

# Weighted Zone based Kannada Character Recognition Using Crack Feature Extraction Technique And Support Vector Machine Classification

Roopananda M K

*M.Tech Student*

*Department of Computer Science & Engineering  
UBDT College Of Engineering Davanagere Karnataka, India*

## Abstract

Many methods and technologies evolved in recognizing optical character but very less research made on recognizing handwritten character because of the complexity of variations of individual way of writing style. Many systems are available for recognizing optical character. In this paper, I have worked recognizing the character using crack feature extraction and well-known SVM classifier but worked on new concept of WEIGHTED zone based. More weight need to be given on the zone which has maximum number of character overlapping on that zone thereby increasing the efficiency and maximum probability of finding a part of character.

**Keywords:** Crack Code, Support Vector Machine, Optical Character Recognition

## I. CHARACTER RECOGNITION

Efficient methods involved in recognizing character that might be either optical character or handwritten character and also differentiated into online and offline character recognition. Online recognizing is the one where as and when a character is written it is recognized. Offline is the one where a digitalized image is fed as an input and thereafter preprocessing step is carried out and then fed its output to feature extraction technique and SVM to classify the image and in recognizing a character.

## II. A SURVEY

Handwritten Character Recognition is one of the emerging technologies in the field of archeology and crime detection of recognizing the handwritten of either victim and suspects. Few are said below. Handwritten character recognition of Devangiri Script using gradient feature technique [1]. Sobel and Robert operators are used for extracting gradient features and fed as input to SVM to find and study accuracy. Modified quadratic discriminant function used for recognizing Offline Bangla Handwritten Character [2]. Extracting Feature is purely on directional basis. For extraction technique, crack code concept is used where number of vectors is used for extracting features. Fuzzy model based can be used for recognizing the Bangla Character[3].

Very Less attention has made in work and research of Kannada Handwritten Character Recognition. Pal et. al [4] has proposed a technique of quadratic classifier based for recognizing offline handwritten character recognition of popular three south Indian languages Kannada, Tamil and Telgu. Kannada Character Recognition using FLD based technique is yet another way of popular way of recognizing a character [5].

## III. SYSTEM ARCHITECTURE

Zone based handwritten character recognition has four steps.

- 1) Database Creation.
- 2) Pre processing
- 3) Feature Extraction
- 4) Classification

All the above four steps are very important and have equal importance that will contribute in recognizing character. All the four steps involved above will be discussed in detail one by one.

### A. Database Creation:

Before carrying out preprocessing step first we need to create a database for making preprocessing step easier. Creating database for handwritten character recognition is very difficult and it is no inbuilt way or tool to create database. One need to collect different samples from people belonging to different department and should make database. Collecting database is difficult job as their may be chances of unconstrained handwriting , broken strokes, poor binarization. Since there is no standard database for handwritten kannada characters, we need to create our own database and store in a database



Fig. 1: Handwritten Kannada Characters

### B. Pre-processing:

In preprocessing step, we need to read a digitalized image i.e. raw image. Usually raw image contain noise and inaccuracy. Noise can be removed from median filtering thrice. After removing noise we need to do segmentation where we go for doing histogram equalization, plotting histogram and image binarization is carried out using well-known Otsu's method.

### C. Feature Extraction:

Feature extraction is one of the difficult step in overall recognizing a character. As there are different techniques for extracting a feature but the one technique I used is very famier one i.e... Crack Code

#### 1) Crack Code:

It is one of the way to encode boundry by connected straight line of length and direction. Boundry line between foreground and background that form or give raise to crack and encoding this leads to crack code concept.

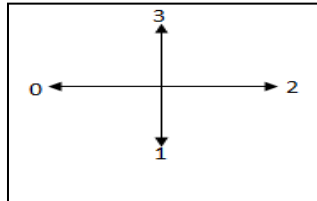


Fig. 2:

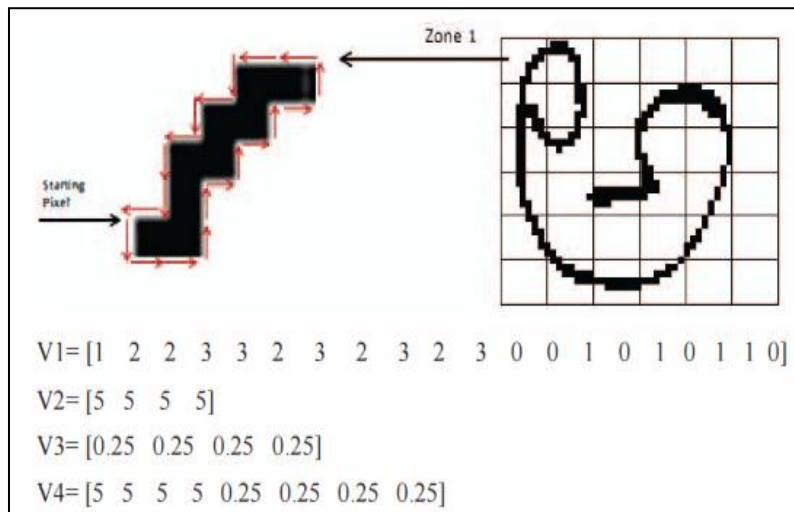


Fig. 3

#### 2) Steps Involved In Feature Extraction:

- 1) Step 1: Divide input image into 8x8 zones (thereby we will get 64zones)
- 2) Step 2: For each zone, we need to calculate or compute crack code by traversing the character portion in zone in anticlockwise direction.
- 3) Step 3: For each zone, we need to calculate feature using eq(1) and eq(2)
- 4) Step 4: Repeat steps 1 and 2 till we get feature vector of size 512

