



# INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

## LIBL BILL AUTOMATION

Muneebah Mohyiddeen<sup>1</sup>, Athira A S<sup>2</sup>, Ramseena V S<sup>3</sup>, Silpa K R<sup>4</sup>, Sruthy E S<sup>5</sup>

Assistant Professor, Department of Computer Science And Engineering, *Universal Engineering College, Vallivattom, Thrissur, India.*<sup>1</sup>

B.Tech Student, Department of Computer Science And Engineering, *Universal Engineering College, Vallivattom, Thrissur, India.*<sup>2,3,4,5</sup>

**Abstract:** The project deals with day to day account settlement for every user. It's a simple application that makes expense management easier and quicker. It uses OCR methodology to read and course receipts which can be turned to spread sheets, documents, or CSV. It will provide an interface for the user to take/upload a receipt image which is used for recognition. The bill images will be captured through camera by image processing and then the data are extracted and also analyses the text by recognizing. All these steps are performed by using an OCR tool, which performs text detection and recognition on the pre-processed image using the Two-Step CNN Framework model. Once the image is clicked, it is then concealed and sent to the server. The extracted text data then uses a simple translating technique to extract the consistent information (such as date, company name, items, total, etc.). This data is then sent back to a desired pattern. The extracted data which is sent to the main server can be used where an accountant task is shortened thus only need to certify the data without having to carefully repeat it back into the database.

**Index Terms – OCR, CNN ,Text Recognition, Character Segmentation**

### I. INTRODUCTION

The monetary transactions are managed traditionally with paper bills to store informations. Although this is a reliable way , it is very time-consuming to go through the information and also definite transaction is tedious. These types of works are done easily by a computer, but the paper bills cannot be stored digitally without physically typing their contents. So from these bills we can extract and store the informations undoubtedly. This can be attained by using optical character recognition. In this day and age, a very accurate output from images of printed text are obtained by using an optical character recognition, so as in the same order and with the slotting of the words as found in the images. So an optical character recognition can be used as an approach in a computer which is mainly used to store this type of printed information as machine-process able text. To store these bills in an organized manner, any android phone user can use this app to convert images of the bills into text pattern. This further reduces the time to retrieve text from an image of a bill. It is smooth to use because the user need only to click a picture of the bill from a fair distance and the information in the bill will be stored. According to their date, these informations on the bills can be easily acquired.

### II. THEORY

#### 2.1 Optical character recognition

Optical character recognition or optical character reader (OCR) is the electronic or mechanical transformation of pictures of composed, written by hand or printed text into machine-encoded text. It can be from a filtered record, a photograph of an archive, a scene-photograph or from portrayed content covered on a picture. This can be electronically altered, looked and put away more minimalistic ally. During OCR handling word processors like Microsoft Word or Google Docs can be utilized to alter the archives .The improvement interaction generally encompasses 6 stages with the assistance of optical character recognition. The initial step is to get pictures of paper archives, that is a unique picture can be caught and put away. The paper archives are for the most part high contrast, and an OCR scanner can limit the images. The objective of pre-processing is to make essential information usable by PCs. The commotion level is reformed and territories outside the text is ousted. Preprocessing is explicitly vital for perceiving transcribed records that are more delicate to commotion. Preprocessing gives a spotless character picture to yield better consequences of picture recognition. To bunch the characters into significant squares division is performed. To cut the information into a bunch of highlights and to discover fundamental attributes for design unmistakable element extraction is finished. Each character gets distinguished in a specific class .Once every one of the highlights are removed, these are gotten to a neural organization to prepare it to perceive characters. Post preprocessing stage is the way toward refining a few amendments in an OCR model. The ID of characters thickly relies upon the specific circumstance. The confirmation of the yield requires a human-in-the-bend approach.

