

Virtual Mock Interview Assistant

-Sheradha Jauhari¹, Diksha Awal², Chetan Aggarwal³, Apoorv Gautam⁴

¹Assistant Professor, Computer Science Engineering, ABESIT, Uttar Pradesh, India

² Student, Computer Science Engineering, ABESIT, Uttar Pradesh, India

³ Student, Computer Science Engineering, ABESIT, Uttar Pradesh, India

⁴Student, Computer Science Engineering, ABESIT, Uttar Pradesh, India

Abstract— In this paper, we are exploring state-of-the-art models in multimodal emotion recognition. We choose to investigate video inputs and construct an ensemble model that collects data from all of these sources and presents it in a clear and understandable manner. The project's goal is to extract useful information from a variety of sources, including video inputs. The goal is to compile this data and present it in such a way that its meaning is apparent and understandable to others. The goal of this project is to help the user discover their strengths and flaws in order to prepare for an interview. The purpose is to help the user succeed in an interview by educating them about their personality and coaching them on the precise personality features and adjustments they need to succeed (of organisation of their choice.). The paper describes a method for creating artificial conversational agents and bots that employ actual human video. This approach differs from others in that it starts with real human examples and builds artificial behaviours from there, rather than artificial agents building the real interaction. In terms of behaviour and emotion, Video Bot might be considered a substitute for human beings; it is conceivable to create a video that behaves like a individual human for all the practical reasons. This method also reveals some of the difficulties encountered during the Video Bot's preparation for the interview.

Keywords: Emotion recognition · Video · Affective Computing · multimodal emotion recognition · interview assistant · affective computing

I. Introduction

As the population and knowledge of education has grown, so has the number of people seeking employment. As a result, businesses and educational institutions have begun interviewing potential employees and screening out the top applicants as needed. To easily get selected in these competitive interviews, a candidate needs have awareness of some ethics, etiquettes, and conventions, as well as personality skills.

Chatbots are conversational interfaces in software. Emotional emotion is a unique feature of humans that distinguishes them from machines. Hence to make chatbots more humanized, an emotional chatbot to analyse the user's sentiment is developed.

Affective computing is a study of Machine Learning and Computer Science that studies the recognition and the processing of human affects.

Virtual Interview Assistant is a Chatbot Based platform that allows users to interact with a Chat-bot and receive a full report on their interview performance, allowing them to better understand their personality and providing a means to improve for the real interview.

The multi-turn dialogue approach and emotive analysis are combined for constructing a chatbot that categorises statements entered in terms of emotional states and selects an appropriate response

Virtual mock interview assistant is a new field that tries to incorporate video into the process. The rise of this field has coincided with the growth of social networks, which have provided academics with access to large amounts of data. Recent research has looked into potential measures for measuring the coherence of emotions across different mediums. Depending on the input, we'll look at a few different categorization goals. Table 1 summarises the entire category targets we're assessing based on the data type.

Video	Happy, Sad, Neutral and Disgust , Angry, Fearful, Surprise.
-------	---

A. Computer vision for emotion recognition

Challenges such as emotion recognition can typically not be solved though classical machine learning techniques. All the recent research papers focus on several deep learning techniques, some of which include Artificial Neural Networks (ANN), Convolution Neural Network (CNN), Region-CNN (R-CNN), Fast R-CNN, Recurrent Neural Network (RNN) or Long Short-Term Memory (LSTM). The aim of the following section is to develop the bases that lead to Convolutional Neural Networks (CNN).

A. I .Architecture

1. Image processing

The alteration of digital images in order to extract more information than is evident on the original image is known as image processing. Picture processing is a method of improving or extracting information from an image by performing operations on it. It's a type of signal processing in which a picture serves as the input and the output is either that image or its characteristics/features.

2. Facial emotion detector

The practise of detecting human emotions from facial expressions is known as facial emotion recognition. The capture of face expressions is comparable to the capture of facial motions. It's a method of manipulating characters created by a computer using data from human faces or recognising a user's emotions via visual or mechanical means. Emotion Analysis seeks to discover and distinguish emotions such as anger, disgust, fear, pleasure, sadness, and surprise through the expression of words.

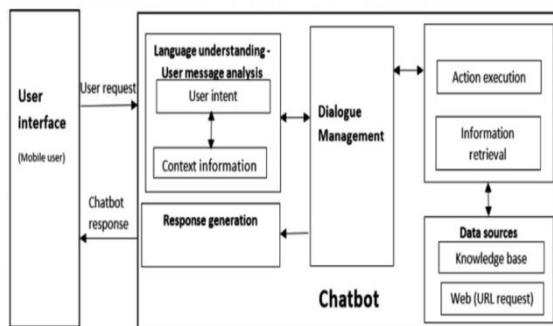


Figure: data flow diagram of VMIA

II. Proposed Methodology

A. Entity used

1. User: candidates own no accounts in the system. Candidates are provided with system accounts after authentication of email id cards by database. Candidates can then use their account to log into the system during the preparation. A list of candidates who run for election (including their personal information) is stored in the system database.

