," in 2009 Seventh International Conference on Advances in Pattern Recognition.

[18]. Aoba, Aramaki, Aoba-ku, Sendai-shi, "Character pattern extraction from colorful documents with complex backgrounds," Object recognition supported by user interaction for service robots, Quebec City, 2002, vol.3.

[19]. Chun-Ming Tsai, "An efficient and effective background determination for color document images," 2009 International Conference on Machine Learning and Cybernetics.

[20]. M. Brisinello, R. Grbic, M. Pul, and T. Andelic, "Improving optical character recognition performance for low quality images," in 2017 International Symposium ELMAR.