**Interpreting Doctor’s Handwritten Presciption Using Deep Learning Techniques**

Rizwanullah Mohammad, Ajay Kumar Varma Nagaraju and Suneetha Manne

[mailtorizwanullah@gmail.com](mailto:mailtorizwanullah@gmail.com)

[ajaykumarvarma@gmail.com](mailto:ajaykumarvarma@gmail.com)

[hodit@vrsec.ac.in](mailto:hodit@vrsec.ac.in)

**Abstract:** A Doctor’s Handwriting Recognition model can predict (recognize) the text present in the doctor’s prescription, by feeding image of that medicine name as an input to the model and the model processes the image with deep neural network and it predicts the text present in the image and it gives the final medicine name as digital text.This model is suitable only for Text written in English Language and not suitable for other languages of texts written in prescription. The model based on training dataset the output it produce may get varied and based on images training count. Both convolution layers and Bi-LSTM layers can be used for feature extraction and recognizing text respectively.

**Keywords:** Bi-LSTM Layers, Convolution Layers, Adam optimizer, Relu Activation Function.

**1 Introduction**

It is most common that people can’t understand and interpret the doctor’s handwriting. The calligraphy they follow which is always challenging for ordinary people and even for pharmacist to understand doctor’s handwriting. Until they understand correctly then cant give correct medicine to Patient. Due to usage of wrong medicines they may face severe consequencies with respect to their health. This problem need to be solved with the latest technologies we are having at present. The solution for this is deep learning models. A deep learning model can take large input of data and can process with help of neural network and layers. They can give high accuracy and more reliable. Now with the help of deep learning techniques involving all the terms in deep learning to provide a optimal solution for this. The Bi-LSTM model can provide a solution which can predict text present doctor’s prescription’s image which we passed as input to our model.

**2 Literature Study**

Here are the available models, That Researchers proposed and developed. They Proposed different methodologies to recognize text in an image and showcasing their performance.

Handwriting recognition for medical prescriptions using a cnn-bi-lstm model (2021) [1]:

Authors presented a model employing the BI-LSTM Model for the recognition of a doctor's handwriting. They have only created a model; no mobile or web applications have been created to execute the model in real time. In order to minimise overfitting and increase the model's resistance to noise, they employed data augmentation approaches. The feature extraction is carried out using Convolutional Neural Networks with many layers, and the decoding of the extracted features into English letters is assisted by Bi-LSTMs. We employ Connectionist Temporal Classification to get past the fact that the true alignment between the input and the output is unknown. To properly identify language specific to prescriptions provided by the doctors, more bias is applied to terms that are present in a manually produced corpus.

Recognition of doctors’ cursive handwritten medical words by using Bidirectional lstm and srp data augmentation (2021) [2]:

Authors presented a model for the identification of doctors' handwriting (Bangla Handwriting). After employing the SRP Augmentation approach, they achieved an accuracy of 89%. Some of the participating physicians' prescription pictures were made available. A collection called the "Handwritten Medical Term Corpus" was produced with 17,431 handwritten examples of 480 medical terms in English and Bangla. The introduction of SRP, a data augmentation technique, increased the size of the data sets. For predicting the handwriting of doctors, an online character recognition system utilising Bi-LSTM was employed. As long as characters are gathered as time-series data of coordinates, the SRP approach may be used to other datasets. These datasets can also be utilised to extend offline characters' data if they are enlarged and converted to picture data.

Doctor’s cursive handwriting recognition system using deep learning (2019) [3]:

Authors presented a model for interpreting the doctor's handwriting was developed. They employed CRNN Model. They created a smartphone application that allows users to enter images and receive digital text as output. Of the 540 input photos, 389 have been successfully recognised.The accuracy of the tests conducted using the mobile application was 72%. The model is implemented through the use of a mobile application called DCHRS, which stands for "Doctors' Cursive Handwriting Recognition System" and aims to recognise the name of the medication and usage instructions inside the image of doctors' cursive handwriting that has been captured, as well as to provide the normal text version of the handwriting.

**3 Proposed Work**

**3.1 Dataset**

The proposed model was trained with 86800 images in gray scale. We used IAM Dataset which contains handwritings of different people with different fonts. For training we splitted dataset into 90:5:5 Ratio. 90% of training data, 5% of validation data and 5% of testing data.

