EDU

EAST DELTA
UNIVERSITY

Banknote Classification using Machine Learning

Course Title: Digital Image Processing
Course Code: CSE 439
Section: 01

Submitted To:

Muhammad Asif Hasan
SoSET, East Delta University

Submitted By:

Mishkat Hossain (Id: 201012012)
Abrar Hossain Chowdhury (Id: 201005512)
Md Ahnaf Akif (Id: 201012112)
Tasfia Tabassum (Id: 201001512)
Mushpika Karnayn (Id: 201005412)

Abstract*—Automatically recognizing and classifying different categories of Bangladeshi currency notes. This project was achieved by traditional Machine Learning (ML) using image processing and classification techniques. It was conducted on a dataset called “Bangla Money” which was obtained from Kaggle. The dataset contains 1637 training images, 333 testing images and 9 categories. A number of preprocessing techniques such as – Gray scaling, denoising, edge detection, binary thresholding and contour detection were applied on the dataset. The preprocessed** images **were classified with classifiers such as – K-Nearest Neighbors algorithm (KNN), Support Vector Machines (SVM), Random Forest and lastly Gaussian Naïve Bayes. The project shows the highest accuracy resulting 98% using the SVM classifier. A Graphical User Interface (GUI) was also developed that takes an input image and classify it using SVM classifier. Other than all these, no deep learning technique and their respective libraries were used.

***Keywords*—SVM, KNN, Random forest, Gaussian Naïve Bayes, Traditional Machine Learning, preprocessing**

I. INTRODUCTION

The official currency of Bangladesh is Bangladeshi Taka (BDT). The latest series of Bangladeshi currency notes was issued in 2011. In this project, we created a system for classifying Bangladeshi currency notes. To achieve this system, we have used a set of preprocessing techniques.

II. WORKING ENVIRONMENT

We used python programming language and google colab (a data science Cloud based research IDE) instead of Local support machine. We used Visual Studio Code for building the GUI from scratch.

III. IMAGE PREPROCESSING

The most challenging and crucial step of the project was the preprocessing phase. It took two weeks to design the step by step logic of preprocessing. Before classification it is necessary for the images to go through the preprocessing phases. All the preprocessing steps are shown below:

A. Grayscale conversion: In this step all the colored images are converted to grayscale for simplification of the images. Grayscale conversion was applied to the colored images of the currency notes. This enhances the contrast of the images while reducing noise.

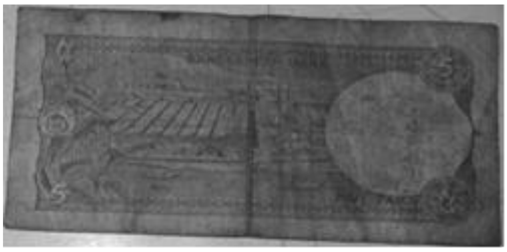


Fig: Gray-scaled Image

B. De-noising images: In order to further minimize noise, a denoising filter known as the Median blur was used to the grayscaled images. This is done for smoothening of the gray scaled images. Here, Median Blur has been used. Because it showed better result than the Gaussian Blur.

Fig: Denoised Image

C. Edge detection: The edges of the currency notes were extracted using an edge detection method. For this, we employed the Sobel filter. This makes it easier in identifying the various characteristics of the notes.

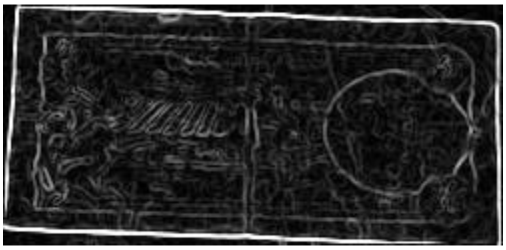


Fig: Sobel Filtered Image

D. Binary Thresholding: The grayscale pictures were converted to binary images using a binary thresholding technique. This simplifies the images and makes them easier to classify. In order to get a better threshold value we used histogram analysis.

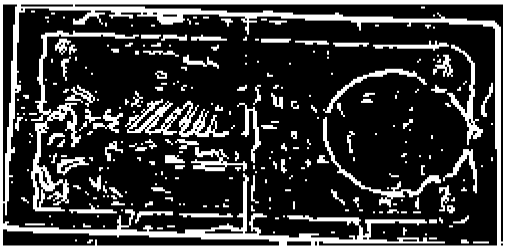


Fig: Thresholded Image

E. Contour detection: The contours of the currency notes were identified using a contour detection technique. We've also highlighted the contours found by the algorithm. This helps in identifying each category of notes uniquely.

