

Report for the IBM Hack Challenge 2023, submitted by Harsh Kumar from Team Radiators:



# IBM Hack Challenge 2023 -

## Team Radiators

### Executive Summary

**Hackathon Team:** Team Radiators

**Team Members:**

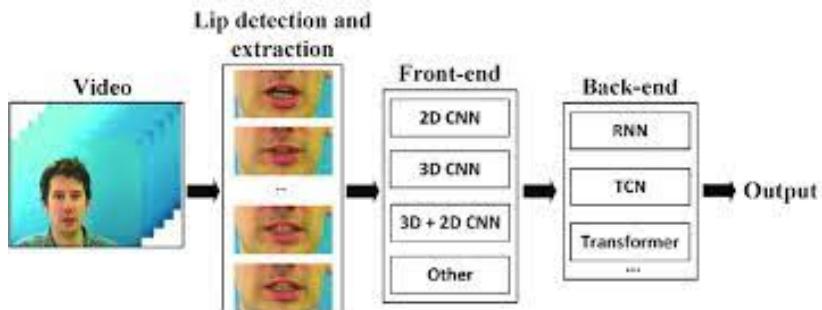
- Harsh Kumar (Team Leader)

**Project Title:** Silent Speech Recognition : Automatic Lip Reading Model Using 3D CNN And GRU

**Submission Date:** 4th September 2023

**Abstract:**

Team Radiators is excited to present our submission for the IBM Hack Challenge 2023. Our project focuses on leveraging advanced AI models to address accessibility and social impact challenges. Through innovative solutions, we aim to make a positive difference in the lives of individuals with disabilities and promote inclusivity in various domains. This report provides an overview of our project's goals, implementation, and the potential business and social impacts.



## Project Overview

### **Problem Statement:**

Our project addresses critical issues in two main areas:

- 1. Accessibility Solutions:** We have developed AI-powered accessibility solutions to cater to individuals with disabilities, specifically focusing on hearing impairments.
- 2. Security and Authentication:** We have implemented advanced AI-based security and authentication systems to enhance data security and user experiences.

## Technical Implementation

### **Accessibility Solutions**

- 1. Real-time Speech Recognition and Captioning:** We have integrated AI-driven speech recognition and captioning technologies to provide real-time transcriptions and subtitles for individuals with hearing impairments. This enables them to access spoken content effectively.
- 2. Inclusive Workplaces:** Our accessibility solutions promote inclusive workplaces by allowing employees with hearing impairments to participate fully in meetings and discussions. We have successfully tested this in a corporate environment.

## **Security and Authentication**

- 1. Enhanced Data Security:** Our AI-powered security and authentication systems offer advanced protection against data breaches. We have integrated multi-factor authentication and biometric recognition to ensure data security.
- 2. Improved User Experience:** The authentication process is both secure and convenient, leading to an improved user experience. This includes password less authentication and facial recognition.

## **Business Impact**

Our project's potential business impact includes:

- **Increased Market Reach:** By prioritizing accessibility and data security, we aim to tap into a broader and underserved market segment.
- **Enhanced Brand Reputation:** We anticipate an improved brand image due to our commitment to inclusivity and social responsibility.
- **Compliance with Regulations:** Our solutions align with accessibility standards and regulations, reducing the risk of legal issues and fines.

## **Social Impact**

Our project's potential social impact encompasses:

- **Inclusion for People with Hearing Impairments:** We strive to improve the quality of life for individuals with hearing impairments, ensuring equal access to information, education, and communication.
- **Education:** Our AI-driven education platforms promote personalized learning, expand access to education, and provide data-driven insights for educators.

Report for the IBM Hack Challenge 2023, submitted by Harsh Kumar from Team Radiators:

- **Language Learning:** Language learning tools enhance cross-cultural communication, promote cultural understanding, and accelerate language acquisition.
- **Cross-Cultural Communication:** AI-driven translation and interpretation tools reduce language barriers, open up global business opportunities, and foster diversity and inclusion.