## 2.2 Image processing

Image processing is a technique to play out certain procedure on a picture, to get an upgraded picture or to remove some valuable data from it. It is a sign handling where an information is a picture and yield might be picture or features related with that picture. These days, image preprocessing is among expediently developing technologies. Image handling essentially incorporates the accompanying three stages, right off the bat bringing in the picture by means of picture procurement apparatuses then investigating and controlling the picture. The yield will be an amended picture or report which depends on picture analysis. The two sorts of strategies utilized for this are simple and advanced image processing. The printed version like printouts and photos utilizes analogue picture handling. While utilizing these visual procedures the picture investigators use basics of understandings. Computerized image handling strategies help in control of the advanced pictures by utilizing PCs. The three general angles while utilizing computerized strategy are pre-handling, improvement, and show, data extraction.

## 2.3 Text recognition

Text recognition stage changes over images of text into series of characters or words. It is essential to change over pictures of text into words as word is a basic people utilized by human for his visual acknowledgment. Various methodologies of acknowledgment are character acknowledgment and word recognition. The text picture is separated into numerous patterns of single characters in character acknowledgment. Partition between neighbouring characters is vital for these strategies. Character recognition approach utilizing Optical Character Recognition module (OCR) is utilized where initially pictures are partitioned into  $k$  classes. Prior to getting taken care of to OCR, the twofold content picture theory is created to go through associated parts examination and dim scale consistency imperative module. Support Vector Machine (SVM) based classifier is utilized for character acknowledgment. Word acknowledgment utilizes character recognition yields alongside language models to notice words from text picture.

## 2.4 Character Segmentation

Character division includes changing over text into different arrangements of single characters. It is reasonable while on account of corrupted content or associated characters. It initially decides competitor cut pixels from the characters and afterward two pass way discovering measure is utilized. Assists with discovering possible cuts in forward pass and refute genuine cuts and expel cuts in reverse pass. It is said that an activity that looks to broke down a picture of a grouping of characters into sub images of individual images. It is one of the choice cycles in a framework for optical character recognition (OCR).

## III. RELATED WORK

Here we introduce each papers based on the technologies which has been used for the recognition of images and this are arranged in technologies manner.

The aim of this paper <sup>[1]</sup> OCR (Optical Character Recognition) for decades of years, it 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; the training samples are make 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 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 <sup>[2]</sup> 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 complexity of the OCR systems by font style recognition can be reduced .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. 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 <sup>[3]</sup> a preprocessing method is presented for improving Tesseract Optical Character Recognition (OCR) 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 image which took from STB devices and are cropped in a way that they contain only one line of text presented with single colour. After image clustering into  $k$  clusters, identification of the image that contains text is performed in which for that an appropriate input variables are choosed.The advantages include Tesseract OCR performance improvements and gives 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 <sup>[4]</sup> 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. After that all the segmented text characters become basic input that requires labelling. After labelling, the next step is to train using a deep neural network. In this work, the deep neural network used as a pre-trained Inception V3 model with transfer learning. When all the labelled images have been processed through the Inception V3 model and the resulting values rescued to a cache file, then these transfer-values can be used as the input to another neural network. Test dataset is given to the trained network for optical

character recognition. The advantages include reduction in training time and also improves the recognition accuracy for poor quality images. And also there is an overall reduction in error rate. The drawback of this paper is that there will be more confusion in case of character recognition for the quality of image and also the shortage of sufficient training data.

