# Assistant Robot for Academic Fraternity: A survey on methodologies and challenges

**Lekhana Kalyan Raj, Chirag Sanjeev Bijapur, Gururaja H S**
Student, Student, Professor
Information Science and Engineering,
BMS College of Engineering, Bengaluru, India

**Abstract** - Robots are machines that can perform a variety of tasks on their own or with minimal human supervision. They have a diverse range of applications, spanning multiple industries such as manufacturing, healthcare, transportation, and education. Living in the twenty-first century, we have witnessed extreme technological advancements in the field of education, but there is still room for improvement through robotics. Various robot series can be employed in a classroom setting, in accordance with the specific priorities and objectives of the institution. Some examples include: [1] Tutoring robots: These robots can help students with specific subjects, such as math or reading, and provide personalized instruction. [2] Collaborative robots: Robots that can work with students on group projects or activities, helping to facilitate teamwork and communication. [3] Social robots: Robots that are intended to engage with students socially or emotionally, such as by offering support or fostering connections. [4] Educational robots: These robots are designed to teach specific curriculum or subjects with the help of various sensors, motors and other attachments. [5] STEM robots: These robots are designed to teach science, technology, engineering, and math concepts to students. All of these robots can be controlled and programmed by teachers or students to customize the learning experience. In this paper, we provide a critical survey of the factors that are important in building a successful assistant robot for the academic fraternity.

## I. INTRODUCTION

As everyone is aware, artificial intelligence is booming in the marketplace, and it is becoming increasingly dependent on artificial intelligence to handle multifaceted tasks. Robotics is a widely known division in the field of manufacturing and sciences in which all engineers work hard to create a robot that can do a specific task and produce appropriate results.

The paper's aim is to evaluate different methodologies to come up with an AI-based robot for the Academic fraternity. This robot would assist both teachers and students with various tasks. For example, by developing a face recognition model, we hope to automate the attendance monitoring during teaching hours and also initial screening process prior to internals and exams. Students would benefit from the portal's convenient access to teachers as well by easy scheduling of meetups. Face Recognition, NLP and other vision modules will be equipped into the robot. The robot can also be designed to patrol in the classroom for interactions. The hardware is where we can achieve a humanoid version of the software, attempting to make the robot more appealing and friendly to use.

## II. LITERATURE SURVEY

- **Natural Language Processing**

Designing realistic interaction with social robots is a very challenging process due to the vast design space of robots in terms of visual appearance and behavior, as well as the difficulties that arise when utilizing facial detection and speech recognition in the field. More organic and highly autonomous interaction models like natural language processing are needed between users and robots in order to encourage trust and engagement and to create a long-term social relationship. Natural language processing is the extraction of text from spoken or written natural language input. The goal of NLP is to teach computers to generate sentences from natural language in the same way that humans do.

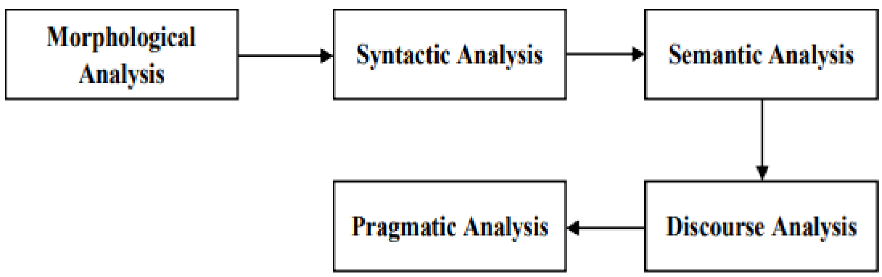


Fig.1. NLP Flow

The tasks included in the NLP application are depicted.
[1] Lexical Examination: It involves recognizing and examining word structure. The diction of a language is the volume of expressions and words. The complete text is broken down into phrases using lexical analysis.[2] Syntactic Evaluation: It requires checking the linguistics of each word in a sentence and arranging the words in a way that clearly shows how they relate to one another. An English syntactic analyzer will not accept a statement like "The class goes to girl." [3] Semantic Evaluation: It captures the text's precise or dictionary meaning and retrieves it. The meaning of the passage is examined. The task domain's grammatical structures and elements are mapped to do this. The semantic analysis ignores phrases like "hot ice cream". [4] Integration of Discourse: Each and every sentence's meaning is influenced by the meaning of the preceding sentence that comes before it. [5] Pragmatic Evaluation: During this duration, what's been spoken is re-evaluated to determine its true meaning. It involves deciding the language features that demand practical experience.