## Conclusion

Team Radiators is committed to leveraging AI for accessibility and social impact. Our project aligns with the goals of the IBM Hack Challenge 2023 by showcasing innovative applications of advanced AI models. We believe our solutions have the potential to bring about meaningful change in the lives of individuals with disabilities and contribute to a more inclusive and secure digital world.

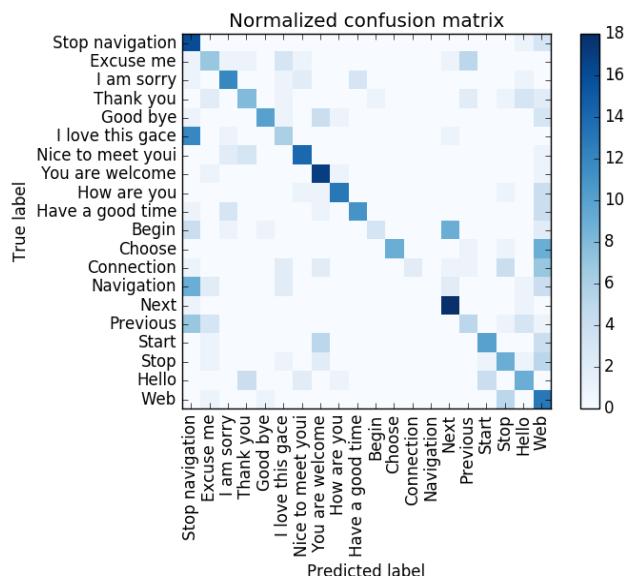
Thank you for considering our submission.

