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 TECHNOLOGYV R SIDDHARTHA ENGINEERING COLLEGE

(AUTONOMOUS - AFFILIATED TO JNTU-K, KAKINADA)

Approved by AICTE & Accredited by NBA

KANURU, VIJAYAWADA-7ACADEMIC 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)**

Professor & Head

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**
4	**Name of The Mentor**	**Y.Sandeep**
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. **Md Rizwan**
2. **N Ajay Kumar Varma**

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 to our 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 :

ACKNOWLEDGEMENT 3

LIST OF FIGURES 6

ABSTRACT 7

CHAPTER-1 Introduction 8

1.1 Origin of the Problem 9

1.2 Basic definitions and Background 9

1.3 Problem Statement 10

1.4 Socetial applications of proposed work 10

CHAPTER-2 Review of Literature 11

2.1 Description of Existing Systems 11

2.2 Summary of Literature Study

CHAPTER-3 Proposed Method 13

3.1 Design Methodology 13

3.2 System Architecture Diagram 15

3.3 Description of Algorithms 25

3.4 Description of datasets,Requirement specification 26

CHAPTER-4 Results and observations 35

4.1 Stepwise description of Results 35

4.2 Test case results/Result Analysis 37

4.3 Observation from the work

CHAPTER-5 Conclusion and Future Study 38

5.1 Conclusion

5.2 Future Study

REFERENCES...3

LIST OF FIGURES

Figure	Page
Fig-3.2: Architecture Diagram	13
Fig-3.3.1: Bi-LSTM NOTATION	14
Fig-3.3.2: Structure of Bi-LSTM Model	15
Fig-3.3.3 Layers Present in Neural Network of our model	17
Fig-3.4: Images present in dataset	19
Fig- 4.1.1: Accuracy value	20
Fig- .1.2: Losses vs epoch Graph	20
Fig- 4.1.3: Output of images in form of text	22
Fig- 4.1.4: Training Dataset Sizes and corresponding epochs	23
Fig- 4.1.5: Epoch value and corresponding Accuracy	24

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.

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

The inability of recognizing Medicine name present in the prescription. Even pharmacists also can't understand some doctor's handwriting,if this happens then they may give wrong medicine which leads to some negative consequencies and fatal causes.So ,Here is the main problem if we fail to recognize and interpret handwriting properly.we need a solution for this so we can avoid this problem.

1.2 Basic Definitions and Background

1.2.1 CONVOLUTIONAL NEURAL NETWORKS:

The convolutional layer is the core building block of a CNN, and it is where the majority of computation occurs. It requires a few components, which are input data, a filter, and a feature map. A convolutional neural network (CNN or ConvNet), is a network architecture for deep learning which learns directly from data, eliminating the need for manual feature extraction. CNNs are particularly useful for finding patterns in images to recognize objects, faces, and scenes.

1.2.2 LSTM:

Long short-term memory (LSTM) belongs to the complex areas of Deep Learning. It is not an easy task to get your head around LSTM. It deals with algorithms that try to mimic the human brain the way it operates and to uncover the underlying relationships in the given sequential data.

1.2.3 ACTIVATION FUNCTION:

The activation function compares the input value to a threshold value. If the input value is greater than the threshold value, the neuron is activated. It's disabled if the input value is less than the threshold value, which means its output isn't sent on to the next or hidden layer.

1.2.4 OPTIMIZER:

An optimizer is a function or an algorithm that modifies the attributes of the neural network, such as weights and learning rate. Thus, it helps in reducing the overall loss and improve the accuracy.

1.3 Problem Statement

- Interpreting Doctors hand written prescription with deep learning methodologies.

1.4 Societial Applications of Proposed work

An ordinary person and also pharmacists can recognize the medicine name and
They can accurately recognize the medicine name. An individual can avoid negative consequencies of recognizing medicine name in wrong way. So this model can help pharmacists to recognize the medicine name and provide right medicine to patient.