- **Face Recognition**

Face recognition is an algorithm that verifies or identifies a person by using their facial features. It is most commonly used in security systems, such as unlocking a smartphone or gaining access to a building, but it can also be used in other applications such as tagging people in photos or tracking attendance at an event. The technology compares the unique characteristics of a person's face, including the distance between their eyes, nose, and mouth, to a database of known faces.

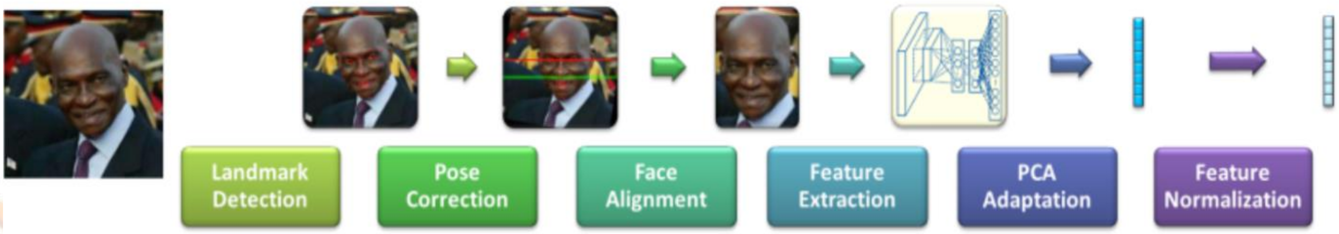


Fig.2. Facial Representation Pipeline

There exist several approaches to perform face recognition, including:

- **Deep Multi-Pose Representations**

A face image is processed by multiple deep CNN models that are tailored to specific poses, generating various pose-specific features. Using 3D rendering, many face positions are created from the input image. The system's sensitivity to pose fluctuations was decreased because they applied an ensemble of pose-specific convolutional neural network features.

- **An Embedded GPU System**

In an embedded GPU system, there are numerous face recognition frameworks in use. The framework includes a cutting-edge deep CNN face recognition algorithm as well as face tracking and CNN based face identification.

- **An upgraded quicker RCNN approach**

By combining various techniques, such as feature concatenation, multi-scale training, hard negative mining, pre-training, and accurate alteration of crucial parameters, they enhance the cutting-edge Faster RCNN framework.

- **Deep neural network-based face recognition using multiple views**

The regions of the face are deeply encoded using a deep NN, and the critical points within the faces are localized using a face alignment algorithm. Then, they employ the well-known PCA to eliminate the extraneous and tainted visual features while also bringing down the dimensionality of the deep features. Then, to examine the similarity of feature vectors and achieve highly competitive face recognition accuracy, they propose a joint Bayesian framework.

- **Face Recognition and Identification using Deep Learning**

The researchers developed a deep learning-based face recognition system using OpenCV in Python. Deep learning is a promising technique for face recognition due to its high accuracy. The performance of the system is verified through experimental results.

