Chapter 4

IMPROVED FUZZY SUPPORT VECTOR MACHINES FOR FACE RECOGNITION

4.1 Introduction
The chapter will describe the new membership function for fuzzy SVM. The proposed membership function combines distance feature, correlation and fuzzy clustering. The distance concept is using to identify the overlapping between classes, while the correlation is using to represent the similarity between points and fuzzy clustering membership is using to handle outlier points.

4.2 Membership Function
A membership function is used to quantify linguistic term and represent a fuzzy set graphically. A membership function for a fuzzy set x on the universe of discourse A is defined as $\mu(x): A \rightarrow [0, 1]$. Here, each element of A is mapped to a value between 0 and 1. It is called membership value or degree of membership.

For example, weight is used linguistically when its value are described as “heavy,” “not too heavy,” “light,” quantitatively it we take its value to be measured in kilograms. Linguistic variable can be replied by function. For example if weight between 0 and 100 kg are possible we can extent to which an object x is heavy by function $\mu$ that take value between 0 and 1 for x between 0 and 100, the value $\mu(x)$ indicating to that degree the object is heavy. Then $\mu(x)$ is called membership function.
Thus, a membership function is a curve that defines how each point in the input space is mapped to a membership value or degree between 0 and 1.

4.2.1 Membership Function using Distance

One of the most popular bases for measuring contributions of sample point is a function of distance between each sample point to its class center in fuzzy membership. Here, this concept is applied to recognize the overlapping between classes. This concept is shown in Figure 4.2.

![Figure 4.2 The Distance Criteria](image)

Figure 4.2, for class+ the Point A is the outermost point of class+, and point B is the nearby point from class- to class+ center. Thus, \( r_1 \) is maximum radius distance between the class+ center (x+) and point A whereas \( r_2 \) is minimum radius to point B. If \( r_1 > r_2 \) then we can say there is overlapping between classes. So, the fuzzy membership function based on distance \( \mu_1 (x) \) is defined as in equation (40):
\[
\begin{cases}
1, & \text{if } \|x_+ - x_i\| < r_2 \\
1 - \frac{\|x_+ - x_i\|}{r_1}, & \text{if } r_2 \leq \|x_+ - x_i\| \leq r_1
\end{cases}
\] (40)

Here, Hodges-Lehmann (HL) method used for calculating the class centers. As mention below:

\[ \text{HL} = \text{median } x_i - x_j / 2, \ 1 \leq i \leq j \leq n \] (41)

n- The number of data

HL is using median, it is more trustworthy than using the mean value. Because the mean calculates the average value of an entire class it is very sensitive to outliers.

For example, if there is a point which have a wide range value as compared to other data, then the mean value of this class can be greatly unusual. But, HL calculates the location of the center of classes by using ranges of data value in a class.

### 4.2.2 Membership Function using Similarity

The similarity is the second variable used for generating the fuzzy membership. To show the similarity, we need to calculate correlation between variables. The correlation is calculated by using covariance. Covariance indicates, how two variables are related. According to B. Sarwar [106], the correlation using covariance as given below equation (42):

\[ \text{Sim} (x_i, x_j) = \text{corr} (x_i, x_j) = \frac{\text{cov} (x_i, x_j)}{s_x s_y} \] (42)

Where

\[ \text{Cov} (x_i, x_j) = \frac{\sum (x_i - \bar{x}_i)(x_j - \bar{x}_j)}{n - 1} \]

\[ s_x = \text{sample standard deviation of the random variable x} \]

\[ s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x}_i)^2}{n - 1}} \]

\[ s_y = \text{sample standard deviation of the random variable y} \]
Generally the similarities are compared as data with its class center. Data with low similarity considered as outlier. Here, we also used correlation to find the similarity between each point to its class center. The fuzzy membership function based on similarity \( \mu_2(x) \) is defined as in equation (43):

\[
\begin{align*}
1 & \quad \text{if } \text{abs}(\text{corr}(x_i, x_j)) \geq 0.9 \\
\text{abs}(\text{corr}(x_i, x_j)) & \quad \text{if } \text{abs}(\text{corr}(x_i, x_j)) < 0.9
\end{align*}
\]

(43)

In the equation (43), if any point that have correlation coefficient greater than 0.9 than we can say strong relationship. Then, the membership degree for those points is 1. Otherwise, correlation will be close to 0 when there is no relationship exists.

