International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

Personalized Smart Mirror with User Detection

Deep Mehta[1], Mithil Jain[2], Preet Thakkar[3]

Abstract - *It is a growing need for one to organize one's day in an efficient manner. One of the ways is to use a smartphone which is proving to be more destructive than productive. In order to reduce screen time on a smartphone, we can use a daily use object, a mirror, to help organize the day while not having to reach out to your smartphone for basic needs. The mirror can be used as a smart device that can offer services like displaying weather, date and time, latest updates of news headlines, user personalized daily tasks and reminders, some motivational quotes to get the day going positively with voice assistance. This use of the mirror can help the user save time and multitask by planning their day. The smart device uses Raspberry Pi as the central controller to dictate the actions to perform. When a user stands in front of this smart mirror, the controller recognizes the person and it displays the tasks that a particular user has to perform. It can also be used to drop messages for other users registered under the same smart system. The mirror allows to command the mirror using voice and choose what is to be done, e.g., Display the user tasks or display messages. The smart device comes with a mobile application that enables other users to check their tasks or messages when the mirror is unreachable.*

***Key Words*: Smart Mirror, Ambient Home, Smart Home, Personalized Assistant, Raspberry Pi**

I. INTRODUCTION

We live in a world where we rely heavily on our smartphones for our daily activities. We rely on our smartphones for almost every task to be performed. While these devices are designed to facilitate our schedules efficiently, they often tend to do more harm due to their excess usage. Teens are highly affected due to the adversity of these devices. A Smart Mirror is an approach to reduce our dependency on mobiles devices for trivial tasks that can be accomplished rather effectively. A Smart Mirror is an extension to a traditional mirror. It consists of a two-way mirror with an electronic display behind the glass. In today's world, we can fairly assume that everyone uses a mirror to groom themselves. Some may use it once or twice a day while some watch themselves in it for hours. This mirror can be used to display a wide range of information —weather, time, date, and news updates— in the form of widgets. This product allows busy individuals to multitask and stay informed on the go. The goal of this mirror is to reduce the need for constantly pulling out a device and thereby reducing the unnecessary screen-time on mobile phones. The proposed device can help to conquer the day and utilize the mirror to receive all the information one needs to start their day.

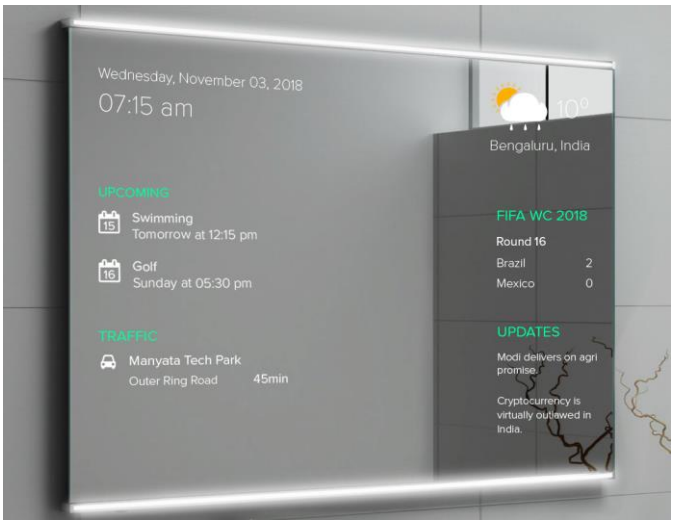

Fig - 1: Smart Mirror Example [1]

The Smart Mirror will be powered with Machine Learning models that will enable User Recognition using their Facial [2] and Speech inputs. This will allow the mirror to work without having the user to explicitly specify their identity. Every mirror will provide a personalized experience based on the user's specified interests. The main purpose of technology is to make our life easier and focus on what it's really important and the proposed mirror does that effectively.

