Chapter: 4
Proposed System

Introduction

4.1 Methodologies

4.1.1 Pre-processing

4.1.2 Feature Extraction

4.1.3 Classification

Conclusion

Introduction:

Bangla is the seventh most widely spoken language with about 250 million speakers in the Indian subcontinent still there are no trustworthy OCR engines for Bangla handwriting recognition. Complex natures of writing styles and a high degree of variability have made handwritten Bangla character is a big challenge for the researchers. These problems can be illuminated with the amelioration of feature extraction techniques. This paper presents a feature fusion approach using Histogram of Oriented Gradients (HOG) and Gabor filter to leverage the performance of the classification rate. HOG descriptor is used to extract gradient directional features by encoding the shape information of the object to be detected by intensity gradient distribution. Features delivered by Gabor filter specify a set of strong responses for locations of the target images which have structures in the given directions. Finally, these features are fed jointly to Extreme Learning Machine (ELM) classifier.

4.1 Methodologies

To amend the recognition of handwritten Bangla characters, a feature fusion based system has been adduced which has three major parts such as Preprocessing, Feature Extraction, and Character Recognition. The phases of the proposed system are depicted in Fig. 4.1.

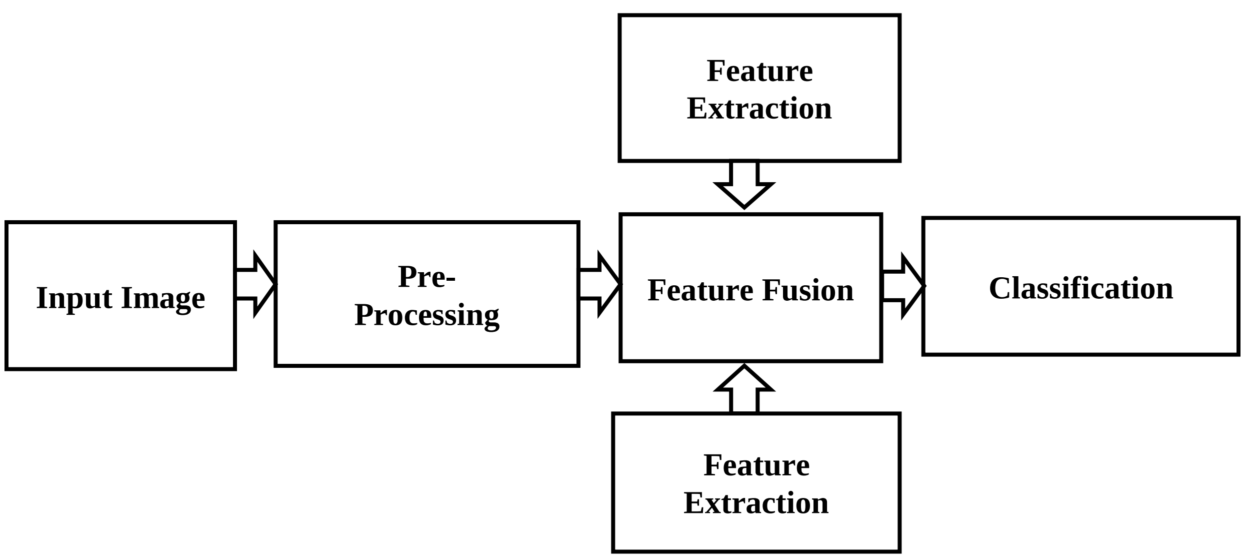

Fig. 4.1: Overview of proposed system

4.1.1 Preprocessing

Preprocessing of the scanned images are performed to lessen varieties in the writing style of individuals. The sequence of preprocessing steps consists of binarization, noise removal, skew detection and correction, segmentation and scaling.

