LANGUAGE RECOGNITION AND TRANSLATION FROM DOCUMENT

BY
Md. Mehedi Hassan Khan
ID: 151-15-5154

A.B.M Abdullah Al Noman
ID: 151-15-4853
And
Swapnil Rajbangshi
ID: 151-15-5478

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Md. Tarek Habib
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By

Mr. Md. Sadekur Rahman
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH
DECEMBER 2018

APPROVAL

This Project/internship titled **"Language Recognition and Translation from Document"**, submitted by Md. Mehedi Hassan Khan ID No: 151-15-5154; A.B.M. Abdullah Al Noman, ID No: 151-15-4853 and Swapnil Rajbangshi ID No: 151-15-5478 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **December 09, 2018**.

BOARD OF EXAMINERS

_____

Prof. Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Dr. Sheak Rashed Haider Noori — **Internal Examiner**
Associate Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Md Zahid Hasan — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that this research has been done by us under the supervision of **Md. Tarek Habib, Assistant Professor of Department of CSE**, Daffodil International University and co-supervision of **Mr. Md. Sadekur Rahman Assistant Professor of Department of CSE**, Daffodil International University. We also declare that neither this research nor any part of this research has been submitted elsewhere for the award of any degree.

Supervised by:

_____

Md. Tarek Habib
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

_____

Mr. Md. Sadekur Rahman
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

_____

Md. Mehedi Hassan Khan
ID: - 151-15-5154
Department of CSE
Daffodil International University

_____

A.B.M Abdullah Al Noman
ID: - 151-15-4853
Department of CSE
Daffodil International University

_____

Swapnil Rajbangshi
ID: - 151-15-5478
Department of CSE

ACKNOWLEDGEMENT

First we would like to express our heartiest thanks and gratefulness to Almighty **Allah** for his divine blessing which has made us possible to complete the project successfully. We are grateful to our parents for their motivation and support.

We are grateful an indebted to our respected supervisor **Md. Tarek Habib**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. His endless patience, scholarly guidance, continual encouragement, constant and energetic correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Professor Syed Akhter Hossain**, Head Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE Department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patience and believe of our parents which help us to complete this long journey till childhood.

ABSTRACT

Global communication nowadays is common phenomenon. There are diverse purposes behind global communication. It could be business, tourism etc. People want to connect with people that live overseas. And to communicate with them an individual has to know multiple languages because each nation has its own speaking language. But for an individual learning multiple languages won't be easy. One can hire translator or use a translating book to communicate. But now we live in the age where technology made human life more comfortable and easy than ever.so we are here to propose an application that can be used as translate source language from target language. The application can scan input text from a document, segment the text and calculate the confidence value then detect the types of input language and convert the text into target language.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Identifying the language(s) of a document is a key step of a document reading system since recognition engines require the integration of a language model to increase the transcription performance. We address this task in a very difficult context where documents are unconstrained, single variable writing type and different alphabets.

According to the proposed approach for identifying the language of a documented image first text blocks are extracted by a segmentation stage. Then the writing type of each text block is identified through an analysis of the connected components using codebooks of contour fragments. A similar approach is then used to identify the script [1]. This second stage takes advantage of the writing type information to choose an optimal codebook configuration. If a Bengali script is decided, the block language is considered to be Bengali. The language identification is performed by exploiting the outputs of a recognition engine [2] [3]. A statistical analysis is carried out analyzing the transcription of printed blocks.

1.2 MOTIVATION

Nowadays there are so many languages running frequently at a time. And the amount of spoken language will be 6500 more or less. However among those, 2000 languages have fewer speakers than 1000 .Based on country or culture it became normal for everyone. But for an individual with communication purpose learning multiple languages will be much difficult. So what if they can use a mobile application that can make their life much easier about communication. By using single application and with one click they can have their expected meaning from any unknown language.

Besides most running translate applications have no problem translating European languages. They can provide accurate performance without any time consumption. But when it comes to Asian languages they often struggle to give accurate results. This usually happens because most of the applications are created in United States and they have insufficiency of extensive data in the database of Asian languages. So in these continental we can provide more accurate and extensive data that can gives more accurate results on our continental languages.

1.3 RATIONALE OF THE STUDY

Technology is now more available and so mobile device are being transforming to the digital computing device. We can do almost everything in smart phone; there have lots of apps, lots of services. This proposed application is able to process documents with mixed printed text and various layouts.

The methods for the writing type recognition and the script discrimination are based on the analysis of the connected components while the language identification approach relies on a statistical text analysis, which requires a recognition engine**.**

1.4 RESEARCH QUESTIONS

1. How much accurate result one can get translating text from images or documents?

1.5 EXPECTED OUTPUT

From a documented image the system should scan the text format and separate the text from image and create a new text image. Then analyze that image and compare with the linked database for expected match result. After analyzing certain document the system should performs the following sub-tasks:

1. Language identification.

2. Confidence value of the language

3. Translate into target language.

1.6 REPORT LAYOUT

This report contains five chapters. Summarization is given below:

Chapter 1

Introduction, motivations, rationale of studies, research questions and expected outcome has been discussed in details.

Chapter 2

In Background chapter we covered related works, research summery, scope of the problems and challenges we faced is discussed been discussed.

Chapter 3

In this chapter we covered research subject and instrumentation, data collection procedure and implementation requirements in details.

Chapter 4

In this chapter we covered the Experimental Results, Descriptive Analysis and summary of our proposed system.

Chapter 5

This chapter will have conclusion and Future work.

CHAPTER 2

BACKGROUND

2.1 INTRODUCTION

There are multiple application that can do fetch and translation from document like Google translate. When we start our research on this particular area Google translate was our idol application to follow. We also follow some other application to do some background work for better development purpose. At starter we tried to understand the working process how it should be done and which model we should use to analyze data for language identification. Then we start on the model and methods for statistical analysis.