**3.2 Pre-processing**

In this pre-processing phase the images will get reshaped to width of 128 and height of 32 and padding to 99. And then the datatype will be changed to float32 which we call it as casting. Thıs may increase the model performance.This is distortion free image processing.

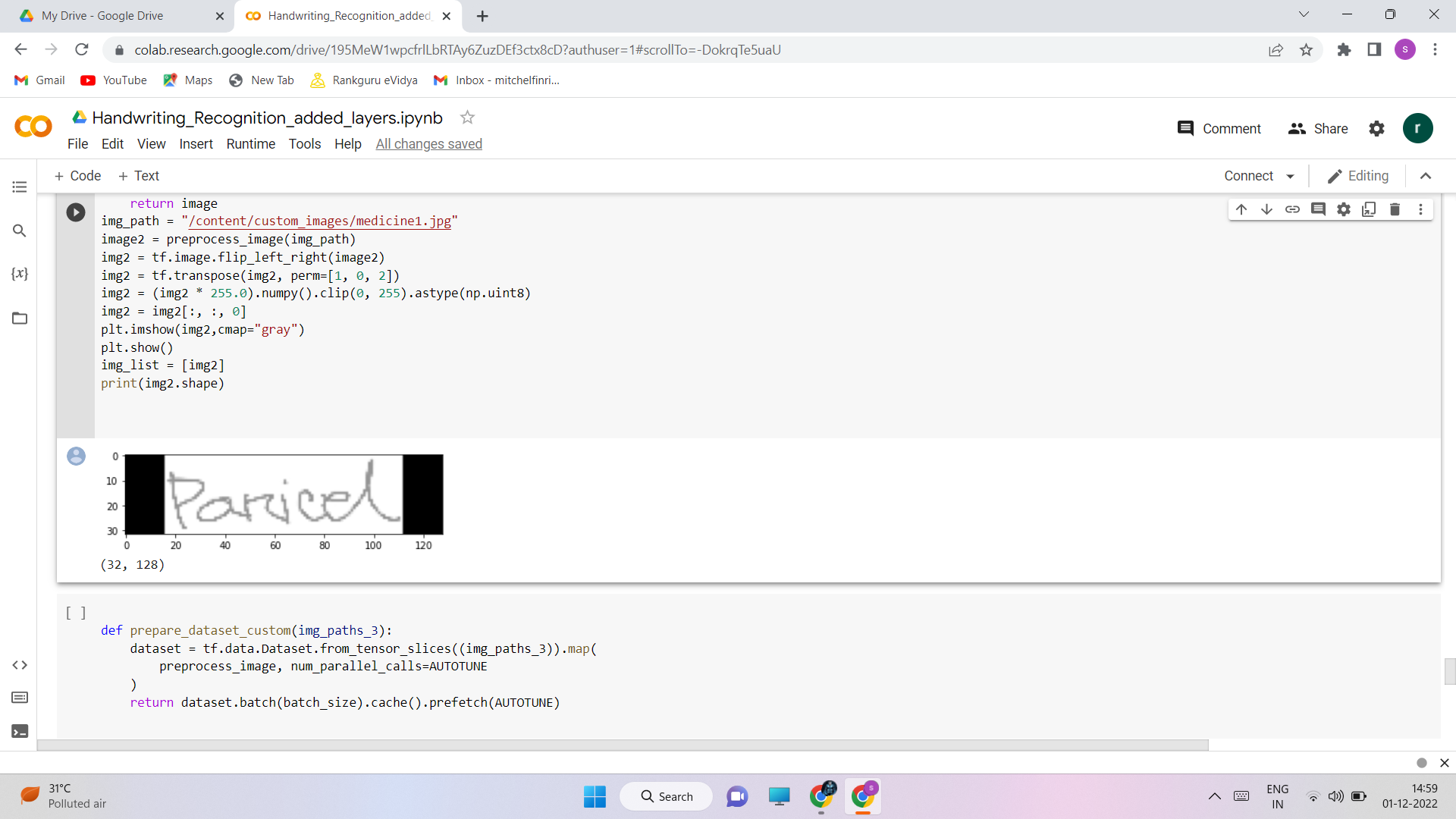


Fig. 1. Distortion free image pre-processing

**3.3 Design Methodology**

The below diagram describes architecture of our work.