2. Virtual video bot: virtual assistant mock interviewer is a virtual interviewer in which user interacts with the chat bot and practices for the upcoming interviewers

B. Phases

1. User choose the question according the question no. which available on the screen after then video is start.
2. Video captures the facial expression on the basis of sections.
3. Recording stores the facial expression and on the basis of face capture it will show the result like: How you behave in the particular question so that user can easily identifies his/her strength on particular question.
4. Artificial intelligence based screen check: here comes the assistance asking the question.
5. Everything is set to go: the chatbot window begins and the mock interviewer is ready to begin with the learning procedure.
6. Project report: At the conclusion of the interview, the candidate will be able to check his or her score and determine whether or not he has to prepare further. Essentially, the candidate is given the opportunity to develop himself in preparation for subsequent interviews.

C. System design

The methods involved in creating the application for vmia, namely the architecture, its various components, and the data that passes through the system, are described in depth in this chapter.. In addition, the minimum hardware requirements, target users, and software requirements will also be discussed.

1. Requirements that are both functional and non-functional. The virtual mock interview phase and the virtual mock interview assistant administrators' phase are the two phases of the system's functional requirements. They will be able to chat with an assistant, request interview preparations, request some of the most important questions to be asked in the interview, have a visual discussion with the assistant, and receive the result of the discussion in the form of a report card during the virtual mock interview phase. They will be able to log in via the administrators' portal, update the chatbot database with the

current qualification, and receive a list of questions during the virtual mock interview assistance administrators phase.

The following are examples of non-functional requirements:

- (1) security: unauthorised users will not be allowed access to the system
- (2) usability: the proposed system is simple to operate, enter data, and interpret the output for the user.
- (3) Scalability: the system should always run well, regardless of any modifications.
- (4) Browser compatibility: the proposed approach works with all web browsers.

2. Hardware requirements are minimal. The minimal hardware requirements pertain to the physical characteristics of the computer that are necessary to run the virtual mock interview helper. The specifications are as follows: at least 250 GB hard drive, 4 GB RAM, and an Intel Pentium dual-core processor.

3. Software specifications. The virtual mock interview helper requires the following computer programmes and procedures to be implemented.

Table 1 indicates the minimum software requirements.

Requirements	Software
Operating system	Microsoft windows
Dbms	Excel
Programming language used	Python
Python Libraries Used	Numpy, Keras, Tensorflow, Xlsxwriter, OpenCv
Development tool	Visual studio code /Jupyter

- 4. Candidates preparing for interviews at different companies for recruitment (placement) are the key target users of the vmia, notably students who are looking for a career after completing any study. Furthermore, because students are the most frequent interviewees during placement activities, the development of a virtual help for candidates will be welcomed.

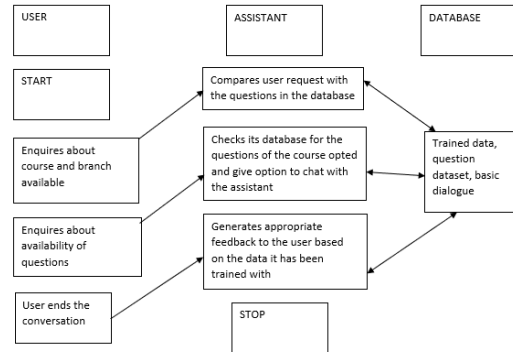


Figure: activity diagram of VMIA

III. Technology Used

1. Numpy

Numpy is a Python library that adds support for multi-dimensional arrays, large and matrices, as well as a large collection of high-level mathematical functions for array operations.

2. OpenCV

OpenCV contrib is a Python programming language specialised module that is required for the system to run SURF feature descriptions alongside the OpenCV module in the open-source library.

3. Keras

Keras is a Python interface for artificial neural networks that is open-source software. Keras serves as a front-end for the TensorFlow library.

4. TensorFlow

TensorFlow is a free and open-source machine learning and artificial intelligence software library.

5. XlsxWriter

XlsxWriter is a Python module that allows you to write files in the XLSX file format.

6. Flask

Flask is a Python-based micro web framework. It is classified as a microframework because it does not necessitate the use of any specific tools or libraries. It lacks a database abstraction layer, form validation, or any other components that rely on pre-existing third-party libraries to perform common functions.

IV. Conclusion

A virtual assistant mock interviewer is one of the simplest ways to convey data from a computer without having to think of appropriate keywords for a student to search or browse multiple web pages to gather material for an interview.

A virtual assistant mock interview is simply a chatbot. A chatbot is an excellent tool for interacting with users quickly. They helped us by providing entertainment, saving time, and answering difficult-to-find questions. The chatbot should be straightforward and conversational.

