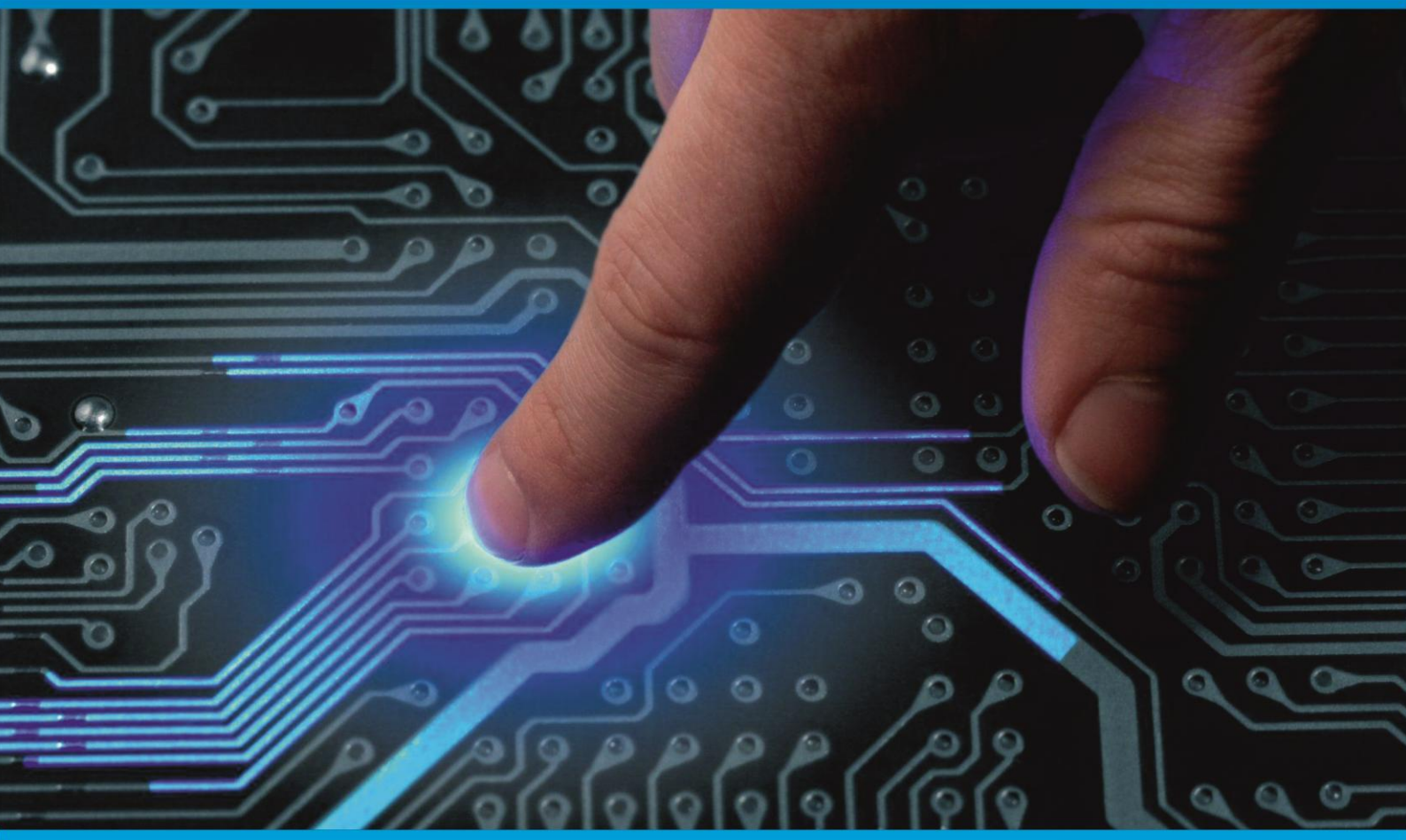
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905196 |

End-to-End Language Translation

Shubhangi Gund, Gayatri Gawali, Shweta Chindhe, Shivani Dasarwar

Department of Computer Engineering, Bharati Vidyapeeth's College of Engineering for women, Pune, India