II. SIMILAR WORKS

In "Design of the Smart Mirror Based on Raspberry Pi" [3], the project attempts to include a mirror into the list of smart devices for a smart home. The Smart Mirror, based on traditional household mirror comes along with home automation control, environmental monitoring, information acquisition and remote operation. The paper points out that the lack of progress in Smart Mirrors is due to the less compatibility and integration along with other smart devices. Built on Raspberry Pi, the Mirror is designed to be suitable with various smart home functions including security. The Mirror consists of various modules including but not limited to: Image Processing, Voice Conversion, Information Display, Speech Communication, Security Monitoring, Home Appliance Control and Environment monitoring. The Mirror is fit into an infrared frame that is followed by a unidirectional mirror along with a display screen at its back.

In "IoT Based Home Automation by Using Personal Assistant" [4], the device built is aimed at the development of a personal assistant that will help to make the life of the user easier than before. It uses the IoT devices and uses gesture control and speech recognition to initiate the function of the

smart mirror. A Raspberry Pi is used, which acts as the brain of the smart mirror and perform various functions such as collecting the data, displaying the output, accepting the input, etc. IoT - The Internet of Things (IoT) can be described as a network of physical objects or "things" embedded with software, electronics, sensors and network connectivity that allows these objects to communicate with each other and exchange digital data. It was observed that personal assistants like Amazon Echo failed to provide the user with a sense of control as it remained unresponsive when given with an invalid voice command.

As stated in "Automatic Speech Recognition" [5], speech-to-text refers to the conversion of the input speech into another format e.g., text and then react appropriately. Speech recognition starts with creating an utterance with sound waves. These sound waves are converted to digital signals. This signal is then converted to a series of feature vectors which contains only relevant information. For feature extraction, it is important to remove irrelevant information. Finally, the best match is found from the knowledge base for a particular feature vector. A Decoder is a model that decides recognition of a Speech utterance using the combination of the information conveyed by the acoustic and language model.

In "A Simple but Efficient Real-Time Voice Activity Detection Algorithm" [6], the paper focuses on Voice Activity Detection (VAD), a major process of Speech Detection. VAD is the process of detecting whether the incoming sound consists of a voice or is only filled with noise. It also tries to reduce the amount of noise from the input using multiple features. The paper proposes the use of short-term features like Spectral Flatness and Short-Term Energy to detect a voice activity. The paper covers a brief history of previous attempts and possible but computationally heavy solutions for noise removal. Later, SFM is discussed and various samples of noisy and clean corpora are tested with the proposed method. The resulting accuracy shows and improvement of 9.26% compared to the current standard G.729B.

The paper "A Smart Mirror to Monitor Children by Using Raspberry Pi Technology" [7] features having a parental control to monitor and send daily schedules (to-do lists) to their children. It also includes a feedback system where the children can send updates on the work. Based on the idea, we propose the addition of messages and sticky notes for users registered under the system.

In "Smart Mirror a Novel framework for interactive display" [8] Spending a lot of time in front of the mirror, it can be used for some productive stuff or to realize some day-to-day essential tasks. It would be helpful if your mirror could detect you and let you know that you have an important business meeting at 4 pm today or if the mirror could tell you that it's cold outside and then recommend you to wear a sweater. An interactive mirror or a glass table can be used for this purpose. Infrared sensors can be used to detect the presence of somebody in front of the mirror or someone sitting beside the table.

III. Problem Description

A. Problem Statement

The goal of this project is to convert a traditional mirror into a futuristic Smart Mirror. We will create a smart mirror that provides daily updates to an individual along with other features aimed at assisting personalized schedule management and communication within a family. The mirror will be powered by various Artificial Intelligence models that will allow enhancing the experience of the users. Furthermore, integration of user recognition will open up a lot of doors for future scope in the mirror.

B. Objectives

- Provide an alternative to the functionalities of mobile devices.
- Enhance daily life by assisting the users to multitask and provide on-the-go updates.
- Recognizing user on a low-power device such as Raspberry Pi.
- Build a prototype to enable future developments on a mirror.