Face recognition model for Video

| approach | Feature used | Mathematical term used for | Test database | Result/conclusion on technique |
|---|---|---|---|---|
| Gabor Filter [38] | Skin Texture | Calculation of edge detection and rotation value. Then frequency and simulation are used to extract features. | pubfig+10 dataset, Movie trailer face dataset | Feature Vector is extracted. |
| HOG Filter [39] | Skin Texture | The magnitude and orientation gradient of the localized region is calculated. | pubfig+10 dataset, Movie trailer face dataset | Feature Vector is extracted. |
| Principal Component Analysis-PCA [40] | Extracted feature vector from hog, Gabor. | Removal of redundant feature vectors with the help of eigenvectors of the covariance matrix. | pubfig+10 dataset, Movie trailer face dataset | Removal of redundant feature vectors. |
| SRC-Sparce Representation based Classification [41] | Feature vectors optimized by PCA. | Calculating residuals using minimum coefficient vectors and to create prediction classes. | pubfig+10 dataset, Movie trailer face dataset | Predicting classes using residuals |
| RSRC | Feature vectors optimized by PCA. | Used to calculate sparsity index fro single coefficient vector for a single frame to get an improved average precision. | pubfig+10 dataset, Movie trailer face dataset | Predicting classes using residual with much more precision than SRC |

Fig.3. Approaches of Face recognition

- ## Attendance Monitoring

Traditionally, professors call out students' names and keep attendance records. It takes a while to record attendance. Assume that a single subject class lasts about 55 minutes and that recording attendance takes about quarter of an hour. This is a squander of time for each lecture. To avoid many such losses, an automatic process that utilizes image processing is proposed. The module will be linked to the camera, and the student database has been assembled. The names, images, and roll numbers of the candidates are stored in the repository. This module will be placed in front of the class in order for us to capture the entire class. As a result, recording attendance with this system is extremely simple. The process begins with detecting faces, which is then followed by recognizing them. The Viola Jones method is used to detect faces, while a combination of PCA and LDA algorithms are used for face recognition.

- **Viola-Jones algorithm**

The algorithm is composed of three integral components: Intrinsic Images, the Ada-Boost function, and the Cascade. The integral image determines an estimate for each single pixel (PX, PY) which is a summation of the pixel values above and to the left of (PX, PY) by scanning the image only once. Haar-like features are employed within the Viola Jones algorithm. This merely refers to the scalar product of the image and certain structures that resemble hears. Ada-boost is used to choose the feature. Ada-Boost offers a powerful learning method and reliable performance bounds for generalization.

- **Protocol Flowchart**

The system includes a camera that captures pictures in a classroom setting. The images are then sent to an picture enhancement algorithm. The enhanced pictures proceed through the algorithms, where we can detect and recognize faces. Then the attendance is logged in the respective database. During the time of enrolment, the Face database stores individual student face photograph templates. The program recognizes every face in the input image and evaluates each one against the face database one by one. When a face recognition is completed, the attendance is recorded on the database, which is publicly accessible and can be utilized for multiple purposes. The system also features a timetable module that fetches the subject, class, date, and time automatically. The technology automatically registers attendance without the involvement of students or teachers. This approach is both time-saving and extremely secure. The server stores attendance so that everyone has access to it, including the administration, parents, and students.