2.2 RELATED WORKS

Language identification from document is kind of solved problem nowadays. Reliable system with high accuracy is available. As an example, to the best of our knowledge Google plugin and Wikipedia works with 53 languages by using n-gram of characters and language databases. And social media like twitter they work with 6 languages by using 3-gram of characters.

Some other systems like iTranslate, Translator Foto, camera translator they all work with multiple language at a time. And they are mostly work with common language like English, Spanish, Chinese, and Arabic and so on. They have rich language profiles on those languages and perform with high accuracy.

2.3 RESEARCH SUMMARY

After doing background check on existing system one common thing was they gave their major thoughts on common languages as we talked in [2.2].And other languages like Bengali, their database isn't rich enough to give high accurate performance. They mostly work with single word or a short speech. But when it comes to long speech they gave random or inaccurate meaning.

Another observation is all of them we have select the language name for input. When the selected type and input type is same the can perform the translations. If not they can't/won't. In other words they don't have the automatic language identification option.

Table 2.1: Study on related work

SL	APPLICATION	SERVICE PROVIDER	METHODOLOGY	DESCRIPTION	LANGUAGE
1	Translation API	GOOGLE CLOUD	AutoML translation method [14].	Fast and dynamic translation from source to target language. Language recognition is also available in case source language is unknown.	103
2	Translator text API	MICROSOFT	Neural Machine Translation (NMT) and caused Radical Shift technology [15].	Take input from source language normalize the input language for more appropriate translation and then translate through text translation engine.	60
3	iTranslate	iTranslate [17]	Optical character recognition (OCR).	It's also another most accurate application with audio input support system. It also has the facility to share the translated phrase over E-mail, SMS, Twitter and other social media and saved all the history so that can be	100(16 offline)

				accessible via offline.	
4	Papago Translate [18]	Naver Search engine [19]	Unknown.	It much more effective for Asian languages. It also offers the same facilities like other application. Though it suffers from inaccurate photo recognition.	10
5	Camera Translator	App world Studio [20]	Unknown.	It has also image input along with voice input. And also has bookmark and saved history facilities	50

2.4 SCOPE OF THE PROBLEMS

1. Our first intention was to enrich the data base for Bengali language to get higher accuracy on translation. So who people doesn't know the Bengali language can get accurate and meaningful translation.

2. We added the automatic identification of input language so that if anyone doesn’t familiar with the input language then our system can automatically identify the language and let the user know.

3. If we can make this system available for everyone then vogue of Bengali language can be increase all over the world.

2.5 CHALLENGES

1. An extensive database was needed to perform accurate result for a particular language.

2. Run time duration was a huge consideration while making the system. After added automatic identification while input a text system has to check every database every time. And it might make the system too slow to use.

3. Image resolution was another challenge for us. If the image is noise free then the system can easily recognize the language. So we had to set minimum clarity limit for input image.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 INTRODUCTION

In the full project there are some steps to translate the language from the image documents. We are working for four languages. Which is Bangla [4], English, Spanish and Arabic. Here will be describe how we can process the image. How we collect and process data and implementation requirements.

3.2 RESEARCH SUBJECTS AND INSTRUMENTATION

3.2.1 RESEARCH

The project detects and converts the language from an image by some trained data. For this we needed test data and trained data. As input we give test data. On the other hand the language is detected and converted by those train data. This project work in three steps. Firstly the system detect the language. Secondly it separate the text. Finally it convert the text into English as [7].

For our system we used Recurrent Neural Networks (RNN). In the system there will be using text data, sequence of data and one language to another. For this we use LSTM which is one kind of RNN.

3.2.2 INSTRUMENT

For this we needed some instrumentation to make the project.

- Install python 3.
- OpenCV
- Tesseract
- Textblob
- Numpy
- Argparse

3.3 DATA COLLECTION PROCEDURE

For the project we had to two types of data. Test data and Trained data. For collecting and processing data we followed some step which is given below.

3.3.1 TEST DATA

As known there will be needed two types of data. To test the system we need some data. For this we collected four type of language documents which are Bangla, English, Spanish and Arabic [11]. Because the system takes only image input. For collection we follow some articles online newspaper etc. From there we have our input data for our system. This documents we take as an image. In the system this image will be pre-processed.

3.3.2 TRAINED DATA

In this step we need that type of data which can be compared between input data and the database. By this the system can detect the language. To learn Neural Networks uses the data which is represented in the image. There are many methods in neural networks. But now a days Recurrent Neural Networks are most using and this one is most promising. Because it is the only one which is working with the internal memory. As they are working with internal memory so they can remember the input they receive. For this it makes them easy to predict what is coming next. This is the reason they are most preferred algorithms. The information are through in a loop in RNN. When RNN make a decision in considers two things. Current input and other one is the previous inputs which one is he learned from the previous data. RNN can predict sequential data that other algorithm cannot do.

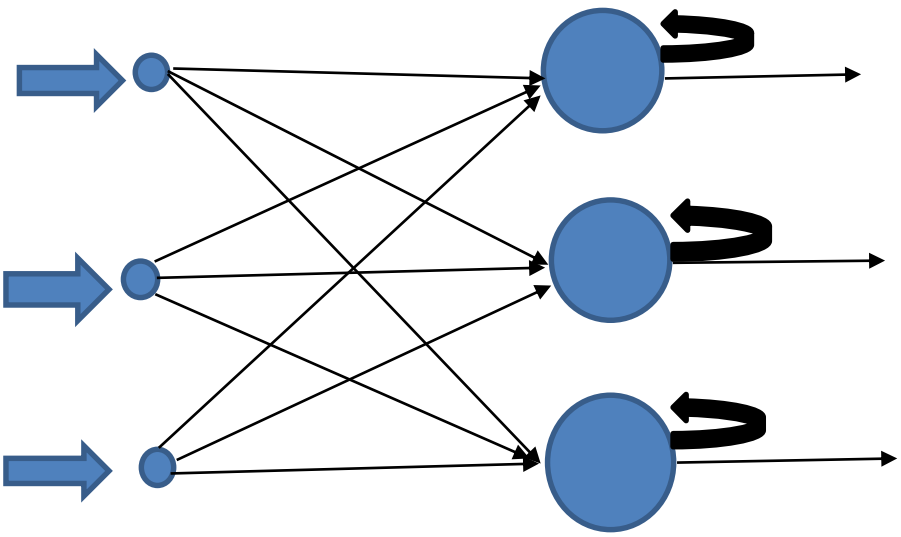

Figure 3.1: Recurrent Neural Network

Though RNN are performing well at context processing but there are some problems. The main problem is when they are performing at OCR [12] they are getting trouble for vanishing the gradient. For this problem neural networks are have been shown at Long Short Term Memory.