- **Binarization:** At first, a scanned image is converted into a grayscale image from RGB image. Then the binarization process is performed on the grayscale image using Otsu's method [15] of global thresholding which finds the threshold that attenuates the intra-class variance.
- **Noise Removal:** Scanned images may contain noise that emerges as a result of the printer, scanner, print quality, age of the document, etc. Therefore, it is imperative to filter this noise before dealing with the image. Lowpass filter is generally used to smoothen the characters for later processing.
- **Skew Detection and Correction**: The text line of the acquired document image makes an angle with the horizontal direction is known as skew angle. Skew correction can be performed in two steps. At first, estimating the skew angle θ and then, rotating the image by θ, in the counter direction.
- **Segmentation:** Segmentation is a mechanism that disintegrates an image of sequence of characters into separate characters .The binary image is segmented into three different levels: line segmentation, word segmentation and character segmentation. The input image is scanned horizontally to detect text lines by calculating the frequency of black pixels in each row which forms row histogram. Word segmentation is performed by scanning each line vertically and the no. of black pixels in each column creates column histogram. Finally, the removal of headline of the word namely 'Matra' provides isolated characters [16]
- **Scaling:** Bangla script has various sized characters. The characters need to be uniformed before extracting features for more appropriate results. In this paper, each character is scaled into a 20 x 20 matrix.

4.1.2 Feature Extraction

Feature extraction plays the most crucial role in character recognition. Each character is represented as a feature vector in this phase. Here we have used two feature extraction manners namely Histogram of Oriented Gradients (HOG) and Gabor filter.

- **Histogram of Oriented Gradients**

The appearances of gradient orientations are calculated in confined parts of an image by HOG descriptor. In this approach, simple convolution kernel is used to convolve with each pixel as (4.1) & (4.2):

$$G_x = f(x+1, y) - f(x-1, y) \tag{4.1}$$

$$G_y = f(x, y+1) - f(x, y-1) \tag{4.2}$$

Where, G_x and G_y represent the horizontal and vertical components of the gradients respectively. Each image is partitioned into overlapping blocks. Each block contains a group of cells and the cell size is chosen according to the number of features required to identify the character. Then histogram of gradient directions is calculated by weighing the gradient and orientation computations in each cell [17]. Finally, gradient strength of each cell is normalized. Different cell size parameters are shown in Fig. 4.2. In this paper, 4×4 as the cell size has been chosen since it contains enough shape information to identify a character.

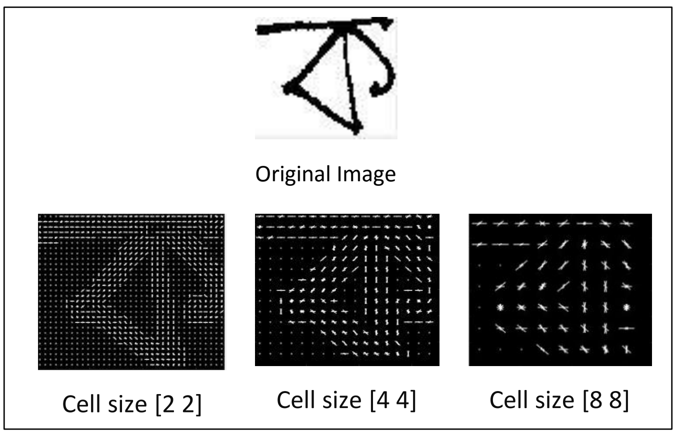

Fig. 4.2: HOG features of a Bangla character for different cell size

- **Gabor Filter**

Gabor filter has been extensively used in feature extraction. It is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. That is, Gabor filter can be viewed as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope as shown in (4.3)

$$g(x, y) = s(x, y)\, h(x, y) \tag{4.3}$$

Where, $s(x, y)$ is a complex sinusoid, known as a carrier, and h(x, y) is a 2-D Gaussian-shaped function, known as the envelope. The complex sinusoid is defined as (4.4),

$$s(x,y) = e^{-j2\pi(u_0 x + v_0 y)} \quad (4.4)$$

The 2-D Gaussian function is defined, in (4.5), as follows,

$$h(x,y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2})} \quad (4.5)$$

Thus the 2-D Gabor filter can be written as (4.6)

$$g(x,y,\varphi,\sigma_x,\sigma_y) = exp\left[-\frac{1}{2}\left\{\frac{x'^2_1}{\sigma_x^2}+\frac{y'^2_2}{\sigma_y^2}\right\}\right] \times (cos2\pi f x' + jsin2\pi f x') \quad (4.6)$$

