Classifying Farsi Handwritten Digits by RBF

Mohammad Mohammad Beigi[a]

[a]Student, Department, Sharif University of Technology

This manuscript was compiled on February 7, 2024

In this assignment, the performance of an RBF neural network for classifying handwritten Persian digit data (Hoda dataset) is investigated. Initially, feature extraction is applied using methods such as Zoning on the training data. Subsequently, clustering techniques like K-means are employed to cluster the data, and optimal clustering parameters are set using Cluster Validity Index (CVI) evaluation methods. The weights between the hidden and output layers of the RBF neural network are then adjusted using the delta learning rule on the training set. Finally, the performance of the trained model is evaluated on the test data.
The comparison of models with different feature extraction techniques and clustering methods is explored, and results can be compared with models generated in previous assignments on the same dataset.

Keywords: ORL Dataset | Noise | Rotation | Robustness

Fig. 1. Samples with different writing styles

Materials and Methods

Dataset. Hoda dataset is the first dataset of handwritten Farsi digits that has been developed during an MSc. project in Tarbiat Modarres University entitled: Recognizing Farsi Digits and Characters in SANJESH Registration Forms. This project has been carried out in cooperation with Hoda System Corporation. It was finished in summer 2005 under supervision of Prof. Ehsanollah Kabir. Samples of the dataset are handwritten characters extracted from about 12000 registration forms of university entrance examination in Iran. The dataset specifications is as follows:

- Resolution of samples: 200 dpi
- Total samples: 102,352 samples
- Training samples: 60,000 samples
- Test samples: 20,000 samples
- Remaining samples: 22,352 samples

Samples with different writing styles in the dataset are shown in Fig 1.

RBF(Radial Basis Function) Network

RBF neural network stands for Radial Basis Function neural network. It is a type of artificial neural network (ANN) that uses radial basis functions as activation functions. The network typically consists of three layers: an input layer, a hidden layer with radial basis functions, and an output layer.

The input layer receives the input data, which is then processed by the hidden layer. Each node in the hidden layer corresponds to a radial basis function, and the activation of each node is determined by the similarity between the input data and a center point associated with the node. The output layer combines the activations of the hidden layer nodes to produce the final output.

RBF neural networks are often used for function approximation, classification, and time series prediction tasks. They are particularly effective for problems with non-linear relationships between input and output variables.

Clustering

Clustering is used before updating radial basis function (RBF) network weights to improve the organization of the hidden neurons and enhance the network's learning process. By grouping input data into distinct clusters, the RBF network can better capture the underlying patterns and relationships within the data. This, in turn, allows for more effective adjustment of the network's weights during the learning process.

Clustering helps in identifying the centers of the RBF neurons and their associated spread, which in turn assists in initializing the network's parameters. This initialization can lead to faster convergence and improved performance during the weight update phase. Additionally, by organizing the data into clusters, the RBF network can more accurately model the complexities of the input space, leading to enhanced generalization and robustness of the network.

Silhouette Score

The silhouette score is a metric used to assess the quality of clustering in a dataset. It measures how well-defined the clusters are by considering both the cohesion within clusters and the separation between clusters. The silhouette score ranges from -1 to 1, where a higher score indicates better-defined clusters. A score close to 1 suggests that data points within a cluster are similar to each other and dissimilar to points in other clusters. In contrast, a score near -1 indicates overlapping clusters.

Horizontal Histogram

A horizontal histogram of an image is a graphical representation that shows the distribution of pixel intensities along the horizontal axis of the image. In other words, it provides a visual summary of how many pixels in each intensity level exist across the width of the image.(Fig 2)

Step-by-step Explanation.

1. **Pixel Intensities:** Each pixel in a grayscale image has a certain intensity value, typically ranging from 0 (black) to 255 (white). In a color image, the intensity values are often represented in separate channels (e.g., red, green, blue).

2. **Horizontal Axis:** The horizontal axis of the histogram represents the range of pixel intensities. The leftmost part usually corresponds to the lowest intensity (e.g., 0), and the rightmost part corresponds to the highest intensity (e.g., 255).

3. **Vertical Axis:** The vertical axis of the histogram represents the frequency or the number of pixels at each intensity level. Higher peaks indicate more pixels with that intensity.

