

**Title of the Project**

**Mini-Project Synopsis**

*submitted to*

Manipal School of Information Sciences, MAHE, Manipal

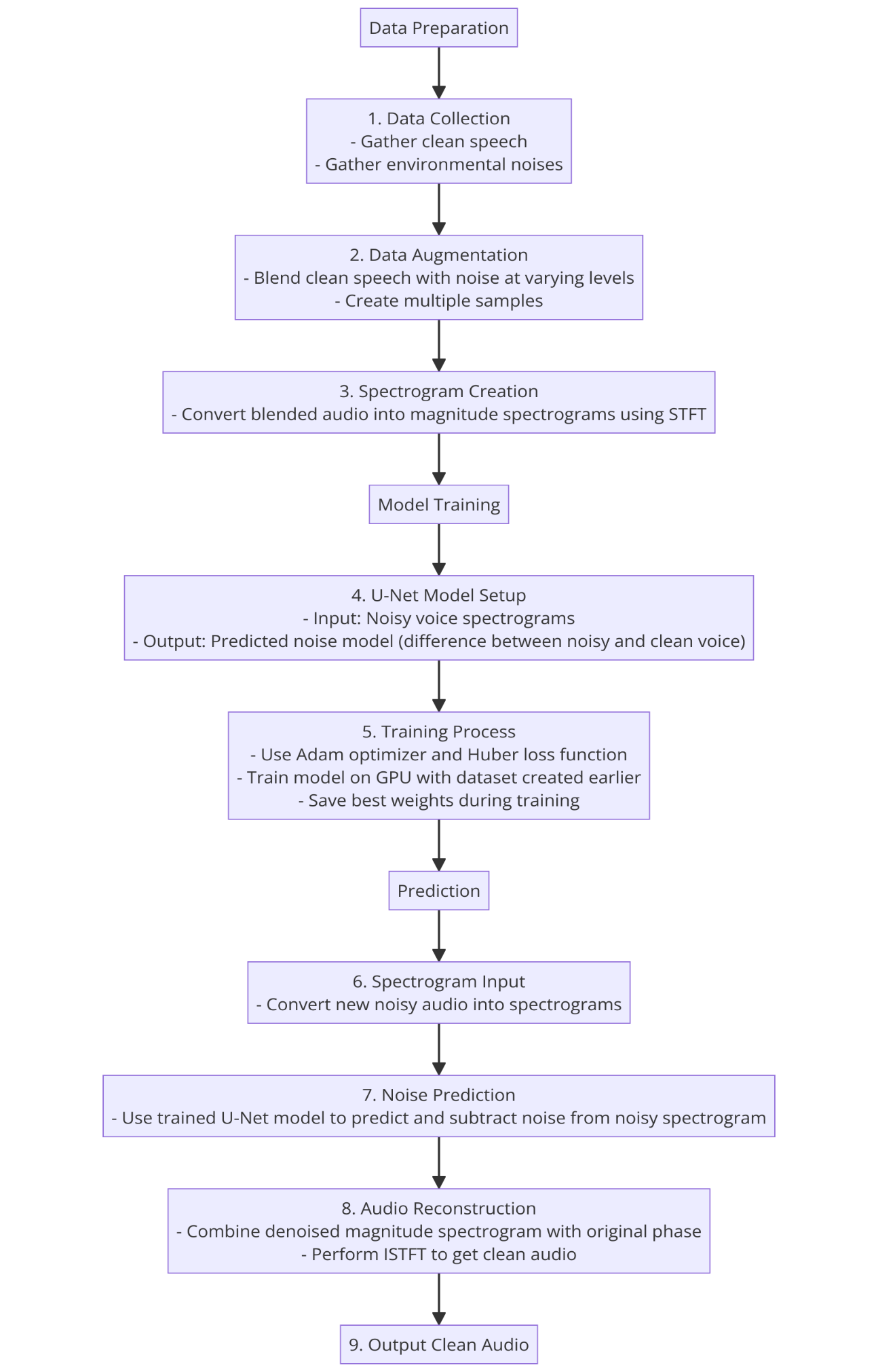
| **Reg. Number** | **Name** | **Branch** |
| --- | --- | --- |
| **241058020** | **Nikhil M** | **Big Data Analytics** |
| **241058024** | **Nikhil S G** | **Big Data Analytics** |
| **241058030** | **Vinayashree M Shet** | **Big Data Analytics** |

**14/08/2024**

****

1. **Objectives**

The goal of this project is to develop a **speech enhancement system** that reduces environmental noise from audio recordings. The system uses **magnitude spectrograms** as a representation of sound, leveraging a U-Net deep learning model to predict and subtract noise from noisy voice spectrograms. This results in clearer speech audio. The process involves creating a dataset of blended noisy and clean voices, training the model to recognize noise patterns, and applying the trained model to new noisy audio for real-time noise reduction.

1. **Block Diagram/Flowchart**
2. **Applications**

**3.1. Noise Reduction in Voice Recordings**

**Application:** The system can be used to enhance speech clarity by removing background noise from voice recordings, making it easier to understand spoken words in environments where there is significant noise, such as crowded public spaces or noisy workplaces.

**Examples:**

**Call Centers:** Reducing background chatter to improve customer service interactions.

**Lecture Recordings:** Cleaning up recordings of lectures or presentations held in noisy environments to ensure the speaker's voice is clear.

**Public Safety:** Enhancing audio from police or emergency responder recordings, where clear communication is critical.

**3.2. Improved Audio Quality in Telecommunication and Assistive Devices**

**Application:** This technology can be integrated into telecommunication systems, such as VoIP services, mobile networks, or video conferencing tools, to ensure that voice communication remains clear, even in environments with high levels of ambient noise. It can also be used in assistive hearing devices to filter out unwanted background noise.

**Examples:**

**Video Conferencing:** Enhancing voice clarity during video calls, especially in remote work settings where participants might be in less-than-ideal acoustic environments.

**Hearing Aids:** Implementing real-time noise reduction in hearing aids to help users better focus on conversations in noisy settings like restaurants or busy streets.

**Smartphones:** Integrating noise reduction features in mobile phones to improve call quality in noisy environments like subways or city streets.

**3.3. Audio Processing in Media Production**

**Application:** In the media and entertainment industry, the system can be used during the post-production phase to improve the quality of audio recordings. This is particularly useful when the original audio is recorded in less-than-ideal conditions, where noise cannot be completely avoided.

**Examples:**

**Film and Television Production:** Enhancing dialogue clarity in scenes shot in noisy outdoor locations or crowded places.

**Podcasting:** Cleaning up interviews or discussions recorded in public spaces or with imperfect equipment, ensuring the final product sounds professional.

**Music Production:** Reducing noise in vocal tracks recorded in home studios or during live performances to produce cleaner audio tracks.

1. **Software & Hardware Requirements**

**Software:**

* Python: Main programming language.
* TensorFlow/Keras: For implementing and training the U-Net model.
* LibriSpeech, ESC-50, SiSec datasets: For obtaining clean speech and noise data.
* Google Colab: For using free GPU resources.

**Hardware:**

* GPU (Graphics Processing Unit): For faster training of deep learning models.
* High-Performance CPU: Supports general data processing and model training tasks.
* Storage: Enough space for storing and processing audio datasets, preferably over 5GB.

1. **Review of existing literature**