HR-Chat bot: Designing and Building Effective Interview Chat-bots for Fake CV Detection

Diaa Salama AbdElminaam[1], Noha ElMasry[2], Youssef Talaat[3],
Mohamed Adel [4], Ahmed Hisham[5], Kareem Atef[6], Abdelrahman Mohamed[7] , Mohamed Akram[8]

Faculty of Computer Science
Misr International University, Cairo, Egypt
diaa.salama[1],noha.elmasry[2], youssef.talaat[3],
mohamed1710193[4], ahmed1709091[5], kareem1703594[6], abdelrahman1708919[7], mohamed1700772[8]{@miuegypt.edu.eg}

Abstract—The Human Resources (HR) departments in every company can face challenges regarding the number of job applicants that are applying for a job . So ; The main idea of this paper is to help HR employees in collecting information about applicants to a specific job through identifying applicant's personal skills and through writing their CV in an organized manner and filtering or rearranging job applicants throw a chatbot interview that can discover fake CV skills and find applicant's personal skills. Afterward, the proposed model will recommend this applicant according to the personal skills the job requires. The model based on creating a chatBot to extract these skills through an interview. Also, the proposed model will provide facial expression recognition techniques to track the applicant's expressions throughout the interview. This virtual interviews could highly minimize the amount of effort the HR employee have to deal with.Also it rearranges job applicants to the right track.

Index Terms—Chat-bot , Face emotional recognition , Machine learning , Fake CV

I. INTRODUCTION

Have you ever wondered why applying for a job required too much effort from applying to multiple companies, attending multiple interviews or cruising around trying to advertise yourself around companies. This problem is common for the past years trying to communicate with one another. Besides this problem, during interviews HR interviewers could make mistakes judging or addressing some job applicants as of humans nature of addressing appearance first. Today technological trends currently include artificial intelligence, advanced machine learning, human machine interaction and Natural language processing. Chatbots nowadays are common in our daily use like cortana, Siri or Alexa in addition to virtual personal assistance there is another type of smart application that can improve our user experience and makes our lives easier and better. Chatbot is an AI concept that interacts with humans through voice or text chatting. The chatbot communicates on behalf of the companies in order to simplify online communication. Chatbots are typically used in dialogue systems for various practical purposes including customer services and information acquisition. Chatbots with the help of facial expression recognition techniques could help accessing better results for detection throw conversa-

tions. Facial expressions are the facial changes in light of an individual's interior passionate states, expectations, or social interchanges.

The system will be a simple web application like any other website a web surfer interacts with at least once a day, all the applicant must do is to login into the system using Google accounts and then attempt the interview. And then, the system will establish the connection between the chatbot and the applicant to make the interview. The important data such as the extracted data will be saved in attributes during the interview and after finishing the interview, the chatbot will send those attributes using Webhook to the database (MongoDB Atlas). Web-hook is used to connect two applications using a URL. The HR employee will check all applicants' CVs and an individual report for each one of them. Also, he can see an individual report about his interview (e.g. Answers, Grades, etc...). There are several challenges in the chatbot implementation .The extraction of key words throw every answer and comparing them to each synonym .Also most of the question are based on situation wise so the ability to extract certain skill based on many different situations could be highly challenging . Merging facial expression with applicants answer to get a certain score and constructing multiple flow of questions all that is challenging and many others.

This HR chatbot system could highly minimize the time and the amount of human processing on job applicants, filtering all applicants to see who is suitable to carry on to the next step of job applying .Also , it helps the job applicants to connect with multiple companies throw just one interview.

The major contributions of this paper are as follows:

- The construction of a chatbot that can assess job applicants and extract applicants' skills by interview.
- Applying Facial extraction techniques that can get applicant's expression during the interview to achieve a better decision.
- Minimize the time of interviewing that HR takes on deciding the right man while decreasing Human error and judging
- Verifying to the HR employees that applicant's skills are real or fake.

