# CHAPTER 7

# 7. WRITER IDENTIFICATION TOOL

As identification of individual plays vital role in signature verification, biometric authentication, forensic analysis, the promising results obtained in this research motivated to develop an interactive writer identification tool. This chapter briefly elucidates the development of tool and outlines the work flow of the tool. The writer identification tool has been developed using MATLAB for predicting the identity of writer based on their handwritten Tamil text. The prototype of writer identification model described in chapter 6 based on the CNN is incorporated with MATLAB GUI for developing the tool.

## 7.1 DEVELOPMENT OF TOOL

Two independent classifiers (i) word level trained model (ii) paragraph level trained model, are integrated into a single writer prediction system, which can accept a Tamil handwritten text (either a single word text or a paragraph) as input and can predict the writer of the handwriting. Here the tool is developed based on offline text dependent writer identification.

The tool is designed and developed with four modules namely input, preprocessing, feature extraction, training or prediction. The Tamil handwritten text of a person whose identity is to be predicted is converted into image in offline mode and stored. The input module is designed to browse the input image and passed into the next component. The preprocessing module is intended to preprocess the input image by invoking either word text image processing function or paragraph text image processing function (described in chapter 3). Various preprocessing tasks are carried out according to the type of the text.

The third module is considered to capture the features of the preprocessed text image. According to the type of the text i.e., word or paragraph, the respective feature extraction process is called and various discriminative global and local features are extracted. The feature vector is formed for which the class label is to be predicted. So the dummy class label is assigned to the feature vector enabling supervised classification. This feature vector is then fed into the classifier based on the CNN for training or prediction.

This tool also has the facility to include new persons. The fourth module is designed for two purposes, one is training and the second is prediction. (i) In case of training, the new writer is added into the existing database and his / her writing patterns are updated with existing corpus. The tool can be used to carry out preprocessing and feature extraction. The corresponding class labeled feature vectors are generated through the above modules and augmented with current training dataset. The learning of the updated training dataset is done by calling the respective routine and the new model is built. (ii) In case of prediction, this component takes the feature vector with dummy class label (normally 0) as input to the learned model based on the CNN (with accuracy 95.6% ) and the classifier then predicts the class label or identity of the writer. Since it is observed from the experimental results about 95.6 % of prediction accuracy, the prediction model and the writer identification tool can very well predict the identity of writer accurately and efficiently.

The sample screenshots showing the opening screen and the components of the tool are shown in Fig. 7.1 and Fig. 7.2