ABSTRACT: Traditional speech-to-speech translation approaches concatenate automatic speech recognition (ASR), text to-text machine translation (MT), and text-to-speech synthesizer (TTS) by text information. The current state-of-the-art models for ASR, MT, and TTS have mainly been built using deep neural networks, in particular, an attention-based encoder-decoder neural Network with an attention mechanism. Recently, several works have constructed end-to-end direct speech-to-text translation by combining ASR and MT into a single model. However, the usefulness of these models has only been investigated on language pairs of similar syntax and word order. For syntactically distant language pairs (e.g., English-Japanese), speech translation requires distant word reordering. Furthermore, parallel texts with corresponding speech utterances that are suitable for training end- to-end speech translation are generally unavailable. Collecting such corpora is usually time consuming and expensive. This article proposes the first attempt to build an end-to-end direct speech-to- text translation system on syntactically distant language pairs that suffer from long-distance reordering.

KEYWORDS: NLP, Language translator, Python

I.INTRODUCTION

Human beings are able to communicate with each other and it is the most fundamental part. Worldwide there are around 7000 different languages. As the world is getting connected, language translation helps in bridging different cultures and economy between people from different countries and groups

Directly translating spoken utterances from a source language to a target language is challenging because it requires a fundamental transformation in both linguistic and para/non-linguistic features. Traditional speech-to-speech translation approaches concatenate automatic speech recognition (ASR), text-to-text machine translation (MT), and text-to-speech synthesizer (TTS) by text information. The current state-of-the-art models for ASR, MT, and TTS have mainly been built using deep neural networks, in particular, an attention-based encoder-decoder neural network with an attention mechanism. Recently, several works have constructed end-to-end direct speech-to-text translation by combining ASR and MT into a single model. However, the usefulness of these models has only been investigated on language pairs of similar syntax and word order (e.g., English-French or English-Spanish). For syntactically distant language pairs (e.g., English-Japanese), speech translation requires distant word reordering. Furthermore, parallel texts with corresponding speech utterances that are suitable for training end-to-end speech translationaregenerallyunavailable.Collectingsuchcorpora is usually time-consuming andexpensive.

Speechisaneasiestwaytocommunicatewitheachother. Speech processing is widely used in many applications like security devices, household appliances, cellular phones, ATM machines and computers. The human computer interface has been developed to communicate or interact conveniently for one who is suffering from some kind of disabilities. Speech-to-Text Conversion (STT) systems have a lot of benefits for the deaf or dumb people and find their applications in our daily lives. In the same way, the aim of the system is to convert the input speech signals into the text outputforthedeafordumbstudentsintheeducationalfields.

Python provides an API called Speech Recognition to allow us to convert audio into text for further processing. In this article, we will look at converting large or long audio files into text using the Speech Recognition API in python. Speech recognition involves audio input, pre-processing of audio input to de-noise it, extraction of important features and background removal and finally converting to text.

II.LITERATURE SURVEY

Article "End-to-End Speech Translation with Transcoding by Multi-Task Learning for Distant Language Pairs", proposed the first attempt to build an end-to-end direct speech-to-text translation system on syntactically distant language pairs that suffer from long-distance reordering. Takatomo Kano et al train the model on English (subject-verb-object (SVO) word order) and Japanese (SOV word order) language pairs. To guide the attention-based

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905196 |

encoder-decoder model on this difficult problem, they construct end-to-end speech translation with transcoding and utilize curriculum learning (CL) strategies that gradually train the network for end-to-end speech translation tasks by adapting the decoder or encoder parts. We use TTS for data augmentation to generate corresponding speech utterances from the existing parallel text data. The proposed approach provides significant improvements compared with conventional cascade models and the direct speech translation approach that uses a single model without transcoding and CL strategies.

Automatic Speech Recognition is the method in which speech signals are translated into the related sequence of characters into text (word). The research on speech recognition has been done for many years ago. After a lot of research work ShayanikaHazarika et al [2] got the idea of RNN (Recurrent Neural Network) algorithm is more effective and helpful algorithm for speech recognition. The Recurrent Neural Network is the type of Neural Network techniques in which we observe the difference of alphabet. The application or purpose of the RNN is to make people comfortable for the use of Hindi, English and Assamese languages. With the help of this research, people can easily understand languages and get its insight. Languages pose different challenges for speech recognition. This paper discusses about the use of RNN and Statistical Machine Translation algorithms that can be used for translation systems.