4.2.3 Membership Function using Clustering

We used fuzzy c-means clustering methods to decide clusters. These clusters comprise both normal and outlier data points [103]. Set 1 fuzzy memberships of these data points and fuzzy memberships of other data points are calculated by using closest cluster. For optimized result set clustering termination as maximum number of iteration is 25 and the objective function improves by less than 0.001 between two consecutive iterations.

Fuzzy c-means (FCM) is a clustering method that allows each data point to belong to multiple clusters with varying degrees of membership.

FCM is based on the minimization of the given objective function \( J_m \)-

\[
J_m = \sum_{i=1}^{D} \sum_{j=1}^{N} \mu_{ij}^m \|x_i - c_j\|^2
\]

(44)
Where

1. $D$ is the number of data points.
2. $N$ is the number of clusters.
3. $m$ is fuzzy partition matrix exponent for controlling the degree of fuzzy overlap, with $m > 1$. Fuzzy overlap refers to how fuzzy the boundaries between clusters are, that is the number of data points that have significant membership in more than one cluster.
4. $x_i$ is the $i$th data point.
5. $c_j$ is the center of the $j$th cluster.
6. $\mu_{ij}$ is the degree of membership of $x_i$ in the $j$th cluster. For a given data point, $x_i$, the sum of the membership values for all clusters is one.

FCM performs the following steps during clustering:

1. Randomly initialize the cluster membership values, $\mu_{ij}$.
2. Calculate the cluster centers:

   \[
   c_j = \frac{\sum_{i=1}^{D} \mu_{ij}^m x_i}{\sum_{i=1}^{D} \mu_{ij}^m}
   \]

3. Update $\mu_{ij}$ according to the following:

   \[
   \mu_{ij} = \frac{1}{\sum_{k=1}^{N} \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{m-1}}
   \]

4. Calculate the objective function, $J_m$.
5. Repeat steps 2–4 until $J_m$ improves by less than a specified minimum threshold or until after a specified maximum number of iterations.
4.3 Integration of Membership functions

Briefly, the algorithm of membership function generation can be written as the following stages:

Algorithm-

Step 1: First of all, find the membership degree $\mu_1(x)$ based on distance. By using equation (40).

Step 2: Find the membership degree based on Pearson correlation $\mu_2(x)$. By using equation (43).

Step 3: To find out membership degree based on clusters $\mu_3(x)$. Perform clustering on the training data set.

a) Apply clustering algorithm.

b) Perform clustering on the training data set.

c) Determine a subset containing clusters that contain both normal and abnormal data. Denote this subset as ALLCLUS.

d) For each data point $x \in$ ALLCLUS, set its fuzzy membership to 1.

e) For each data point $x \in$ ALLCLUS, do the following:

1. Find out the cluster whose center is closest to $x$.
2. Calculate fuzzy membership of $x$ with this cluster.

Step 4: Find the membership degree $\mu(x)$ based on different observations, by taking mean as:

$$\mu(x) = \text{mean} \left[ \mu_1(x), \mu_2(x), \mu_3(x) \right]$$

(46)
### Table 4.1 Membership Value Generation

<table>
<thead>
<tr>
<th>$\mu_1(x)$</th>
<th>$\mu_2(x)$</th>
<th>$\mu_3(x)$</th>
<th>$\mu(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.3333</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.3333</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.6666</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.3333</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.6666</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.6666</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

$$\mu(x) = \begin{cases} 
1 & , if \mu(x) \geq 0.6666 \\
0 & , Otherwise 
\end{cases} \quad (47)$$