2. REVIEW OF LITERATURE

2.1 Description of Existing Systems

<table>
<tr><th>S.No</th><th>Title</th><th>Authors</th></tr>
<tr><td>1</td><td>Recognition of Doctor’s Cursive Handwritten Medical Words</td><td>Shaira Tabassum</td></tr>
<tr><td colspan="3">Description: Developed model for Recognition of Doctor’s Handwriting (Bangla Handwriting).She got 89.5% accuracy after using SRP Augmentation method(Stroke, Rotate, Parallel shifting).Some doctors involved in this provided their prescription images.</td></tr>
<tr><td>2</td><td>Medical Handwritten Prescription Recognition Using CRNN</td><td>Lovely Joy Fajardo, Mideth B. Abisado</td></tr>
<tr><td colspan="3">Description: Developed a model for Interpreting the Doctor’s Handwriting. They used CRNN Model (Neural Network).They developed a mobile application so a user can give image input and can get Digital Text as output.</td></tr>
<tr><td>3</td><td>Handwriting Recognition for Medical Prescription</td><td>Tavish Jain</td></tr>
<tr><td colspan="3">Description: Has developed a model for Doctor’s Handwriting Recognition using LSTM Model. They just built a model but haven’t developed any Mobile Application or web application to implement model in Realtime.</td></tr>
</table>

2.2 Summary of Literature Study

1) **Shaira Tabassum** has developed a model for Recognition of Doctor’s Handwriting (Bangla Handwriting).She got 89% accuracy after using SRP Augmentation method(Stroke, Rotate, Parallel shifting).Some doctors involved in this provided their prescription images.

2) **Lovely Joy Fajardo with Mideth B. Abisado** has developed a model for Interpreting the Doctor’s Handwriting. They used CRNN Model (Neural Network).They developed a mobile application so a user can give image input and can get Digital Text as output.

3)**Tavish Jain** has developed a model for Doctor’s Handwriting Recognition using BI-LSTM Model. They just built a model but haven’t developed any Mobile Application or web application to implement model in Realtime.

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

1) The model takes an image input which is a cropped image of medicine name. The user can scan image, crop it and Feed as input to the model.

2) This model can take both Gray scaled images or RGB images as well. The image gets resized as per values we specify.

3) The type of value gets converted into dfloat 32

4) Then the processed image is Fed as input to input layer and then the image based on initial weights and bias we fix it process it.

5) The Cnn layer with 128,64,1 weights, bias value trains model and extract features from model.

6) The extracted features are decoded by Bi-LSTM layer and then it extract text present in image and predict it as an output.

7) The output of Bi-LSTM Layer is further converted from dense to sparse and then decode with tensorflow decode and then we can print the predicted Text present in image as Digital Text.