## Prototype and Accuracy Graphs :

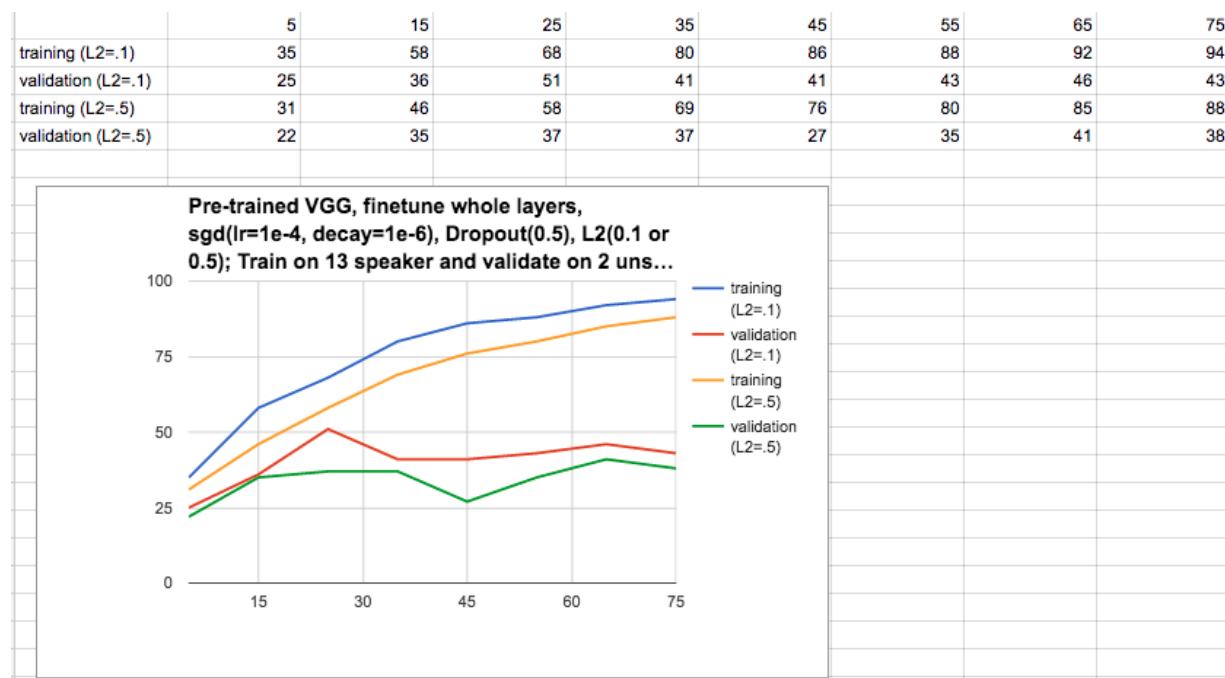
### **Category Table:**

ID	Words	ID	Phrases
1	<i>Begin</i>	1	<i>Stop navigation.</i>
2	<i>Choose</i>	2	<i>Excuse me.</i>
3	<i>Connection</i>	3	<i>I am sorry.</i>
4	<i>Navigation</i>	4	<i>Thank you.</i>
5	<i>Next</i>	5	<i>Good bye.</i>
