INTERPRETING DOCTOR'S PRESCRIPTION

EPICS PROJECT REPORT submitted in partial fulfillment of the requirements

Submitted by

MOHAMMAD RIZWANULLAH (208W1A1299)

NAGARAJU AJAY KUMAR VARMA (208W1A12A1)

Under the Guidance of

Dr M.SUNEETHA, Ph.D

PROFESSOR AND HEAD

For the award of the degree

BACHELOR OF TECHNOLOGY IN

INFORMATION TECHNOLOGY

DEPARTMENT OF INFORMATION TECHNOLOGY V R SIDDHARTHA ENGINEERING COLLEGE

(AUTONOMOUS - AFFILIATED TO JNTU-K, KAKINADA)

Approved by AICTE & Accredited by NBA

KANURU, VIJAYAWADA-7 ACADEMIC YEAR

(2022-23)

V.R.SIDDHARTHA ENGINEERING COLLEGE

(Affiliated to JNTUK: Kakinada, Approved by AICTE, Autonomous)(An ISO certified and NBA accredited institution)
Kanuru, Vijayawada – 520007

CERTIFICATE

This is to certify that this project report titled **"INTERPRETING DOCTOR'S PRESCRIPTION"** is a Bonafide record of work done by **MOHAMMAD RIZWANULLAH (208W1A1299)** and **NAGARAJU AJAY KUMAR VARMA (208W1A12A1)** and under my guidance and supervision is submitted in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Information Technology, **V.R. Siddhartha Engineering College** (Autonomous under JNTUK) during the year **2022-2023**.

(Dr. M.SUNEETHA)	**(Dr. M.SUNEETHA)**
Professor & Head	Professor & Head
Dept. of Information Technology	Dept. of Information Technology

EXTERNAL EXAMINER

DEPARTMENT OF INFORMATION TECHNOLOGY

V.R.SIDDHARTHA ENGINEERING COLLEGE

PROJECT SUMMARY

S.No	Item	Description
1	**Project Title**	**INTERPRETING DOCTOR'S PRESCRIPTION**
2	**Student Names & Numbers**	**MOHAMMAD RIZWANULLAH (208W1A1299)** **N. AJAY KUMAR VARMA (208W1A12A1)**
3	**Name of The Guide**	**Dr.M.Suneetha** **Professor and Head**
4	**Name of The Mentor**	**Y.Sandeep** **Asst Professor**
5	**Research Group**	**AI/ML/DL**
6	**Application Area**	**Health Care**
7	**Aim of the Project**	**Recognize Doctor's Handwriting**
8	**Project Outcomes**	**Providing the digital text of handwriting**

Student Signatures

1. **Signature of Guide**

2.

ACKNOWLEDGEMENT

First and foremost, I sincerely salute our esteemed institution **V.R SIDDHARTHA ENGINEERING COLLEGE** for giving me this opportunity for fulfilling my project. I am grateful toour principal **Dr. A.V.RATNA PRASAD**, for his encouragement and support all through the way of my project.

On the submission of this Project report, I would like to extend my honor to **Dr.M. Suneetha**, Head of the Department, IT for her constant motivation and support during the courseof my work.

I feel glad to express my deep sense of gratitude to my project guide **DR.M.Suneetha, Professor and Head** for her guidance and assistance in completing this project successfully.

I would also like to convey my sincere indebtedness to all faculty members, including supporting staffof the Department, friends and family members who bestowed their great effort and guidance at appropriate times without which it would have been very difficult on my part to finish the project work.

TABLE OF CONTENTS:

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

Equation	Brief Description
$\hat{y}^{\langle t \rangle} = g(W_y[\underset{a}{\rightarrow}^{\langle t \rangle}, \underset{a}{\leftarrow}^{\langle t \rangle}] + b_y)$	This equation is mathematical Notation for Bi-Directional LSTM $\hat{y}^{\langle t \rangle}$ represents output y at time t. W represents weights and b represent Bias at time t.

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, Batch Normalization, 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.

1.1 ORIGIN OF THE PROBLEM

Being unable to recognise the medicine name on the prescription. Because some doctors' handwriting is difficult for even pharmacists to see, they might prescribe the wrong medication, which could have catastrophic results. So, if we don't read correctly and understand handwriting, this is the major issue. In order to avoid this issue, we need a solution.

1.2 BASIC DEFINITIONS AND BACKGROUND

1.2.1 CONVOLUTIONAL NEURAL NETWORKS:

The major constituent of a CNN is the convolutional layer, where the processing takes place. It needs input data, a filter, and a feature map, among other things. The CNN learns directly from input, doing away with the requirement for human feature extraction. CNNs are very helpful for recognising objects and scenes in photos by looking for patterns in the images.

1.2.2 LSTM:

Long-term short-term memory It belongs to the class of recurrent neural networks. LSTM is a difficult concept to comprehend. It deals with algorithms that attempt to replicate the way the human brain functions and find the correlations in the sequential data provided.