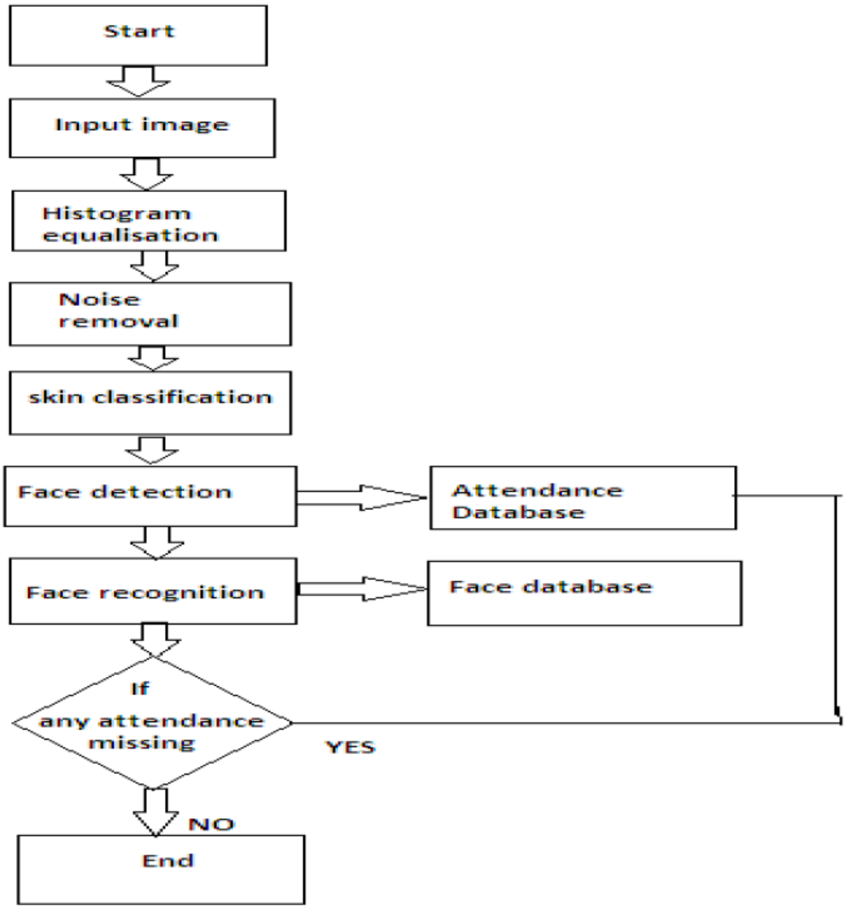


Fig.4. Face Detection and Recognition flowchart for Attendance.

- ## Human Robot Interaction

In a world where robots and humans coexist and cooperate, to predict future actions and support decision making, robotic systems must understand human behavior. It is difficult to interpret human activities in the context in which they occur because when an individual's activity is powered by objectives, and needs that may be contradictory. Whether stationary or moving.

- **A Review of Techniques for Human-Robot Interaction**
  There are two separate ways by which a robot might pose a risk to a human. The first method entails making direct physical contact. Simply put, no undesirable or unwanted interactions between the human and the robot is permitted for HRI to be safe. Harm-full HRI is a potential source of harm because it can have serious adverse consequences. Furthermore, any of the other above factors and robotic violations of conventions and social norms during contact, can have long-term negative consequences for humans. The main goal is to combine and summarize a large body of research on facilitating secure human-robot interaction. This study in the papers describes the strategies and methods that are created so far, classifies them into subcategories, describes the connections, and identifies any knowledge gaps that might require additional study.

- **Human–Robot Interaction for Supporting Learning Process**
  A scenario involving physical human-robot interaction in which a robot acts as a teacher within a predetermined framework for dance training. The idea is to provide performance feedback that combines cognitive and physical aspects to aid in the learning of new skills. An adaptive an impedance-based controller has been developed for direct contact cooperation that adapts based on how well the partner performs the task. A scoring system for evaluating performance has been created using the idea of progressive teaching (PT). Rooted on the performance of users and the number of practices, the system modifies the difficulty. Comparative experiments using the suggested method and a baseline constant controller have demonstrated that the PT exhibits better performance in the early stages of skill learning.

- **Human-Robot Interaction: Insights and Difficulties**

Discussions include the state of human-robot interaction (HRI) today as well as the main issues that the community of researchers are currently working to solve. The limitations of human factors research are discussed in this mini-review, along with the developments in HRI in four application areas. It was determined that the study of human-robot interaction (HRI) is a rapidly growing field with a high demand for the inclusion of human factors in research and design, particularly as robots are put under more demanding conditions to perform more complex tasks. In any event, it is much simpler to replace humans with robots.

- ## Robotics in Education

The majority of affluent nations now provide popular educational robotics programs, and these programs are now spreading across the developing globe. Students at various levels of schooling can learn problem-solving skills, programming, design, physics, arithmetic, even music and art through the usage of robotics. Students will gain a more interesting (and enjoyable) perspective of science and engineering in both high school and college when robots are used in the classroom. They will also be able to see firsthand how theoretical ideas in the fields of technology and mathematics are applied in real-world settings.