IV. Raspberry Pi

The Raspberry Pi 4B is a small computer with a powerful Quad core Cortex-A72 64-bit processor and 4GB LPDDR4 SDRAM powered via 5V DC USB-C connector. It can plug into any monitor or TV. It supports output display connection of upto 4K. With standard mouse and keyboard, it can be used like a mini desktop. It has support for wireless and/or bluetooth connection. To load the Operating System and for data storage, the raspberry pi uses a micro-sd card slot. The device is powerful to support the processing of various python modules [9].

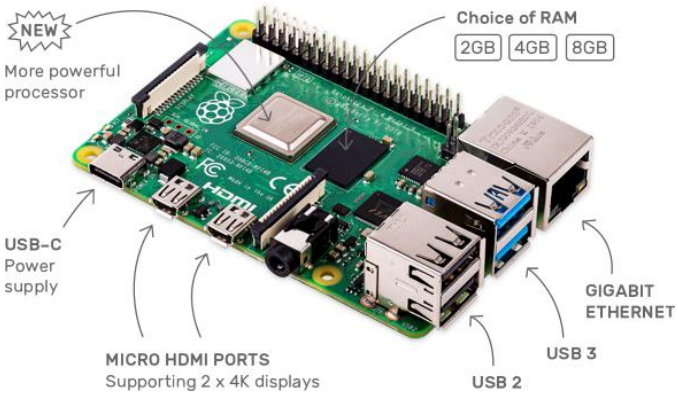

Fig - 2: Raspberry Pi [10]

V. METHODOLOGY

We propose an assistant-based solution to build the mirror. The mirror will provide all the basic information required through voice-based inputs. All the inputs will be captured through devices, processed using Natural Language Processing to extract the meaning and accordingly call the corresponding algorithms to generate the output or perform an action.

A. System Flow

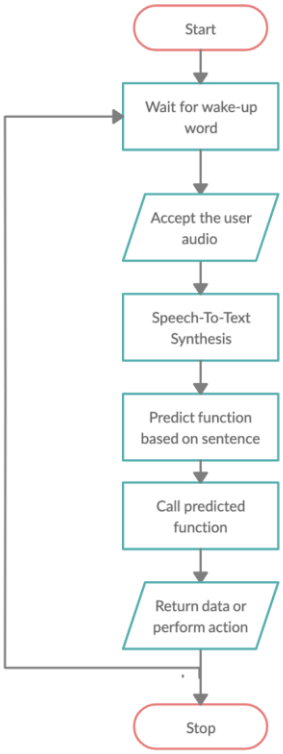

Fig - 3: Flow Chart for Smart Mirror

- **Input**: The main input trigger is the user's voice. The mirror will remain on stand-by waiting for the statement "Hey Cassie". When the mirror recognizes the statement, it'll initialize the corresponding modules.
- **Sentence Processing**: Every statement followed after "Hey Cassie" will be temporarily stored and mapped to available functions. Once the match is found, the required modules will be loaded and the function will be set for execution.
- **Execution**: The modules will be executed in specific orders. In case the module requires user recognition, the mirror will try to detect the face using a camera. If the face is not found or detected the mirror will perform Speaker Verification. Moreover, for sensitive information, the mirror will compulsorily perform the verification of the face as well as speech.
- **Output**: The corresponding response or the output will be provided by the mirror. The output can be in the format of Speech through speakers or on-screen GUI display.