1.2.3 ACTIVATION FUNCTION:

The activation function assesses the input value against a predetermined threshold. The neuron is engaged if the input value exceeds the threshold value. If the input value is below the threshold value, it is disabled, which prevents its output from being delivered to the next or Hidden layer.

1.2.4 OPTIMIZER:

Optimizer is used to update Learning rates, Weights and Bias values in Network. Some Optimizers uses Summation and Some uses Tangent Function and they vary between value Ranges. 0 to 1 for sigmoid, -1 to 1 for Tanh and 0 to infinity for RELU.

1.3 PROBLEM STATEMENT

Interpreting Doctors hand written prescription with deep learning methodologies.

1.4 SOCIETIAL APPLICATIONS OF PROPOSED WORK

The name of the medication may be recognised by both a regular person and pharmacists with accuracy. By correctly identifying a drug name, one can prevent the unpleasant effects of incorrect recognition. In order to identify the medication name and give the patient the appropriate medication, this approach can aid pharmacists.

2. REVIEW OF LITERATURE

2.1 DESCRIPTION OF EXISTING SYSTEMS

2.1.1 RECOGNITION OF DOCTORS' CURSIVE HANDWRITTEN MEDICAL WORDS BY USING BIDIRECTIONAL LSTM AND SRP DATA AUGMENTATION [1]

Authors: Shaira Tabassum, Ryo Takahashi, Md Mahmudur Rahman.

Year of Publishing: 2021

Observation:

Developed 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.

2.1.2 HANDWRITING RECOGNITION FOR MEDICAL PRESCRIPTIONS USING A CNN-BI-LSTM MODEL [2]

Authors: Tavish Jain, Rohan Sharma, Ruchika Malhotra

Year of Publishing : 2021

Observation:

Developed 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.

2.1.3 DOCTOR'S CURSIVE HANDWRITING RECOGNITION SYSTEM USING DEEP LEARNING [3]

Authors: Lovely Joy Fajardo with Mideth B. Abisado.
Year of Publishing: 2019

Observation:

Developed 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.

2.2 SUMMARY OF LITERATURE STUDY

Table: 2.1 Summaries of Literature Study

Sno	Paper	Methodology	Year	Algorithm	Result
1	Recognition of Doctor's Cursive Handwritten Medical Words	Has developed a model based on Bi-Directional LSTM Model and SRP data augmentation to recognition of Bangladeshi Doctor's Hand Writing Recognition.	2021	Bi-Directional LSTM with SRP Data Augmentation.	A model that can Output the text which is present in the Doctor's Prescription(Bangladeshi Handwriting).
2	Handwriting Recognition for Medical Prescriptions	Have developed model based on Bi-LSTM model for hand writing recognition, They used CTC loss functions for Normalization. They passed input to 7 Convolution Layers.	2021	Bi-LSTM Model.	A model for predicting Doctor's Handwriting.
3	Medical Handwritten Prescription Recognition Using CRNN	Have Developed model based on Convolutional Recurrent Neural Networks(CRNN).They Trained model with some short texts which contain Alpha- numeric Characters, Spaces etc.	2019	Convolutional Recurrent Neural Networks (CRNN).	A model for Interpreting Doctor's Handwriting, With a Mobile Application.

A new model with some added layers can be developed to get better accuracy and an Api can be made to connect model with application.

3. PROPOSED METHOD

3.1 DESIGN METHODOLOGY

- Data collection: Importing Dataset from git repository (IAM dataset)
- Importing all required Deep learning related python libraries like TensorFlow, Keras, Matplotlib, Numpy, CV2.
- Read the metadata txt file and store the words with “ok” into new array. And then splitting the dataset into 3 parts i.e Training, Validation, testing samples with ratios(70:15:15 & 80:10:10 & 90:5:5)
- Find maximum length and the size of the vocabulary in the training data. Then build character vocabulary.
- Resizing the images without distortion in a rectangular size.
- Set image width to 128, height to 32, Batch size to 64 and padding to 99.
- Code for pre-processing image, Vectorizing image labels, process image labels and for prepare dataset.
- Build model starting with an input layer and next convolution base add convolution layer with different filter size (1024,512,128,64,32)
- Add 2 pooling layers after every 2 convolution layers. Its recommended to use ReLu Activation function for Convolution Layers.
- Set Optimizer to “Adam” and then build the model.
- Now add Evaluation metrics, validation images and validation labels. And then create a callback to monitor the edit distances.
- Now, Build model and then set epoch value and fit model. And train the model.
- Now take images from testing set and pass them to input layer and predict them. Then, we can get output of predicted text by model.
- Now plot a graph for Value loss function Loss and Validation Loss Vs Number of Epochs.