In this paper <sup>[5]</sup> 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 preprocessing, select an image. So the advertising image is the input. It extracts all the text from image through preprocessing which involves three parts; Edge detection, Erosion and dilation, Binarization. So the images are splitting into discrete regions called edge detection. Then it transform it into gray scale image using sobel filtering algorithm. The next step is smoothing which is the process of removing noise. After that seek contours using sobel filter is analyzed for shape analysis. Next binarization is done, which is the process of taking a grayscale image and switching it into black and white. The next step is erosion and dilation used to process and analysis the shapes in an image. After removing the noise we can obtain the outer contour of text area. After image preprocessing has been completed, the contours of the text are determined. Character recognition is performed on the text area cached in the test images. It uses Tesseract OCR. 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 <sup>[6]</sup> 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. And also the problem of overfitting is avoided by using CNN. A Long Short-Term Memory is used to implement the text extraction module whenever once the images are classified in which the LSTM networks are the units of Recurrent Neural Network. As the text data are extracted from the classified images it will be then implemented using Tesseract OCR package which contains an optical character recognition (OCR) engine - libtesseract. The extracted details from the images are then stored in the database. Since the convolutional layer maps the image to a matrix pattern, to reduce the dimensions for faster processing time the pooling layer is used and the flattening converts the image into a linear array format for delivering it into the neural network. Then, the features will be extracted from the images. After all steps, the extracted features of the image are compared to that of the training set images and are classified into various layers. 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 overfitting. 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 <sup>[7]</sup> presents image order of the proposed framework utilizing CNN. The proposed framework utilizes a pytesseract bundle of Tesseract OCR. Tesseract OCR shows an improved performance. So the proficiency and execution of the framework can be enlarged. The point of pre-handling is to improve the picture quality by eliminating unfortunate twisting and the info picture for additional preparing is raised. The camera sensor distinguishes the commotion in food images. The trimming of external part of Indian food pictures is required to improve the casing and the condition proportion of the image. Hence to get the assurance in the picture, the district is edited around the pictures which assists with discovering the edge. People face trouble in picking the best food with high supplements to remain solid, subsequently an openCV based food ID framework is utilized to discover, restrict and notice Indian food pictures. The convenient cameras or cameras associated with a wearable glass, cap or cap with high goal are utilized to catch the food pictures. The framework can likewise uphold the outwardly weakened see the food on the platter with its highlights like tone, surface and shape. The advantage is that it give higher productivity and accuracy. The downside of this paper is that it is computationally costly and furthermore low quality of pictures.

The image characterization task is <sup>[8]</sup> one of the significant subjects in different PC see tasks. The Naive Bayes classifier depends on Bayes' hypothesis of likelihood. In Bayes' hypothesis, the restrictive likelihood that a thing has a place with a class can be calculated. The CNN is taken care of as the nearby classifier. By utilizing Naive Bayes classifier a characterization model for picture information is proposed. The proposed classifier handles the Naive Bayes classifier for limiting the preparation time returning exact characterization results. Information grouping implies a sort of information into various classification as indicated by rules. The point is to separate a sort of configuration from an example of objects. The characterization calculation gets the preparation set and assembles a model and that model is utilized to decide new articles. Choice tree is utilized, which is a tree structure, where a test on a trait addresses an inward hub, and a class name is holded by a leaf hub. A way is followed from the root to a leaf hub that holds the class forecast for that tuple. The benefits incorporates great precision and results inside least preparing time. The disadvantage is that when the size of the dataset is little then the accuracy will diminish.

AI is a field of study<sup>[9]</sup> that enables PCs to concentrate without being particularly modified. The machine gets grouping the given information and gets a choice limit and classifier for doing arrangement and forecast issues. Calculated relapse and neural network strategies for a face recognition procedure dependent on double pictures are utilized. These techniques move a shading picture to dark picture and afterward denoised utilizing a low pass channel. So the neighbourhood power deviations around eyebrows, eyelids, nose and mouth are caught by applying neighbourhood window standard varieties to the denoised picture. At that point, to binarize the image the versatile thresholding is utilized to get a great quality double image. The size of picture is standardized and decrement to its unique size utilizing closest neighbour interjection strategy. A face information base is made for each decreased size pictures. An identified face locale is passage by eliminating the non face area in the first picture. The paired picture got is utilized as a preparation dataset. The benefits incorporate minimization of computational space and preparing time. The downside is that occasionally it will be hard to distinguish and perceive text characters.

Transcribed character recognition <sup>[10]</sup> is consistently a space of exploration in the field of example recognition. There is a huge case for OCR on transcribed archives in Image handling. Highlight extraction and grouping are the means of character recognition measure. This by and large uses computerized picture preparing procedures, for example, Feature Extraction, Image Restoration and Image Enhancement. Optical recognition is a disconnected interaction for example subsequent to composing or printing has been finished the recognition will be started. The issue of character recognition can be characterized dependent on



two standards. One depends on the sort of the content which is printed or manually written and other is obtaining measure which can be on-line or disconnected. Disconnected character recognition for typewritten or written by hand character is analyzed as a paper archive and feasible as a parallel or dim scale image to recognize calculation. Due to high adequacy counterfeit neural organizations are generally utilized. The benefits incorporate picture quality improvement, featuring foundation and closer view. It is in reality tedious and expensive which is a downside.