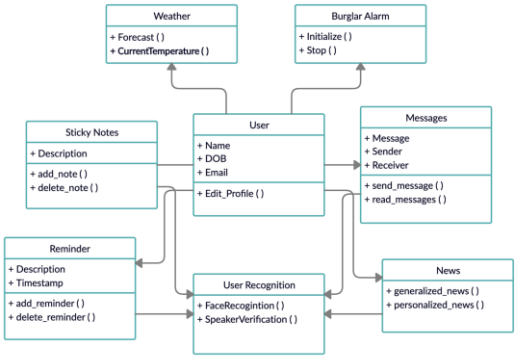

Fig - 4: Class Diagram for Smart Mirror

B. Framework

- **OpenCV**: Open Source Computer Vision is a widely used computer vision package used in Python for real-time facial detection. It is used to detect the presence of faces and pre-process it before passing as an input to the Machine Learning Model for recognizing the user. It acts as an interface to capture frames from the connected camera.
- **Scikit-Learn**: Scikit-Learn is a simple and efficient tool for building lightweight Machine Learning Models. It is used to build a model for sentence-to-function mapping.
- **Django**: Django is a Free, Open-source, High-level Python Web Framework that helps in rapid development and clean pragmatic design. It also provides an optional administrative create, update, read and delete interface. The Django application is hosted as a PWA (Progressive Web App) to allow installation on multiple devices like Windows, Mac, iOS, Android, etc.
- **Raspbian**: Raspbian is a Debian-based computer operating system for Raspberry Pi [7]. It is used as a platform for all the software executions related to the mirror.
- **MongoDB**: MongoDB is a NoSQL database server that allows to store JSON-like collections. We use MongoDB to store all the user details required for personalization.

C. Algorithms

- **Multinomial Näıve Bayes**: Naive Bayes is a family of algorithms based on applying Bayes theorem with a strong (naive) assumption, that every feature

is independent of the others, to predict the category of a given sample [11]. MultinomialNB provides an efficient way to create a very light model for mapping sentences to functions.

VI. Facial Recognition

Facial Recognition is one of the most crucial part in the mirror as all the personalized features rely on accurate user recognition. Any inaccurate prediction will mean a threat to security, confidentiality and provide a bitter user experience. Thus, the accuracy of the Face Recognition Model was the at most priority for the mirror. Moreover, a user may use personalised features multiple times in a day, which demands for fast user detection. To ensure these criteria, various Facial Recognition models were tested before selecting the one to integrate with the mirror.

A. OpenCV Face Detection

OpenCV's Deep Neural Network-Based Face Detector relies on SSD (Single Shot Detector) [12] with ResNet-10 as backbone [13]. It is a pre-trained model used to detect faces. Compared to haar cascade, this model provides more accurate results and works with faces of different alignment: front, left, right, up, down and side faces. It is picked over its alternative, yoloface [14] (provides better accuracy than OpenCV DNN), as it is a very light model and can be loaded efficiently into memory while working with RPi. This model is used to capture images while registration of the user as well as during the Face Recognition to capture the test image.

B. Experimental Models

- **KNN**: Facial Recognition using KNN (K-Nearest Neighbors) is common for applications with a smaller number of classes. So as a starting point, a KNN model was implemented and it did not perform well with an increase in the number of users [15]. Moreover, with every registration of a new user, the model would take a couple of minutes to train. The lower accuracy in the detections made the mirror susceptible to false positives, and depending upon the usage, it could possess a high security threat.
- **SVM**: In an attempt to overcome the accuracy drawbacks of KNN, an SVM (Support Vector Machine) model was later implemented [16]. The SVM provided better results compared to KNN; however, the accuracy was still not up to the mark as it would provide wrong predictions very often.
- **CNN Based Recognition**: A CNN (Convolution Neural Network) model takes even longer to predict an image, but the accuracy was the greatest compared to other models [17]. However, after testing rigorously, it was found that the accuracy was highly affected by the lighting conditions of the test environment. Also, raining a CNN model on RPi is a high resource-oriented task that discourages on-the-go setups.

C. One-Shot Training

A major drawback of the previous is the amount of data required to train along with the time to register a new user. One-Shot Training overcomes this drawback by using an embedding of single image to compare with available embeddings [18]. A Siamese Network then tries to find the similarity between the test image and each train image.

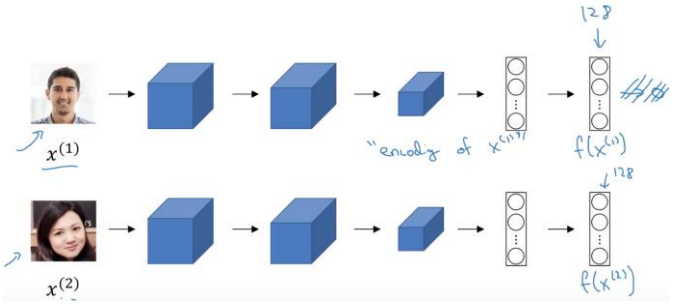

Fig - 5: One-Shot Training [19]

- **FaceNet**: FaceNet is an approach to generate embedding for a given facial image. [20] It converts an image into a vector of features. A FaceNet model learns from the mapping of the faces to Euclidean Space. With this metric, similarities or dissimilarities can be compared between any 2 given images.
- **Siamese Network**: Once the vectors of the images are generated, a Siamese Network then finds similarity between each image from the train set and the test image. This quantifies the differences between the images into numeric values.