3.3.3 LONG SHORT TERM MEMORY

Long Short term memory concept was introduced by Hochreiter and Schmiduber. This architecture overcomes from many limitations. LSTM over overcame from the vanishing gradient problem. This one is the problem in the RNN. LSTM uses gates to solve the problems [9].

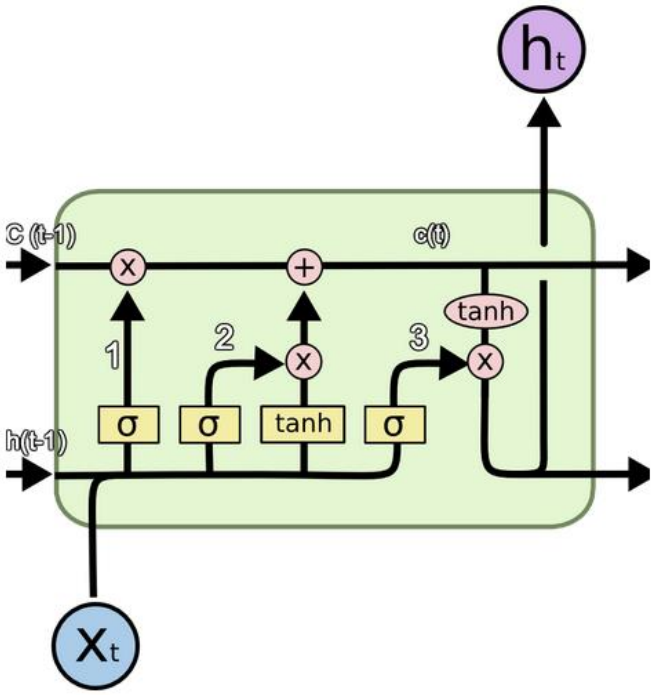

Figure 3.2: LSTM gate

The meaning of symbols:

- X: Scaling of information
- + means adding information
- tanh means tanh layer
- h(t-1) means output of last LSTM unit
- c(t-1) is the memory of LSTM
- X(t) current input

- C(t) is the new updated memory
- h(t) current updated output

Here this algorithm is using tanh to solve the vanishing gradient problem. tanh is that function whose second derivation can sustain for a long range before it going to zero. The sigmoid is used because it can output 0 or 1. It helps for to remove or remember any information. There are three components in LSTM unit as the diagram shown.

1. It has the architecture that enables it to remove or forgot any information.
2. Decide and store is the next step.
3. Last step is the output. The sigmoid layer decides the part of cell states are going to output.
4. The ability of LSTM to forgot, remember and update the information push in one step ahead.

3.3.4 IMAGE PRE-PROCESSING

It is one of the major parts of this project. It is a common name of image for operations. It works at lowest level of abstraction. It works in the input and output image [5]. The aim of image pre-processing is improve the image. There are three things to do:

3.3.5 NOISE REMOVAL

Most printed records are probably going to contain noise to some degree. The primary explanations behind this noise may vary, it makes it harder for program to recognize characters. Noise on pictures can be removed by utilizing a couple of various procedures consolidated. We applied converting image to gray scale, dilation, erosion for this matter.

3.3.6 GRAYSCALE IMAGE

It is otherwise called an intensity, gray scale, or gray level picture. Cluster of class uint8, uint16, int16, single or double pixel values determine intensity values. For single or double cluster, values extent from [0, 1]. For unit8, values go from [0, 255]. For uint16, values run from [0, 65536]. For int16, values run from [-32768, 32767].Gray scale reduces complexity from a 3D RGB value to a 1D value. Edge recognition task is not suitable with 3D pixels.

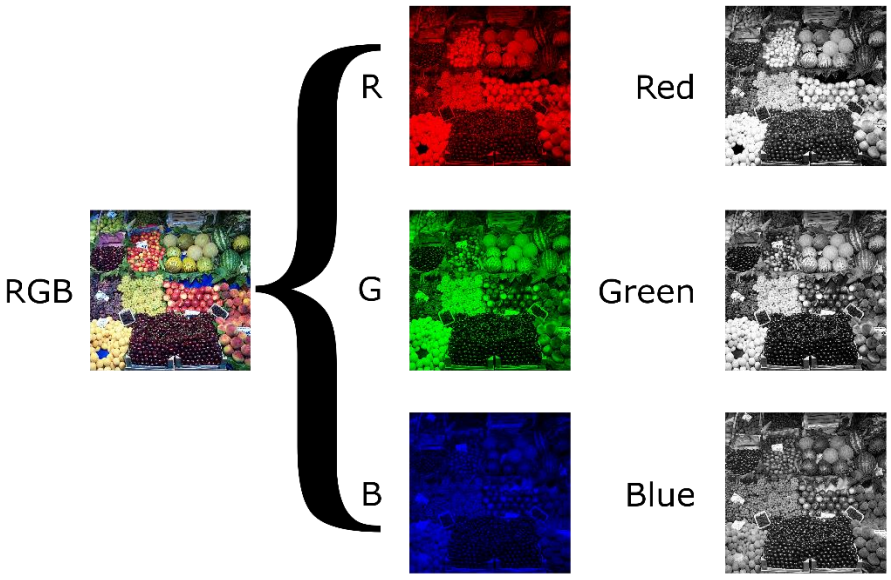

Figure 3.3: RGB to Gray scale.