3.2 SYSTEM ARCHITECTURE DIAGRAM

Architecture diagram is displayed in the figure 3.2

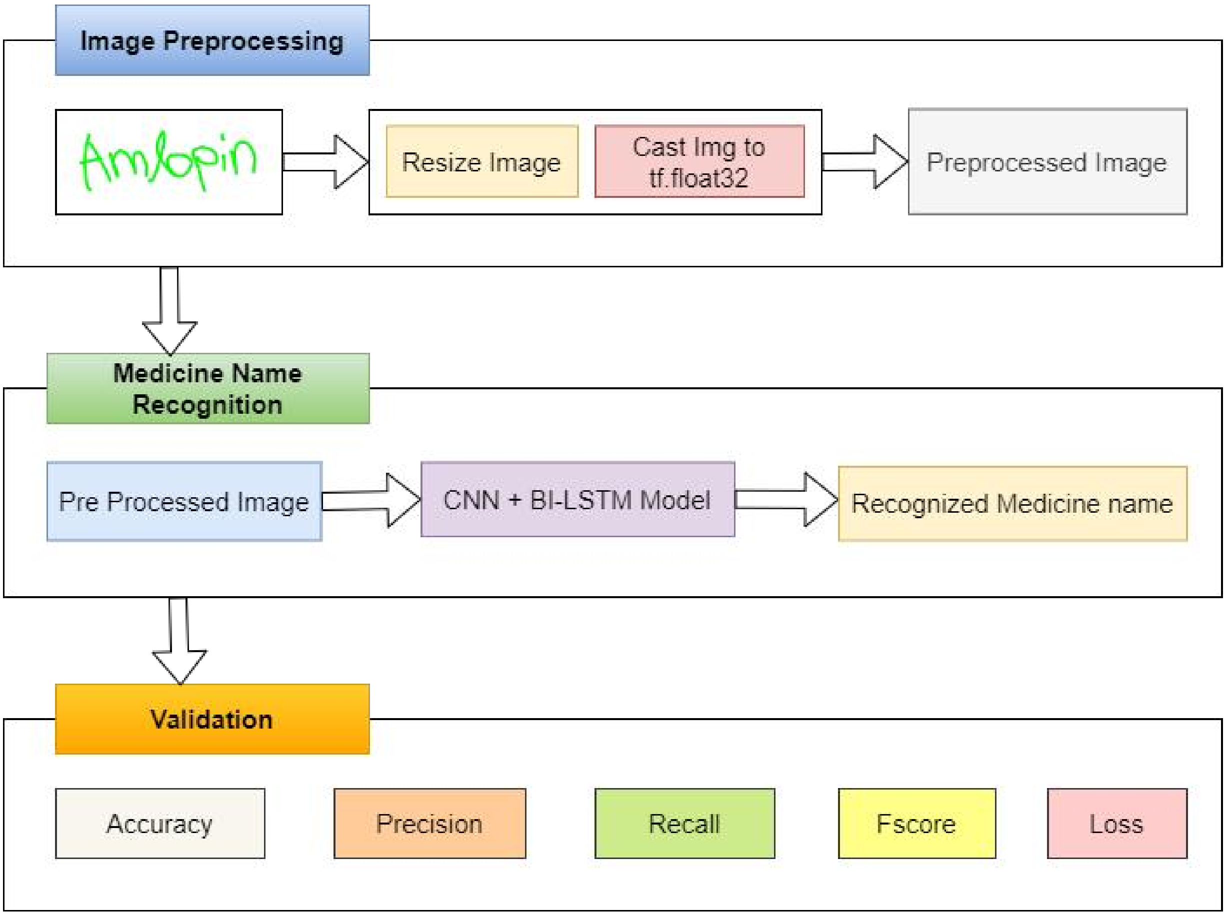


Figure 3.2 Architecture Diagram for Doctor's Handwriting Recognition**.**

Fig 3.2 describes the architecture of our work. The cropped image of doctor's prescription is Fed as input to the model. Then we preprocess image with respect to our layers which we are using in model.

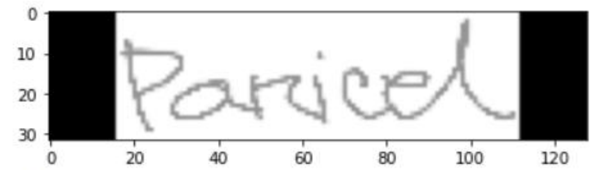
3.2.1 RESIZE IMAGE

In this Resizing Image Process, the image size gets Resized to Image sizes which we fixed that is image width to 128 and image height to 32 and also set padding to 99. We created a function to perform this Image Resizing.