3.2 SYSTEM ARCHITECTURE DIAGRAM

Architecture diagram is displayed in the figure 3.2

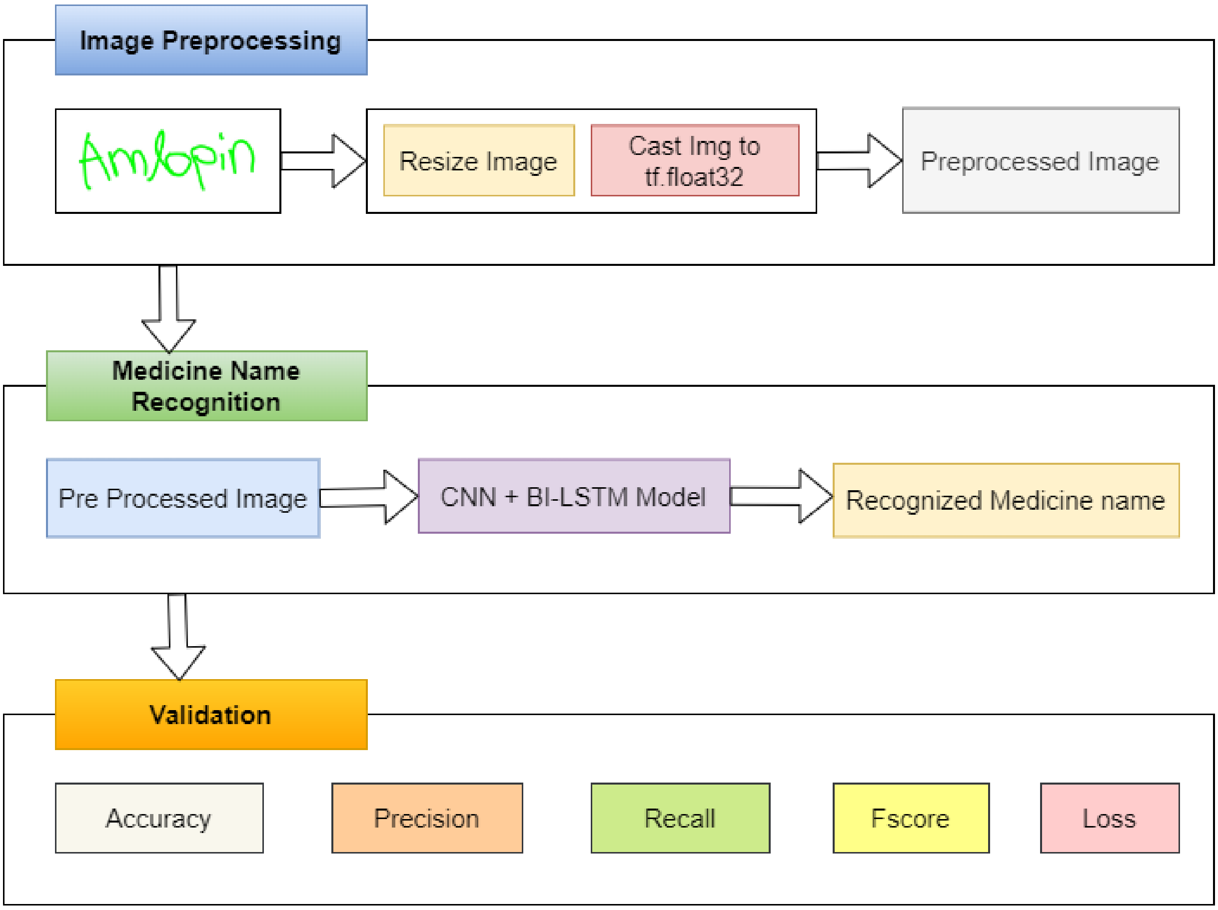

Figure 3.2 Architecture Diagram

Lets Understanding the above architecture diagram,

The cropped image of doctor's prescription is feeded as input to the model. Then we preprocess image with respect to our layers which we are using in model. We do image resizing and cast image to float 32. Then our image is ready to pass to input layer of neural network. The convolution layer which is used for feature extraction.

According to our model, we used IAM Dataset combined with some medical dataset cropped images and used them for training. We splitted the training dataset, validation dataset, Testing datasets in 90:5:5 Ratio.

The network uses a complex architecture, using seven convolutional layers, along with optional batch Normalization layers, Max Pooling layers, Relu activation functions for the Encoder, and a Bi-directional LSTM layer and a CTC layer.

In First convolution layer, The operation changes number of channels from 1 to 64. Which is increased to 128 in further layers. After every layer there will be Max pooling layer with RELU Activation function. Further we can also add more convolution and pooling layers in order to make model more accurately predict the output. And more parameters can be added to get better results in terms accuracy and loss can be reduced to very lesser extent.

The Bidirectional LSTM layer helps to decode the features that are produced by convolution layers. Further we will add a layer for input of labels of respective images followed by a dense layer. Finally, last layer would be CTC layer to find the CTC loss (Connectionist temporal classification).

Next step is to calculate call back distance, In this phase we will perform ctc_decode with predictions. Then, sparse the predictions from dense with help of tf.sparse.from_dense and dtype as int64. Now, Build model train the model with different increasing epoch values and set checkpoints if needed and identify a point where increase in epoch values increasing loss value and at this point the model will not improve its performance so we can stop training. In order to find the accuracy of this model we can take count variable and then pass some data as input and check the total correct predictions and find percentage of correct predictions. We can get Accuracy, Precision, Recall and F-score as well.

The initial weights we are fixing in input layer, the accuracy will be proportionate to this. It is necessary to set suitable initial weight values in input layer and for the further layers and select suitable activation function for this. Generally, for models like CNN and LSTM's the activation functions choosen are Relu and Softmax inside a convolution layer and Dense layer. Usually, for Convolution layers the 'Relu' and 'Sigmoid' can be used. But, preferred is Relu activation function. With this we built our model successfully.

3.3 Description Of Algorithms

Mathematical notations of algorithms:

1) Here we apply forward propagation 2 times , one for the forward cells and one for the backward cells
2) Both activations(forward , backward) would be considered to calculate the output y^ at time t

$$\hat{y}^{<t>} = g(W_y\,[\,\overrightarrow{a}^{<t>}, \overleftarrow{a}^{<t>}\,] + b_y)$$

Fig: 3.3.1 Bi-LSTM NOTATION

Structure of Bi-Directional LSTM Model:

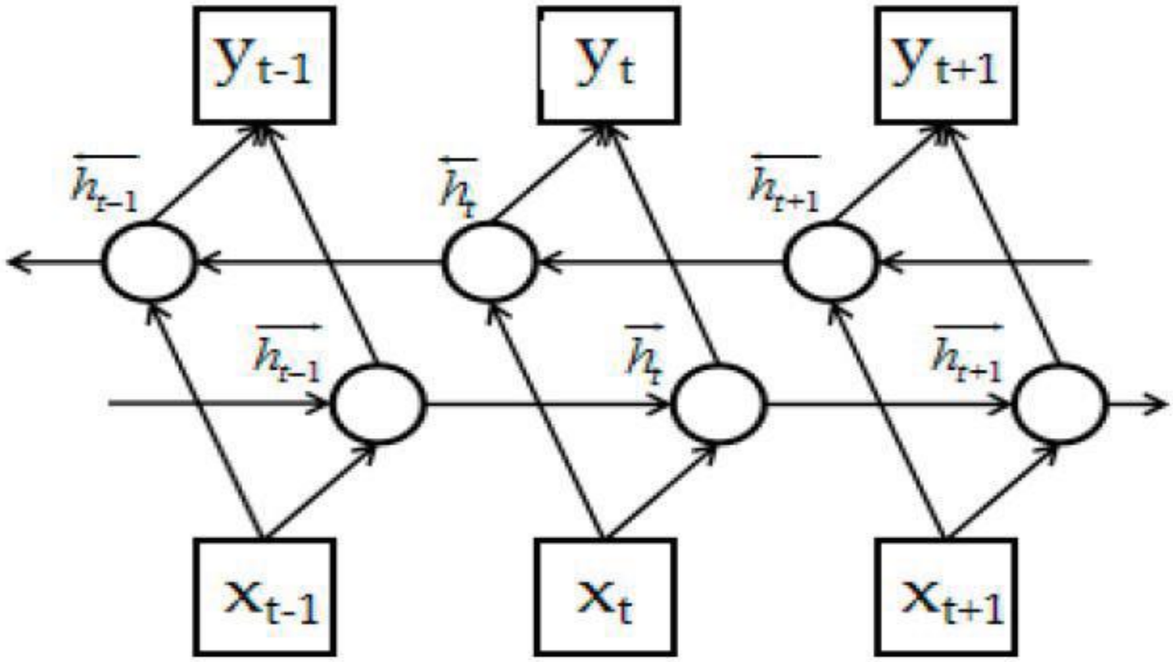

Fig 3.3.2 Structure of Bi-LSTM Model

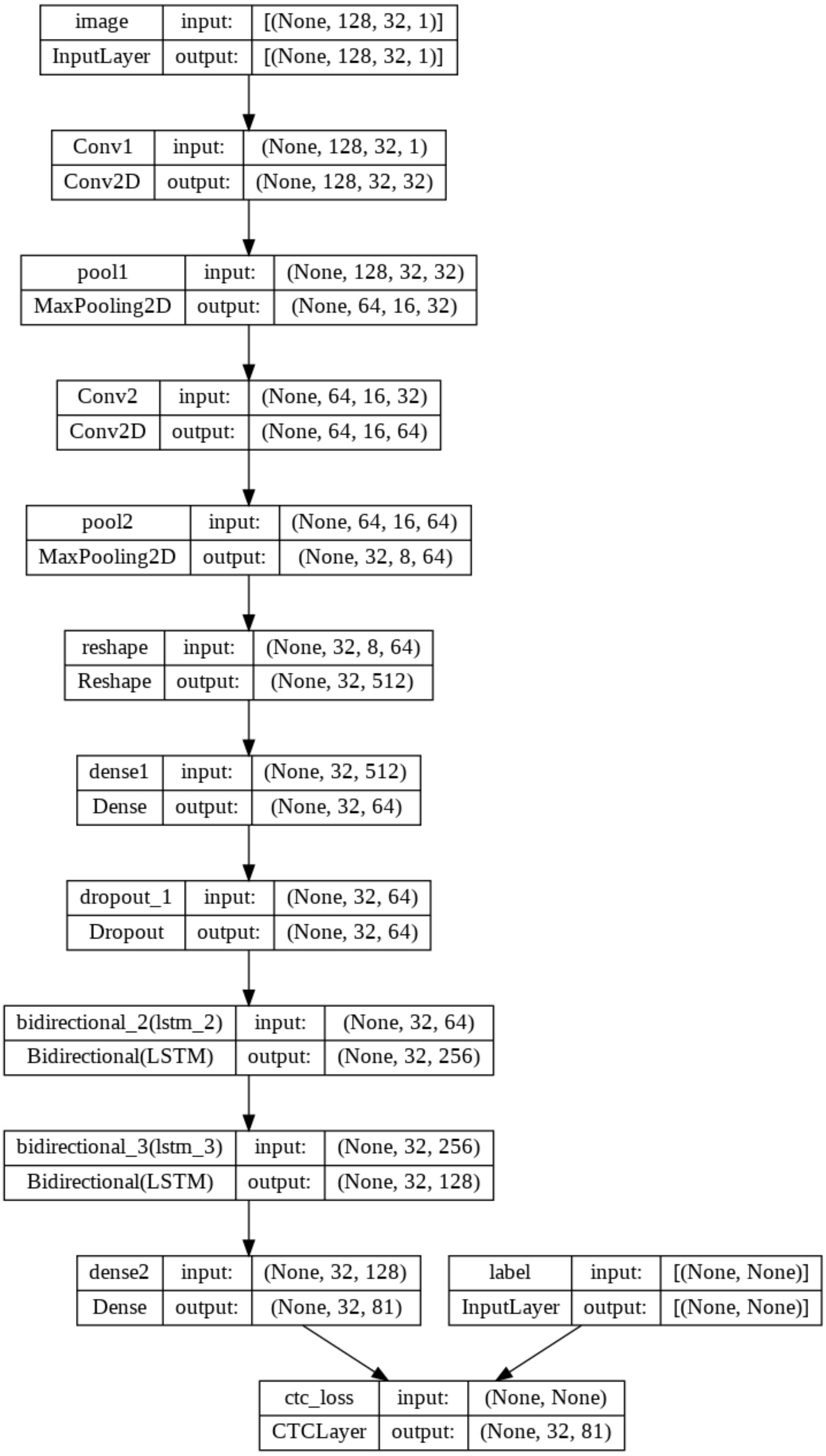

Fig: 3.3.3 Layers Present in Neural Network of our model.

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

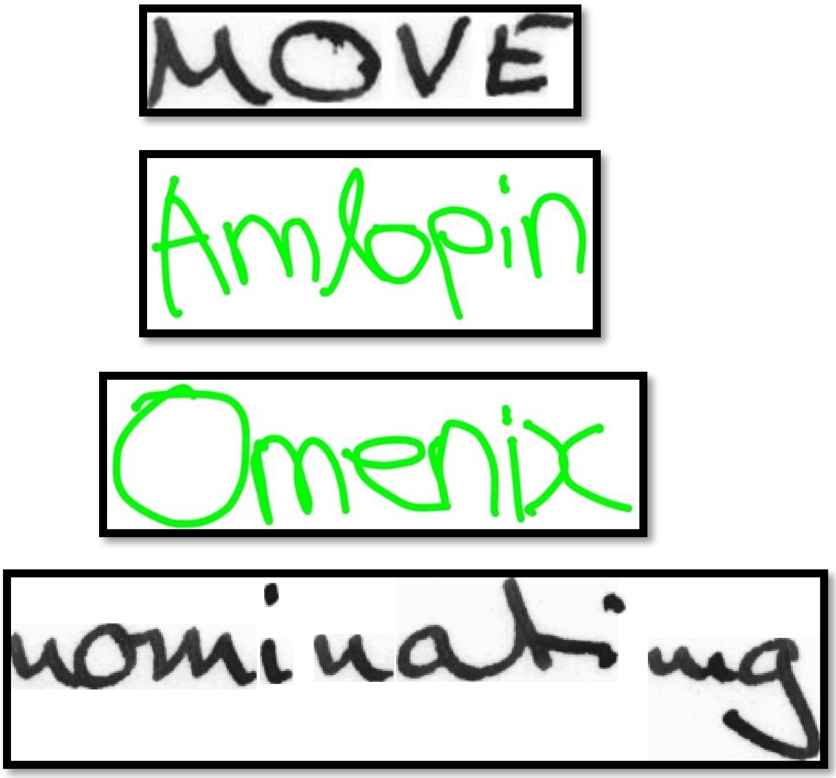

Fig: 3.4 Images present in dataset

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

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

Recommended Specifications: A PC with minimum 8gb ram and support 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.