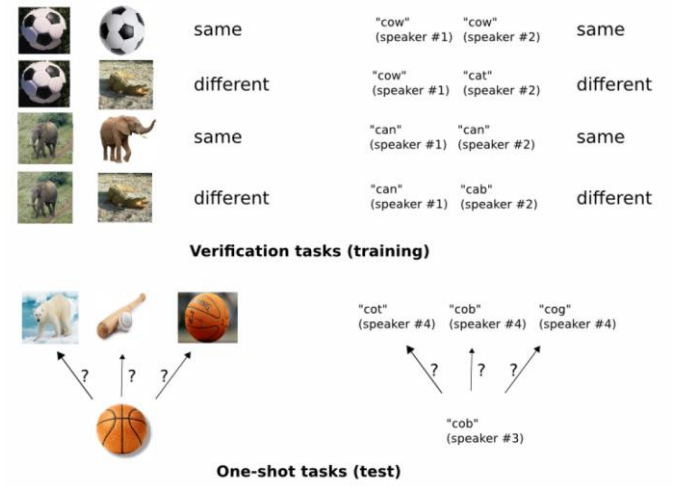

Fig - 6: Siamese Network [21]

- **Binary Classification**: Instead of searching for "Who the person in the image is?", One-Shot Training tries to answer the question of "Do these 2 images belong to the same person?". That is, the network compares the new test image with every image present in the train set. The prediction results can be provided by either the person with the lowest difference to test image or the person with a similarity above a given threshold (avoids detecting random person) or a combination of both.

VII. SPEECH RECOGNITION

A. Text-To-Speech

Text-To-Speech Synthesis deals with the problem of synthesizing a text format into its equivalent audio format. [22] In recent years, there has been a tremendous development in this field and we have achieved the synthesis of English language very efficiently. These tasks can be achieved by various models that are provided as open-sourced projects or libraries. For this Smart Mirror, we have used Python's pyttsx3 library for TTS Synthesis.

B. Speech-To-Text

As opposed to TTS, Speech-To-Text (STT) tries to predict the text present from given audio. This is done by training models with a large amount of audio data along with its text equivalent. [23] The process can be divided into 2 parts:

- Capturing the audio from an input device, e.g microphone
- Interpreting the recorded audio to find it's text

Such models require very huge data. Some of the models provide APIs (free as well as premium) that accept the sequential audio data in binary format and returns the text present in the given audio. Some of the models include but are not limited to: Speech Recognition, google-cloud-speech, apiai, wit. For this Smart Mirror, we use Speech Recognition that acts as a wrapper to Google's recognize [24]. Along with Google, it provides other options like CMU Sphinx, Microsoft Bing, IBM Speech to Text. However, for an Indian accent, Google's Web Speech API works very well compared to the alternatives.

Another alternative to using APIs is using Mozilla's DeepSpeech [25] that provides an offline standalone STT model. The limitation is that the pre-trained model doesn't quite work well with Indian Accents; however, if it supports your accent, that's a great tool to consider as a replacement for APIs.

C. Wake-Up Word Detection

With STT integrated into the mirror, it is neither feasible nor secure to send every audio data to Google's API to be converted back to text. For offline models (DeepSpeech), running STT in background would only increase the utilization of Raspberry Pi causing it to overheat in short period of time. Wake-up words are words that can be used to initiate a particular program or script. [26] Few common examples of Wake-up words are "Hey Siri", "OK Google", "Hey Cortana", "Alexa", etc. Snowboy is a library that provides the ability to listen for such wake-up words or hotwords. When these words occur in the audio, it will record the following sentence and call Google's web speech API (or invoke DeepSpeech) to convert to its equivalent text. Since the library works offline, it reduces the usage of an internet connection.