3.2.2 CAST IMAGE TO FLOAT32

We do image resizing and cast image to float 32. Inorder to get floating point output values we are converting uint to float32 so there will be no loss of precision.

```
img_path = "/content/custom_images/medicine1.jpg"
image2 = preprocess_image(img_path)
img2 = tf.image.flip_left_right(image2)
img2 = tf.transpose(img2, perm=[1, 0, 2])
img2 = (img2 * 255.0).numpy().clip(0, 255).astype(np.uint8)
img2 = img2[:, :, 0]
plt.imshow(img2,cmap="gray")
plt.show()
img_list = [img2]
print(img2.shape)
```




```
(32, 128)
```

Fig 3.3 Distortion Free Image Preprocessing

3.2.3 CNN + BI-LSTM MODEL

We utilised cropped photos 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. There will be a Max pooling layer with RELU Activation function after each layer. In order to improve the model's ability to predict the outcome, we can add extra convolution and pooling layers. Additionally, adding additional parameters will improve accuracy and allow for a much reduction in loss.

The Bidirectional LSTM layer aids in the decoding of the convolution layer-generated features. We will next create a layer for label input for the appropriate photos, followed by a thick layer. To find the CTC loss, the last layer would be the CTC layer (Connectionist temporal classification).

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. Typically, the terms "Relu" and "Sigmoid" can be used to convolutional layers. However, the Relu activation function is favoured. This allowed us to successfully construct our model.

3.2.4 VALIDATION:

The performance of our model is assessed using the validation metrics. The proportion of photographs that are correctly categorised indicates how accurate the outcome is. The percentage of tuples where the model's projected right text really occurred will be the model's precision. How many tuples the model accurately predicted is measured by the recall. The harmonic mean of Precision and Recall is the F-Score.

3.3 DESCRIPTION OF ALGORITHMS

Mathematical notations of algorithms:

- We are applying propagations two times, one propagation for forward cells and another for backward cells.
- Forward and Backward Activations would be considered and we can calculate Output

$$\hat{\boldsymbol{y}}^{\langle t \rangle} = \boldsymbol{g}(\boldsymbol{W}_{\boldsymbol{y}}[\underset{\boldsymbol{a}}{\rightarrow}^{\langle t \rangle}, \underset{\boldsymbol{a}}{\leftarrow}^{\langle t \rangle}] + \boldsymbol{b}_{\boldsymbol{y}})$$

The above equation is mathematical notation for Bi-Directional LSTM’S.

Structure of Bi-Directional LSTM Model:

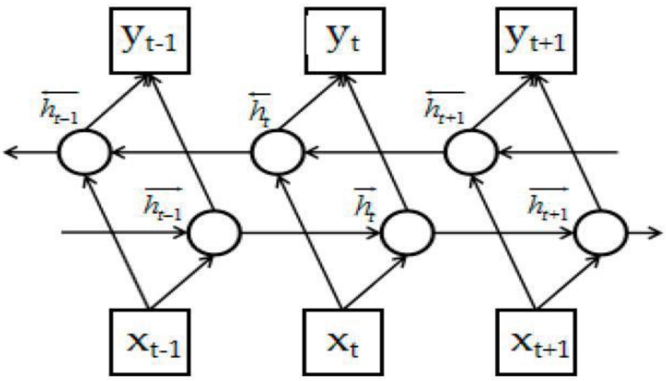


Fig 3.4 Structure of Bi-LSTM Model

3.4 DESCRIPTION OF DATASET, REQUIREMENT SPECIFICATION

Dataset we used is IAM Dataset which contains words images of different people with different writing styles and fonts. The images present in dataset are below

Fig:3.5 Images present in dataset

Doctors hand writing dataset: https://github.com/rizwanrockzz/epics/tree/main/dataset

IAM hand writing dataset: https://git.io/J0fjL

Recommended Specifications: PC or Laptop with minimum 8gb ram and supports GPU.

4. RESULTS AND OBSERVATIONS

4.1 RESULT ANALYSIS

This chapter consists of results and observations obtained by execution of the project. We usually get some accuracy when we train our model. And then after adding one hidden layer example to extract more features if we want to add an extra convolution layer, we changed layers added a extra convolution layer and the accuracy for improved with lesser number of epochs. We trained our model with different quantities of training dataset and each at different epoch values. As the training data is increasing, the accuracy of model is getting increased or gain percentage getting increased. We passed two medicine names to model in form of an array and it predicted the output. By observing the outputs, we can conclude that the amount of training data and number of epochs decides the output (prediction) and also the initial weights we give to input layer, further hidden layers and the output layer.