The remainder of this paper is organized as follows. Section 2 discusses related work. In Section 3, we describe our

methodology for this study. In Section 4, we briefly discuss The results we reached with our models and it's experiments. Finally, our conclusions and future work are discussed in Section 6.

II. RELATED WORK

In [1] ,the authors are discussed the a problem of unorganized text data that people are making are huge. So, the authors decided to create a process that extracts useful features from any big text data by using two algorithms in the deep learning field and they are CNN and RNN. So, they combined CNN and RNN in one model that has been applied to their Natural Language Processing(NLP) system. NLP had been used in text classification like spam filtering, retrieving information, classifying documents, and a lot of things. The results were so good because of the advantages of CNN in extracting features and RNN in capturing long-term dependencies and to reduce the loss of details in the information. They tested their model on SSTB and IMDB datasets, it achieved a good competitive accuracy despite its small size with CNN-LSTM architecture.

The main problem in [2], is talking about is detecting fake news. So, the authors build a model based on Naive Bayes classifier that detects fake news. They tested their system on a data set that contains Facebook news posts and achieved 74% accuracy during testing. In this paper, they mentioned some ways to improve the classifier like making the dataset bigger or treating words as one word like ('read' and 'reading').

The main problem in [3] ,is discussing on the higher order thinking skills assessment of students.They used A technique called NCI which is a hybrid technique a combination between Natural language processing (NLP) and CNN-seq and Information theory.For the data set they used all Exams questions that used assess people according to higher order cognitive competency stated in Bloom Cognitive Competency Taxonomy.They compared the results against the baselines method such as Vector Space Model and Latent Semantic Analysis and NCI Proved better results.

In [4] ,the main problem is that there are large volumes of applicants, communicating with many candidates and that are already busy with many other duties making a high impact to the recruitment problems.They tried to solve this problem by making a chatbot called JARO .They used it replace or help in minimizing the number of applicants also for accelerating the interviewing process .The system analyzes the applicant's CV and then starts a quiz system by using the chatbot to analyze his knowledge.For the result of this assessment they print out a full report about the applicant and a Score for him.

In [5] proposed an Artificial intelligence and natural language processing (NLP) are used by chat bots and it helps robots think well to address applicants. It proposed to understand the concept, applications of chat-bot, and to identify the benefits of chat-bot for HR and businesses. Moreover,to determine the problems that HR faces when implementing chat-bots.

In [6] proposed a chat-bot system to simulate human conversations to help companies employees with their needs and give more personal experience. Users of this system will be able to ask any question related to HR such as view salary and others. The chat-bot will understand what will understand the question then generates an appropriate response. The chat-bot will respond immediately and will advice to workers to call or visit their HR assistant for more help when required. Artificial intelligence and machine learning techniques where used for decision making. Python programming language is used to create chat-bot and pattern matching is used to match the data with the inputs saved in database. The problem in this paper is that they didn't mention the data-set used to train and test the chat-bot.

In [7] , proposed a chat-bot design with 3D avatar, voice interface and facial expression. The system will listen to the user's speech and then find the appropriate reply by showing 3D avatar and voicing its text to speech sound. The chat-bot API has variety of features like creating a conversation session. The system will use Text to Speech (TTS) API to make the computer talking. The APIs reached a level of confidence below 80% due to using HTML5 speech recognition API.

In [8] proposed a chat-bot system to help provide accurate travel information by taking information needed from the user and then predicts the accurate answer but first the system identifies the missing information then starts to ask questions until it fulfills the missing information. The system architecture consists of Five modules: Flight, Rental car, Hotel, Web Application and Recommendation module. The system is developed using machine learning algorithms and restricted Boltzmann machine (RBM). The RBM reached 83.2% accuracy.