3.3.7 DILATION AND EROSION

Dilation and Erosion: Dilation adds pixels to the limits of items in a picture, while erosion expels pixels on protest limits. The quantity of pixels included or expelled from the items in a picture relies upon the size and state of the organizing component used to process the picture. In the morphological dilation and erosion activities, the condition of some random pixel in the yield picture is dictated by applying a standard to the comparing pixel and its neighbors in the info picture. The standard used to process the pixels characterizes the activity as a dilation or an erosion. The both require a kernel matrix to work with. We use numpy array to create kernels. Since the text is in small fonts so we use small kernel size value 1, 1 that's fits all.

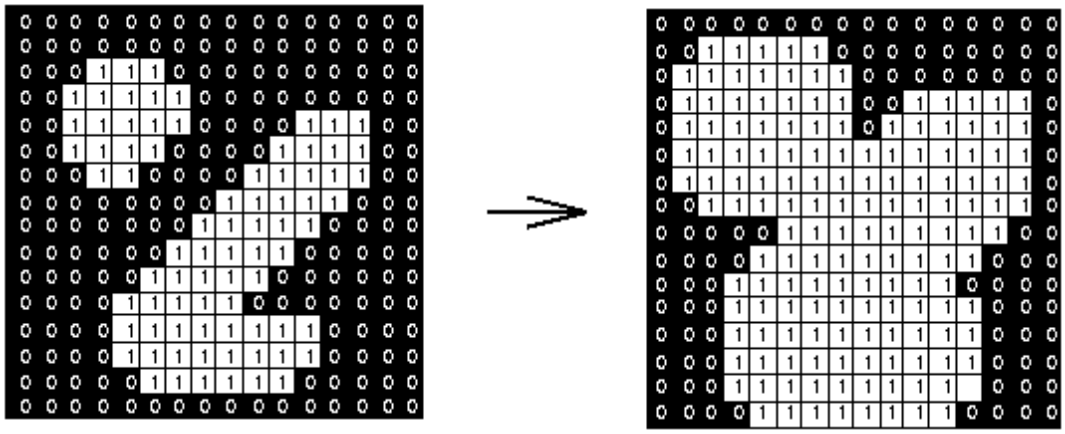

Figure 3.4: Dilation

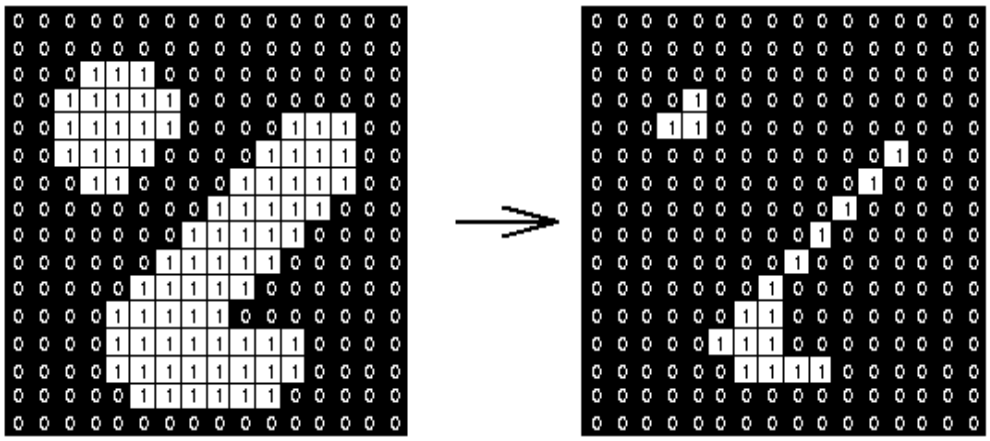

Figure 3.5: Erosion

3.3.8 BINARIZATION

The Binarization Method changes over the dark scale picture (0 up to 256 dark dimensions) in to dark and white picture (0 or 1). The result of OCR exceptionally depends upon the binarization. The high caliber binarized picture can give more exactness in character acknowledgment as thought about original picture since clamor is available in the unique picture. IN this project we use adaptive threshold method and binary threshold type where the size of a pixel neighborhood that is used to calculate a threshold for the pixel is 31.

Figure 3.6: Binarization

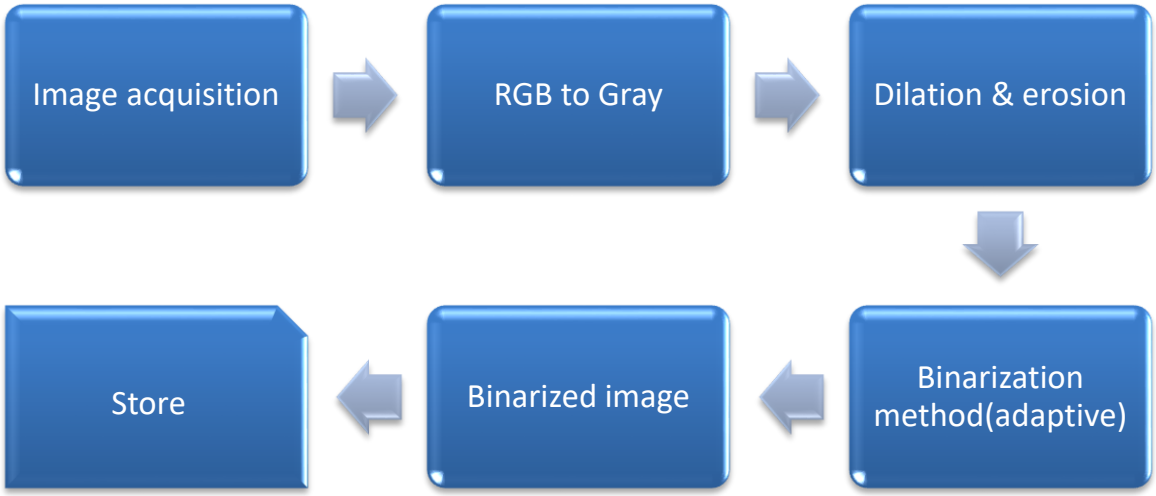

Figure 3.7: Proposed pre-processing

3.9 IMPLEMENTATION REQUIREMENTS

For implement the projects users need to have some requirements. After fulfil the requirements they can use the system

- CPU
- Python will be installed in operating system
- OpenCV

- Document will be converted into image

In this project we are translating a language from a document. We are Using four language. Which is Bangla, English, Arabic, and Spanish. In this project the process is done by three steps.