- **A Review of In-field Studies on the Employment of Social Robots in an Educational Settings**

The studies showed that robots could be used in regular classroom settings, but they also showed how challenging it was to establish enduring, highly independent interactions between Robots and Children. The primary factors in the studies were very different. Research did not show social robots as better than human teachers and only briefly touched on major ethical and security concerns. It also demonstrated how challenging it is for robots to replace the emotional and human space needed for learning.

- **Neural Machine Translation in Classrooms**

Neural Machine Translation (NMT) technology has the potential to be a valuable resource in the classroom. It can facilitate language learning by providing real-time, accurate translations of texts, spoken words, and written material. This can help students overcome language barriers, improve their language skills, and broaden their understanding of different cultures.

While NMT has the potential to greatly enhance language learning and facilitate communication in the classroom, it is important to also consider its limitations. NMT can struggle with idiomatic expressions, cultural references, and nuances in language, which can result in incorrect or incomplete translations. Furthermore, there may be privacy and security concerns associated with using NMT technology in the classroom, such as the protection of student data and the potential for misuse.

# III. PROPOSED SYSTEM

Robots can undertake a series of actions automatically, most notably by being programmed by a computer. They can perform tasks ranging from simple repetitive actions to more complex operations, such as those performed by humans. The growth in robots is the result of advancements in technology, especially electronics, computing, and AI. The integration of robots in the classroom offers several advantages to students' education. As a result, we suggest the implementation of the architecture depicted in Figure 5. The hope is to build this platform to reach the academic fraternity and establish robots as an educational technology.

- **Hardware Requirements**

To construct a robot, one must have a thorough understanding of its purpose and requirements, as well as the necessary electrical components such as micro-controllers, sensors, actuators, etc. and mechanical components like motors, wheels, casing, etc. and the tools needed for assembly and soldering.

Electronic components in robots are crucial for several reasons: [1] The micro-controller serves as the brain of the robot and is responsible for processing information from sensors and deciding on the appropriate actions. It also communicates with other components, such as motors and actuators, to control their behavior. [2] Sensors are used to gather information about the environment and provide inputs to the micro-controller. This information is used to make decisions and control the behavior of the robot. [3] Actuators, such as motors and servos, convert the electrical signals from the micro-controller into physical motion. [4] Electronic components, such as batteries and power regulators, are used to provide power to the robot and ensure stable operation.

- **Software Implementation**

When it comes to the software aspect of a robot, there are several modules involved, including User Interface Application Development, Deep Learning modules like Face Recognition, Dialogflow for interactions, an Attendance Monitoring system, NMT and Database Connection programs. The key modules are Attendance Monitoring and Face Recognition for screening purposes. An Attendance monitoring is used to automatically track and monitor attendance. It typically uses computer vision and deep learning algorithms to identify and recognize individuals and their attendance status in real-time. The model is developed using a substantial data-set of images and information, such as facial recognition or bio-metric data, to learn the patterns and characteristics of individuals. The deep learning model can then be deployed in real-world scenarios, where it can automatically detect and record attendance in a reliable and efficient manner. This can improve the accuracy and efficiency of attendance tracking, while also reducing the workload and manual labor required to manage attendance. It is important to note that the accuracy and effectiveness of the attendance monitoring depends on selection of algorithms, together with the quality and distinctiveness of the training data. A face recognition model can be used for initial screening in examinations and cheating detection. The model uses computer vision to identify and recognize individuals based on their facial features. By integrating this model into an examination setting, it can be used to screen students before they enter the examination room to ensure that they are the correct individuals and to detect any instances of cheating, such as impersonation or collaboration. This can help improve the fairness and integrity of the examination process, while also reducing the workload and manual labor required to manage the screening and cheating detection process. A Neural Machine Translation system can enhance inclusiveness in the classroom by providing support for students who speak different languages. By breaking down language barriers, all students will have equal access to educational resources and opportunities. The model can also aid in understanding difficult concepts, technical jargon and specialized vocabulary, even for students who are not fluent in the language of instruction. To summarize, deep learning models have the potential to significantly transform the classroom environment, offering new and innovative ways to enhance learning and engagement for students. By incorporating these models into the classroom, teachers and students can have access to advanced tools that can improve learning outcomes and motivation. However, it is important to carefully consider the ethical and privacy implications of using deep learning models in the classroom, and to ensure that these models are used in ways that promote equity and inclusiveness for all students. With careful planning and implementation, the models being developed can play a significant role in the future of education and help to create a more dynamic and engaging learning environment for students.