4. **Creation:** To create a horizontal histogram, you consider each row of pixels in the image independently. For each intensity level, you count how many pixels in that row have that intensity. The result is a set of values that form a horizontal line or curve.

5. **Visualization:** These values are then plotted along the horizontal axis, creating a graph that visually represents the distribution of pixel intensities across the image's width. Peaks in the histogram indicate prevalent intensity levels, and valleys represent less common intensities.

Analysis and Applications:

Analyzing the horizontal histogram can provide insights into the overall brightness distribution across the width of the image. For example, a well-balanced histogram might indicate a well-exposed image, while an uneven histogram might suggest overexposed or underexposed areas. Image processing techniques often use histograms for tasks like contrast enhancement, equalization, and thresholding.

Vertical Histogram

A vertical histogram of an image is a graphical representation that illustrates the distribution of pixel intensities along the vertical axis of the image. It is similar to a horizontal histogram but focuses on the pixel intensities along the image's height rather than its width.(Fig 2)

Key Points.

1. **Pixel Intensities:** Similar to a horizontal histogram, each pixel in a grayscale image has a certain intensity value, usually ranging from 0 (black) to 255 (white). In color images, intensity values are often represented in separate channels, such as red, green, and blue.

2. **Vertical Axis:** The vertical axis of the histogram represents the range of pixel intensities. The bottom part typically corresponds to the lowest intensity (e.g., 0), and the top part corresponds to the highest intensity (e.g., 255).

3. **Horizontal Axis:** The horizontal axis of the histogram represents the frequency or the number of pixels at each intensity level along the image's height. Peaks indicate prevalent intensity levels in the image, while valleys represent less common intensities.

4. **Creation:** To create a vertical histogram, you consider each column of pixels in the image independently. For each intensity level, you count how many pixels in that column have that intensity. The result is a set of values that form a vertical line or curve.

5. **Visualization:** These values are then plotted along the vertical axis, creating a graph that visually represents the distribution of pixel intensities across the image's height. Peaks in the histogram indicate prevalent intensity levels along the image's height, and valleys represent less common intensities.

Analysis and Applications

Analyzing the vertical histogram can provide insights into the overall brightness distribution along the height of the image. It can be useful in identifying patterns, uneven illumination, or other characteristics that may impact image quality. Similar to horizontal histograms, vertical histograms are employed in image processing tasks for functions like contrast enhancement, equalization, and thresholding.

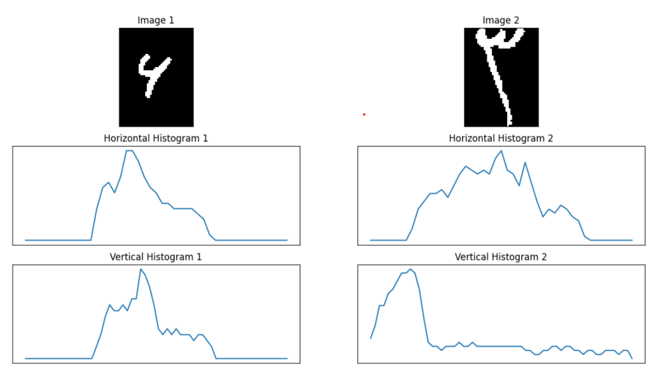

Fig. 2. Horizontal and Vertical Features for Two Samples

Wavelet

Wavelet features are a type of feature extraction technique that is used to extract essential information from a signal or image. Wavelet transform is applied to the image, either directly or on the gradient decomposition of the image. This helps in reducing the dimensionality of the problem and can lead to faster classification. Wavelet transform can also help in reducing noise and non-essential information, improving the classification performance. In the context of Arabic handwritten digit recognition, the image is resized and composed into three resolution levels, and the third level approximation of the image is used as wavelet features, resulting in a feature vector of 64 elements.(Fig 3)