- Detect
- Separate
- Translate

Detect:

Firstly when we give the input document the program detect the language first. Firstly it takes words an array. For detecting the language it calculate a confidence value for each word of the documentation. After calculating value the confidence value for every language it pulls out an average for each language. The confidence value of document language will be highest. By this the language is detected [8].

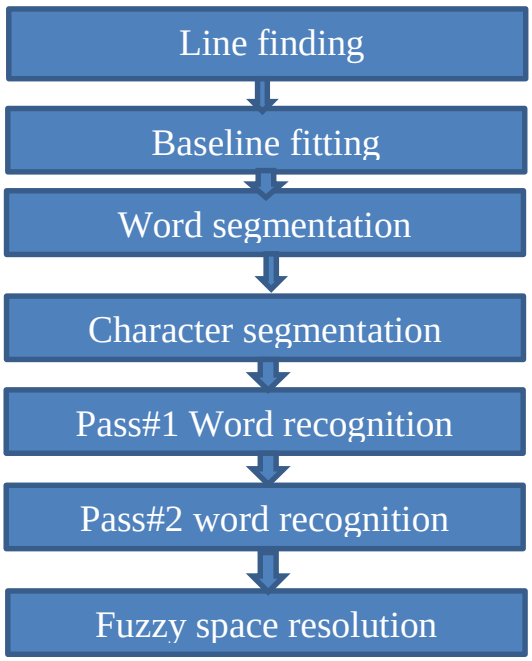

Figure 3.8: Tesseract component processing

Separate:

After detecting the language the lines will be separated from the document. IT is done by the image to string methods. By this the lines are totally separated from the document.

Translate:

For translating we are using google translator. After recognition and separation of the lines the next move is translating. For this we use google translator. After getting all the lines of the documentations this line will translate by google translator from online. This is done by the Tesseract OCR [14]. Currently the most accurate OCR engines is Tesseract. This software is free. It was developed by Hewlett-Packard from 1985-1995. Now it is maintained by Google. Tesseract can recognize only English, French, Italian, German, Spanish and Dutch [15]. If we train Tesseract it will recognize other languages. It can be train with the requisite data. It helps for image preprocessing.

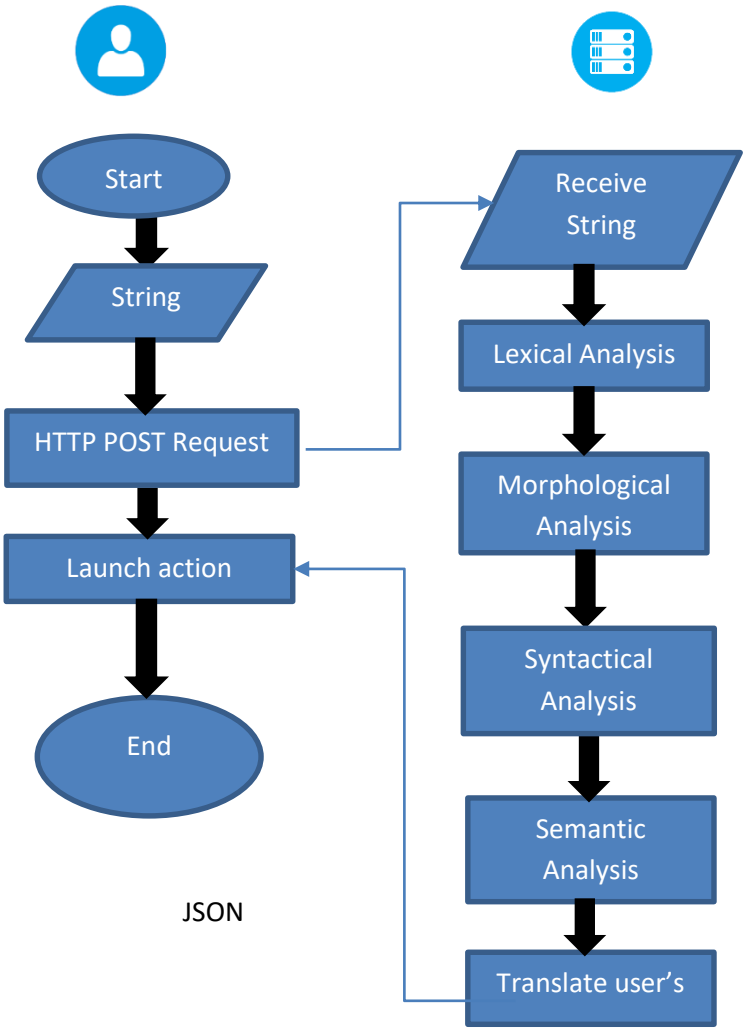

Figure 3.9: Google Cloud Translate API.

Terminology:

- Pixel:

  Pixel is the smallest graphic element of the image. With many pixel a picture is build. There are combinations of the RBG intensity to display different type of color combinations. 1080*720 pixels means that there are 1080 pixels for every row and 720 pixels for every column. Together formatting we can get 1080*720 pixel of image.