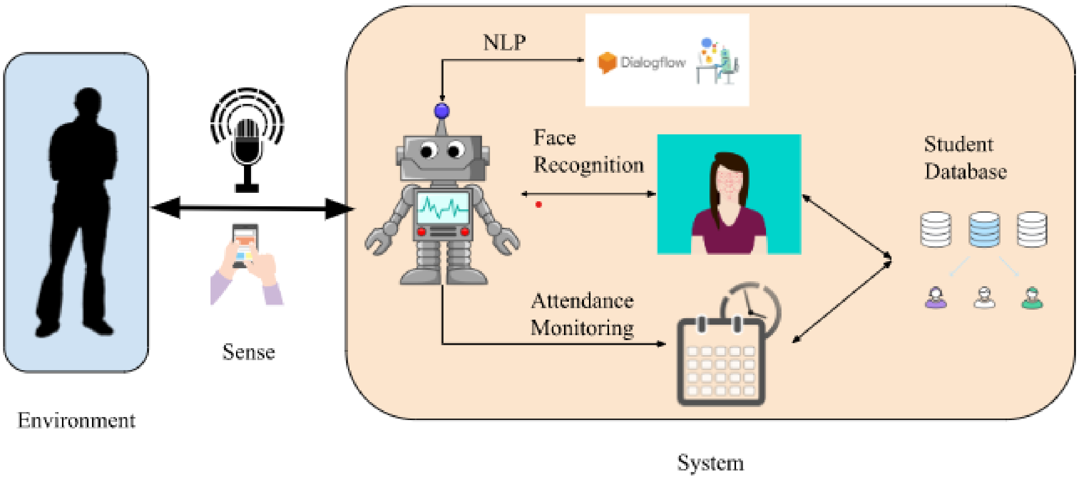


Fig.5. System Architecture

## IV. CONCLUSIONS

Robotics is continuously advancing with ongoing discoveries and implementation of new advancements and applications. Robotics has a wide range of applications, including in manufacturing, healthcare, transportation, and many other industries. With advancing technology, robots are likely to have a growing significance in education and society. It's crucial to contemplate the ethical and social consequences of the expanding implementation of robots and automation in different industries.

Face recognition technology is a rapidly advancing field that has a wide range of applications, including security, marketing, and personal identification. The technology uses algorithms to analyse and match facial features, allowing it to accurately identify individuals. But there are also worries about the possibility of abuse and invasion of privacy with widespread usage of facial recognition technology. It is important for institutions to implement regulations and guidelines to ensure that the technology is used ethically and responsibly. Additionally, research should continue to be conducted on the potential biases and inaccuracies in the technology.

Natural Language Processing has various applications, such as language translation, text summarization, sentiment analysis, and chat-bots. NLP has made significant progress in recent years, thanks to the advancements in machine learning and deep learning. However, NLP still has many challenges to overcome, including dealing with context, understanding idiomatic expressions and handling the nuances of human language. Also, the use of NMT in the classroom can have influential impact on students' learning experiences and outcomes, and help to create a more inclusive and supportive learning environment.

As the use of NLP continues to grow in different industries and applications, it is essential to consider its great potential in education. There are concerns about the potential for misuse of the technology. Therefore, it is important for researchers, educationalists and practitioners to work towards developing more inclusive and NLP models and for academic fraternity to implement regulations and guidelines to ensure that the technology is used responsibly.