D. Speaker Verification

Speaker Verification works as an alternative to Face Recognition. In case the Face Recognition fails to detect a user with the required threshold or if the user is not in front of the mirror (a user may intend to use the mirror in situations when they are not right in front of the mirror), a Speaker Verification model can be used to detect the user. The Smart Mirror is integrated with Microsoft's Text-Dependent Speaker Verification API [27] with the mirror. A text-dependent verification relies on a single sentence to detect accurately. However, for a limited set of sentences and less number of users to detect from, it is found to provide very accurate results for every sentence detected by the mirror. One may consider building their own Text-Independent Speaker Verification; however, it is outside the scope of this paper.

VIII. DESIGN

A. Smart Mirror

The mirror is based on a GUI with a black theme and bright font colors. Any display screen can be connected to the Raspberry Pi and the display can be fit into a frame, preferably wooden. To increase the size of the mirror, a black colored paint can be applied on the region outside the screen on the frame. The whole frame will have a thin acrylic sheet (or expensive glasses designed to serve 2-way displays) that will work as a two-way mirror. When the display is turned off, the sheet will reflect every light and act as an ordinary mirror. When the display is turned on, the bright font will show up on the mirror and the black display will still act as the mirror providing an ambient feel to the user.

The primary mode of interaction with the system is the visual interface of the mirror. The smart mirror provides a user specific experience to every member by providing

personalized information to the respective user. The general features include a dashboard with news and weather updates. The user-specific features are listed as follows:

- **Personalised News Feeds**: Based on the news feed preferences selected by the user, personalized news headlines are displayed on the feed.
- **Dashboard**: The dashboard provides a brief overview of the user's information. In case the user has any new unread messages or an upcoming reminder, the dashboard briefly notifies the user about the respective information.
- **Messages / Reminders / Tasks**: When asked for, the mirror can provide a list of user's messages, reminders and tasks on the graphical interface.

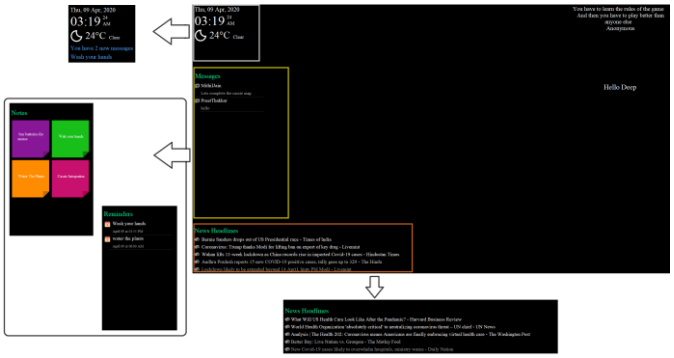

Fig - 7: Smart Mirror GUI

B. Mobile Application

To remotely access the features of the smart mirror, an easy-to-use application is provided to stay in touch while the user is away from home. The application provides a user with various features like:

- **Messaging**: Allows the user to check for new messages or send messages to other users registered under the system.
- **Reminders**: Helps the user to add or delete reminders as per convenience. These reminders are displayed on the smart mirror.
- **Sticky Notes**: It enables users to save quick notes or important text on the fly.
- **Additional Features**: Allows the user to manage News Feed Preferences for the Smart Mirror and toggle the Burglar Alarm