Some facial recognition methods are not accurate those days, so, this paper [9] presents one of the efficient methods for facial emotion recognition (FER) Via SVM-based Classification combined with the GA-based parameter optimization. Two geometric features: landmark curvature (LC) and vectorized landmark (VL), Also, this method makes the following three concurrent processes for achieving optimal landmark selection and SVM parameters: 1) Landmark detection and geometric feature extraction. 2) GA-based parameter selection. 3) • SVM-based classification. The researchers used the CK+ and MUG dataset to train and test this method. As a result of the system after testing three cases: 8-class CK+, 7-class CK+, and 7- class MUG, the result was highly efficient and accurate as the result was only between 93% to 97%.

Those days and in the last few years the interest for facial recognition was increased, so, this paper[10] shows a high-performance facial expression recognition method based on facial action unit. The facial action unit research studies the movement of facial muscles and describes facial movement changes and it's one of the most representatives' methods. Moreover, it uses CK+ dataset and it is composed of image sequences of eight expressions video recorded by 118 subjects. But the result was a bit low from the CK+ dataset but on real-time performance proves that this method can realize video or camera facial expression recognition in real time.

Also , numerous meta-heuristic optimization algorithms and deep learning have been proposed in recent years to

address difficult optimization problems such as clustering text documents, data mining,video mining,image segmentation, computer vision, and sports video mining .swarm intelligence including mainly Tunicate Swarm Algorithm (TSA) [11] , Gradient-Based Optimizer (GBO) [12] , Turbulent Flow of Water-Based Optimization (TFWBO) [13] ,Owl search algorithm (OSA) [14] ,Fitness-Dependent optimizer (FDO)[15] , and Squirrel Search Algorithm (SSA) [16] , and deep learning algorithms [17, 18, 19]

III. METHODOLOGY

The main idea of the system is to help HR interviewers, Job Applicants, and HR employees collecting information about applicants to a specific job. The system will help to build an easy link between HR employees and their job applicants through identifying applicant's skills and through writing their CV in an organized manner. Afterward, the system will recommend this applicant according to the skills the job requires. Each company will have its own Recommendation skills. The System will use a chat-Bot to extract these features with the help of facial gesture recognition techniques.The system will use different machine learning and deep learning techniques for extracting the HR skills. The system will also include different tests that can discriminate between applicants according to the job desired skills to ensure that the right person is in the right place

Our system is based on three essential stages and they are input & preprocessing stage, processing stage, and finally the output stage.

A. Input & Pre-Processing Stage

In this stage, the applicant will start inserting his CV in two ways. The first one is by giving him an empty CV form and the second one is by making him upload his cv as a soft copy. At the same time, the computer's camera will start to monitor his emotions using machine learning algorithms such as Random Forest or Support Vector Machine or using deep learning such as Convolutional neural network. Then will be saved in the database called MongoDB Atlas.

We are creating a chatbot using a platform called 'Snatchbot', we will use HR employee's questions to ask them to the applicant and collect the answers in attributes and send them using webhook to the database. The flow of interactions will be based on the applicant's reply using natural language processing. Natural Language Processing is using Bernoulli Naive Bayes model algorithms to train Intents and List & Conditional Random Field to train Entities.

B. Processing Stage

In this stage, the system will establish a connection with the created chatbot. The chatbot will start collecting some data about the applicant and then it will start to ask questions from the dataset that we made. The chatbot will be using NLP or waiting for specific keywords or any synonym to this keyword in the applicant's response to redirect to the next interaction. At the same time, the system will be using a machine learning model or deep learning model to extract his emotions during his interview.

C. Output Stage

Finally, the HR employee will be able to check all applicants' CVs using our system. Also, he can see an individual report about his interview (e.g. Answers, Grades, etc...). This individual report contains all questions, answers, extracted personal skills and his emotions extracted.

D. Facial Emotion

In our facial emotion recognition , the target of this system is to know the applicant's emotion The system start working by deecting the face in the image first and as shown in step a in the figure, afterwards it sets 68 facial landmark on the face ,Facial landmarks are used to localize and represent salient regions of the face, such as Eyes, Eyebrows ,Nose , Mouth and Jawline . Facial landmarks have been successfully applied to face alignment, head pose estimation, face swapping, blink detection and much more and then compare the landmark points to the emotionlabels that we got them from our dataset(I-bug 300) then send the emotion with the timestamp.