This paper <sup>[11]</sup> suggests that after the happening to advanced PCs, blending human capacities to PCs. Productive calculations have progressed to execute the machines to perceive characters. Such a framework is named as Optical Character Recognition, produced for gathering character-based records from digitized images of printed or typewritten reports or manually written original copies. Level bed scanners or advanced cameras are utilized for digitizing. In the event of info gadget, there are two classes of character recognition: on-line and disconnected recognition. The principal bunch of frameworks utilizes gadgets like digitizer tablets for information procurement and keeping in mind that composing the acknowledgment will be finished. Static gadgets, for example, scanners and cameras are utilized to develop information for late frameworks. On-line recognition framework is for simultaneous information assortment design though disconnected recognition framework required clear strategies for making the image. It incorporates commotion expulsion and recuperation of the info picture to expell the harm created during the procurement cycle. Minimal expense and colossal preparing force and its benefits. In any case, the precision rate is low which turns into its inconvenience.

Neural network is assuming a significant part in transcribed character recognition <sup>[12]</sup> a space of example recognition. Written by hand characters are filtered and translated into 1024 twofold pixels. The skeletonization interaction will be utilized to double pixel picture and the extra pixels will be erased that are not existing to the assurance of the character and the overgeneralized terms has been diminished to thin lines. Characters are standardized into 30X30 pixel and moved to the upper left corner of pixel window. It utilizes neural Feed Forward Multi-player Perceptron network for English transcribed character recognition. For preparing, back-engendering calculation can be utilized. The contributions of MLP are skeletonized and standardized paired pixels of English characters. The eight-neighbour nearby technique has been kept up to remove the data of a limit transcribed character. The filtering is done until it discovers the limit and looking is in the clockwise direction. During the limit following interaction, the program will consistently check the principal directions of the limit are equivalent to the last facilitates or not. If got then the entire limit will be followed and limit following cycle is completed. Good precision is given which is a benefit. However, for that a great deal of exertion is required which turns into its downside.

The paper presents a profound learning match to no-reference quality evaluation of archive images <sup>[13]</sup>. Direct effect on the OCR execution is given by the archive quality. The corrupted record picture quality is accurately identified with the presentation of optical character recognition. Initially, the archive picture is parted into patches and non-educational patches are filtered out utilizing Otsu's binarization method. Furthermore, utilizing a Convolutional Neural Network (CNN) quality scores are achieved and the fix scores are adjusted to acquire the archive score. The CNN contains two layers of convolution, area dazzle max-min pooling, and Rectified Linear Units in the completely associated layers. With nearby standardization a dark scale record is pre-prepared then harvest the picture into patches, and the CNN think about quality scores for chosen fixes, and normal the quality scores to get a score for the picture. Partitioning the picture into patches brings weighty examples for preparing CNN. The benefit is that it can accomplish craftsmanship execution. Be that as it may, there will be huge varieties in a single valuation which is a hindrance.

The point of this paper <sup>[14]</sup> is to accomplish significant blunder decrease for different order issues organizations. A DNN is a feed-forward neural organization which has more than one secret nonlinear layer is proposed utilizing teachable initiation functions. The MNIST written by hand digit acknowledgment space is utilized for strength. It is principally used to address nonlinear enactment capacities in teachable structures and retrain utilizing a blunder back-spread system. Most importantly crop the patches from the pre-prepared pictures and check patches on twofold guide. On the off chance that the fix on the parallel guide is customary, fix is ignored. Isolating a picture into patches has two significant advantages. In the first place, it is simpler for the CNN to deal with patches of the full picture. Second, by isolating pictures into patches, the quantity of tests is intensely expanded. By editing pictures into patches get sufficient preparing tests. Since the preparation patches are totally named. Utilizing a stochastic slope plunge algorithm DNN for 50,000 ages was trained which utilized cross-entropy for the misfortune function. The Convergence rate is a lot quicker so concurrent preparing should be possible which is an advantage yet the exhibition isn't unreasonably much fulfilling.

