Vivekanand Education Society's Institute of Technology

Department of Computer Engineering

Group No.: 31

Date :- 02/08/2024

Project Synopsis (2024-25) - Sem VII

AeroVoice : Automatic Speech Recognition for ATC Communication

Ms. Lifna C.S.
Assistant Professor
CMPN

Anurag Shirsekar
V.E.S.I.T
2021.anurag.shirsekar@ves.ac.in

Dhruva Chaudhari
V.E.S.I.T
2021.dhruva.chaudhari@ves.ac.in

Preethika Shetty
V.E.S.I.T
2021.preethika.shetty@ves.ac.in

Sneha Tanna
V.E.S.I.T
2021.sneha.tanna@ves.ac.in

Abstract

Given the continuous rise in global air traffic, it is necessary to make advancements in air traffic control (ATC) systems to ensure optimal performance and safety. Precise transcription and interpretation of ATC messages provide a significant challenge in this domain, frequently worsened by factors such as varying dialects, professionals aviation terminology, and radio interference. These issues attract attention to the inadequacies of present ATC systems coupled with the need of real-time data integration and analysis.

With a view toward altering the field of air traffic management, this project offers the establishment of a revolutionary ATC Radio Conversation Analyzer. The system attempts to provide a robust solution for the inherent complexity of air traffic control (ATC) communication by leveraging the capabilities of modern technologies such as speech recognition, natural language processing (NLP), and machine learning. Apart from assisting accurate transcription, the analyzer will provide increased contextual understanding—which is important to improve situational awareness.

By means of these advancements, the ATC Radio Conversation Analyzer marks a key progress in the evolution of air traffic control, hence tackling important operational challenges and clearing the route for a more integrated and responsive ATC system.

Introduction

The aviation industry is experiencing unprecedented growth, with increasing volumes of air traffic creating new challenges for air traffic controllers (ATCs). These professionals are tasked with ensuring the safe and efficient movement of aircraft, requiring them to manage a constant flow of real-time communications with pilots. The diversity of accents, variations in speech clarity, background noise, and the use of specialized jargon make these communications complex to transcribe and analyze accurately. Traditional methods often fall short in handling these challenges, potentially leading to misunderstandings, delays, and even safety incidents.

To address these issues, this project proposes the development of an advanced ATC Radio Conversation Analyzer. This system aims to leverage state-of-the-art technologies in speech recognition, natural language processing (NLP), and machine learning to provide accurate and real-time transcription and analysis of ATC communications.

It is designed to handle the complexities of diverse accents, jargon, and noisy environments, extracting and contextualizing critical information from conversations. The integration of this data will enhance situational awareness and support ATCs in making informed decisions.

Furthermore, the project includes the implementation of real-time monitoring and anomaly detection features, offering timely alerts to potential issues. Through these innovations, the project seeks to transform the management of ATC communications, ultimately contributing to a safer and more efficient air traffic control system.

Problem Statement

The increasing volume and complexity of air traffic present significant challenges for air traffic controllers, who must manage a high volume of real-time communications between pilots and controllers. This task is complicated by the variability in accents, speech clarity, background noise, and the use of specialized jargon, making accurate transcription and analysis of ATC communications difficult with conventional methods. Miscommunications or delays in recognizing critical information can lead to operational inefficiencies and safety incidents. Additionally, current systems often lack the capability to provide real-time insights and alerts, limiting the ability to proactively manage and mitigate potential issues. Integrating ATC communications with other relevant data sources, such as flight tracking and weather information, remains a challenge, hindering comprehensive situational awareness.