IV. EXPERIMENTAL RESULTS

The aim and goal is to get from the applicants enough data using our system to explore his personal skills with grades and provides this data to HR employees which will help them to reduce HR employee's time and to make it easier for them to hire people.

A. Dataset

In our facial emotion recognition we are using IBug-300 as our dataset that cover a wide range of different subjects, poses, illumination, occlusions, among other things , The given annotations tend to have a number of limitations,It contains 300 Image with multiple dimensions.

Questions and Answers data set By contacting and communicating with the HR department in Misr international university (MIU) we constructed a data set of multiple situations wise questions and some pre-processing were added to construct the questions and answers flow that is needed to give a certain score for each question multiple-choice questions where added just to direct the applicant to the right flow of answering. Some questions were rearranged and reconstructed to fit the assessment criteria of the system as the system is highly dependent on keywords that the applicant says to proceed in a certain flow or change it or give a certain score.

The interview main aim is to reach a point which certain skills is present within the applicant's answers so the interview will be a set of flows or a tree of questions and answers that for example, the applicant is going to be assessed for team working it starts by asking situation wise question if the system got it's desired or related expected answer it will stay in this branch or flow to know more about his level and give him/her a score if not then the system will try to ask the question in a different way to decide if it reached a dead end like
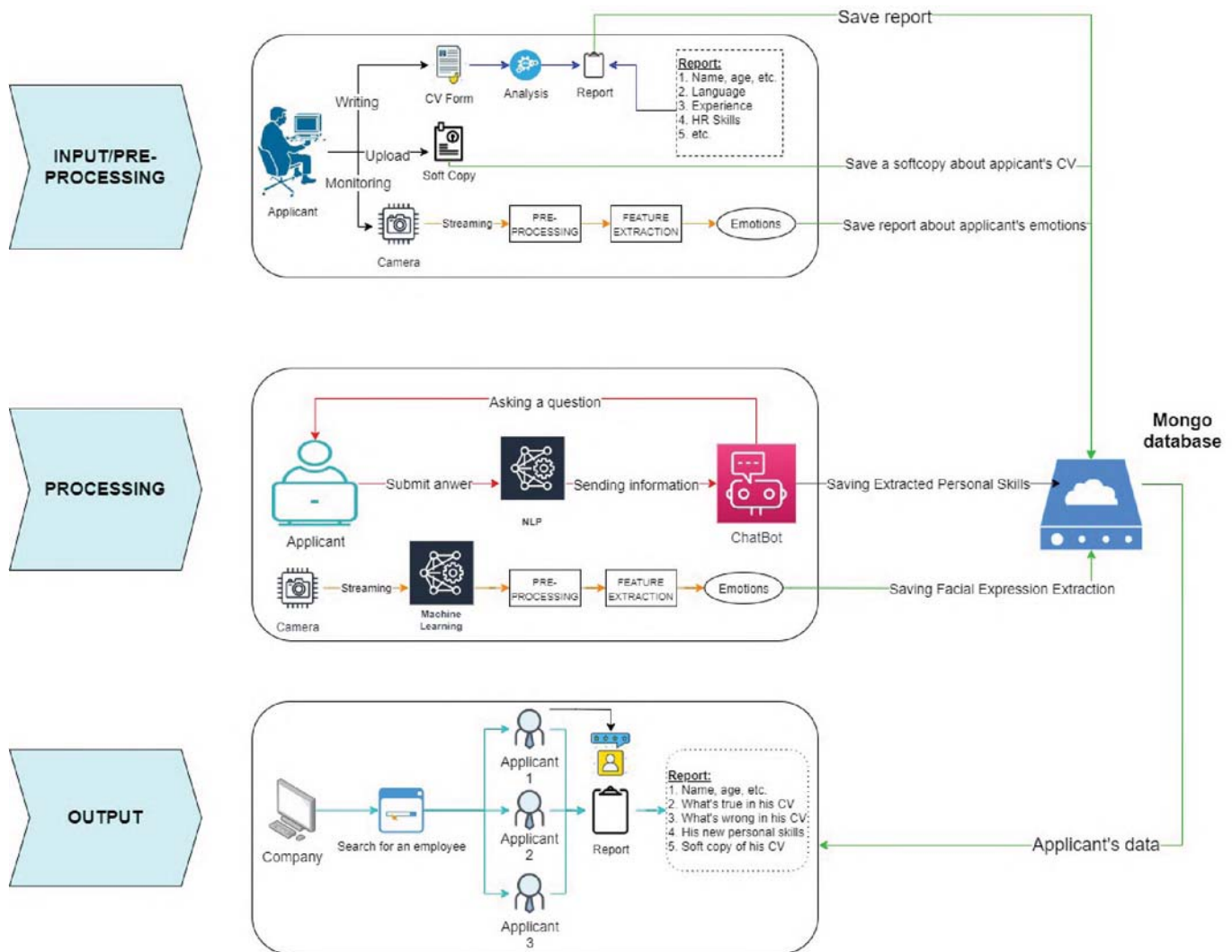
Figure 1. Proposed system overview