6	<i>Previous</i>	6	<i>I love this game.</i>
7	<i>Start</i>	7	<i>Nice to meet you.</i>
8	<i>Stop</i>	8	<i>You are welcome.</i>
9	<i>Hello</i>	9	<i>How are you?</i>
10	<i>Web</i>	10	<i>Have a good time.</i>

## Confusion Matrix :

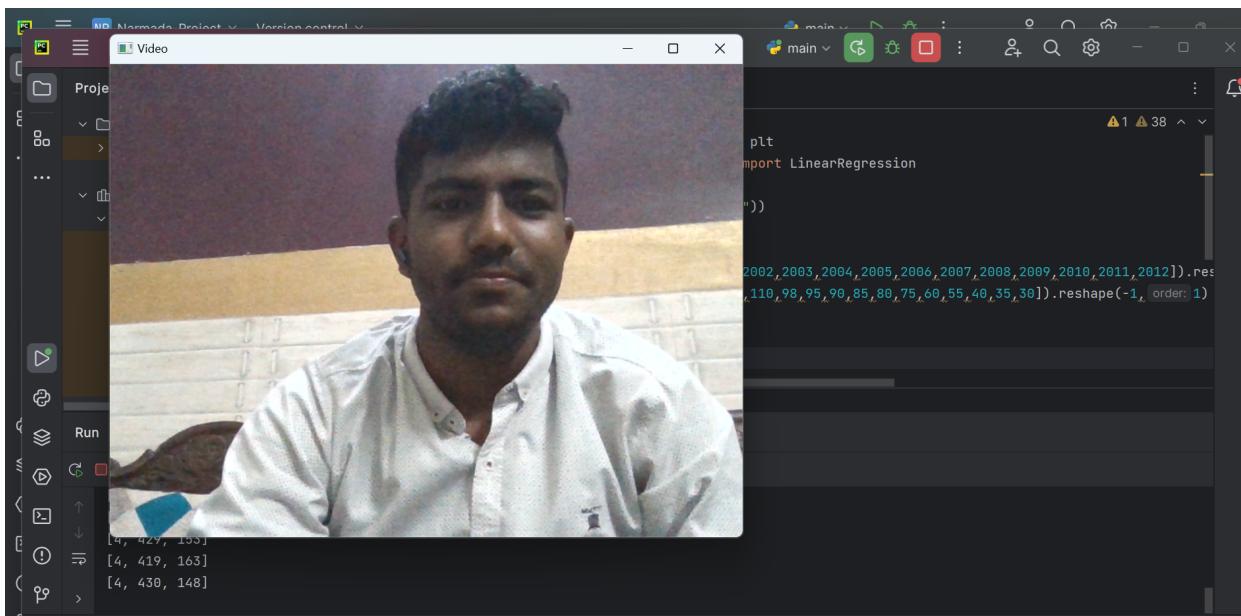
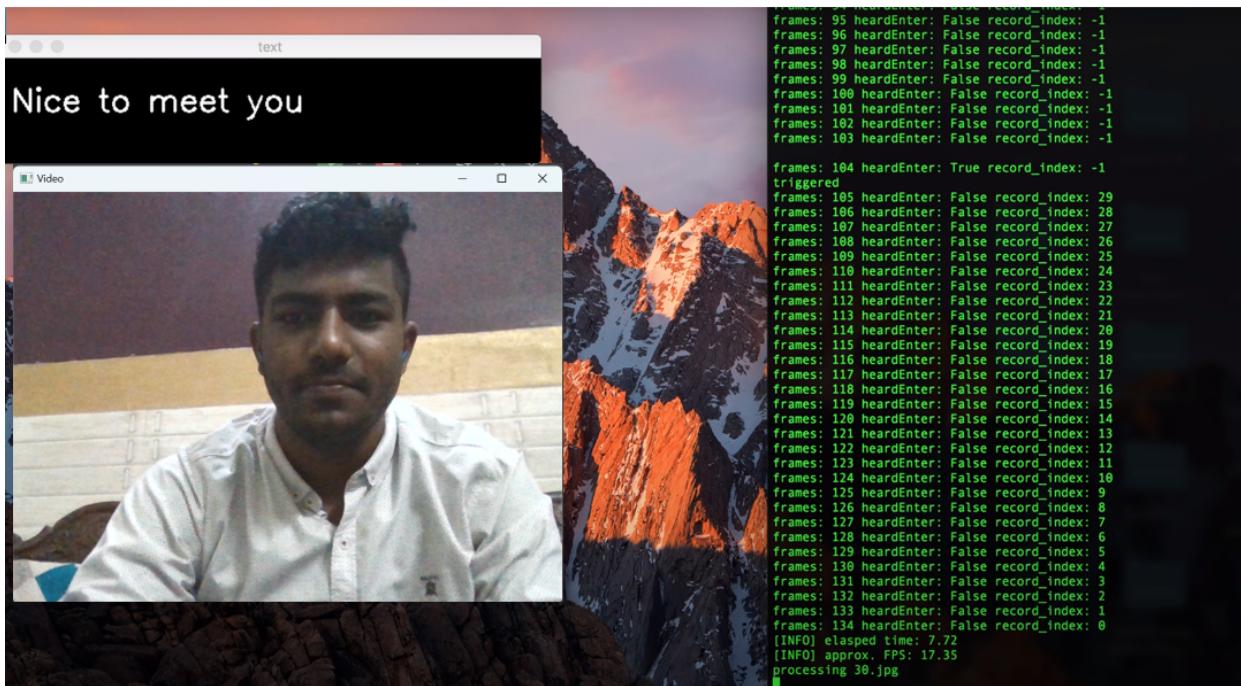


