Chapter 5
MULTILEVEL CLASSIFICATION
Chapter 5
Multilevel classification

5.1 Introduction
5.2 Multilevel Classification
5.3 Discussion of Results

In this chapter, a multilevel classification approach is described for handwritten Marathi character recognition. In this method we have divided the character set into six groups depending on special properties of the characters. This process of classification is carried out in four phases.

5.1 Introduction:

Marathi characters have more interclass and intra-class similarities. By experiment it has been observed that single feature is not sufficient for classification because recognition rate using single feature is very low. Also classifying and recognizing 41 characters is time consuming task. In multilevel classification approach Marathi characters are sub-classified into six groups depending on their special properties such as presence of bar, presence of enclosed region, presence of one component etc. Marathi 41 characters are sub-classified into six subclasses using four phases which takes tree structure and the
Chapter 5: Multilevel Classification

tree has four levels. Since the approach of sub-classification has four levels we are calling it as multilevel classification.

Chavan S. V. et.al (2013) has reported pre-classification approach for handwritten Devanagari character recognition based on location of vertical bar and number of components present in the character. Kale K. V. et.al (2014) has reported local structural sub-classification for Marathi compound character recognition. M. Hanmandlu et. al.(2007) has reported coarse classification approach for handwritten Hindi characters using presence of vertical bar, location of vertical bar and character is open to right or left side. Sushma Shelke and Shaila Apte (2011) has reported structural classification approach for handwritten Marathi compound character recognition.

It has been observed from literature that using single feature for classification of large number of classes is difficult. In the next section 4.2 we are describing multilevel classification approach where total numbers of Marathi characters are sub-classified into six subclasses using their special properties.

5.2 Multilevel classification:

The outcome of segmentation process is 41 isolated full characters, half characters, top modifiers and lower modifiers. These 41 isolated characters are further classified into six sub-classes in phase I to phase IV as shown in the Fig. 5.1.
Chapter 5: Multilevel Classification

Fig. 5.1: Phases in Multilevel classification

Following is the discussion on phase I to phase IV sub-classification where 41 characters are classified into six classes:
5.2.1 Phase I Sub-classification:

Phase I sub-classification is based on bar characters and no bar characters. Marathi consonants are broadly classified into two major categories bar characters and no bar characters. Bar characters are those characters having presence of vertical bar. In order to verify whether a bar is present in the character, vertical projection of image was taken and if any column contains more than 70% of black pixels, then label it as bar character otherwise label as no bar character. As shown in the Fig. 5.2 column number 57 contains 86% black pixels hence labeled as bar character and as shown in the Fig. 5.3 all columns contains less than 70% black pixels hence labeled as no bar character. Thus 41 characters are classified into two subclasses, first contains 28 bar characters and second contains 13 no bar characters as shown in phase I of Fig 5.1.

Fig. 5.2: Bar character
5.2.2 Phase II Sub-classification:

Phase II sub-classification is based on presence of enclosed region. In phase II, 28 bar characters are broadly classified into two major categories having enclosed region or not. To verify whether enclosed region is present, the numbers of holes are counted in the character using eight connectivity as shown in the Fig. 5.4. If one or more than one enclosed region exists in the character then label it as enclosed region character otherwise not enclosed region character as shown in Fig. 5.5. Now 28 bar characters are sub-classified into 18 enclosed region bar characters and 10 not enclosed region bar characters as shown in phase II of Fig. 5.1.

Similarly 13 no bar characters are classified into six enclosed region no bar characters and seven not enclosed region no bar characters as shown in phase II of Fig. 5.1.
5.2.3 Phase III Sub-classification:

Phase III sub-classification is based on number of components present in character. In phase III, 18 bar enclosed region characters are classified into two subclasses, depending on whether number of component is one or more as shown in Fig. 5.6 and Fig. 5.7. Presence of a component can be verified using region properties for each labeled region.
In all 18 bar enclosed region characters are sub-classified into 14 characters having a component and four characters having two components as shown in phase III of fig. 5.1.