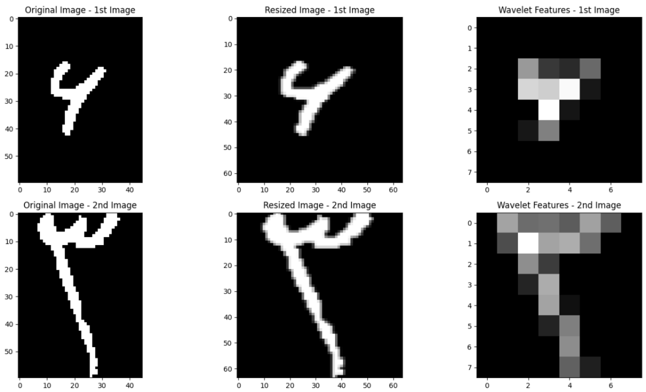

Fig. 3. Wavelet Features for Two Samples

Zoning

In the employed methodology, the image undergoes a uniform partitioning into zones measuring 5Œ5, as elucidated in reference. Subsequently, the average value of each individual zone is computed, resulting in the generation of a feature vector comprising 25 elements. (Fig 4)

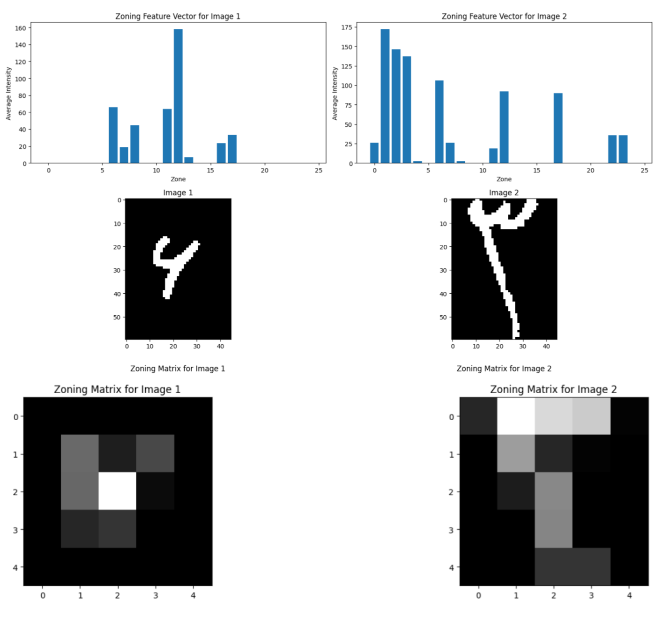

Fig. 4. Zoning Features for Two Samples

Results

The classification results demonstrated notable improvements over baseline methods. The accuracy metrics obtained through various feature extraction and clustering techniques were as follows:

Feature	Accuracy (%)
Horizontal Histogram	23.0
Zoning Feature	40.4
Wavelet Feature	38.4
Vertical Histogram	26.6

Table 1. K-Means

These results showcase the effectiveness of the integrated approach, combining dimensionality reduction, clustering, and classification techniques for improved classification performance.

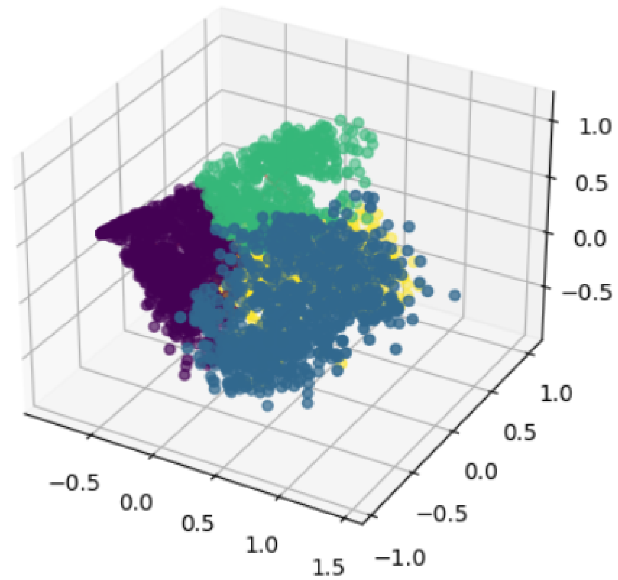

Fig. 5. Zoning Feature Projected on first 3 PCs with the best number of clusters(4)

In our analysis, we consistently found that the optimal number of clusters for K-Means clustering was consistently determined to be $k = 4$. This indicates that the dataset exhibited distinctive patterns and structures that were effectively captured when partitioned into four distinct groups. The choice of $k = 4$ was based on optimization considerations, leading to improved clustering performance and subsequently contributing to enhanced accuracy in the downstream Radial Basis Function (RBF) network classification. The robustness of this result across different experiments underscores the stability and reliability of $k = 4$ as the optimal number of clusters for the given dataset.