Convolutional Neural Networks (CNN) is a well-known profound learning technique <sup>[15]</sup>. Deep learning carry out motorization of highlight extraction work. Perceiving transcribed content is harder than perceiving printed text. To separate features CNN is utilized. The character picture is prepared for eliminating every one of the undesirable substances from a picture which is done in the preprocessing stage. The information pictures are resized to an appropriate organization. The quantity of convolution layers, max pooling layers, ReLu layers and completely associated layers are picked. It is important to outline another dataset without any preparation on the grounds that no open source dataset is there for manually written Malayalam characters. Interpretation, Scaling, Sheering and Rotation are the four significant relative changes are interpretation, scaling, sheering and rotation. Test pictures are achieved by cutting the increased dataset. First it preprocesses the info picture and afterward breaks down the unlabelled test information. In the post handling stage, the whole number marks will be acquired as yield for the classifier. This whole number name will be changed into character Unicode. The benefits are it help to eliminate undesired characteristics of picture and furthermore diminishes overfitting. But hand tailored highlights are needed for discovering highlights in the content which is a downside.

Many record pictures are fixed with foundation images <sup>[16]</sup> like checks, store books, drive licenses, identifications, testaments, etc. Background pictures can be eliminated straight forwardly without losing the nature of text characters. To eliminate the shading foundation without changing the content in the forefront, the initial step is to guarantee each pixel has R, G

and B three qualities. On the off chance that the first pictures are not addressed in the shading space of RGB, at that point we ought to do the change. In a RGB picture, the pixel esteem is made out of R, G and B three qualities. For pixels of shading pictures, the three qualities are not quite the same as one another altogether. We utilize a non-straight change to grow the variety of each channel image, which can broaden the distinction of shading esteems for bright pixels while keeping up the dark pixels. After broadening, pixels from shading foundation champion and hence can be separated from text pixels adequately. It gives improved execution and furthermore powerful in eliminating foundation pictures. Yet there is an issue for high contrast pictures which is a downside.

The fundamental reason for the paper <sup>[17]</sup> is to isolate text data from complex foundation in shading archive pictures. A few records cause entanglements in perusing the archive contents. Thresholding technique is utilized for removing text from foundation. For the partition of text from the unpredictable foundation crossover approach is used. Thus depicts the up-and-comer text locales dependent nervous detection. Finally the limit an incentive for each distinguished content area is gotten naturally from the information. So it can deal with report pictures with changing foundation of different colourist and forefront text of any tone, textual style and size. Shrewd edge locator to identify edges. Each unified part is restricted in a bouncing box and standard deviation of the picture section encased in the jumping box is determined. The picture portions containing text are thresholded locally to isolate the forefront pixels from the foundation pixels. It gives great precision and better execution yet additionally neglected to identify single letter word which doesn't contain an opening which is really a downside.

The point of the paper <sup>[18]</sup> is that an augmented form of the strategy, which applies the neighbourhood shading division and the district developing. Techniques dependent on the shading grouping are proficient of isolating character designs from pictures with complex foundations. It can separate little characters with dainty strokes just as enormous characters with thick strokes from pictures with complex foundations. It is lenient toward the shadings of pictures. The k-closest neighbour technique is utilized. A grayscale picture is made from the shading picture briefly, and the k-closest neighbour pixels are found in the grayscale picture. To discover neighbourhood model tones in each sub picture, a histogram-based strategy is utilized. Every one of the shading segments is isolated by the scale factor and the shading histogram is determined. At that point, all peaks are identified from the histogram. Each named sub picture is examined corner to corner, and the edge benefit is performed on each sweep line. The advantage is that it can deal with even complex format however the inconvenience is that it comes up short on the capacity of separating shading.