![Fig. 5.6: Two component character.](image1)

![Fig. 5.7: One component character](image2)

### 5.2.4 Phase IV Sub-classification:

Phase IV sub-classification is based on number of rows containing at least one black pixel. We have 14 bar enclosed region characters having one
component and further classified into two subclasses, depending on whether character's 80% rows contains at least one black pixel in first 75% columns as shown in phase IV of Fig. 5.1. Out of 14 bar enclosed region characters having one component we get 10 characters satisfying above condition as shown in Fig. 5.8 and four characters does not satisfy the condition as shown in Fig. 5.9.

Using above sub-classification method problem of handwritten Marathi character recognition is simplified into six sub-classes as follows:
Chapter 5: Multilevel Classification

**Sub-class I:** Bar not enclosed region (10 characters)

![Fig. 5.10: Consonants having bar and enclosed region](image)

**Sub-class II:** Bar enclosed region with two components (4 characters)

![Fig. 5.11: Consonants having bar, enclosed region and having two components.](image)

**Sub-class III:** Bar enclosed region with one component and having 80% rows contains at least one black pixel in first 75% columns (10 characters)

![Fig. 5.12: Consonants having bar, enclosed region, one component and black pixels.](image)

**Sub-class IV:** Bar enclosed region with one component and less than 80% rows contains at least one black pixel in first 75% columns (4 characters)

![Fig. 5.13: Consonants having bar, enclosed region, one component and not black pixels.](image)

**Sub-class V:** No bar enclosed region (6 characters)

![Fig. 5.14: Consonants does not have bar and having enclosed region.](image)
Sub-class VI: No bar not enclosed region (7 characters).

![Marathi characters]

Fig. 5.15: Consonants does not have bar and enclosed region.

The problem of classification of 41 isolated Marathi characters is now simplified to small problems. Total 41 Marathi characters are divided into 6 different sub-classes as discussed above. Feature extraction methods suitable for different sub-classes are discussed in the next chapter 6.

5.3 Discussion of Results:

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Level</th>
<th>Phase</th>
<th>Special Property</th>
<th>Result in Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>I</td>
<td>Phase I</td>
<td>Presence of BAR</td>
<td>100%</td>
</tr>
<tr>
<td>2.</td>
<td>II</td>
<td>Phase II</td>
<td>Presence of Enclosed Region</td>
<td>100%</td>
</tr>
<tr>
<td>3.</td>
<td>III</td>
<td>Phase III</td>
<td>Presence of Number of components</td>
<td>100%</td>
</tr>
<tr>
<td>4.</td>
<td>IV</td>
<td>Phase IV</td>
<td>Presence of at least one black pixel in 80% rows</td>
<td>100%</td>
</tr>
</tbody>
</table>
Chapter 5: Multilevel Classification

Table 5.2: Outcome of Multilevel classification

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Subclass</th>
<th>Characters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Subclass I</td>
<td>फर्ख पण न प्रत्यक्ष कः</td>
</tr>
<tr>
<td>2.</td>
<td>Subclass II</td>
<td>क्रिया वाण</td>
</tr>
<tr>
<td>3.</td>
<td>Subclass III</td>
<td>कः श्च ष्कः श्च</td>
</tr>
<tr>
<td>4.</td>
<td>Subclass IV</td>
<td>नः खः तः बः</td>
</tr>
<tr>
<td>5.</td>
<td>Subclass V</td>
<td>सः ठः टः ठः</td>
</tr>
<tr>
<td>6.</td>
<td>Subclass VI</td>
<td>टः डः रः ठः अः</td>
</tr>
</tbody>
</table>

Experimental results for multilevel classification are shown in Table 5.1. We got 100% accuracy for experimental results for all four levels. The outcome of multilevel classification approach described in this chapter is shown in Table 5.2. After application of multilevel classification approach total 41 handwritten Marathi characters are sub-classified into six subclasses in four phases. Subclass I contains 10 characters, Subclass II contains four characters, Subclass III contains 10 characters, Subclass IV contains four characters, Subclass V contains six characters and Subclass VI contains seven characters. Further on these subclasses various feature extraction techniques will be applied for classification purpose.

In the next chapter various feature extraction techniques and algorithms to extract features are elaborated.