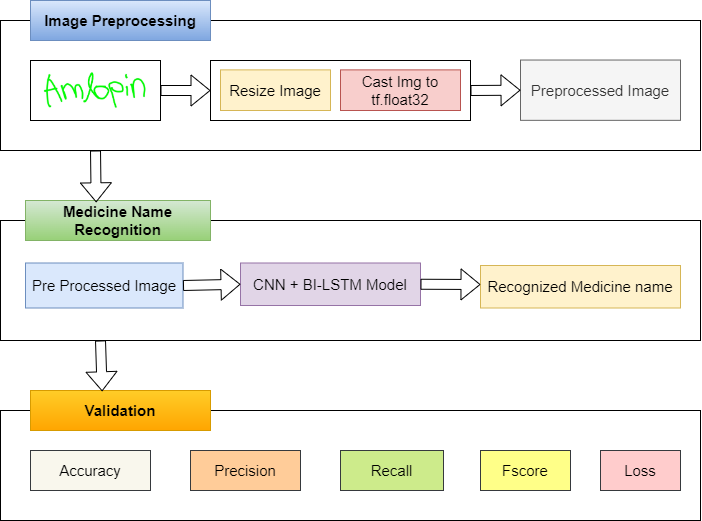


Fig. 2.Architecture Diagram for our work

We utilised cropped images from the IAM Dataset and a few medical datasets in conjunction with our model to train. The training dataset, validation dataset, and testing datasets have been divided in a ratio of 90:5:5.The network has a sophisticated design that includes seven convolutional layers, optional batch Normalization layers, Max Pooling layers, Relu activation functions, a Bi-directional LSTM layer, and a CTC layer.The procedure increases the number of channels in the first convolution layer from 1 to 64. which, after several layers, is raised to 128. The Bi-directional LSTM layer aids in the decoding of the convolution layer-generated features. We will next create a layer for label input for the appropriate images, followed by a thick layer. To find the CTC loss, the last layer would be the CTC layer. The next step is to determine the callback distance. During this stage, ctc decoding will be done using predictions.Then,using tf.sparse.from dense and dtype as an int64, sparse the predictions made from dense. Identify a point when an increase in epoch values leads to a rise in loss value and the model will no longer improve performance at which point training may be stopped. Next, build the model and train it with various rising epoch values, setting checkpoints as necessary.We can use the count variable to determine the model's accuracy by passing some data as input, checking the total number of correct predictions, and calculating the percentage of right predictions. Accuracy, Precision, Recall, and F-score are all available.The accuracy will be proportional to the initial weights we are establishing in the input layer. Setting appropriate starting weight values for the input layer and subsequent layers, as well as choosing an appropriate activation function, are required. Relu inside a convolution layer and Dense layer are typically utilised as activation functions for models like CNN and LSTMs.

**Bi-Directional LSTM Mathematical Notation as follows:**

, ] +

where x is output at time t, W is weights, B is bias.

**4 Result and Observations**

**4.1 Test case Results**

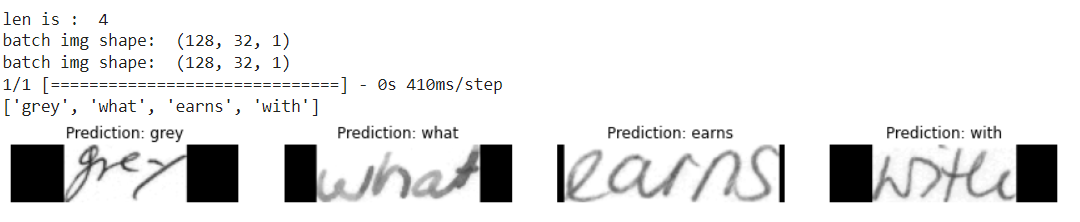


Fig. 3. Output Test Case predicted by model

The model was trained for 25 Epochs and we got this predictions. We passed 4 images custom inputs from testing data in IAM Dataset and we got 4 images predicted correctly.

**4.2 Observations and Analysis**

We noted our observations and displayed in below table.These are metrics we got with initial neural network topology.