The Figures of execution and analysis is below:

```
File  Edit  View  Insert  Runtime  Tools  Help  All changes saved
+ Code  + Text
2/2 [==============================] - 0s 12ms/step
2/2 [==============================] - 0s 11ms/step
2/2 [==============================] - 0s 11ms/step
2/2 [==============================] - 0s 14ms/step
2/2 [==============================] - 0s 13ms/step
2/2 [==============================] - 0s 10ms/step
2/2 [==============================] - 0s 11ms/step
2/2 [==============================] - 0s 12ms/step
2/2 [==============================] - 0s 10ms/step
2/2 [==============================] - 0s 11ms/step
2/2 [==============================] - 0s 10ms/step
2/2 [==============================] - 0s 10ms/step
2/2 [==============================] - 0s 10ms/step
2/2 [==============================] - 0s 10ms/step
Total count :  4800
Correct count :  3462
Percentage : 72.125% .
```

Fig: 4.1.1 Accuracy value.

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

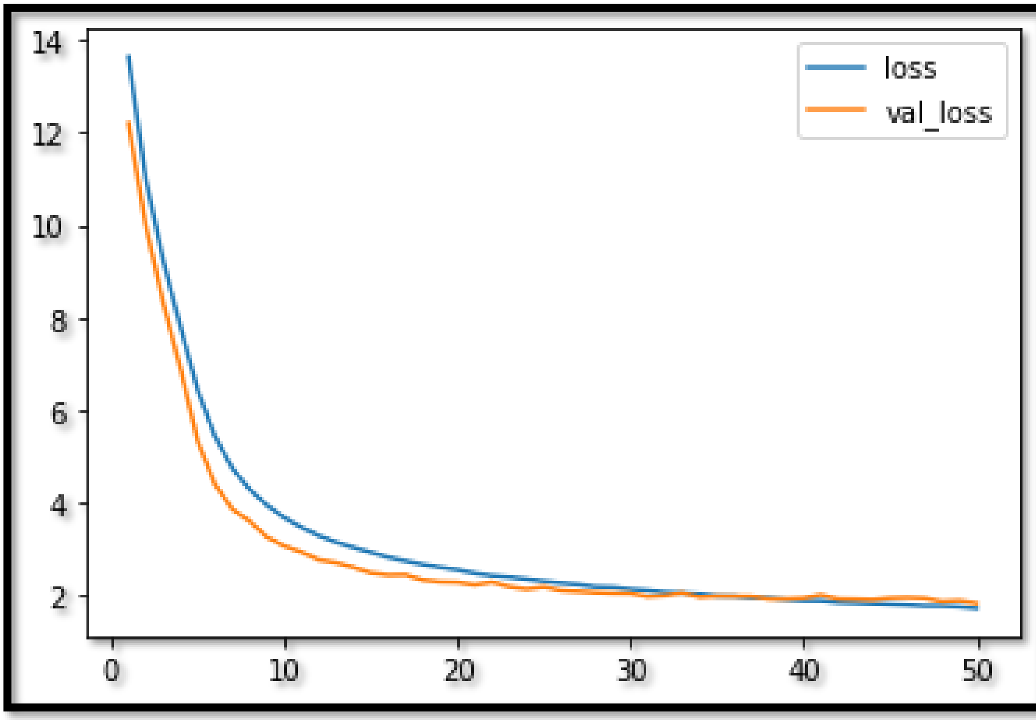

Fig 4.1.2 Losses vs epoch Graph

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: 4.1.3 Output of images in form of text.

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

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

Fig 4.1.4 Training Dataset Sizes and corresponding epochs

Epoch	Accuracy (%)
30	70.67
40	71.93
50	72.12

Fig 4.1.5 Epoch value and corresponding Accuracy(Highest recorded)

Metrics:
Accuracy: 0.74
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.

5 . CONCLUSION AND FUTURE WORK

5.1 CONCLUSION

This model helps pharmacists and normal people to recognize the medicine name accurately in the doctor’s handwritten prescription. This efficiently provides text for all fonts of handwritings. Multiple layers of Convolutional Neural Networks perform the feature extraction, and bi-LSTM's help in decoding the extracted features to English characters. Since the actual alignment between the input and the output is not known, we use Connectionist Temporal Classification to get around not knowing that alignment. More bias is given to words that are present in a manually created corpus to accurately recognize text specific to prescriptions offered by the doctors. The more the size of training data the accuracy can be maximised.

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 f 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 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 (HNICEM), 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

[6] P. S. Dhande and R. Kharat, "Character Recognition for Cursive English Handwriting to Recognize Medicine Name from Doctor's Prescription," 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, 2017, pp. 1-5, doi:10.1109/ICCUBEA.2017.8463842.