The paper proposes a productive and compelling foundation assurance strategy for shading archive pictures <sup>[19]</sup> to tackle the issue in shading documents. The RGB shading space is transferred to dim level is done as the initial step. Furthermore, the limit between the foundation and the closer view is steadfast by examine line separation calculation. Third, the foundation and the frontal area dispersions entered are broke down to decide the foundation ranges. At long last, the report picture is changed over into double picture by the foundation range. The technique works in both dim level and shading record pictures. Likewise works behind the scenes circulate over the low, centre, and high dark level. It is more effective in light of the fact that it checks just the line pixels in the report picture which is a benefit yet since it contains some messed up characters as well.

Proficient Optical Character Recognition <sup>[20]</sup> in pictures caught from Set-Top Boxes <sup>[20]</sup> assumes a significant part in STB testing. OCR is performed with Tesseract 3.5 and Tesseract 4.0. The first preprocessing technique rehearsed for improving OCR execution for bad quality pictures is picture resizing. Resizing is applied just if the stature of the first picture is beneath 100px. The resizing technique utilizes rustic interpolation. The second preprocessing strategy is a picture honing strategy. Gaussian low-pass channel is utilized to get smoothed picture. The next technique is picture blurring, reduces high recurrence data and eliminates commotion from the images. The last preprocessing strategy is applied when picture has striking background. The thought is to seclude the content from the foundation. K-implies bunching calculation is utilized. It eliminates commotion and makes OCR execution more effective. In any case, if there should be an occurrence of more intricate foundations messes some up and besides troubles in perusing text which turns into a disadvantage.

#### IV. IMPLEMENTATION

We propose a system for serving the people who deals with day to day account settlement. It's a simple machine learning application in which a person with a smartphone can easily use. The objective of a computerized charging framework is to convey more noteworthy efficiencies by smoothing out your charging measures, permitting you to both set aside time and money. With independent companies framing a sizeable part of the Indian economy, there is a need to guarantee that they coordinate their bookkeeping framework with current instruments and applications for better administration of charging. Moreover this proposed system is totally depending upon the OCR technology even to read and process the bill images. It will provide a GUI interface for the user to take the image of a bill for recognition. Here we uses CNN - BiLSTM architecture for the character recognition. The python language is used for coding which helps in extracting the details.

As a first step we extract the details like name, items, cost, date, total, etc. from the bills using some python commands. Then we go for the preprocessing step, which is a crucial one. As we know that most of the bill images are noisy, for the good performance of an OCR and extraction of data it is very necessary to preprocess the images. This actually includes greyscaling, binarization and noise removal. The next step in the process is using an OCR. It is used to read text from the images. After that line segmentation and character recognition is performed. For this we uses an open source library called Tesseract which is actually a definite OCR library. The OCR utilizes a Long Short Term Memory (LSTM) network. The system model includes training line segmentation network, training character recognition network and billing. Once the informations are extracted which we think are necessary can be then stored into a CSV file, XML files, Excel sheets, etc. Thus we take each segments and

then convert to text. From the information, we actually sum up the total amounts in which each bill holds and then we can compare the total amount with that of present month and the previous month. So that we can make sure how much is gained between the months in the form of a graphical representation.

