

Abstract: a document
file. this technique
can be

[Design](#), [Architecture](#)



Abstract: It is difficult to recognize handwritten characters since every individual has a unique handwriting in terms of size, shape and angle. In the field of Image Processing, Character Recognition is one of the most challenging area. Many different methods have been established to recognize the handwritten as well as the printed text and convert them to a document file. This technique can be used in various applications like bank form, patient entry forms in hospitals, record sheets in schools, loan forms, tax forms and many other cases. This paper gives an overview of the various features used to recognize characters as well the different algorithms that are used to carry out optical character recognition (OCR). Keywords: Optical Character Recognition, offline and online handwriting recognition, Segmentation, Feature Extraction, Neural Network.

I. INTRODUCTION Optical Character Recognition(OCR) is a mechanism that converts the printed text into an editable text format and allows to save the document. While handwritten character recognition is recognition of handwritten text. The latter is more complex due to high variations in human writing styles. Hence handwritten character recognition requires sophisticated recognition techniques. It can be classified into offline and online handwriting recognition 2.

Offline recognition involves conversion of text to image. Various processes are applied to image like pre-processing, segmentation of individual characters from image, feature extraction, classification and post-processing. In online recognition two dimensional coordinates of successive points are represented as a function of time and order of strokes 2 3. This paper discusses various feature extraction and classification techniques required

<https://assignbuster.com/abstract-a-document-file-this-technique-can-be/>

for offline handwriting recognition. Accuracy of recognition is different when different classifiers are used and the accuracy percentage is approximately calculated. The acquisition and pre-processing of images is common for Fig 1: Key steps in handwriting recognition. all feature extraction and classification techniques.

II. PHASES OF RECOGNITION

1. Image Acquisition: The input image containing handwritten text is captured using a scanner or a camera in a specific format Eg. jpeg, . bmp.
2. Image Pre-processing: This step is essential as input image may not be in the form which is required for further analysis and classification.

1 Noise Reduction: The scanned input image is subject to various types of noise. This may happen due to many reasons like improper settings of camera/scanner, slow shutter speed, low light etc. Due to this there may be disconnected line segments, large gaps between lines, etc., so it is important to remove noise

2. Paper 2 has described the following methods of pre-processing for noise reduction.

1. 1 Filtering: Filtering is done to sharpen the edges and also to remove the various noise like salt and pepper noise, gaussian noise, etc.

Various spatial and frequency domain filters are used for removing noise and spurious points which may occur due to uneven surface or poor rate of sampling of acquisition device

2. In paper 4, median filter has been

mentioned to remove salt and pepper noise. The unwanted regions in the image can be removed using “ bwareaopen”. 2.

1. 2 Morphological Operations: These are used to remove noise due to low quality of ink and paper as well as erratic hand movement. It is usually performed after binarization. Morphological opening is used to remove noise in background, whereas morphological closing is used to remove gaps/holes in image. This is done to facilitate segmentation.

Paper 7 discusses removal of duplicated points only if the length of the sub-trace connecting them is smaller than 10% of the length of the diagonal of the bounding box that contains the whole trace. 2. 1.

3 Binarization: Binarization technique is used for conversion of image from grayscale to binary, this facilitates segmentation of individual characters 2. This is done by thresholding, i. e.

assigning minimum value (0) i. e black to a pixel below a certain value of threshold and a maximum value (white) to pixel value above threshold.

Binarization can be global (a threshold for entire image) and local (threshold value for each local area). Methods for Global thresholding are – fixed thresholding method, otsu method and kittler method, whereas local thresholding can be done using niblack method, adaptive method, sauvola method and bernsen method.

2. 1. 4 Skew Detection and Correction: Handwritten documents can have skewness. Many methods are used for detection, some of which rely on detecting connected components and finding average angles connecting

<https://assignbuster.com/abstract-a-document-file-this-technique-can-be/>

their centroids 2. As per paper 2, the skew angle is calculated and the skew lines are made horizontal.

3. Segmentation: It is one of the important phases in document analysis, where the text document is segmented into individual characters or digits. The Connected Component Labelling(CCL) algorithm can be used for segmentation of characters. The algorithm converts a binary image into a symbolic image with each connected component assigned a unique label. 4-connectivity or 8-connectivity can be used for labelling. The advantage of using CCL for segmentation is that it is not affected by skew object 4 10.

The three levels of segmentation as mentioned in paper 9 are: 3. 1 Line segmentation: It includes Baseline detection and skew correction. The image is scanned horizontally i. e pixel-row by pixel-row from left to right and top to bottom 9. The intensity of each pixel is tested and pixels are grouped together depending on the value of the pixels.

Every region indicates different content in the image file. The desired content can then be easily extracted. 3.

2 Word Segmentation: Unlike Line Segmentation, the image is scanned vertically from left to right and top to bottom. Again, the similar pixels are grouped together and the content can be extracted. Slant angle estimation is used to perform skew correction for the extracted word in heavy noise 9. 3. 3

Character Segmentation: It is the last stage of segmentation where every character is separated from the others. Few those are not connected, occur and hence precaution must be taken 9.

Segmentation methodologies as stated in paper 9 are pixel counting approach, histogram approach, smearing approach, waterflow approach and mixed approach. 4. Feature Extraction: There are two important aspects of character recognition: – feature extraction and classifier. In paper 7, the features were divided into 4 main categories which are global features of the symbol, crossing features, 2D fuzzy histograms of points and fuzzy histograms of orientations.