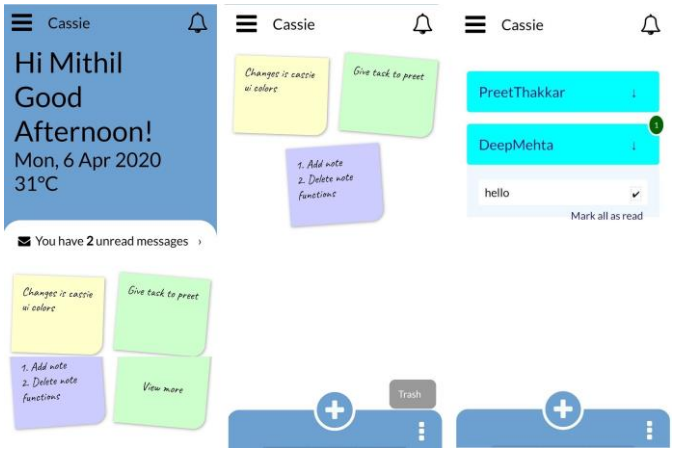

Fig - 8: Mobile Application GUI

IX. FEATURES

A. Burglar Alarm

The mirror can act as a theft surveillance system to closely monitor the house using the cameras connected to the system. When toggled on, the system is set to start recording a video clip as soon as any motion is detected in any of the connected cameras. In case of motion, the users are alerted about the unwanted motion using the recorded video clip that is immediately sent to all users via mail which acts as an alert to all the users.

B. Database Operations

Allowing users to sync their data across the mobile application and the Smart Mirror, the database holds the following data:

- **User Management**: Maintains a list of user and their profile preferences that include their basic details, personalised news feed preferences and login credentials.
- **Messages**: It stores the messages sent or received to a user. The receiving user gets a push notification on their device.
- **Reminders**: To store reminders, a user can select the timestamp for the reminder. As per the user's choice, the reminder can be set to repeat daily. Based on the date and time of all the reminders, the upcoming reminder is displayed to the user on the dashboard section of the smart mirror.
- **Sticky Notes**: Stores the user's plain-text notes that can be used as sticky notes.

C. Requirements

- **Hardware Requirements**
  - Raspberry Pi 3/4 (2GB RAM or higher)
  - Raspberry Pi Camera or Webcam (5MP or higher)
  - RPi compatible Microphone and Speaker
  - Working internet connection (Wi-Fi/ Ethernet)
- **Software Requirements**
  - Raspbian OS
  - Python 3.5 or higher
  - MongoDB Server

X. Conclusion

A Smart Mirror was developed using Raspberry Pi along with accessories. The mirror includes multiple modules such as Facial Recognition, Speech Recognition and personalized User-based experience. The mirror was developed considering family usage where the mirror can act as a device to help everyone in the family manage their daily routines. Other features like reminders, sticky notes and messaging help to stay updated and communicate easily. The mirror consists of hardware devices like a camera, microphone and a speaker. The mirror will prove to reduce our dependence on mobile devices by providing better features in an easy to access manner. Having an online database server allows us to have a centralised database which can be tracked and used for future development. Moreover, it makes it easier to build applications for other platforms. With the addition of any camera for surveillance, the Mirror provides a solution towards the security aspect of the home. Thus, we developed a Smart Mirror with various Machine Learning models using IoT that can be further developed and used in a wide range of industries.

XI. Future Scope

A Smart Mirror was developed using Raspberry Pi along with accessories. The mirror includes multiple modules such as Facial Recognition, Speech Recognition and personalized User-based experience. The mirror was developed considering family usage where the mirror can act as a device to help everyone in the family manage their daily routines. Other features like reminders, sticky notes and messaging help to stay updated and communicate easily. The mirror consists of hardware devices like a camera, microphone and a speaker. The mirror will prove to reduce our dependence on mobile devices by providing better features in an easy to access manner. Having an online database server allows us to have a centralised database which can be tracked and used for future development. Moreover, it makes it easier to build applications for other platforms. With the addition of any camera for surveillance, the Mirror provides a solution towards the security aspect of the home. Thus, we developed a Smart Mirror with various Machine Learning models using IoT that can be further developed and used in a wide range of industries.

XII. References

[1] Building an iot magic mirror with hosted web apps and windows 10.
https://blogs.windows.com/msedgedev/2016/05/31/magic-mirrorhosted-web-app/.

[2] Ivette Garcia, Eduardo Salmon, Rosario Riega, and Alfredo Barrientos. Implementation and customization of a smart mirror through a facial recognition authentication and a personalized news recommendation algorithm. pages 35–39, 12 2017.