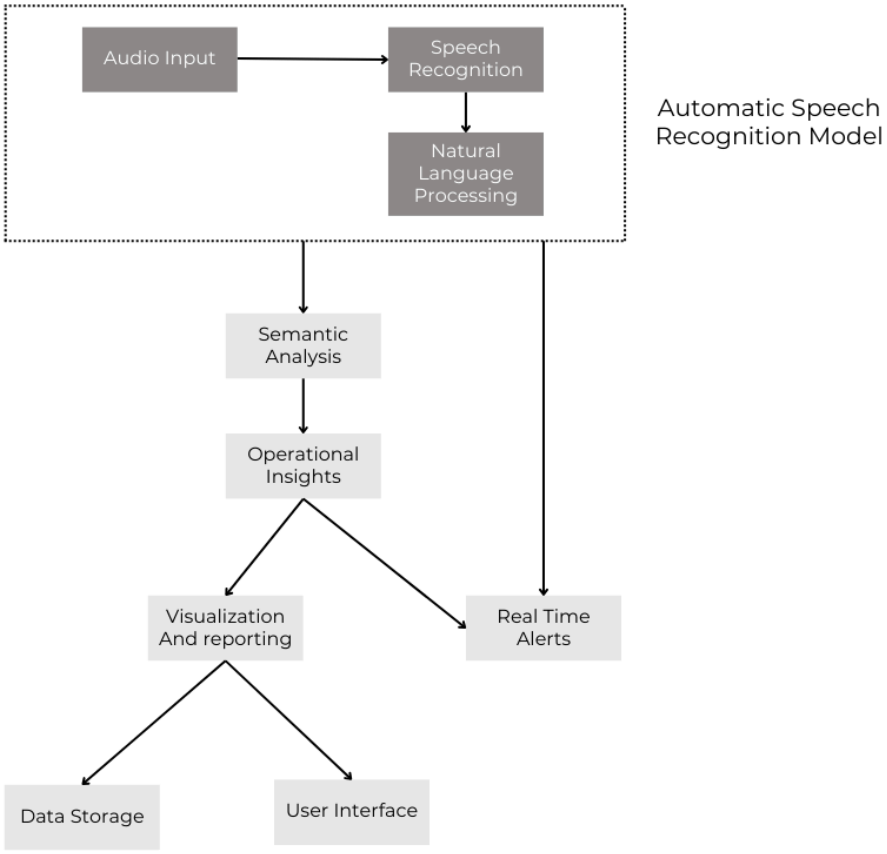
To address these challenges, this project aims to develop an advanced ATC Radio Conversation Analyzer that leverages cutting-edge speech recognition, natural language processing (NLP), and machine learning techniques. The system will provide robust real-time transcription and analysis of ATC communications, tailored to handle diverse accents, jargon, and noisy environments. By extracting and highlighting critical information from conversations and integrating it with flight data and weather conditions, the system will enhance operational efficiency, safety, and decision-making processes. Additionally, the analyzer will include real-time monitoring and anomaly detection to identify and alert controllers to potential issues, facilitating proactive management of air traffic. Through these innovations, the project seeks to transform ATC communications, supporting air traffic controllers in managing airspace more effectively and safely.

Proposed Solution

To tackle the rising difficulties in air traffic control communications, our research proposes a dual-focused method that integrates sophisticated studies with practical application. The system seeks to boost transcription accuracy, situational awareness, and safety with powerful speech recognition, natural language processing (NLP), and machine learning technologies. And with research this application-driven solution focuses on establishing an operational ATC Radio Conversation Analyzer that increases the efficiency, safety, and decision-making capabilities of air traffic controllers.
The primary components of the suggested technique are:

1. Speech Recognition and NLP.:
    - Research: Create and develop Automatic Speech Recognition (ASR) models that are specifically designed for ATC contexts, taking into consideration different accents, loud situations, and specialist jargon. This includes investigating sophisticated noise reduction techniques and developing models for domain-specific datasets.
    - Application: Use these ASR models in a real-time system that can accurately transcribe live ATC messages. The system will also be able to process pre-recorded audio recordings, making it suitable for usage in a variety of operational settings.
2. Natural Language Processing (NLP) and Semantic Analysis:
    - Research NLP systems for recognizing entities, processing commands, and detecting intent in ATC communications. The goal will be to refine models that can accurately grasp complex and context-dependent discussions.
    - Application: Incorporate these NLP skills into the system to automatically detect critical information including aircraft numbers, orders, and probable abnormalities. The system will provide contextual understanding and real-time notifications to help in decision-making.
3. Real-Time Audio Processing and Transcription:
    - The system will enable for real-time processing of live ATC radio conversations as well as pre-recorded audio recordings, transforming voice into text with high accuracy even in loud settings.
    - Implementation: Using updated ASR models, the system will integrate noise filtering and accent adaption to enable accurate transcription across a wide range of accents and audio characteristics.
4. Operational Insights:
    - Feature: NLP algorithms will scan recorded text to extract and highlight important information like flight numbers, instructions, and emergency phrases.
    - Implementation: The system will use entity recognition and intent detection models to parse and tag essential parts in discussions, with an emphasis on operational relevance. Contextual understanding will aid in detecting and flagging anomalies or departures from standard procedures.

Proposed Solution



AeroVoice caters to two main user groups: Air Traffic Controllers (ATCs) and ATC Management Teams.

For Air Traffic Controllers (ATCs):

- **Real-Time Transcription:** Converts live ATC communications into text, enhancing clarity and reducing the chance of miscommunication.
- **Contextual Analysis:** Provides contextual understanding of conversations, highlighting critical information such as aircraft numbers, instructions, and emergency phrases.
- **Anomaly Detection:** Alerts controllers to potential issues or deviations from standard procedures.