- Digital image

  -It is a computer image which is represented by pixels. This thesis will be used in word recognition. It will be used in image processing.

At starter of language identification, text block from an image format need to be identified [10].

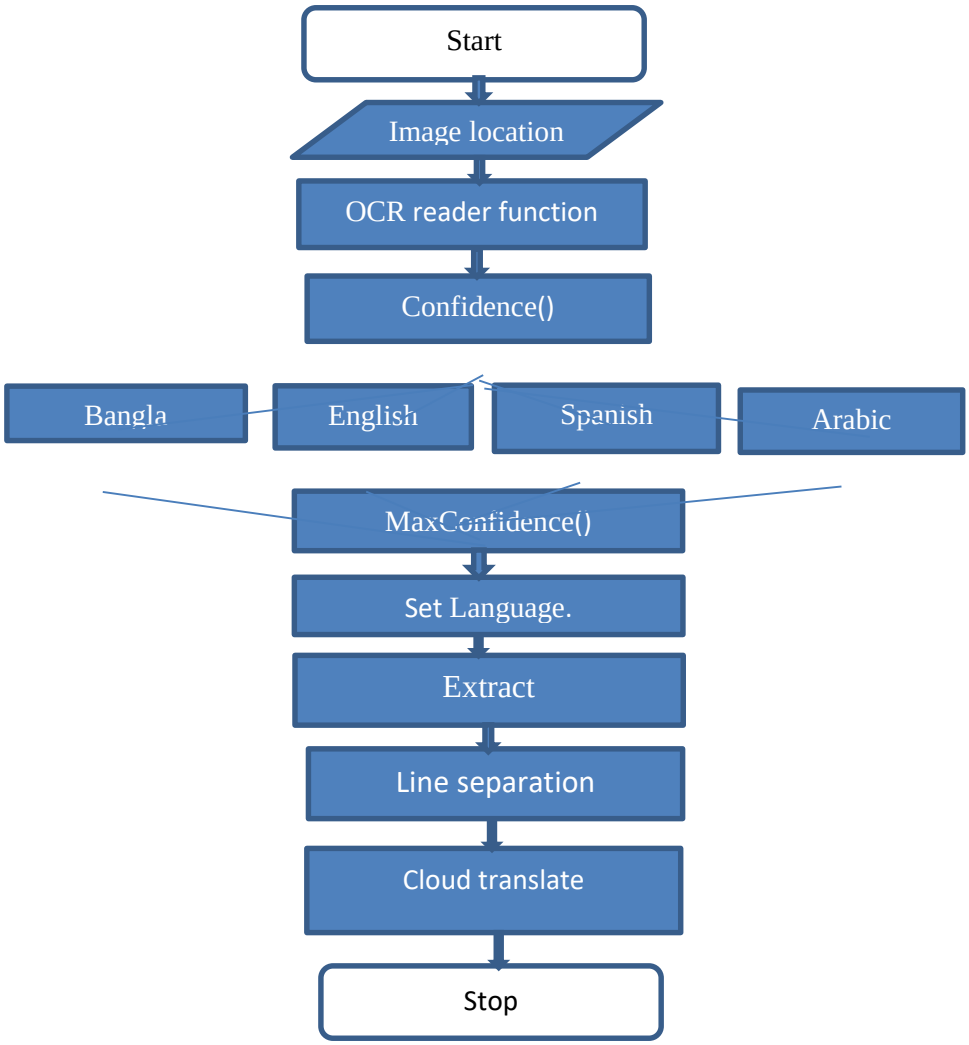

Figure 3.10: Language identification

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 INTRODUCTION

The system is divided into three parts. For the text recognition we use English, Bangla [6] and Arabic and Spanish language database. Using the detected language extracting text from image will start. At First it will scan the document for proper input text and then it will extract the text. Extracted text will be stored as a string. While extracting there is noise removal phase and through these noise will be removed from the text to perform with better accuracy. And then line separated string will be translated into target language.

4.2 EXPERIMENTAL RESULTS

Language recognition: We worked with four languages so far. Four types of input given bellow. Our proposed system first calculates the average confidence of the language and it detects the language that has the maximum confidence value.

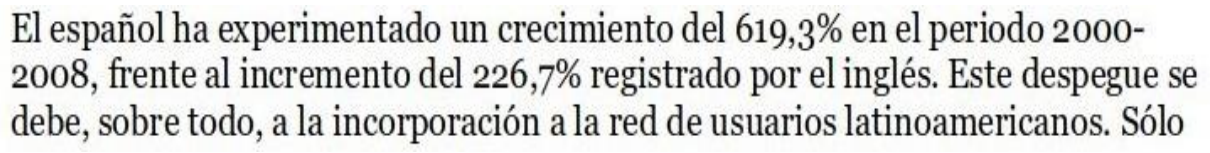
El español ha experimentado un crecimiento del 619,3% en el periodo 2000-2008, frente al incremento del 226,7% registrado por el inglés. Este despegue se debe, sobre todo, a la incorporación a la red de usuarios latinoamericanos. Sólo

Figure 4.1: Spanish input


النساء والرجال من كافة الاعمار سواء اصحاب احتياجات خاصه ام غير ذلك، لهم اولويات وتوقعات و احتياجات شخصيه عندما يتعلق الامر بالحصول الكامل المناسب على المياه وخدمات الصرف الصحي، ان الوضع