[3] Yong Sun, Liqing Geng, and Ke Dan. Design of smart mirror based on raspberry pi. pages 77–80, 01 2018.

[4] Venkataraman Chayapathy, Dr anitha g.s, and B Sharath. Iot based home automation by using personal assistant. pages 385–389, 08 2017.

[5] Sal Benk, Youssef Elmir, and Abdeslem Dennai. A study on automatic speech recognition. 10:77–85, 08 2019.

[6] Mohammad Moattar and Mahdi Homayoonpoor. A simple but efficient real-time voice activity detection algorithm. European Signal Processing Conference, 12 2010.

[7] Raju Nadaf, Rubina M, Sujata P, and Vasudha Bonal. Smart mirror using raspberry pi for human monitoring and intrusion detection. pages 116–121, 07 2019.

[8] S. Athira, F. Francis, R. Raphel, N. S. Sachin, S. Porinchu, and S. Francis. Smart mirror: A novel framework for interactive display. In 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), pages 1–6, 2016.

[9] Raspberrypi 4 tech specs.
https://www.raspberrypi.org/products/raspberrypi-4-model-b/specifications/.

[10] Raspberry pi 4 model b.
https://www.raspberrypi.org/products/raspberrypi-4-model-b/.

[11] Shuo Xu, Yan Li, and Wang Zheng. Bayesian multinomial na¨ıve bayes classifier to text classification. pages 347–352, 05 2017.

[12] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander Berg. Ssd: Single shot multibox detector. volume 9905, pages 21–37, 10 2016.

[13] Face detection – opencv, dlib and deep learning (c++ / python).
https://www.learnopencv.com/face-detection-opencv-dlib-and-deeplearning-c-python/.

[14] Wang Yang and Zheng Jiachun. Real-time face detection based on yolo. pages 221–224, 07 2018.

[15] Prabin J, P. Poornima, and Kukkapalli Kumar. A novel method for color face recognition using knn classifier. 02 2012.

[16] Guodong Guo, Stan Li, and Kapluk Chan. Face recognition by support vector machines. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 02 1970.

[17] Musab Cos¸kun, Ays¸egul Uc¸ar, ¨ Ozal yıldırım, and Yakup Demir. Face ¨ recognition based on convolutional neural network. 11 2017.

[18] Nikhil Thakurdesai, Nikita Raut, and Anupam Tripathi. Face recognition using one-shot learning. International Journal of Computer Applications, 182:35–39, 10 2018.

[19] Neural networks - one shot learning.
https://www.youtube.com/watch?v=r8LLorRACPM.

[20] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. pages 815–823, 06 2015.

[21] Dattaraj Rao, Shruti Mittal, and S. Ritika. Siamese neural networks for one-shot detection of railway track switches. 12 2017.

[22] Ifeanyi Nwakanma, Ikenna Oluigbo, and Okpala Izunna. Text – to – speech synthesis (tts). International Journal of Research in Information Technology, Volume 2, Issue 5, May 2014, Pg: 154-163, 2:154–163, 03 2014.

[23] Prerana Das, Kakali Acharjee, Pranab Das, and Vijay Prasad. Voice recognition system: Speech-to-text. Journal of Applied and Fundamental Sciences, 1:2395–5562, 11 2015.

[24] The ultimate guide to speech recognition with python. https://realpython.com/python-speech-recognition/.

[25] Muhammad Firmansyah, Anand Paul, Deblina Bhattacharya, and Gul Urfa. A.i. based embedded speech to text using deepspeech, 02 2020.

[26] Veton Kepuska and T.B. Klein. A novel wake-up-word speech recogni- ¨ tion system, wake-up-word recognition task, technology and evaluation. Nonlinear Analysis, 71:2772–2789, 12 2009.

[27] Shi-Xiong Zhang, Zhuo Chen, Yong Zhao, Jinyu Li, and Yifan Gong. End-to-end attention based text-dependent speaker verification. 01 2017