4.2 TEST CASE RESULT

The Figures of execution and analysis are below:

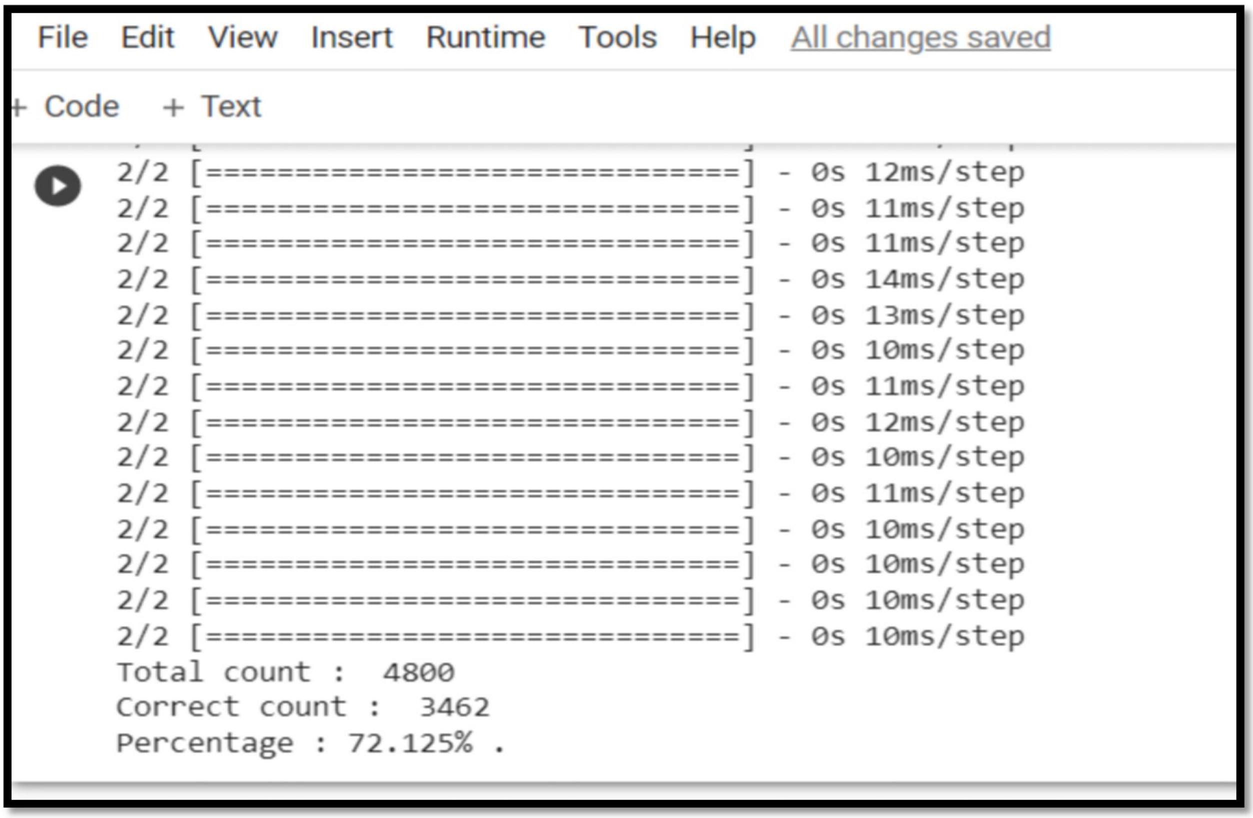


Fig: 4.1 Accuracy metric

Below is the graph plotted between Losses vs epoch (Training loss and Validation loss vs Number of epochs):