Figure 4.2: Arabic input

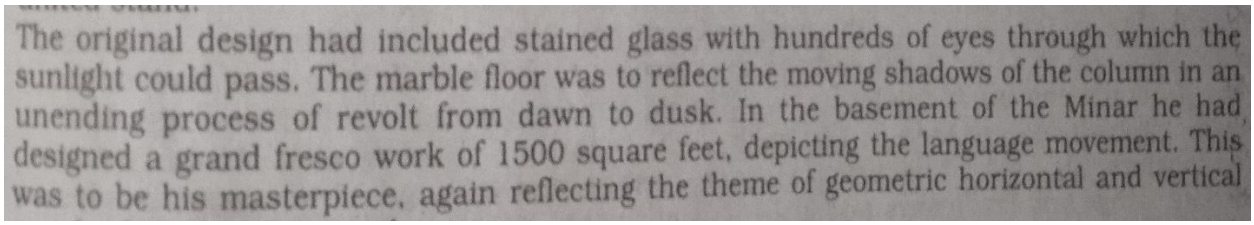
The original design had included stained glass with hundreds of eyes through which the sunlight could pass. The marble floor was to reflect the moving shadows of the column in an unending process of revolt from dawn to dusk. In the basement of the Minar he had designed a grand fresco work of 1500 square feet, depicting the language movement. This was to be his masterpiece, again reflecting the theme of geometric horizontal and vertical

Figure 4.3: English Input

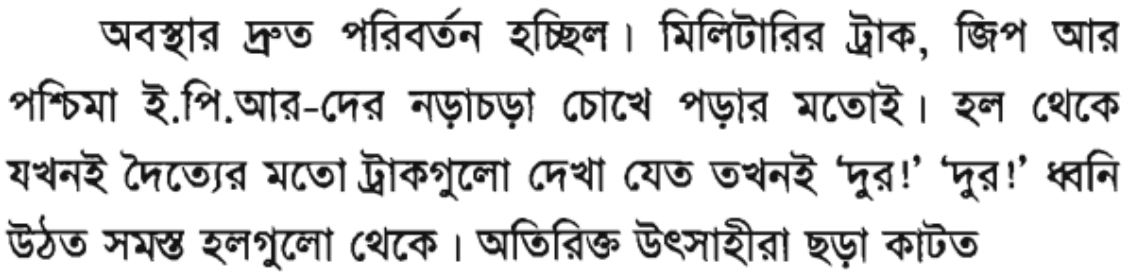
অবস্থার দ্রুত পরিবর্তন হচ্ছিল। মিলিটারির ট্রাক, জিপ আর পশ্চিমা ই.পি.আর-দের নড়াচড়া চোখে পড়ার মতোই। হল থেকে যখনই দৈত্যের মতো ট্রাকগুলো দেখা যেত তখনই 'দুর!' 'দুর!' ধ্বনি উঠত সমস্ত হলগুলো থেকে। অতিরিক্ত উৎসাহীরা ছড়া কাটত

Figure 4.4: Bangla input

4.3 DESCRIPTIVE ANALYSIS

Recurrent Neural Networks have been shown a great promise in recent times due to the Long Short Term Memory (LSTM) architecture. The LSTM architecture differs significantly from earlier architectures and appears to overcome many of the limitations and problems of those earlier architectures.

Traditional RNN, through are good at context aware processing, have not shown good performance for OCR task because of the vanishing gradient problem .The LSTM architecture was designed to overcome this problem. The highly non-linear recurrent network contains with multiplicative gates. Bidirectional LSTM model for accessing context in both forward and backward direction where both layers are connected to a single output layer. To avoid requirement of segmented training data, we use 4 language types to detect word confidence.

$$\text{Word confidence} = \min(100, \max(0, 100+5\times\text{certainty}))$$

Certainty value is returned by "tesseract".To measure the performance we use fore types of input in fore language. For every input, the program computes the word confidence value for each language.

Table 4.3: Confidence counts

Name	English	Spanish	Arabic	Bangla
Spanish	11	56	2	4
Arabic	10	15	99	12
English	40	14	3	3
Bangla	1	1	0	27

Based on the maximum value of word confidence, it will determine the language.

Table 4.4: Total language confidence.

Name	English (%)	Spanish (%)	Arabic (%)	Bangla (%)
Spanish	88	90	42	46
Arabic	23	29	56	22
English	88	87	36	41
Bangla	42	40	30	88

For every detected language, maximum average confidence value counts.

After detecting the language “pytesseract” will start image to string conversion.

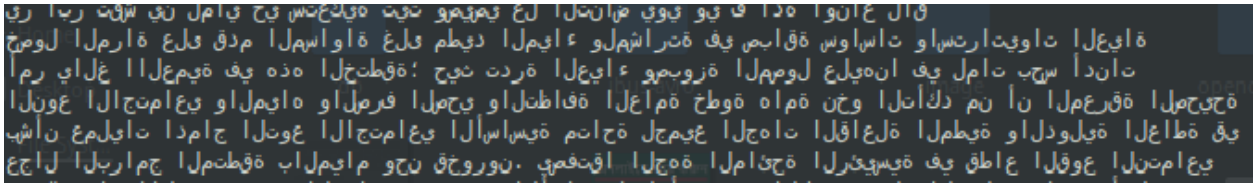

Figure 4.5: Text extract

“Tesseract” language database use iso-639-2 language identifier. The translation process is based on cloud translation API which uses iso-639-1 as a language identifier. After converting detected language to iso-639-1; cloud API will convert the string into English as a target language.

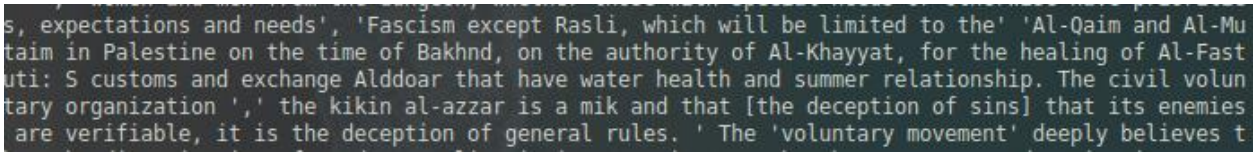

Figure 4.6: Text Translate.

4.4 SUMMARY

Our experimental results confirm that by using neural nets LSTM OCR Engine Model (OEM) in pytesseract, we can detect language from printed document. By using the identified language, we can extract all text in the document into a string using image to string conversion.