Paper [3] illustrates the architecture and working of a proposed multilingual machine translation system which is abletotranslatefromEnglishtoUrduandHindi.Thesystem applies translation rules based approach with artificial neural network. The efficient pattern matching and the ability of learning by examples makes neural networks suitable for implementation of a translation rule based machine translation system. This paper also describes the importance of machine translation systems and status of the languagesin a multilingual country like India. Machine translation evaluation score for translation output obtained from the system has been calculated using various methods such as n-gram bleu score, F-measure, Meteor and precision, recall. Theevaluationscoresachievedbythesystemforaround500 Hinditestsentencesareas:n-grambleuscore0.5903;Metric for Evaluation of Translation with Explicit ORdering (METEOR) score achieved is 0.7956 and Fscore of 0.7916 and for Urdu n-gram bleu score achieved by the system is 0.6054; METEOR score achieved is 0.8083 and Fscore of 0.8250.

Neural machine translation or NMT is a new proposed defined approach to the machine translation. Unlike the traditional SMT i.e. statistical machine translation, NMT focuses on constructing a single neural network that can be jointly aligned to maximize the performance, translation and efficiency. The models proposed in [4] for NMT is belongto agroupofencodersanddecodersandencodeasourcetextor sentence into a vector of fixed length from which a decoder generates an appropriate translation. This paper discussed different approaches for language translation. The HMM based encoder decoder models are used in the survey. The comparison in NMT and SMT is also analysed in this survey paper

Machine Translation using Deep Learning (Neural Machine Translation) is a newly proposed approach to machinetranslation.ThetermMachineTranslationisusedin the sense of translation of one language to another, with no human improvement. It can also be referred to as automated translation. Unlike the traditional statistical machine translation, the neural machine translation aims at building a single neural network that can be jointly tuned to maximize the translation performance. Survey presented in [5] reveals the information about Deep Neural Network (DNN) and concept of deep learning in field of natural language processing i.e. machine translation. It is better to use Recurrent Neural Network (RNN) inMachine Translation.

This paper studies various techniques used to train RNN for various language corpuses. RNN structure is very complicated and to train a large corpus is also a time- consuming task. Hence, a powerful hardware support (Graphics Processing Unit) is required. GPU improves the system performance by decreasing training time period.

The paper [6] deals with the implementation on the recent studies made on using DNN (Deep Neural Networks) to map sequences to sequences and implementing the same to automate Language translation process without the requirement of expert attention or statistical model. The paper deals with the demonstration of production ready end-to-end NMT (Neural Machine Translation) model using multi-layered Bidirectional RNN (Recurrent Neural Network), Long Short-Term Memory (LSTM) to encode the source statement to a sequence of numbers and the train another such RNN to decode the same to the target language statement. Our initial attempt is to build a vanilla NMT model and then further build the best possible NMT model using TensorFlow best practices. We reach this position by: i)Using the recent decoder / attention wrapper API, TensorFlow data iterator, ii) Incorporating Bidirectional LSTM, iii) Making use of TensorFlow methods such as batching, bucketing, beam search as well as deploying the model to multiple GPUs using GNMT attention. The main test bed is on 3 language conversions, English-Vietnamese, German-English and English-French.

In our day to day life the people are facing many problems in understand the languages. For example, if the people move from one state to the other they don't understand their language at that time Mobile Application

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905196 |

presented in [7] would help them. Existing system having a separate application for each and every process like camera, Google translator and Optical Character Recognition (OCR) text scanner. But, people expect the application consists of all the three facilities together. So this proposed application provides a new idea to the people to translate the other language text into their known language. This application contains three steps. 1. Take a photo image of the unknown languagetextwhichyouwanttotranslate(eitherhandwritten or printed material), 2.Tessaract is an open source Optical Character Recognition (OCR) technology, which is used to extract the text from the image then Google API and Bing API is used for translation of language. 3. The translatedtext is generated in PDFformat.