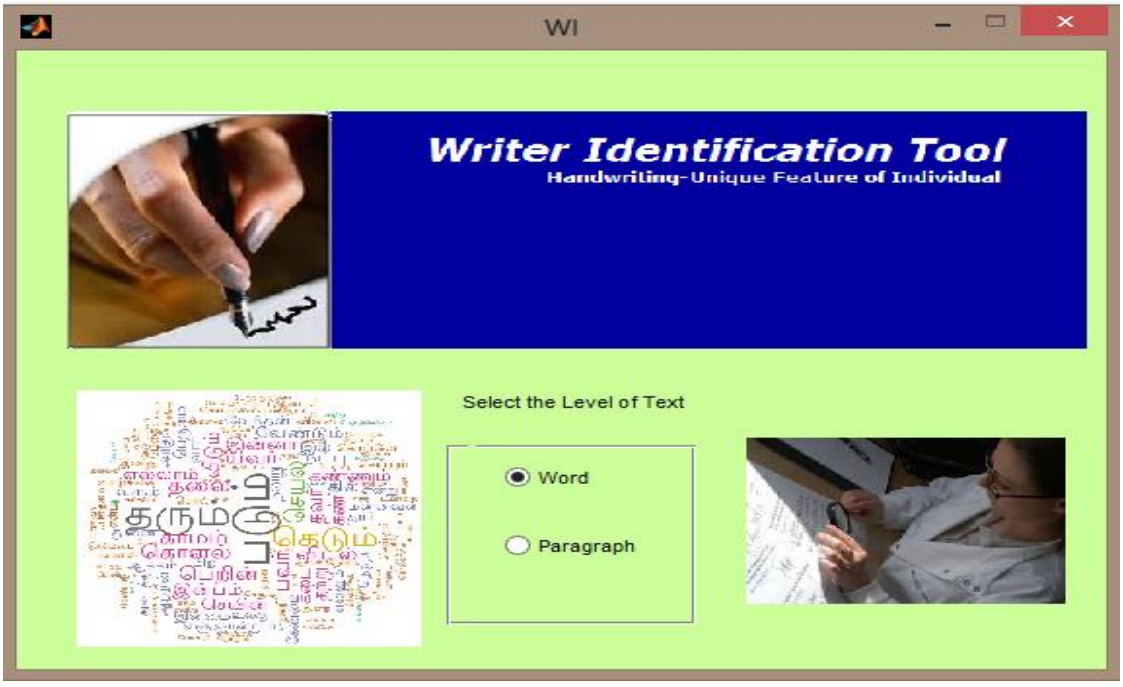


**Fig. 7.1 Opening Screen of the Tool**

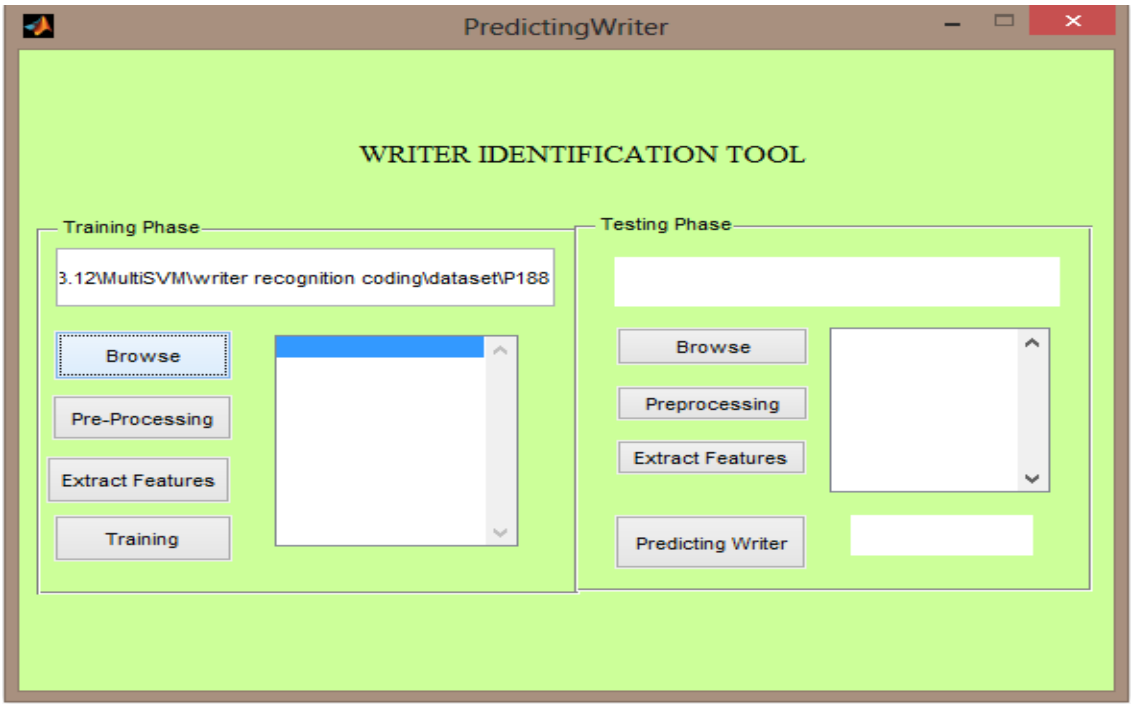


**Fig. 7.2 Components of the Tool**

***Workflow of the tool***

The working process of the tool is outlined below.

**Step 1:** A Tamil handwritten text is obtained as image from the file (offline mode)

**Step 2:** A MATLAB function is used to differentiate the text as word or paragraph.

**Step 3:** When 'preprocessing' option is selected, according to the level of the text, the preprocessing function is invoked and preprocessing is carried out.

**Step 4:** When 'feature extraction' option is selected, according to the level of the text, the feature extraction module is executed and a feature vector is created.

**Step 5:** In case of samples of new writers, the option 'Training' is selected and goto Step 6. In case of prediction, the option 'Prediction' is selected and goto Step 7.

**Step 6:** The feature vector is assigned writer identity as class label, the feature vector is added into the dataset and new CNN based model is built.

**Step 7:** The dummy class label is assigned to the feature vector and the labeled feature vector is then fed into the respective classifier i.e. word level trained model or paragraph level trained model. The model will then predict a class label of the writer. Finally the name of the writer will be displayed.

The sample screenshots depicting the workflow of the writer identification tool are enclosed in Appendix D.

## 7.2 SUMMARY

This chapter demonstrated the development of writer identification tool based on Tamil handwritten document. Since the writer prediction model has produced about 95.6% of prediction accuracy, it is ascertained that the writer prediction system can very well predict the identity of writer accurately and efficiently. Therefore it is highly significant that this writer identification tool can be made as an integral part of forensic document analysis, biometric authentication system and signature verification system in order to provide timely and accurate information to the user.