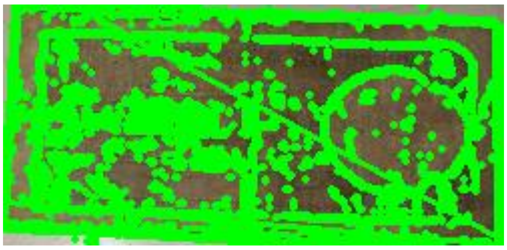
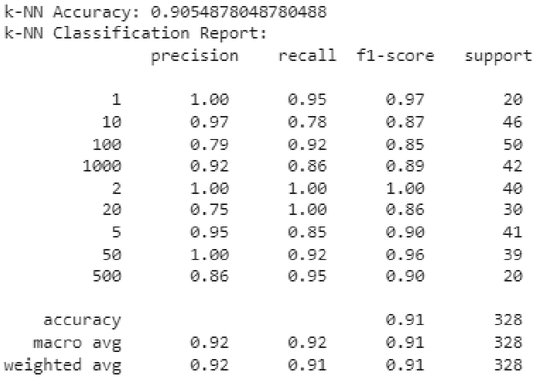


Fig: Contoured Image

IV. CLASSIFICATION MODELS

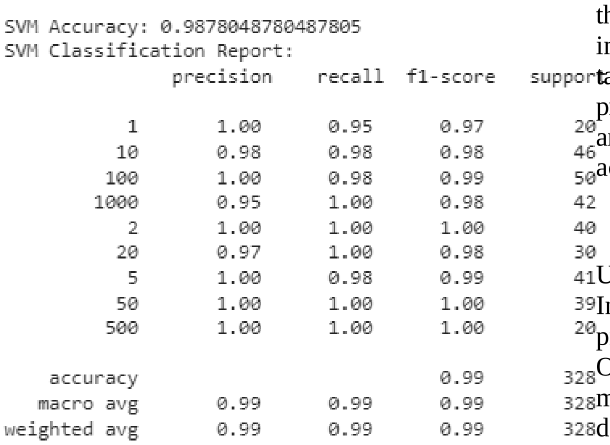
After the preprocessing, comes the classification procedure. With their respective algorithm models are trained. With these trained model classification is done. Accuracy, F1-score and others are prompted. All the classification algorithm and their report are provided.

A. K-nearest neighbor: The K-nearest neighbors' method, often known as KNN or k-NN, is a supervised learning classifier that makes predictions or classifications about how a single data point will be grouped.



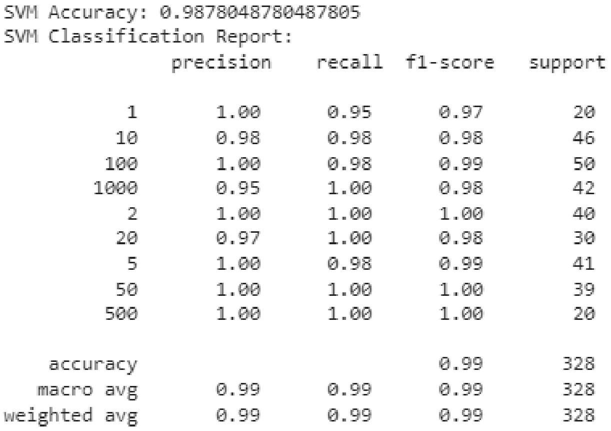
```
k-NN Accuracy: 0.9054878048780488
k-NN Classification Report:
              precision    recall  f1-score   support

           1       1.00      0.95      0.97        20
          10       0.97      0.78      0.87        46
         100       0.79      0.92      0.85        50
        1000       0.92      0.86      0.89        42
           2       1.00      1.00      1.00        40
          20       0.75      1.00      0.86        30
           5       0.95      0.85      0.90        41
          50       1.00      0.92      0.96        39
         500       0.86      0.95      0.90        20

    accuracy                           0.91       328
   macro avg       0.92      0.92      0.91       328
weighted avg       0.92      0.91      0.91       328
```

B. Support Vector Machines: The main objective of the SVM method is to find the best hyperplane in an N-dimensional space that can classify the data points into different groups in the feature space.



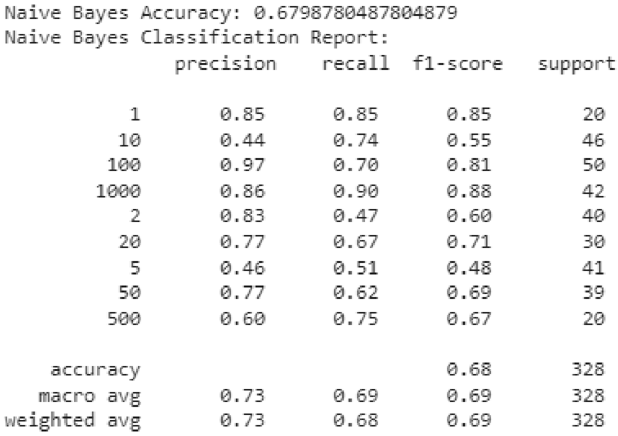
```
SVM Accuracy: 0.9878048780487805
SVM Classification Report:
              precision    recall  f1-score   support

           1       1.00      0.95      0.97        20
          10       0.98      0.98      0.98        46
         100       1.00      0.98      0.99        50
        1000       0.95      1.00      0.98        42
           2       1.00      1.00      1.00        40
          20       0.97      1.00      0.98        30
           5       1.00      0.98      0.99        41
          50       1.00      1.00      1.00        39
         500       1.00      1.00      1.00        20

    accuracy                           0.99       328
   macro avg       0.99      0.99      0.99       328
weighted avg       0.99      0.99      0.99       328
```

