

# **SIT740 Research and Development in Information Technology**

## **High Distinction Task 6.1: Propose/Pitch an Idea**

---

### **IOT voice monitoring application using AI chip:**

#### **Introduction:**

With the fast-growing world, there is an increase in need to monitor and regulate various activities of mankind to ensure safety and privacy. This has led to various surveillance/monitoring systems with Video surveillance system being the most popular. However, video surveillance system has limitations in situations where visual recording is not feasible or simply not allowed. A viable way to address this limitation would be implantation of voice recognition system. The proposed idea dives into the usage of AI on chip (CM1K) combined with IOT technologies for voice monitoring application in security, education and media.

#### **Description:**

The existing neural networks AI algorithms used for IoT applications are software-based and are run on classical processors which are slow for pattern recognition processing. The idea is to train a hardware neural network AI system on CM1K chip designed for fast pattern recognition processing to perform voice recognition task. The concept is to have a standalone AI system connected to an IoT system that receives and process audio signal to determine number of voices in the audio signal in real-time. The system should be able to perform data processing (data segmentation) and classification (clustering and results). The voice training dataset is fed to train the CM1K chip after successful training the model is ready to be tested on the data from IoT system. The data received undergoes segmentation to convert the audio signal to respective coefficients and is sent to the CM1K chip for testing or recognizing in this case. If the voice is recognised the process also checks for if the voice was recognised before or else an increment counter is added to notify and retrain the chip on the new data.