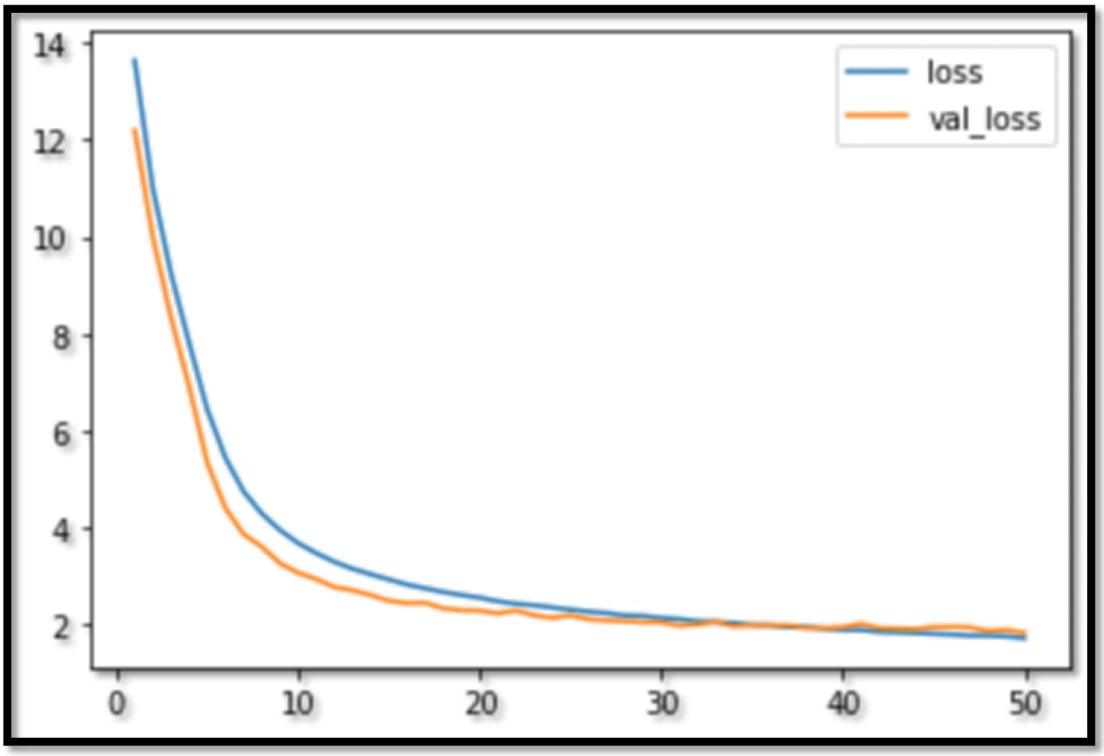


Fig 4.2 Losses vs epoch Graph (Loss on y-axis and Epochs on x-axis)

Its clear that when Number of epochs is high the Loss Values of both Training Loss and Validation Loss is getting Reduced.

Outputs produced by model after giving custom image paths:

```
len is :  6
batch img shape:  (128, 32, 1)
batch img shape:  (128, 32, 1)
1/1 [==============================] - 0s 198ms/step
['fariocl', 'Eaium', 'Aft', 'Aimepin', 'Calcin', 'Omenic']
```

Fig: 4.3 Output of images in form of text.

The letters present in the each word gets changed with respect to epoch we set during training. Quality and Size of the training dataset can also lead to best and accurate output of Text that is present inside an image.

4.3 OBSERVATION FROM THE WORK

Table: 4.1 Ratios and Highest Accuracy recorded Epoch

Ratios (train : validation : testing)	Highest Accuracy epoch
70:15:15	50
80:10:10	50
90:5:5	50

Table: 4.2 90:5:5 Dataset Ratio-Accuracy

Epoch	Accuracy (%)
30	70.67
40	71.93
50	72.12

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

These are results of metrics we choosed and obtained. Here, we did training upto 50 epochs and got different accuracy values at different epochs.