couldn't identify this skill then this skills will be marked with no score. Of course, some skills will take more questions and conversations than others as the diversity of skills. All of the questions and answers will be added to the applicant's report for the companies or HR employees' further assessment or further correction.

B. Setup & Approach

Our system is built using the 'Django' Web-framework. Django is a web framework that's based on Python programming language and it's a free open-source. Django uses HTML, CSS, and JavaScript as front-end and python as a back-end. We can work on the Django project using Visual Studio or PyCharm or any code editor that supports Python Language using Python version '3.7'.

We are using MongoDB Atlas as our database to save and retrieve data. During the interview with the applicant, the chatbot saves the important data in an attribute and send it to the database using webhook.

We are using in our system platform called 'Snatch-bot'. 'Snatch-bot' is a platform that helps in creating a chatbot faster, easier, and professionally. Also, it's easy to merge the bot created with many channels (e.g. Whatsapp, Facebook Messenger, Website, Viber, Telegram). All we have to do in this platform is to create interactions and link them with each other to make an interactive chatbot. This interactions contains attributes that we made to save all data needed (e.g. applicant's name, q&a) and save them in database using webhook.

Not only but also, we made a facial emotion recognition, it's a python code written using Visual Studio or PyCharm. We used 2 machine learning algorithms: Support Vector Machine (SVM) and Random forest to get the best accuracy we can.This function is built to set 68 points around the face to detect the emotion on the face comparing it to our used dataset and return the emotion on the face every time it changes. Our main target is to get the applicant's reaction to our Chatbot question at the same time to evaluate the applicant.

V. RESULTS & DISCUSSION

In this section , the results ad dissuasion from Chat-bots and facial emotion are described in details in the following subsections

A. Chatbot

Our chatbot is build using a platform called 'Snatchbot'. This platform provides many features that help developers to build a good chatbot in a short time and for free. You use this framework to create interactions and link them with each other with help of Natural Language Processing. Snatchbot also provides a ready-made NLP to make the chatbot understands better to act like a real human.
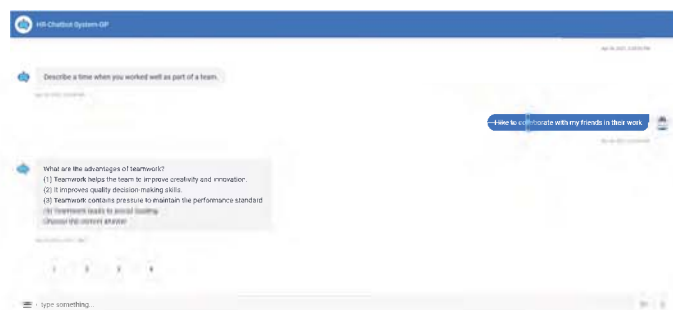

Figure 2. Chatbot

It can be trained using the Naive Bayes model, List model, and Conditional Random Field model. This chatbot is connected with the database. The chatbot sends in the last interaction using Webhook the attributes that we want to save it like his name, email, etc...