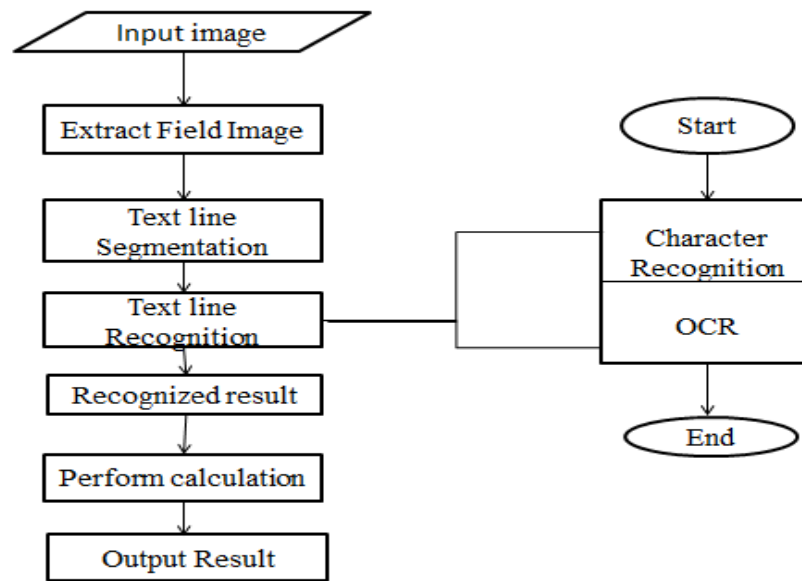


figure1: block diagram

The Smartphone should have at least 8 Megapixel camera specifications for picture quality. In case of processor the smartphone at least should have good processor because the system need to perform image processing, so this processor is minimum specification for this project. The performance of this project is purely based on the speed so the smartphone should have 2GB ram at least. The accuracy of the system totally depend upon the training of the model. The extent of our undertaking utilizing optical character recognition on a network foundation is to give a productive and improved programming apparatus for the clients to perform picture investigation, bill image preparing by perusing and perceiving the characters in business associations that will have a huge pool of bill. It gives an advantageous and simple way to keep up day by day bookkeeping works. It helps in smooth and successful working of a business. Gives a lot of exactness and to wellbeing ahead. It additionally offers heaps of adaptability, simple access and totally secure in numerous ways.

## V. CONCLUSION

Robotized charging frameworks have a possible answer for a wide scope of issues that are brought about by abrupt expansions in the quantity of customers, mistakes made during manual billing, etc. This is the best answer for lessening human intercession in receipt recognition. By decreasing human mediation we can build the uprightness and steadiness of the bill handling. Here we utilize CNN learning and OCR strategies to peruse and handle receipts. We can coordinate our accountings to get ready for future costs. It will help in business examination, can make the charging cycle later on a lot quicker and simpler and furthermore in more advantageous manner. This framework shows us that advancements like optical character recognition can be utilized to remove data from bills precisely. Thus, this project proposes a development which will serve a wide scope of clients and their tough undertakings into very simple ones.

## VI. ACKNOWLEDGMENT

We utilize this chance to convey our feeling towards all those that have helped United States directly or indirectly for the completion of our work. We tend to deeply and wholeheartedly impart adult male. Sreeraj R -HOD, applied science and Engineering for his extreme valuable recommendation and encouragement. We tend to particularly glad to our guide and supervisor adult male. Muneebah Mohyiddeen -Assistant prof, applied science and Engineering for giving Maine valuable suggestions and demanding inputs within the preparation of this paper. We might prefer to extend our sincere feeling to any or all school of applied science and Engineering department for the support and suggestions that helped United States within the development of our work to what it's currently. We tend to impart our folks and friends for the mental support provided throughout the course of our work the days once our energies were very cheap.

## VI. REFERENCES

- [1] YejunTang, L. Peng, Q. Xu, Y. Wang and A. Furuhashi, "CNN Based Transfer Learning for Historical Chinese Character Recognition", IAPR Workshop on Document Analysis Systems (DAS), IEEE, 2016
- [2] V Bharath and N. S. Rani, "A font style classification system for English OCR", International Conference on Intelligent Computing and Control (I2C2), 2017
- [3] Matteo Brisinello, Ratko Grbi, Dejan Stefanovic, Robert Peckai-Kovac "Optical character recognition on images with colourful background", IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin) 2018.
- [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) IEEE 14th International Colloquium, 2018.
- [5] A. Pise and S. D. Ruikar, "Text detection and recognition in natural scene images", International Conference on Communication and Signal Processing, 2014.
- [6] R. Deepa, Associate Professor Department of Information Technology Loyola-ICAM College of Engineering and Technology Chennai, India [ICECA] 2019 "Image Classification 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" - 2018 International Journal of Engineering & Technology
- [8] Shih-chung Hsu, 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", International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017
- [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, International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE, 2014.
- [12] A. Pal, D. Singh, Handwritten English character recognition using neural network, IJCSC (International Journal of Computer science and communication), vol 1, 2010.
- [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", 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", 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", 2017 International Symposium ELMAR.