Where, $x' = xcos\varphi + ysin\varphi$ and $y' = -xsin\varphi + ycos\varphi$. λ and φ represent the wavelength and orientation of plane wave, σ_x and σ_y are the standard deviations of Gaussian envelop along x-axis and y-axis but here $\sigma_x = \sigma_y$. The x-y plane is rotated by an angle φ, which will result in orientations. The value of φ is given by $\varphi = \frac{\pi(k-1)}{m}$, where k = 1, 2…m. Here, m denotes the number of orientations.

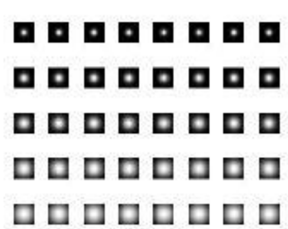

Fig 4.3: Magnitudes of Gabor filter

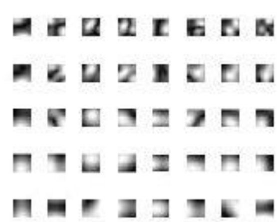

Fig 4.4: Magnitudes of filtered character 'ক'

A filter bank is formulated using Gabor filters with various scales and rotations [18] In this paper, a set of forty Gabor filters having 5 radial frequencies and 8 orientations has been implemented. The sample input images are convolved with the filters and thus creates Gabor space. The magnitudes of Gabor filters and a sample character image of 'ক' are shown in Fig.4.3 and Fig. 4.4 respectively. This process is closely related to processes in the primary visual cortex.

4.1.3 Classification

Classification is the concluding stage of character recognition task in which character images are assigned specific labels depending on the extracted features. The general trend in the current study of automatic character recognition has focused on high accuracy but has not considered the time taken to train the classification models, which should be an important factor for developing an OCR system. Therefore some classification models with high classification accuracy may not be satisfactory when considering the trade-off between the classification accuracy and the time for training the classification models.

In recent years, Extreme Learning Machine (ELM) has been perceptibly popular in classification tasks due to its high generalization ability and fast learning speed. ELM does not only obtain high classification accuracy but also avoid problems such as overfitting, local minima, and improper learning rate. Fig. 4.5 depicts the architecture of ELM algorithm.

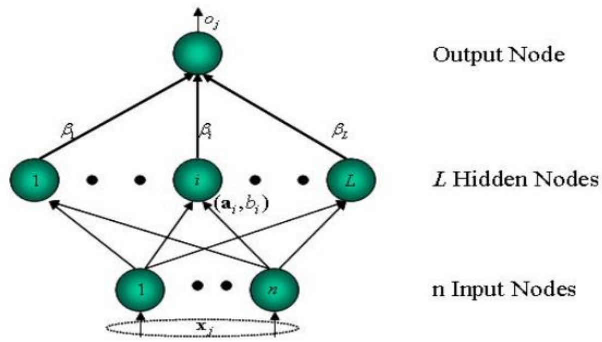

Fig. 4.5: Architecture of extreme learning machine [19].

The procedure of ELM for single-layer feed-forward networks can be expressed as follows:

Given, activation function g, and the number of hidden nodes L

1. Assign randomly input weight vectors or centers a_j and hidden node bias or impact factor b_i, $i = 1 \dots L$
2. Calculate the hidden layer output matrix, H.
3. Calculate the output weight β: $\beta = H^{\dagger}T$ where $H^{\dagger}$ is the Moore-Penrose generalized inverse of hidden layer output matrix, H.

Conclusion:

In this thesis paper, we have demonstrated the effectiveness of feature fusion based model for handwritten character recognition. The work presented here analyses the performance of offline handwritten Bangla character recognition using two feature sets—Gabor filter and histogram of oriented gradient features. These features have been fused and classified using single hidden layer extreme learning machine classifier. A fivefold cross validation scheme has been applied to measure the performance of the system.