B. Facial Emotion

In our 2 used algorithms for the facial emotion recognition , we trained them to check for accuracy and The random forest accuracy 76% and the SVM had an accuracy of 78% and which was higher than the random forest. So, in the end, we used the SVM.This picture shows how the system recognize the face and analyze the 68 points that are set on the face to tell the emotion on the applicant's face.
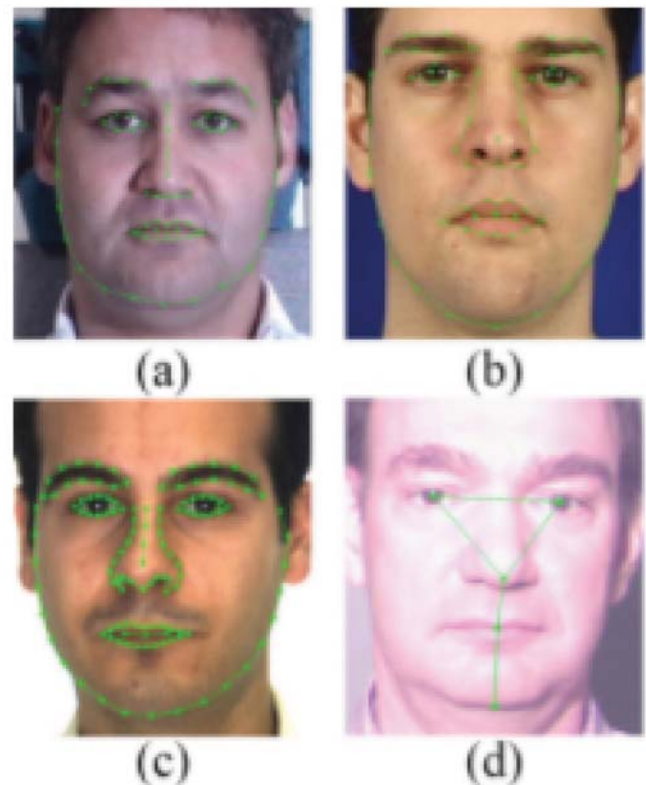

Figure 3. Facial emotion recognition

VI. CONCLUSION

In conclusion,detection of job personal skills is essential to achieve a high level of communication and harmony between company's employees.By helping HR employee's taking a better decision on who's suitable for the job is important to achieve companies aims and goals.Chatbots could highly help facilitate our communications and helps reaching some important decisions with the help of some outer algorithms and techniques a more accurate decision could be reached like facial emotion extraction or speech emotion extraction .As in this case with the usage of the snatchbot platform and construction of a questions and answers dataset and usage of Facial emotion technique SVM reaching an accuracy of 78% and combining all there results together to form a report that could help HR employees achieve a better decision on selecting the right man for the right job.Furthermore, there are many other applications that can be used in the field of chatbots or facial emotions to facilitate decision making in many other fields like psychiatry and business.

VII. FUTURE WORK

The future work is extract more personal skills from the applicant with very high accuracy, enhancing the facial emotion recognition accuracy, and upgrade our system with more features and APIs. Also, by focusing on technical skills and using extra technique of assessment like speech emotion recognition it could lead to better decisions.

407

REFERENCES

[1] S. Zheng, H. Bao, J. Zhao, J. Zhang, Z. Qi, and H. Hao, "A novel hierarchical convolutional neural network for question answering over paragraphs," in *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, vol. 1, 2015, pp. 60–66.

[2] M. Granik and V. Mesyura, "Fake news detection using naive bayes classifier," in *2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2017, pp. 900–903.