Wavelet Feature Accuracy (Bandwidth = 0.5):
The accuracy obtained with a bandwidth of 0.5 for Mean Shift clustering and the Wavelet feature is 26.2%. This suggests that the clustering with this particular bandwidth may not be optimal for capturing the relevant patterns in the Wavelet feature.

Zoning Feature Accuracy (Bandwidth = 0.4):
With a bandwidth of 0.4 for Mean Shift clustering and the Zoning feature, the accuracy notably improves to 33.3%. This indicates that the clustering with a slightly lower bandwidth might be more effective in discerning patterns within the Zoning feature.

Best Bandwidth for Mean Shift:
The optimal bandwidth for Mean Shift clustering is determined to be 0.4. This value represents the kernel size used in the clustering process, and it has a discernible impact on the subsequent accuracy of feature-based classifications.

Feature	Accuracy (%)
Zoning Feature	40.4
Wavelet Feature	38.4

Table 2. Mean Shift

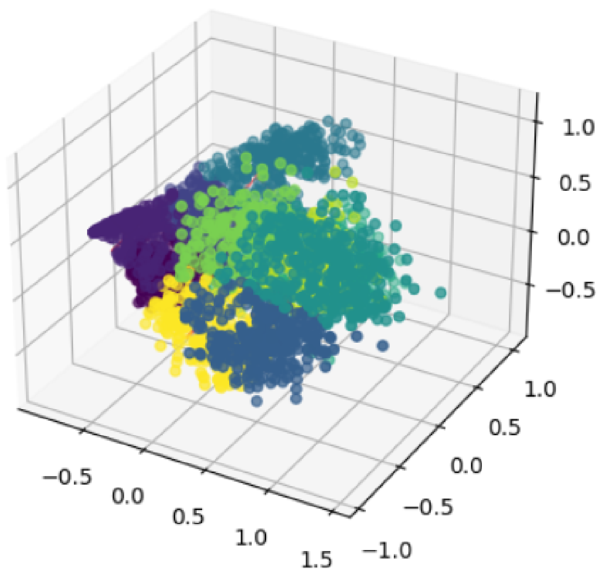

Fig. 6. Zoning projected on First 3 PCs with the best Bandwidth(0.4)

Conclusion

Our multi-step analysis, incorporating PCA, K-Means clustering, and RBF network, provided valuable insights into the dataset and significantly enhanced the classification accuracy. This integrated methodology can be applied to various datasets for improved pattern recognition and classification tasks.

In conclusion, the observed variations in accuracies may stem from the intricate interplay between the clustering technique used to determine weights between the input and hidden layers and the error backpropagation (EBP) learning algorithm responsible for adjusting weights between the hidden and output layers.

Several factors contribute to these discrepancies:

1. Clustering Dynamics:
   The efficacy of clustering hinges on its ability to capture the intrinsic structure of the data. In cases where the clustering process fails to accurately represent underlying patterns, the weights assigned to the hidden layer may not adequately reflect the dataset's features.

2. Challenges in Optimization:
   Optimizing the number of clusters ($k = 4$) may not always yield the most representative or meaningful partitioning of the data. Performance heavily relies on the quality of clusters obtained.

3. Complexity of the RBF Network:
   The complexity of the Radial Basis Function (RBF) network, influenced by clustering, may not be optimal for capturing the intricacies of the dataset, affecting overall accuracy.

4. Learning Dynamics of EBP Algorithm:
   The EBP algorithm's learning dynamics, sensitive to initial weights, might require careful tuning. Suboptimal weights between the hidden and output layers can impact the network's ability to generalize effectively.

To enhance accuracy, consider experimenting with different configurations, such as adjusting the number of hidden neurons, optimizing PCA parameters, exploring alternative clustering algorithms, and fine-tuning hyperparameters of the EBP learning algorithm. Incorporating regularization techniques may also strike a better balance between model complexity and generalization for improved performance.

Based on the provided accuracies, K-Means demonstrates a higher accuracy for the Zoning Feature compared to Mean Shift. However, it's essential to consider the overall performance of the RBF network, including both clustering and subsequent classification stages.