1) Global features: This is a set containing number of traces, angular change, line length, number of sharp points, mean x and mean y coordinate, covariance coordinates. 2) Crossing Features: This indicates the number of times a line intersects the other line. The shape of the symbol can be determined by dividing the region into set of horizontal and vertical lines and calculating the number of sub-crossings. To minimize the sensitivity to error it is advised to take average of parallel lines for a region. 3) 2D Fuzzy Histograms: This is the method in which a normalized region is divided into number of cells, where a grid of n-by-n cell will have $(n+1) * (n+1)$ corners and the points in each trace in this region are assigned a fuzzy membership. 4) Fuzzy Histogram of orientation: It is a zoning approach for description of shapes for off-line optical character recognition. The symbol is divided into a 2-by-2 grid and a grouping criterion is used. Three weighting criteria were also described.

Convolutional Neural Network: In paper 4, a Convolutional Neural Network (CNN) is used for Feature Extraction. In general, a CNN has three basic ideas: local receptive fields, shared weight and pooling. The local receptive field in

the input layer is connected to one neuron (hidden layer) and each connection has a weight and single-bias which is shared with other local receptive field in the same feature map. Feature map is a map from the input layer to the hidden layer. The weight and bias of each feature map are different thus allowing some of the features to be extracted from any position 8.

CNN also contains a sub-sampling layer used after the convolution layer. It implies the convolution layer output information. The proposed architecture of CNN in paper 4 uses 5 layers, the first 4 layers were two sets of convolution and sub-sampling layer and the fifth layer was the output layer. The output of first convolution layer had 6 feature maps of 24×24 pixels and that of second convolution layer had 12 feature maps of 8×8 pixels. The kernel used for convolution was of 5×5 pixels. The output of first sub-sampling layer was 12×12 with 6 feature maps and second sub-sampling layer had the output of 4×4 with 12 feature maps.

The output of second sub-sampling layer was transposed into a feature vector of 1×192 features. 5. Classification: Classification is the decision-making stage of the recognition system and it employs the features extracted from the previous stage 12.

Then depending on the comparison of all the features that are used, characters are classified in suitable classes and then recognized. 5. 1 Feedforward Neural Network Classifier: Paper 3 describes classification by using a feedforward neural network.

Diagonal feature extraction is used for classification. Using above mentioned feature extraction technique (described in paper 3), there are 54 features for each character image which forms the input layer. There are two hidden layers in this neural network with 100 neurons each. The output node has 38 neurons (26 alphabets + 10 numerals + 2 special symbols). Gradient-descent with momentum is the training algorithm used. 5. 2 SVM classifier: The Support Vector Machine(SVM) can be used as a high-end classifier 4.

The method proposed in paper 4 uses a linear SVM. To improve the accuracy of linear SVM regularization can be used to prevent overfitting during learning. Least Absolute Deviation(LAD) or L1 loss and Least Square(LS) or L2 loss techniques were used to generate the SVM model. The method could achieve an accuracy of 83. 37% 4. Paper 7 uses an adaptation of SVM for multiclass classification with two different kernel: Linear and RBF.

The RBF kernels achieved higher accuracy of 85. 89% for CROHME 2013 dataset and 94. 49% for CROHME 2012 dataset. But to get the highest accuracy the right values of C parameter which defines simplicity of the model and Gamma parameters which controls influence of training sample need to be determined.

The other two classifiers discussed are “ AdaBoost”. M1 with C4. 5 Decision Trees and Random Forests. The “ AdaBoost” method boosts the accuracy by correcting the previously misclassified sample.

The tree is evaluated by dividing the total weights in class by the total weight in current node. In Random Forests the attributes are randomized

also with an advantage that each decision tree is trained and evaluated independently of each other. 5. 3 Template Matching: According to paper 11, Template matching or matrix matching uses individual pixels as features. Classification is performed by comparing an input character image with a set of templates from each character class. A similarity measure is performed between the input image and the template. If the pixels are identical in both images then this leads to an increase in similarity measure.

If the pixels don't match then there is decrease in similarity measure. After all templates have been compared with the observed character image, the character's identity is assigned as the identity of the most similar template.

5. 4 Correlation coefficient: Corr2 function is used and the correlation between the input image and the template images is computed and the maximum correlation match is taken into consideration. The accuracy is different based on the database that is used. 6. Accuracy calculation: The accuracy graph is drawn based on the number of characters correctly recognized. The total number of properly recognized characters in the output image is divided by the total number of input characters.

This ration when multiplied with 100 gives the accuracy percentage.

Accuracy = number of correctly recognized characters / Total number of input characters. Fig 2: Accuracy graphIII.

COMPARATIVE STUDY Author Dataset used Classifier/Matching

Technique Accuracy Darmatasia, MohamadIvan Fanany4 NIST SD19 2nd

edition Support Vector Machine (SVM) 83. 37% Kenny Davila, Stephanie Ludi,

Richard Zanibbi7 CROHME20122013 SVM with Linear and RBF Kernel 94.

<https://assignbuster.com/abstract-a-document-file-this-technique-can-be/>

49%85. 59%J. Pradeep, E. Srinivasan, Himavathi3Feed forward Neural Network98. 54% Rakshana Shetty, Nitin Hiraje12Correlation coefficient78.

87%Table 1: Comparison TableIV. CONCLUSIONDifferent approaches for handwriting recognition were reviewed in this paper. Various pre-processing, segmentation, feature extraction, classification techniques have been discussed. From various studies we have seen that selection of relevant feature extraction and classification techniques play an important role in the performance of handwriting recognition system.

It is observed that with respect to accuracy, SVM and Feedforward Neural Network classifiers provide the best results but SVM has more time complexity as compared to neural network. Further to increase overall efficiency specific feature extraction and classification techniques can be used.