[3] A. Lajis, A. F. B. Mohd, T. A. Mat, H. Nasir, and N. A. Aziz, "Data selection in the assessment of higher order thinking skills: Focus on cognitive competency," in *2018 IEEE 5th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, 2018, pp. 1–4.

[4] J. Purohit, A. Bagwe, R. Mehta, O. Mangaonkar, and E. George, "Natural language processing based jaro-the interviewing chatbot," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, 2019, pp. 134–136.

[5] R. Mohan, "The chatbot revolution and the indian hr professionals."

[6] S. Tadvi, S. Rangari, and A. Rohe, "Hr based interactive chat bot (powerbot)," in *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*, 2020, pp. 1–6.

[7] P. A. Angga, W. E. Fachri, A. Elevanita, Suryadi, and R. D. Agushinta, "Design of chatbot with 3d avatar, voice interface, and facial expression," in *2015 International Conference on Science in Information Technology (IC-SITech)*, 2015, pp. 326–330.

[8] A. Argal, S. Gupta, A. Modi, P. Pandey, S. Shim, and C. Choo, "Intelligent travel chatbot for predictive recommendation in echo platform," in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 2018, pp. 176–183.

[9] X. Liu, X. Cheng, and K. Lee, "Ga-svm-based facial emotion recognition using facial geometric features," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11 532–11 542, 2021.

[10] J. Yang, F. Zhang, B. Chen, and S. U. Khan, "Facial expression recognition based on facial action unit," in *2019 Tenth International Green and Sustainable Computing Conference (IGSC)*, 2019, pp. 1–6.

[11] E. H. Houssein, B. E.-D. Helmy, A. A. Elngar, D. S. Abdelminaam, and H. Shaban, "An improved tunicate swarm algorithm for global optimization and image segmentation," *IEEE Access*, vol. 9, pp. 56 066–56 092, 2021.

[12] S. Deb, D. S. Abdelminaam, M. Said, and E. H. Houssein, "Recent methodology-based gradient-based optimizer for economic load dispatch problem," *IEEE Access*, vol. 9, pp. 44 322–44 338, 2021.

[13] D. S. Abdelminaam, M. Said, and E. H. Houssein, "Turbulent flow of water-based optimization using new objective function for parameter extraction of six photovoltaic models," *IEEE Access*, vol. 9, pp. 35 382–35 398, 2021.

[14] W. H. El-Ashmawi, D. S. Abd Elminaam, A. M. Nabil, and E. Eldesouky, "A chaotic owl search algorithm based bilateral negotiation model," *Ain Shams Engineering Journal*, vol. 11, no. 4, pp. 1163–1178, 2020.

[15] D. S. Abdul-Minaam, W. M. E. S. Al-Mutairi, M. A. Awad, and W. H. El-Ashmawi, "An adaptive fitness-dependent optimizer for the one-dimensional bin packing problem," *IEEE Access*, vol. 8, pp. 97 959–97 974, 2020.

[16] W. H. El-Ashmawi and D. S. Abd Elminaam, "A modified squirrel search algorithm based on improved best fit heuristic and operator strategy for bin packing problem," *Applied Soft Computing*, vol. 82, p. 105565, 2019.

[17] D. S. Abdelminaam, F. H. Ismail, M. Taha, A. Taha, E. H. Houssein, and A. Nabil, "Coaid-deep: An optimized intelligent framework for automated detecting covid-19 misleading information on twitter," *IEEE Access*, vol. 9, pp. 27 840–27 867, 2021.

[18] D. Abdul, "Elminaam, shaimaa abdallah ibrahim,"building a robust heart diseases diagnose intelligent model based on rst using lem2 and modlem2"," in *the Proceedings of the 32nd International Business Information Management Association Conference, IBIMA*, 2018, pp. 5733–5744.

[19] D. Salama AbdELminaam, A. M. Almansori, M. Taha, and E. Badr, "A deep facial recognition system using computational intelligent algorithms," *Plos one*, vol. 15, no. 12, p. e0242269, 2020.