Urdu language of Pakistan has more than 100 million speakers in Pakistan, India, Afghanistan and Middle East. With low English literacy rate average Urdu speakingperson faces barriers in communicating with foreign people interms of accessing information, carrying business. Paper [8] proposes an interactive Urdu to English language speech translator using deep Neural Network. ASR module in proposedpipelineiscomposedofdeepneuralnetworkandis simpler as compared to traditional ASR which requires complex hand engineering like feature extraction and resources like phoneme dictionary. It was clearly seen that the proposed model shows the high accuracy when the input is recorded audio and it shows poor performance with real time input. While one HTTP request per input transcription produced English translation for Text to Text translation using Python Text Blob library. The final output was achieved with a delay of no more than 30 seconds. Furthermore,wehavetestedandprovidedsomestatistical

findings, the result shows that value updating for neural network layer's bias, standard deviation when Adam optimizerparametersaresetasfollows:beta1=0.9,beta2=0.9 and learning rate =0.01 meanwhile dropout rate was kept to 5% to offer regularization and observed value for scalar maximum lies between 0 and 0.08. There is a little deviation at 0.05 step, value decreases and afterwards that bias maximum scalar increases with positive values and finally increases exponentially at later stages of training further results are discussed in experiment section respectively. The proposed speech recognition model out performs traditional automatic speech recognition systems in efficiency, simplicity androbustness.

Paper [9] presents machine translation based on machine learning, which learns the semantically correct corpus. The machine learning process based on Quantum Neural Network (QNN) is used to recognizing the corpus pattern in realistic way. It translates on the basis of knowledge gained during learning by entering pair of sentences from source to target language. By taking help of this training data it translates the given text. own text.The paper consist study of a machine translation system which converts source language to target language using quantum neural network. Rather than comparing words semantically QNN compares numerical tags which are faster and accurate. In this tagger tags the part of sentences discretely which helps to map bilingual sentences.

Leena Jain et.al [10] developed a transliteration tool that translates English to Sanskrit. Transliteration is the process of converting the letters of typed text in one language to the letters of another language. The methodology used is to designanalgorithmthatusesUnicodefortransliteration.The Unicode for English and Hindi are mapped to each other. The input is taken in English and the letters are matched to Hindi through the mapped Unicode. The Output is the textin Hindi. Result and Conclusion All test cases passed 100% accuracy. The tool can be used for ML and Natural Language Translations. The interface isuser-friendly.

III.PROPOSED SYSTEM

Speech recognition is the process of converting audio into text. This is commonly used in voice assistants like Alexa, Siri, etc. Python provides an API called SpeechRecognition to allow us to convert audio into text for further processing. In this article, we will look at converting large or long audio files into text using the SpeechRecognition API in python. Speech recognition involves audio input, pre-processing of audio input to de- noise it, extraction of important features and background removal and finally converting to text. Speech processing flow is shown in figure 1.

Wordsegmentation:Wordsegmentationisdividingword into parts for extracting ROI (region of interest) from it. We can divide image into multiple segments that are set of pixels. The goal of Word segmentation is to change or simplify image which will be easier foranalysis.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905196 |

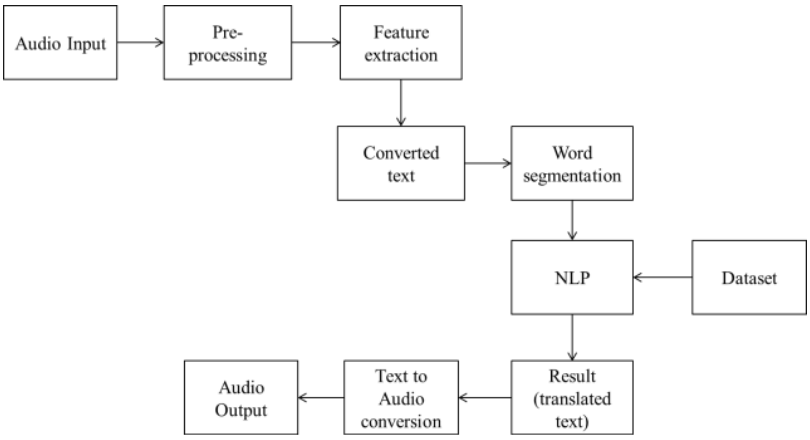


Fig. 1.proposed system

Language translation (from English to Japanese) using NLP:Systemstoredeachandeveryalphabetandgroupthem as a word in text file only when a space is detected. Again after giving a space it resumes grouping of alphabet into words w.r.t. Space. Features are extracted to figure out lines, corners, special patches from segmented image. In this step NLP classifies input alphabet according to extracted features and convert it into Japanese words. Dataset of Japanese sentences is feed to NLP for training. Output result is translated text. Text is converted into audio and audio output can be heard usingspeaker.