Table. 1.Training dataset sizes Ration vs Epoch

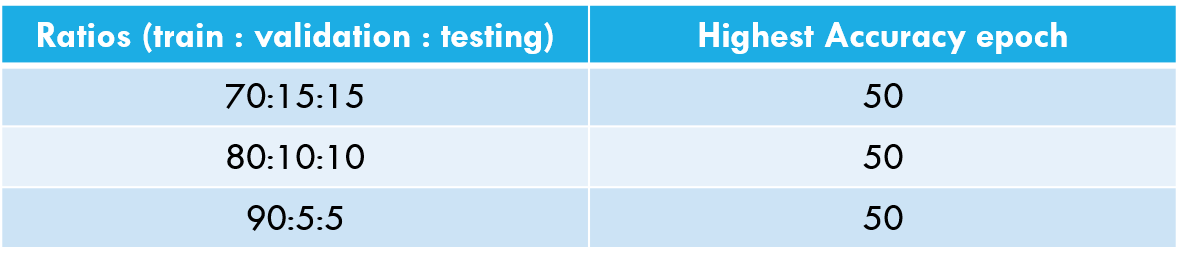
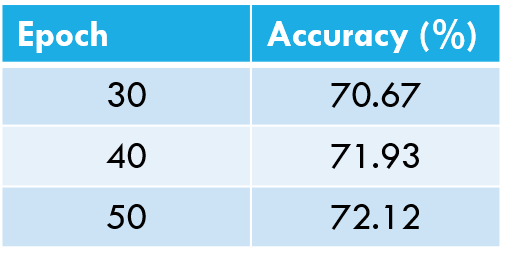


Table. 2. 90:5:5 Training data size vs Accuracy



**Metrics**: Accuracy: 0.72, Precision: 0.95, Recall: 0.728, F-Score: 0.824

The above are observations and analysis of model performance with initial neural network topology.