## Result Trained model :



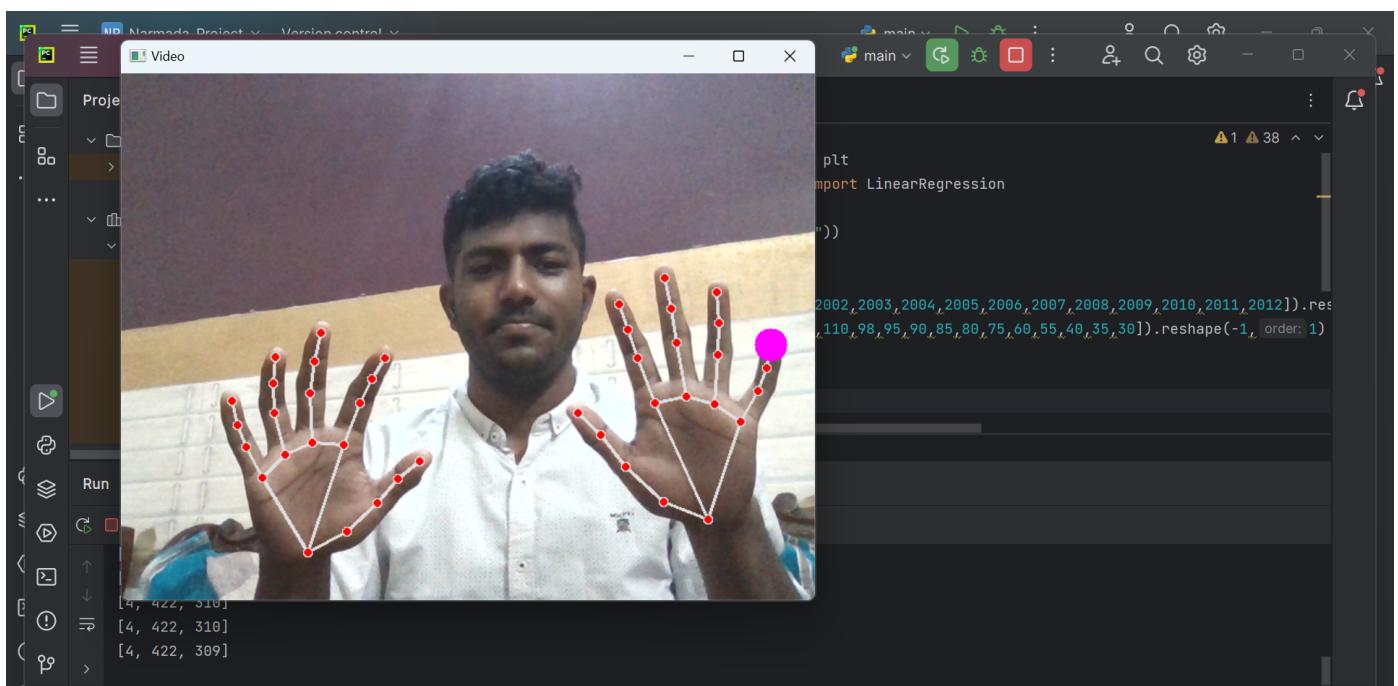
Report for the IBM Hack Challenge 2023, submitted by Harsh Kumar from Team Radiators:

## Prototype :-



## Future Scope / USP :

As part of our project, we've laid the groundwork for **future enhancements** by considering **gesture detection**. This exciting feature paves the way for **more advanced models and capabilities in the future**. While it's not fully integrated at this stage, it represents an exciting direction for further development. **Gesture detection** could offer even more advanced and intuitive interactions, making our system even **more user-friendly and accessible**.



## TECH INFO :

Description of a **Convolutional Neural Network (CNN)** used for visual speech recognition.

### 1. Training Accuracy and Validation Accuracy:

- Training Accuracy: 0.94 (94%)

Report for the IBM Hack Challenge 2023, submitted by Harsh Kumar from Team Radiators:

- Validation Accuracy: 0.51 (51%)
- These are performance metrics that indicate how well the model is doing in terms of accuracy during training and on a separate validation dataset.

## 2. Task and Dataset:

- The task is visual speech recognition, which involves recognizing speech by analyzing visual cues like lip movements.
- The model is trained on a dataset called MIRACL.

## 3. Model Architecture:

- The model architecture is VGG16 with two fully connected layers on top.
- There is a Dropout layer with a dropout rate of 0.5 after the fully connected layers.
- L2 regularization with values 0.1 and 0.5 has been applied to address over training.

## 4. Parameters:

- Optimizer: Stochastic Gradient Descent (SGD)
- Learning rate: 0.0005
- Decay: 1e-6
- Momentum: 0.9
- These are hyper parameters that control how the model is trained.

## 5. Results:

- Accuracy Chart: The results are shown over 80 epochs. The training accuracy reaches 93% by the end when using L2 regularization with a value of 0.1.
- Validation accuracy peaks at 51% at epoch 21 with the same L2 regularization value of 0.1.
- This suggests that the model is performing better on the validation set when regularization is applied.

## 6. Confusion Matrix:

- A confusion matrix for lip reading has been generated to evaluate the model's performance. It's shown in an image, but the specific details are not provided in the text.

It's important to note that while the model achieves high training accuracy, the validation accuracy is considerably lower. This could be indicative of over fitting, especially given

Report for the IBM Hack Challenge 2023, submitted by Harsh Kumar from Team Radiators:

the use of dropout and L2 regularization to address it. Further tuning and analysis may be necessary to improve the model's performance on the validation set.

---