IV.CONCLUSION

We presented the construction of end-to-end speech translation for distant language pairs with transcoding based on CL strategies that gradually trained the network for end- to-end speech translation tasks by adapting decoder or encoder parts. In this paper NLP classifies input alphabet according to extracted features and convert it into Japanese words. Dataset of Japanese sentences is feed to NLP for training. Output result is translated text. Text is converted into audio and audio output can be heard using speaker. Use RNN, LSTM along with NLP for fast and accurate classification.

REFERENCES

[1] Takatomo Kano, SakrianiSakti and Satoshi Nakamura, “End-to-End Speech Translation with Transcoding by Multi-Task Learning for Distant Language Pairs”, IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 28, 2020

[2] ASSAMESE TRANSLATION SYSTEM ShayanikaHazarika, PoonamSuryawanshi, Nikita Shinde Prof. Monika Dangore Multidisciplinary Journal of Research in Engineering and Technology, Volume 6, Issue 3&4, Pg.25-30M5-6-34-10-2019

[3] The International Arab Journal of Information Technology, Vol. 16, No. 1, January 2019 125 A Model for English to Urdu and Hindi Machine Translation System using Translation Rules and Artificial Neural Network Shahnawaz Khan1 and ImranUsman

[4] International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 04 | Apr-2018 www.irjet.net p-ISSN: 2395-0072 © 2018, IRJET | Impact Factor value: 6.171 | ISO 9001:2008 Certified Journal | Page 3720 Survey and Analysis on Language Translator Using NeuralMachine Translation Ms. Neeta Verma1, Abhay Jain2, Animesh Basak3,
KshitijBhartiSaksena

[5] 2018 IJSRSET | Volume 4 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099 National Conference on Advanced Research Trends in Information and Computing Technologies (NCARTICT- 2018), Department of IT, L. D. College of Engineering, Ahmedabad, Gujarat, IndiaIn association with International Journal of Scientific Research in Science, Engineering and Technology 145 Machine Translation Using Deep Learning : A Survey Janhavi R. Chaudhary1,
Ankit C. Patel

[6] LANGUAGE TRANSLATION-USING SEQUENCE TO SEQUENCE LEARNING Arghajyoti Singha1 , A NithyaKalyani2 , Mainak Pal Chowdhury International Journal of Pure and Applied Mathematics Volume 120 No. 6 2018,10299-10310

[7] International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Volume 4,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905196 |

Issue 4, April 2017 22 ISSN (Online) 2394-2320 All Rights Reserved © 2017 IJERCSEExtraction of Text from an Image and its Language Translation Using OCR G.R.Hemalakshmi, M. Sakthimanimala, J. SalaiAniMuthu

[8] Urdu Language Translator using Deep Neural Network Syed Abbas Ali1 *, Sallar Khan2 , Humaira Perveen1 , Reham Muzzamil1 , Mahnoor Malik1 and Faiza Khalid Indian Journal of Science and Technology, Vol 10(40), DOI:10.17485/ijst/2017/v10i40/120273, October 2017

[9] International Journal of Engineering and TechniquL. Jain and P. Agrawal, "English to Sanskrit Transliteration: an effective approach to design Natural Language Translation Tool", International Journal ofAdvancedResearchinComputerScience,Volume8,No.1,Jan-Feb 2017

[10] L. Jain and P. Agrawal, "English to Sanskrit Transliteration: an effective approach to design Natural Language Translation Tool", International Journal of Advanced Research in Computer Science, Volume 8, No. 1, Jan-Feb2017

INNO SPACE

SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details