In this project, we logged how they are and the applications of Chatbot in various fields, and thus chose the field for the candidate to practise for the interviews.

Students can easily ask questions in natural language, and the interviewer will respond with the information.

A simple unified language model architecture can be used to train the model from start to finish. We show that our model, powered by well-defined knowledge grounding, can approach human performance in some ways, but falls short when dealing with detailed knowledge or long-turn consistency.

Candidates must interact with the virtual assistant mock interview in order to prepare for their upcoming interviews.

The gentle purpose of virtual assistant mock interview is simple, user-friendly, and easily understood, and the knowledge base must be compact.

In this study, the model of multi-turn dialogue and sentiment recognition analysis are combined in the

development of chatbots. That is, it will be designed to be used in everyday conversations rather than at specific times. This model will be improved in the future so that lifelong learning can be integrated and intent judgement can be served.

The model can be trained from start to finish using a simple unified language model architecture. We show that our model, powered by well-defined knowledge grounding, can approach human performance in some ways, but still lags behind when dealing with detailed knowledge or long-turn consistency.

V. Future work

Our model appears to be fairly effective. The accuracy was 61 percent, but it was still 14 percent behind current models. Overfitting appears to be a problem. As a result, drop out layers will be included in future model architectures. Furthermore, predicting "disgust" is extremely difficult. Our original model is skewed, and it appears that this has an impact on this class. For example, we require tailored data augmentation for this specific class. Including extracted facial features in the design matrix is another option to improve the model and approach state-of-the-art solutions. In state-of-the-art models, extraction is commonly done using histograms of oriented gradients (Hog) on sliding windows, but face landmarks are also taken into account. Finally, in terms of graphical display, a potential improvement, instead of displaying probabilities on the top corner, would be to display graphs instead. All these improvements will be explored further in the second period of this project.

VI. Reference

- [1]. The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS), <https://zenodo.org/record/1188976/?f=3.XAcEs5NKhQK>
- [2]. The Facial Emotion Recognition Challenge fromKaggle, <https://www.kaggle.com/c/challenges-in-representation-learning-facial-expressionrecognition-challenge/data>
- [3]. End-to-End Multimodal Emotion Recognition using Deep Neural Networks, <https://arxiv.org/pdf/1704.08619.pdf>
- [4]. Facial Expression Recognition using Convolutional Neural Networks: State of the Art, <https://arxiv.org/pdf/1612.02903.pdf>
- [5]. B.Basharirad, and M.Moradhaseli. Speech emotion recognition methods: A literature review. AIP

Conference Proceedings

2017.[https://aip.scitation.org/doi/pdf/10.1063/1.500543](https://aip.scitation.org/doi/pdf/10.1063/1.5005438)

[8](#)

[6]. L.Chen, M.Mao, Y.Xue and L.L.Cheng. Speech emotion recognition: Features and classification models. Digit. Signal Process, vol 22 Dec. 2012.

[7]. T.Vogt, E.Andre´ and J.Wagner. Automatic Recognition of Emotions from Speech: A Review of the Literature and Recommendations for Practical Realisation. Affect and Emotion in Human-Computer Interaction, 2008.

[8]. T.Vogt and E.Andre´. Improving Automatic Emotion Recognition from Speech via Gender Differentiation. Language Resources and Evaluation Conference, 2006.

[9]. T.Giannakopoulos. pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. Dec. 2015<https://doi.org/10.1371/journal.pone.014461>

[10]. Eleni Adamopoulou, LefterisMoussiades,"An Overview of Chatbot Technology" ,International Conference on Artificial Intelligence Applications and Innovations[29 may 2020].

[11]. K. Anitha,V.Shanthi,"A Study on Intervention of Chatbots in Recruitment" ,Innovations in Information and Communication Technologies,spinger[16 july 2021].

[12]. Boris Galitsky,"Adjusting Chatbot Conversation to User Personality and Mood", Innovations in Artificial Intelligence for Customer Relationship Management, springer[24 December 2020].

[13]. Fabio Calefato, Filippo Lanubile, Nicole Novielli, "EmoTxt: A Toolkit for Emotion Recognition from Text", springer.

[14]. Krisztian Balog, "Conversational AI from an Information Retrieval Perspective: Remaining Challenges and a Case for User Simulation".

[15]. Aishwarya B R, Deepthi, Greeshma B G ,Madhumitha Raj C , Jahnavi S, "Sentimental analysis for chatbots to make them emotionally reactive" ,IJARII[2021].

[16]. Doug DeGroot," Maintaining credible dialogs in a VideoBot system with special audio techniques",IEEE,[7 August 2002].

[17].[Jagdish Raheja](#),[RadheyShyam](#) ,[Umesh Kumar](#)," Gesture Capture and Recognition technique from real-time video stream" international conference on "Artificial Intelligence and Soft computing,[January]