In conclusion, robot interaction is a growing field that has the potential to greatly impact various industries and also classrooms. The development of more advanced robots and interfaces for human-robot interaction is allowing for more natural and intuitive ways for humans to interact with robots. By developing the proposed system, teachers and students can have access to cutting-edge tools that can enhance learning outcomes and motivation by implementing these technologies in academic fraternity. As the field of Student robot interaction continues to evolve, it is important for researchers, institutional stakeholders and students to work together to ensure that the technology is developed and used responsibly to provide a better institutional journey for a student.

## V. REFERENCES

[1] Patil, A. and Shukla, M., 2014. Implementation of classroom attendance system based on face recognition in class. *International Journal of Advances in Engineering & Technology*, *7*(3), p.974.

[2] El-Komy, A., Shahin, O.R., Abd El-Aziz, R.M. and Taloba, A.I., 2022. Integration of computer vision and natural language processing in multimedia robotics application. *Inf. Sci*, *7*(6).

[3] Mubin, O., Stevens, C.J., Shahid, S., Al Mahmud, A. and Dong, J.J., 2013. A review of the applicability of robots in education. *Journal of Technology in Education and Learning*, *1*(209-0015), p.13.

[4] Hameed, I.A., 2016, September. Using natural language processing (NLP) for designing socially intelligent robots. In *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 268-269). IEEE.

[5] Sun, X., Wu, P. and Hoi, S.C., 2018. Face detection using deep learning: An improved faster RCNN approach. *Neurocomputing*, *299*, pp.42-50.

[6] AbdAlmageed, W., Wu, Y., Rawls, S., Harel, S., Hassner, T., Masi, I., Choi, J., Lekust, J., Kim, J., Natarajan, P. and Nevatia, R., 2016, March. Face recognition using deep multi-pose representations. In *2016 IEEE winter conference on applications of computer vision (WACV)* (pp. 1-9). Ieee.

[7] Saypadith, S. and Aramvith, S., 2018, November. Real-time multiple face recognition using deep learning on embedded GPU system. In *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)* (pp. 1318-1324). IEEE.

[8] Zhao, F., Li, J., Zhang, L., Li, Z. and Na, S.G., 2020. Multi-view face recognition using deep neural networks. *Future Generation Computer Systems*, *111*, pp.375-380.

[9] Woo, H., LeTendre, G.K., Pham-Shouse, T. and Xiong, Y., 2021. The use of social robots in classrooms: A review of field-based studies. *Educational Research Review*, *33*, p.100388.

[10] Teoh, K.H., Ismail, R.C., Naziri, S.Z.M., Hussin, R., Isa, M.N.M. and Basir, M.S.S.M., 2021, February. Face recognition and identification using deep learning approach. In *Journal of Physics: Conference Series* (Vol. 1755, No. 1, p. 012006). IOP Publishing.

[11] Dubey, A.K. and Jain, V., 2019. A review of face recognition methods using deep learning network. *Journal of Information and optimization sciences*, *40*(2), pp.547-558.

[12] Graterol, W., Diaz-Amado, J., Cardinale, Y., Dongo, I., Lopes-Silva, E. and Santos-Libarino, C., 2021. Emotion detection for social robots based on NLP transformers and an emotion ontology. *Sensors*, *21*(4), p.1322.

[13] Papakostas, G.A., Sidiropoulos, G.K., Lytridis, C., Bazinas, C., Kaburlasos, V.G., Kourampa, E., Karageorgiou, E., Kechayas, P. and Papadopoulou, M.T., 2021. Estimating children engagement interacting with robots in special education using machine learning. *Mathematical Problems in Engineering*, *2021*, pp.1-10.

[14] Lasota, P.A., Fong, T. and Shah, J.A., 2017. A survey of methods for safe human-robot interaction. *Foundations and Trends® in Robotics*, *5*(4), pp.261-349.

[15] Sheridan, T.B., 2016. Human–robot interaction: status and challenges. *Human factors*, *58*(4), pp.525-532.