For ATC Management Teams:

- **Dashboard Access:** Allows monitoring and analysis of communication patterns, identifying areas for improvement.
- **Report Generation:** Facilitates the generation of detailed reports on communication efficiency, accuracy, and system performance.
- **Integration with Flight Data:** Provides insights by integrating with flight tracking and weather data, enhancing situational awareness and decision-making.

Proposed Solution

Hardware , Software and tools Requirements

Software Requirements

1. Programming Languages and Frameworks

- Backend Development: Python (Flask or Django)
- Frontend Development: JavaScript (React.js or Angular)
- Machine Learning & NLP: Python (TensorFlow, PyTorch, SpaCy, NLTK)
- Speech Recognition: Python (Kaldi, DeepSpeech, Google Cloud Speech-to-Text API)

2. Data Integration and APIs

- Flight Data API: FlightAware, ADS-B Exchange
- ATC Communication Sources: Live ATC feeds and audio repositories

3. Development Tools
- Version Control: Git (GitHub)

4. Miscellaneous tools-
VS code, Jupyter Notebook,Google Colab

Hardware Requirements

1.Computer/Laptop - i5 Generation
2.Good internet connectivity

Proposed Evaluation Measures

The main goal of our ATC Radio Conversation Analyzer is to enhance transcription accuracy, situational awareness, and safety in air traffic control by integrating advanced speech recognition, natural language processing (NLP), and machine learning methods. This technology aims to improve decision-making capabilities for air traffic controllers by handling the challenges of various accents, noisy settings, and specialized jargon. As we create and execute this solution, we have devised a number of evaluation metrics to verify the system meets its objectives and performs well in real-world conditions.

- **Speech Recognition Accuracy:** Assess the accuracy of the ASR models in transcribing live and recorded ATC communications, especially in noisy settings and with various accents.
- **Entity Recognition and purpose Detection:** Evaluate the system's ability to correctly identify and tag key entities (e.g., flight numbers, instructions) and accurately determine the purpose behind ATC interactions.
- **Real-Time Processing Performance:** Measure the system's speed in processing and transcribing live audio streams with low latency.
- **Contextual Understanding:** Test the system's ability to understand the context of discussions, including anomaly identification and response creation.
- **Operational Relevance:** Evaluate the practical effect of the system on situational awareness and decision-making, concentrating on its capacity to deliver relevant and actionable data to air traffic controllers.
- **System Usability and Stability:** Assess the overall user experience, including interface responsiveness, system stability, and ease of use in various operating conditions.

Conclusion

The AeroVoice project represents a significant advancement in enhancing air traffic control (ATC) communications through the integration of Automatic Speech Recognition (ASR), Natural Language Processing (NLP), and machine learning technologies. By addressing the complexities of real-time communication—such as diverse accents, noisy environments, and specialized jargon—AeroVoice aims to provide precise and actionable insights for air traffic controllers.

The system's ability to accurately transcribe and analyze ATC communications will greatly improve situational awareness and decision-making capabilities. Real-time transcription, contextual understanding, and anomaly detection features will enhance operational efficiency and safety in air traffic management. Additionally, the integration of flight data and weather information will provide a comprehensive view, supporting proactive decision-making and reducing the likelihood of miscommunications.

Overall, AeroVoice is poised to transform the management of air traffic control communications, offering a more reliable and responsive system that contributes to safer and more efficient air traffic operations.

References

1. J. Zuluaga-Gomez *et al.*, “ATCO2 corpus: A Large-Scale Dataset for Research on Automatic Speech Recognition and Natural Language Understanding of Air Traffic Control Communications,” *arXiv.org*, Nov. 08, 2022. https://arxiv.org/abs/2211.04054
2. E. Pinska-Chauvin, H. Helmke, J. Dokic, P. Hartikainen, O. Ohneiser, and R. G. Lasheras, “Ensuring safety for Artificial-Intelligence-Based automatic speech recognition in air traffic control environment,” *Aerospace*, vol. 10, no. 11, p. 941, Nov. 2023, doi: 10.3390/aerospace10110941.
3. J. Zuluaga-Gomez *et al.*, “Lessons Learned in Transcribing 5000 h of Air Traffic Control Communications for Robust Automatic Speech Understanding,” *Aerospace*, vol. 10, no. 10, p. 898, Oct. 2023, doi: 10.3390/aerospace10100898.

Mentor Signature: 9/8/24.

(LIFNA. C. S)

Group Member Signature:

Sneha Tanva – sneha

Preethika Shetty – P

Dhruva Chaudhari – Dh

Anurag Shirsekar – Anurag...