**Improvements**:We changed our Network Topology and Observations as follows:

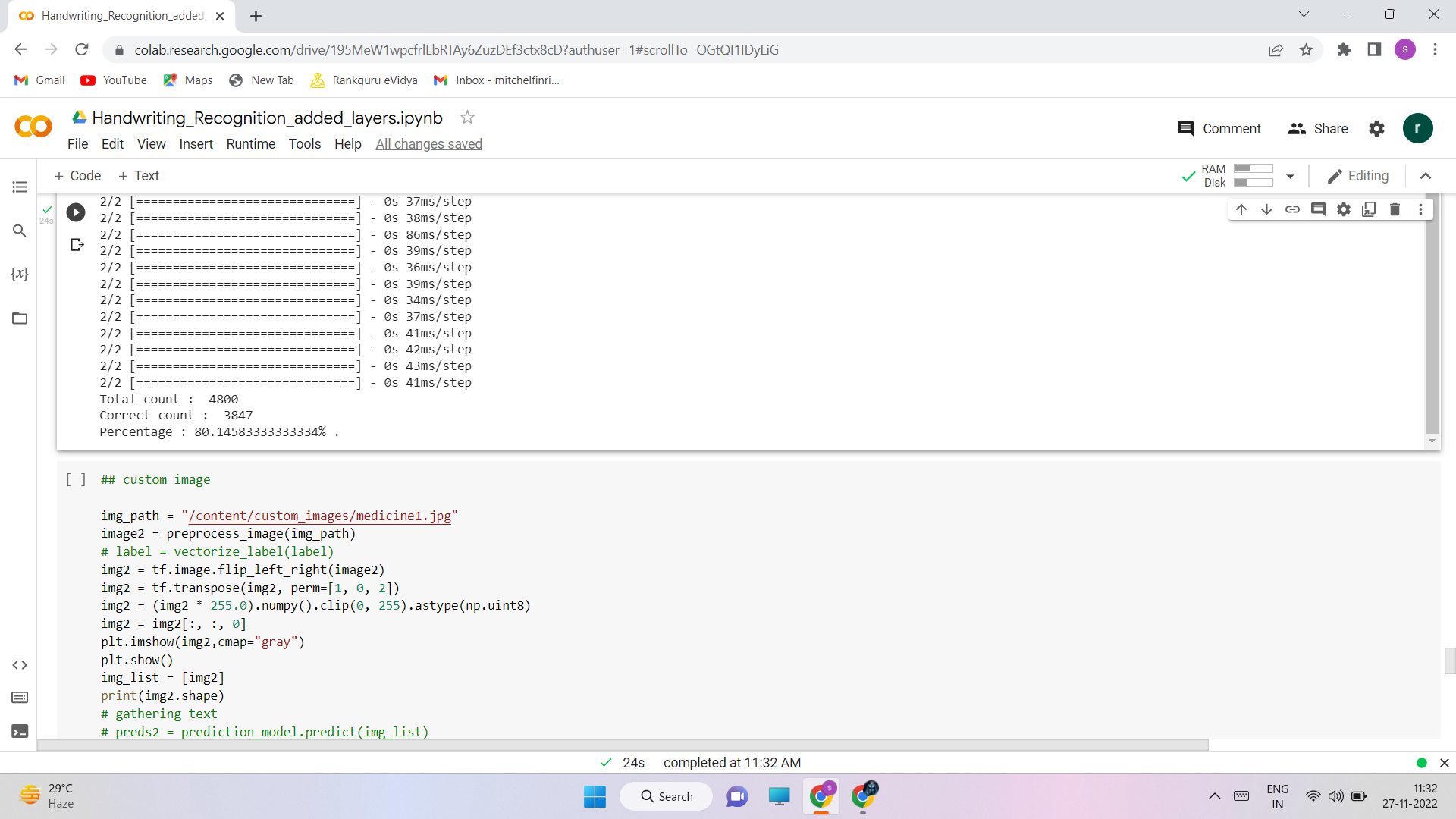


Fig. 4. Increased Accuracy (With new network topology)

We have added 3 more convolution layers and 2 Bi-Directional LSTM layers. The filters in 3 convolutional layers are of sizes 128,256 and 1024 respectively and in Bi-LSTM layer number of hidden units is 1024,512 respectively. For our proposed model after changing network topology we got 80% Accuracy which is better than initial 72% Accuracy. It is proved that when there is no increase in performance of our model then we need to change initial weights and bias values or we should change our model network topology in order to acquire good Performance.

**5 Conclusion and Future Work**

This model helps Pharmacists and normal people to recognize the medicine name accurately which is present in the Doctor’s Handwritten Prescription. This effectively offers text in all handwriting typefaces. The feature extraction is carried out by Convolutional Neural Networks with many layers, and the decoding of the extracted features into English letters is assisted by Bi-LSTMs.We employ CTC to circumvent the fact that the true alignment between the input and the output is unknown. To accurately identify language specific to prescriptions provided by the doctors, more bias is applied to words that are present in a manually produced corpus. The accuracy can be maximised by increasing the size of the training data.

**Future work**: The Accuracy for this model can be further improved by training with more handwritten prescriptions. Further an API can be created for this model which can be useful Mobile Applications or web application to use this model for Recognizing text from Prescription or from input of cropped Medicine Name. And more layers can be added by keeping complexity of model in mind. Training same dataset with new and Advanced Deep Learning Algorithms or Advanced Neural Networks can improve Accuracy.

**References**

1. T. Jain, R. Sharma and R. Malhotra, "Handwriting Recognition for Medical Prescriptions using a CNN-Bi-LSTM Model," 2021 6th International Conference for Convergence in Technology (I2CT), 2021, pp. 1-4, doi: 10.1109/I2CT51068.2021.9418153. (T. Jain, 2021)
2. Shaira Tabassum1, Ryo Takahashi1 , Md Mahmudur Rahman “Recognition of Doctors’ Cursive Handwritten Medical Words by using Bidirectional LSTM and SRP Data Augmentation”,IEEE,DOI: 10.1109/TEMSCON EUR52034.2021.9488622 (2021)
3. Lovely Joy Fajardo1, Niño Joshua Sorillo , Jaycel Garlit , Cia Dennise Tomines , Mideth B. Abisado , Joseph Marvin R. Imperial , Ramon ,” Doctor’s Cursive Handwriting Recognition System Using Deep Learning” DOI:10.1109/HNICEM48295.2019.9073521 (2019)