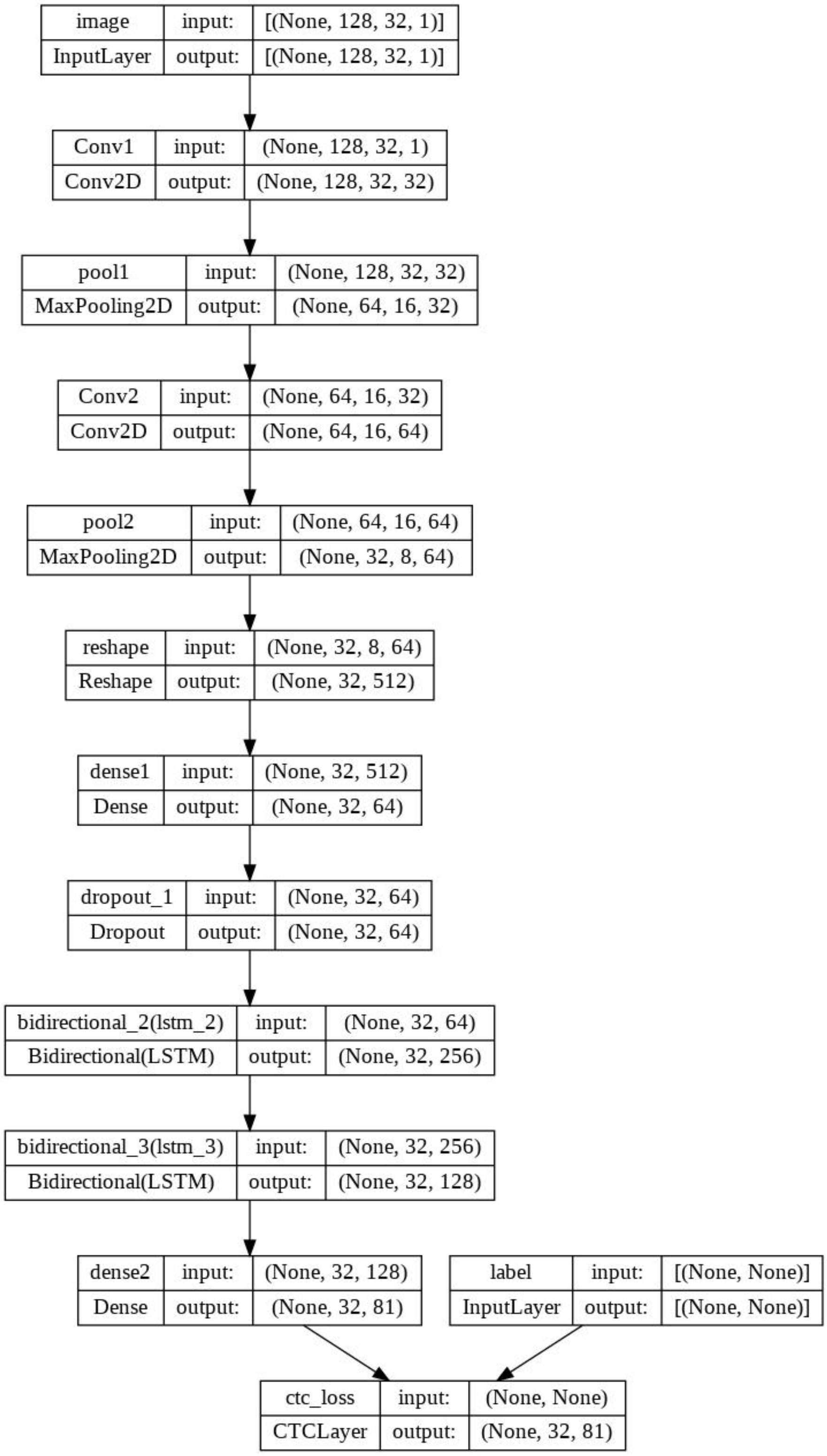


Fig: 4.4 Intial Network Topology

The Accuracy we got after changing Neural Network topology:

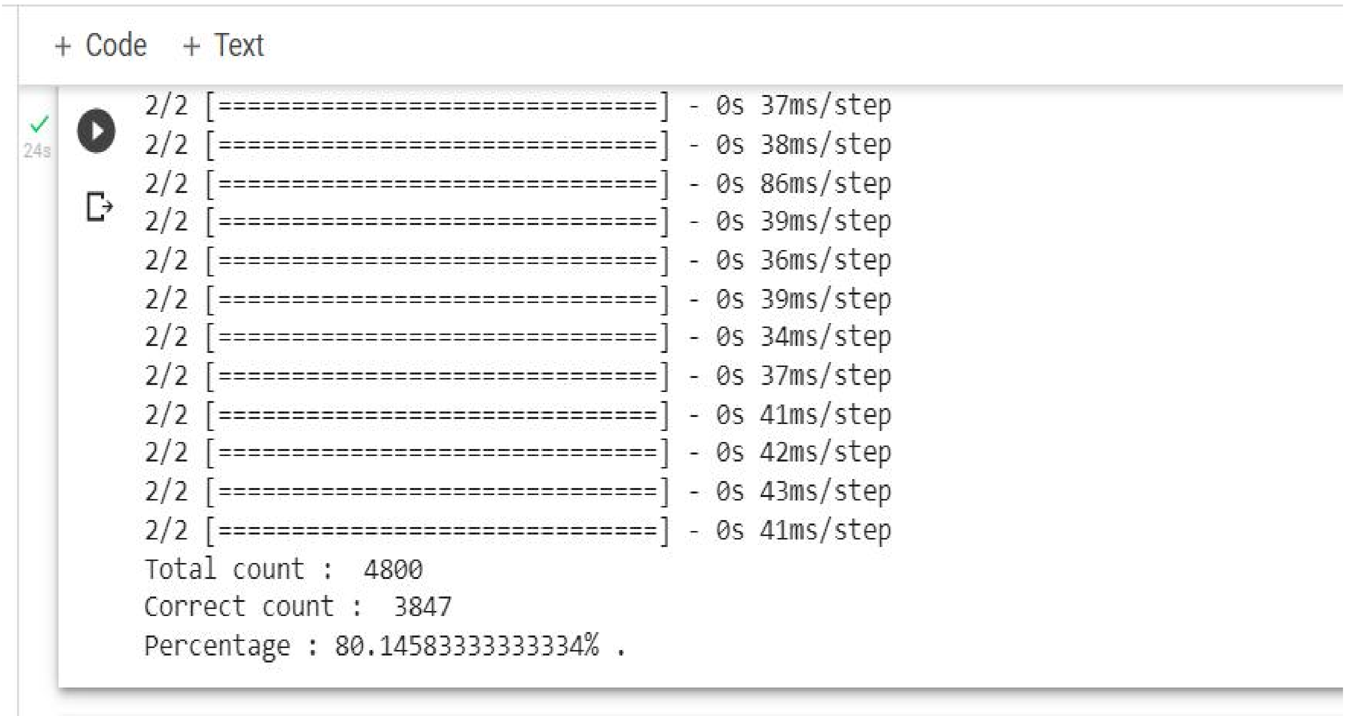


Fig: 4.5 Increased Accuracy

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. We can clearly see the accuracy of model was increased by 8% and training is done with 25 epochs and finally by changing topology we got accuracy increased to 80%.