## Process flow diagram:

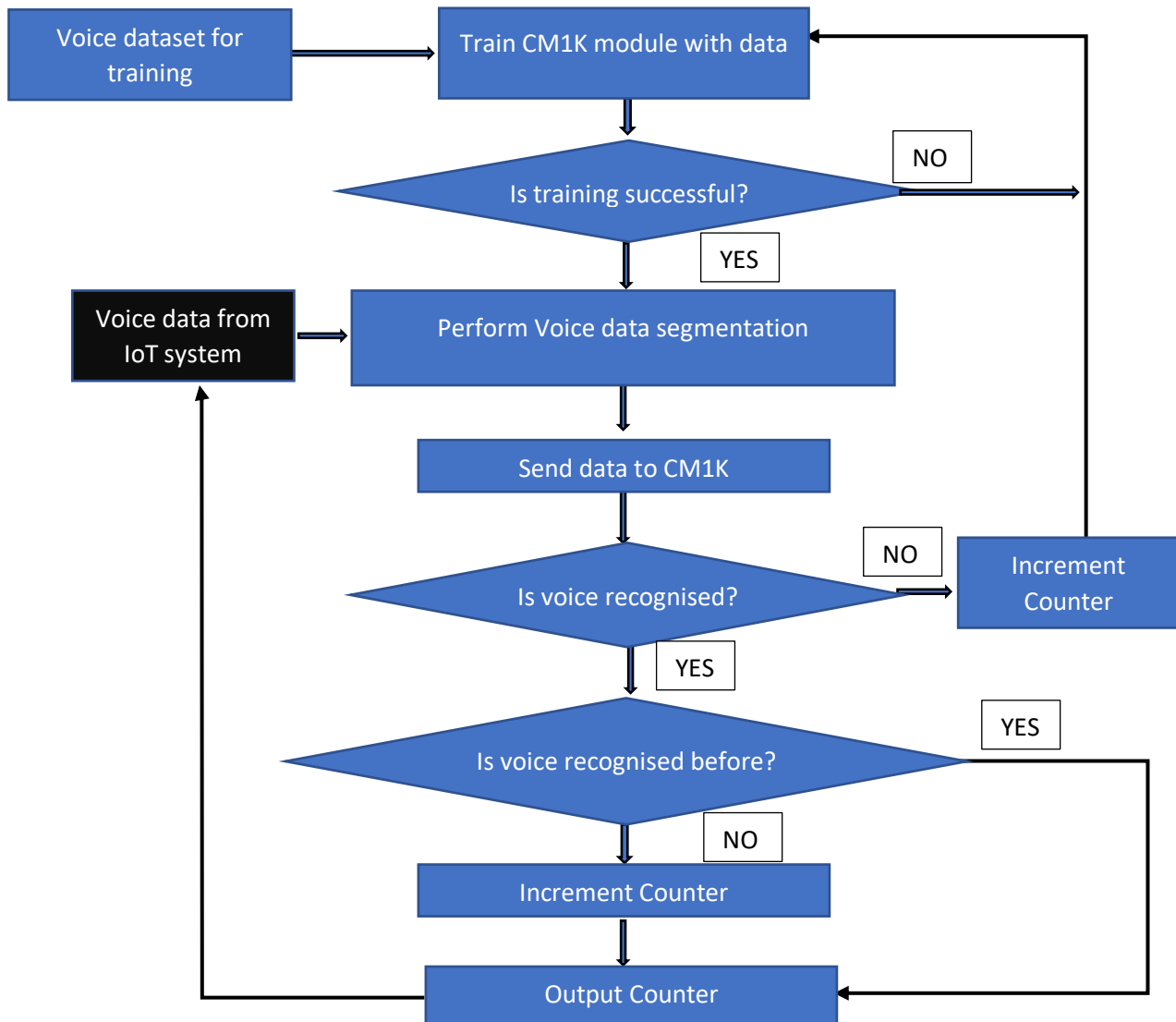
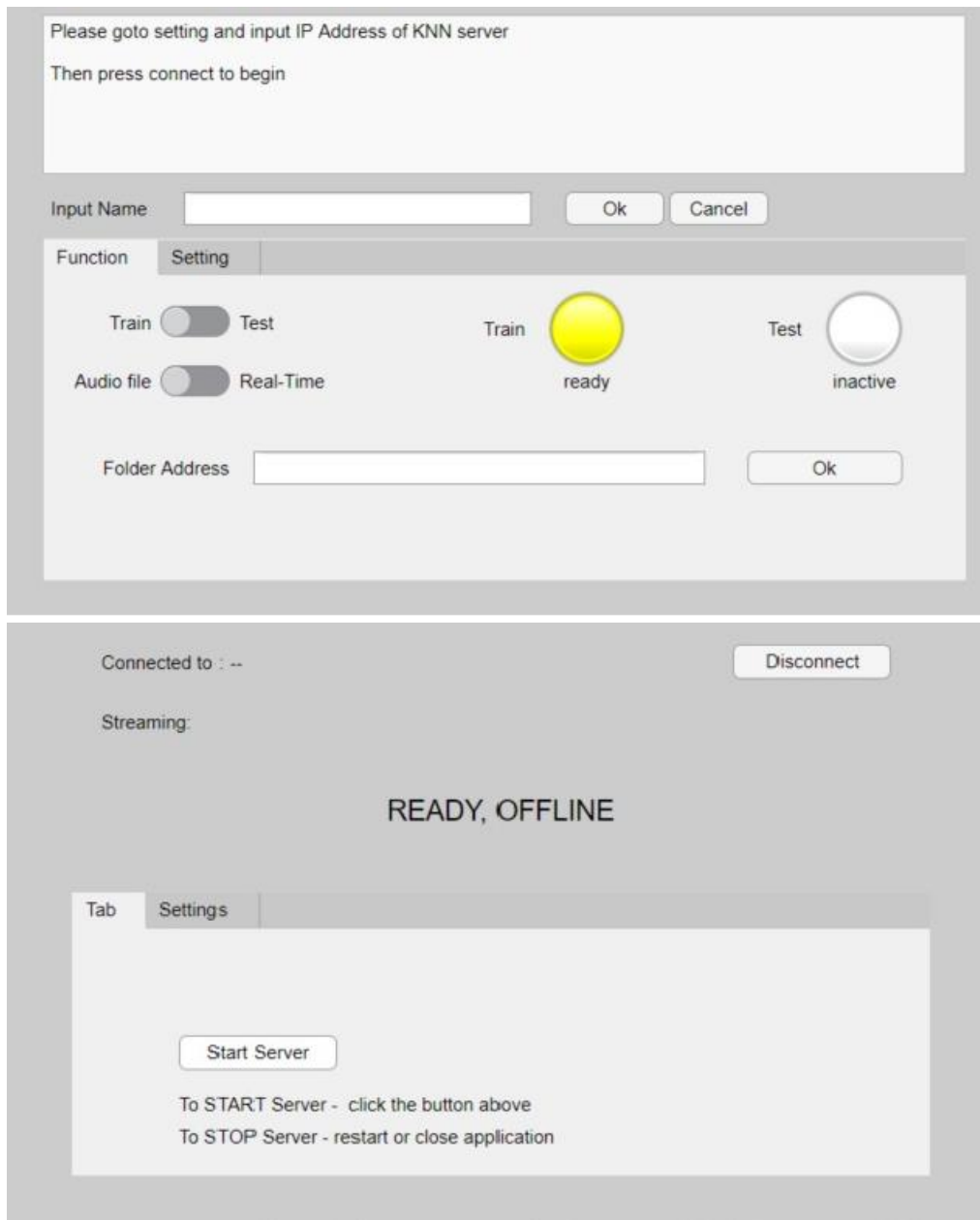