C. Random Forest: Random Forest is a common machine learning strategy that combines the output of numerous decision trees to obtain a single conclusion.



```
SVM Accuracy: 0.9878048780487805
SVM Classification Report:
              precision    recall  f1-score   support

           1       1.00      0.95      0.97        20
          10       0.98      0.98      0.98        46
         100       1.00      0.98      0.99        50
        1000       0.95      1.00      0.98        42
           2       1.00      1.00      1.00        40
          20       0.97      1.00      0.98        30
           5       1.00      0.98      0.99        41
          50       1.00      1.00      1.00        39
         500       1.00      1.00      1.00        20

    accuracy                           0.99       328
   macro avg       0.99      0.99      0.99       328
weighted avg       0.99      0.99      0.99       328
```

D. Gaussian Naive Bayes: In Gaussian Naive Bayes, the continuous numerical attributes are assumed to be normally distributed.



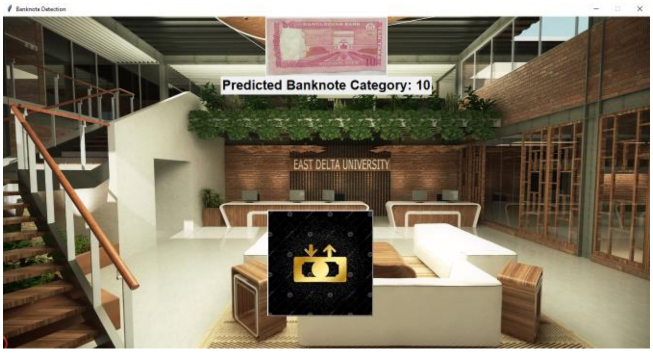
```
Naive Bayes Accuracy: 0.6798780487804879
Naive Bayes Classification Report:
              precision    recall  f1-score   support

           1       0.85      0.85      0.85        20
          10       0.44      0.74      0.55        46
         100       0.97      0.70      0.81        50
        1000       0.86      0.90      0.88        42
           2       0.83      0.47      0.60        40
          20       0.77      0.67      0.71        30
           5       0.46      0.51      0.48        41
          50       0.77      0.62      0.69        39
         500       0.60      0.75      0.67        20

    accuracy                           0.68       328
   macro avg       0.73      0.69      0.69       328
weighted avg       0.73      0.68      0.69       328
```

V. PREDICTION

At first the input image is taken for preprocessing. Then, all the trained models gives prediction of the provided input image, that what type of bank note it is. To predict, we have taken the an image as input, then past it through the preprocess phase and then classify it with Random Forest and SVM models, since both this algorithm showed better accuracy.

VI. GRAPHICAL USER INTERFACE

Using the python Tkinter Library for Graphical User Interphase deployment that runs on local machine, the same preprocessing steps has been applied on the entire Dataset. Only SVM classification Algorithm has been used to train a model. The input image of a Bangladeshi taka is taken from desktop or connective mobile phone and passed through the preprocessing steps and then predicted using the model. The entire system is built on Visual Studio Code.



Predicted Banknote Category: 10

Fig: The GUI deployment

VII. EXTRA

In this project we have preprocessed and classified and got the result accuracy very well. We didn’t need any types of feature extraction for the scenario of our project. Still we have showcased HOG-Feature extraction. Here are the screenshot of the original and the applied HOG feature extraction.

Fig: Original image of the note

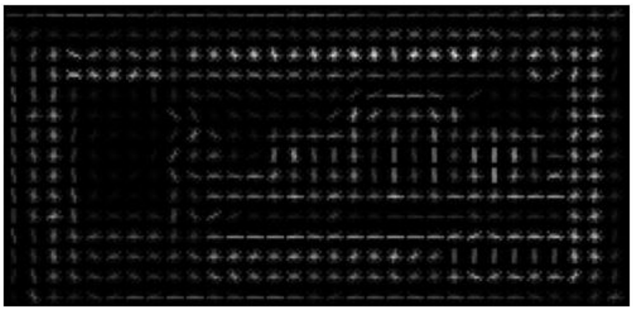


Fig: Applied Hog feature

VIII. CONCLUSION

We all know that features designed by users and experts to handle large and composite datasets have to face limitations. That’s how traditional machine learning techniques is being used to analyze satellite image, medical image etc. For this reason, we have applied this image processing techniques in our project step by step. And the results of the project showed the traditional machine learning and image processing techniques were able to achieve a classification accuracy of 98%. This is a promising result as it suggests that these techniques can be used to develop a reliable system for classifying Bangladeshi Currency Notes.