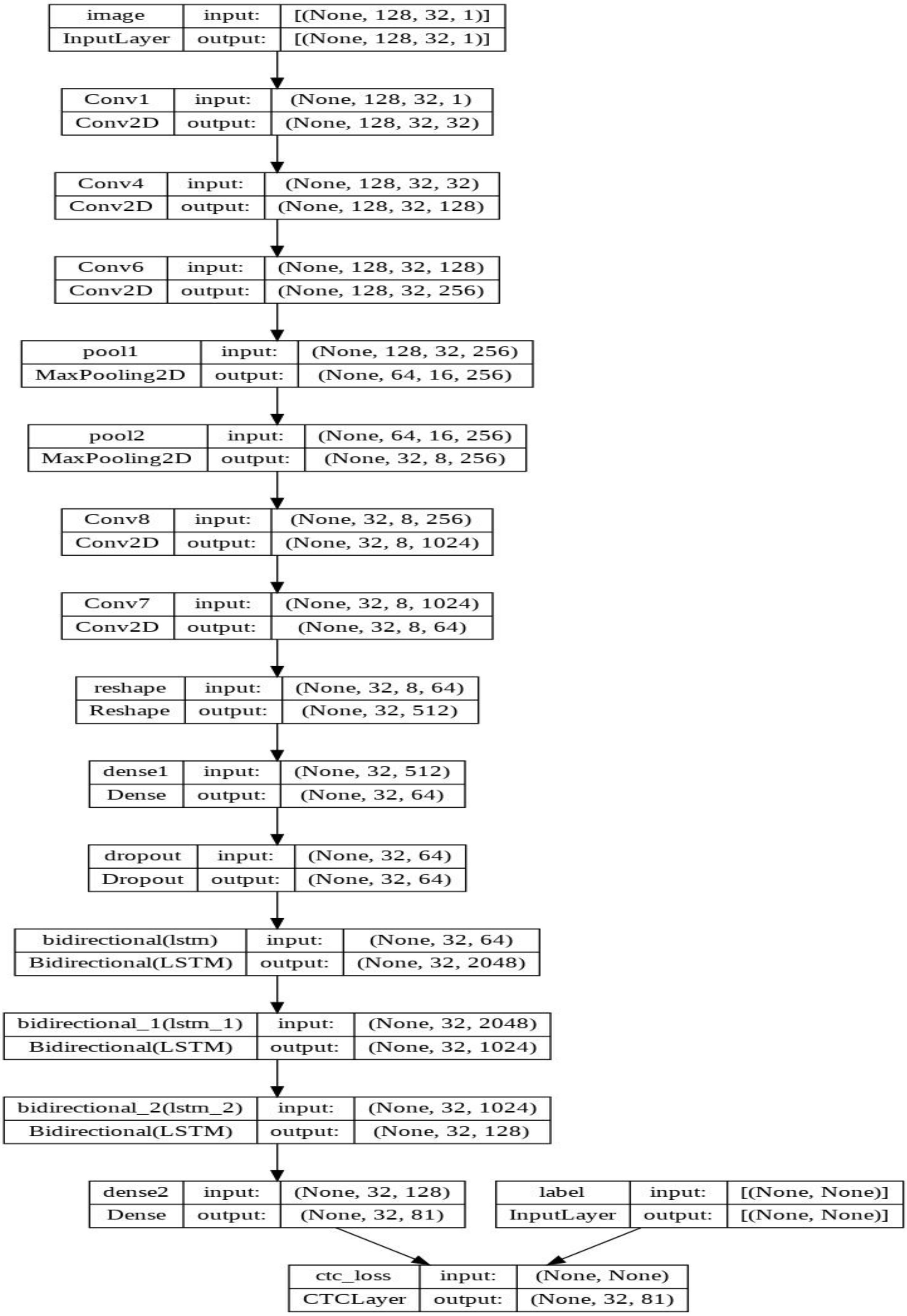


Fig:4.6 New Network Topology

5. CONCLUSION AND FUTURE WORK

5.1 CONCLUSION

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.

5.2 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 even improve Accuracy.

REFERENCES:

[1] 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

[2] Tanvish Jain,Rohan Sharma,Ruchika Malhotra," Handwriting Recognition for Medical Prescriptions using a CNN-Bi-LSTM Model" DOI: 10.1109/I2CT51068.2021.9418153

[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

[4] L. J. Fajardo et al, "Doctor's Cursive Handwriting Recognition System Using Deep Learning," 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management, 2019, pp. 1-6, DOI: 10.1109/HNICEM48295.2019.9073521

[5] Partha Pratim Roy, Ayan Kumar Bhunia, Ayan Das, Prithviraj Dhar, Umapada Pal," Keyword spotting in doctor's handwriting on medical prescriptions", Expert Systems with Applications, Volume 76, 2017, Pages 113-128, ISSN 0957-4174