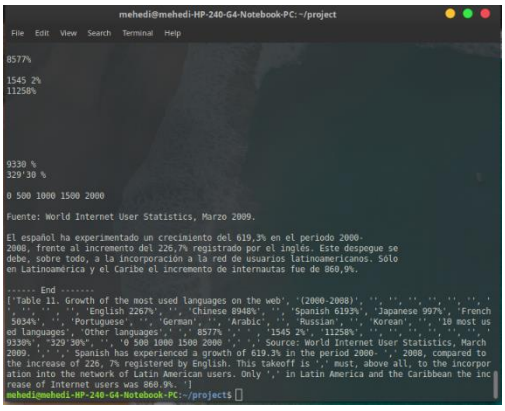

Figure 4.7: Preprocessing, extract and translate.

CHAPTER 5

CONCLUSION AND FUTURE WORKS

5.1 CONCLUSION

In this paper we scan an input language from a document then analyze and detect the language and then translate the source language into target language. To complete this process we have been gone through some pre-processing phase to remove the noise from the text as well. While pre-processing we uses three methods. Dilation and erosion methods to remove the noises and Binarization method to select the text as black area and background of the image as white. After selecting the target text using OCR it recognize the text by matching with the database and compute the average confidence for four languages. And it determines the maximum value among four languages. Consider the maximum value it detects the type of input language and then start translating into target language.

5.2 LIMITATIONS AND DRAWBACKS

It’s critical to understand that no OCR frame work is perfect. There is no such thing as an ideal OCR engine, particularly in real world. Besides expecting 100% accurate optical character recognition is basically unrealistic. As we discovered, our opencv OCR frame work worked well in some images, it failed in others. There are two essential reasons we will see our content acknowledgment pipeline fail:

1. The text is skewed/rotated.

2. The font of the text itself is not similar to what the Tesseract model was trained on.

Despite the fact that Tesseract v4 is essentially more ground-breaking and exact than Tesseract v3, the profound learning model is as yet restricted by the information it was prepared on — if your content contains adorned textual styles or text styles that Tesseract was not prepared on, it's impossible that Tesseract will have the capacity to OCR the content. Preprocessing is needed for this purpose.

Our OpenCV OCR pipeline works best on text that is

1. captured at a 90-degree angle (i.e., top-down, birds-eye-view) of the image

5.3 FUTURE WORKS

Our proposed system can perform only on four specific languages. So we want to develop and enrich our database to perform with more accuracy and add more languages. Besides our system can only take input by scanning the document. But we want to increase the input system facilities like audio input using microphone. Our proposed system can’t identify language from handwriting. So we want to work on that as well. And also we want add some more features like saving the translation history, shared the translation phrase via social media.

REFERENCES

[1] R. Smith, “An Overview of the Tesseract OCR Engine,” in ICDAR, 2007, pp. 629–633

[2]”Python Notes for professionals”: Optical character recognition, pp. 675-676

[3] D. Yadav, Sonia S´anchez-Cuadrado, and J. Morato. Optical character recognition for Hindi language using a neural-network approach. JIPS, 9(1):117–140, 2013.

[4] A. A. Chowdhury, Ejaj Ahmed, Shameem Ahmed, Shohrab Hossain and Chowdhury Mofizur Rahman"Optical Character Recognition of Bangla Characters using neural network: A better approach". 2nd International Conference on Electrical Engineering (ICEE 2002), Khulna, Bangladesh

[5] M. Sonka, V. Hlavac, R. Boyle “Image Processing, Analysis and Machine Vision”.4th edition, pp. 27,106,117

[6] Md. M. Alam, Dr. M. Abul Kashem “A complete Bangla OCR System for printed Characters”ISSN 2218-5114,volume 01,2010

[7] P Barlas, D Hebert, Cl´ement Chatelain, S Adam, Thierry Paquet. Language Identification

in Document Images. Journal of Imaging Science and Technology, Is&t Society for Imaging

Science and, 2016, 60 (1), pp.010407.

[8] Victor Ohlsson, Optical Character and Symbol Recognition using Tesserac.2016

[9] M. Al Azawi, M. Liwicki, and T. M. Breuel, “Combination of multiple aligned recognition outputs using WFST and LSTM,” in Document Analysis and Recognition (ICDAR),2015 13th Int. Conf. on. IEEE, 2015, pp. 31–35.

[10] M. Bacchiani and B. Roark. 2003. Unsupervised language model adaptation. In2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP’03)

.[11] J. H. AlKhateeb, A Database for Arabic Handwritten Character Recognition. ICCMIT-2015.

[12] Learn about OCR, available at: https://en.wikipedia.org/wiki/Optical_character_recognition, last accessed on 06-06-2018

[13] Available application:https://drive.google.com/file/d/0B1xkoc6gGhkMeHRSUUF6S0hESTA/view, last accessed on 12-09-2018

[14] Learn about AutoML, available at: https://www.ml4aad.org/automl/, last accessed on 12-09-2018

[15] Learn about Translation API, available at: https://cloud.google.com/translate/, last accessed on 12-09-2018

[16] Learn about Translator Text API, available at: https://azure.microsoft.com/en-us/services/cognitive-services/translator-text-api/, last accessed on 12-09-2018

[17] Learn about iTranslate, available at: https://www.itranslate.com/, last accessed on 12-09-2018

[18] Learn about Papago Translate, available at: https://papago.naver.com/, last accessed on 12-09-2018

[19] Learn about Naver papago, available at : https://en.wikipedia.org/wiki/Naver_Papago, last accessed on 02-12-2018

[20] Learn about camera translator, available at https://appgrooves.com/android/com.image.translate.cameratranslator/camera-translator/app-world-studio, last accessed on 02-12-2018

[21] Available application: https://www.digitaltrends.com/mobile/best-translation-apps/, last accessed on 02-12-2018

.