Fig 1. Flowchart of the process flow.

## User Interface:





The image displays two screenshots of a software application's user interface. The top screenshot shows a settings window with a title bar and a close button. The main content area has a light gray background and contains the following elements: a text box with the instruction "Please goto setting and input IP Address of KNN server" and a sub-instruction "Then press connect to begin"; an "Input Name" label followed by a text input field and "Ok" and "Cancel" buttons; a "Function" tab bar with "Setting" selected; two rows of toggle switches for "Train" and "Test" (both currently off) and "Audio file" and "Real-Time" (both currently off); a "Train" status indicator consisting of a yellow circle with the word "ready" below it, and a "Test" status indicator consisting of a white circle with the word "inactive" below it; and a "Folder Address" label followed by a text input field and an "Ok" button. The bottom screenshot shows a main application window with a title bar and a close button. It features a "Connected to : --" status bar with a "Disconnect" button. Below this is a "Streaming:" label. The central area displays the text "READY, OFFLINE" in large, bold, black letters. At the bottom, there is a "Tab" bar with "Settings" selected, a "Start Server" button, and two lines of instructional text: "To START Server - click the button above" and "To STOP Server - restart or close application".

Please goto setting and input IP Address of KNN server  
Then press connect to begin

Input Name:

Function: **Setting**

Train ☐ Test ☐ Audio file ☐ Real-Time ☐

Train  ready Test  inactive

Folder Address:

Connected to : --

Streaming:

**READY, OFFLINE**

Tab: **Settings**

To START Server - click the button above  
To STOP Server - restart or close application

## Application:

The system if designed and integrated with IoT system can overcome the limitation of video surveillance systems. The implementation of voice recognition system on neuromorphic chip (CM1K) is a great step to understanding and improving cognitive computing. The neuromorphic chips are faster and has more potentials for cognitive computing than algorithms that run on classical computers.

## **Advantages:**

1. Smart and efficient voice recognition system.
2. Being AI enabled the system adapts to varied circumstances.
3. Most economical than video surveillance system.
4. Faster than ones running on classical computers.
5. Better performance than algorithms implemented on software.

## **Potential Drawbacks:**

Ethical considerations such as intellectual property protection, user data privacy can be few drawbacks when the system is implemented on a commercial scale. Where data and privacy intrusion can be a problem as with everything that is on the internet.

**Video link:** [https://video.deakin.edu.au/media/t/0\\_pvvk144e](https://video.deakin.edu.au/media/t/0_pvvk144e)

## **Reference:**

1. Noah Moroze. "Experiments with the CM1K Neural Net Chip." Medium <https://towardsdatascience.com/experiments-with-the-cm1k-neural-net-chip-32b2d5ca723b>.
2. Y. Chen et al., "Neuromorphic computing's yesterday, today, and tomorrow – an evolutionary view," Integration, vol. 61, pp. 49-61, 2018/03/01/ 2018, doi: <https://doi.org/10.1016/j.vlsi.2017.11.001>.
3. S. S. Tirumala and S. R. Shahamiri, "A review on Deep Learning approaches in Speaker Identification," in Proceedings of the 8th international conference on signal processing systems, 2016: ACM, pp. 142-147.
4. S. Jothilakshmi, V. Ramalingam, and S. Palanivel, "Speaker diarization using autoassociative neural networks," Engineering Applications of Artificial Intelligence, vol. 22, no. 4, pp. 667-675, 2009/06/01/ 2009, doi: <https://doi.org/10.1016/j.engappai.2009.01.012>.
5. B. Juang and L. R. Rabiner, "Automatic Speech Recognition—A Brief History of the Technology Development Abstract," 2005.