National Journal of Antennas and Propagation, ISSN: 2582-2659 Vol. 6, No. 3, 2024 (pp. 93-104)
RESEARCH ARTICLE
HTTP://ANTENNAJOURNAL.COM

Research on the Development and Application of AI-assisted Teaching Robot System Based on Deep Neural Network

Li Wen[1*], **Hongtao Wang**[2]
[1]*Ph.D. Candidate, Department of Music, Krirk University, Bangkok, Thailand*
[2]*Associate Professor, Department of Music, Krirk University, Bangkok, Thailand*

KEYWORDS:
Neural Networks,
Assistive Robotics,
SVM (Support Vector Machine),
Assistive Teaching,
Behavioral Associations.

ARTICLE HISTORY:
Received 23.09.2024
Revised 18.10.2024
Accepted 09.11.2024

DOI:
https://doi.org/10.31838/NJAP/06.03.12

Abstract

This thesis describes an AI-assisted teaching robot system based on a deep neural network, which aims to record the teaching situation in the classroom and continuously improve the teaching method and the quality of education through data analysis. We used the AdaBoost algorithm to implement face detection while using SVM for face classification and recognition. The AdaBoost and SVM algorithms were studied in depth and their effectiveness was verified through experiments. The image acquisition module uses the V4L2 framework to transmit the acquired image data to the client via a wired network, and the network connection adopts the socket network communication mechanism under the Linux system. We also built an Open CV development environment for training the student identification model, realizing student identification, recording the identification, and saving the image data. Based on classroom student behavior detection, this paper achieves more than 90% accuracy by analyzing students' positional features and feature similarity between consecutive video frames. To have a comprehensive understanding of student listening, we quantified students' classroom behaviors using statistical analysis methods, including overall behavioral percentage, behavioral percentage changes, and quantifying students' behaviors into the degree of listening attentiveness by marking scores, to help teachers improve the quality of teaching. This system combines various technologies such as deep learning, image processing, and statistical analysis to bring promising innovations to the field of education.

Corresponding Author: wlzyst@126.com

Author's Orcid id: https://orcid.org/0009-0002-0196-1631, https://orcid.org/0000-0001-7030-7508

How to cite this article: Wen L, Wang H . Research on the Development and Application of AI-assisted Teaching Robot System Based on Deep Neural Network. National Journal of Antennas and Propagation, Vol. 6, No. 3, 2024 (pp. 93-104).

Introduction

AI-assisted teaching has emerged as a vital technology for supporting hearing-impaired children, significantly enhancing the efficiency and quality of their education.[1-3] Nevertheless, traditional teaching methods for computer-aided design courses have presented challenges, including students' lack of engagement and suboptimal learning outcomes.[4-6] Therefore, researching a teaching approach for computer-aided design courses empowered by AI holds immense importance for improving teaching effectiveness and quality.[7, 8]

Our study focuses on developing and implementing a teaching model for computer-aided design courses, leveraging AI technology to benefit hearing-impaired students. We assess the impact of this approach, encompassing an analysis of the current state of computer-aided design course instruction, the design of a computer-aided design course teaching model incorporating AI, the practical implementation of this model, and the evaluation of its teaching outcomes.[9, 10]

In light of the continuous advancements in teaching aids driven by technological progress in our educational institutions, the predominant method primarily involves teachers utilizing multimedia tools for instruction.[11, 12] These tools aim to present various teaching methods that engage students more directly in the educational content, fostering their interest and thereby improving the quality and efficiency of classroom teaching. However, these conventional teaching aids predominantly focus on the teacher's delivery and have limited effectiveness in enhancing students' intrinsic motivation.[14]

In today's academic examination processes, face recognition technology is progressively being employed to verify the identity of candidates. This approach effectively mitigates instances of proxy cheating during exams, upholding the integrity of examinations and ensuring order in examination halls.[15] To enhance the quality of classroom teaching and maintain a conducive learning environment, we can employ face recognition to authenticate students' identities and track classroom attendance, thereby reducing absenteeism.

In this research, we introduce a teaching aid system built on face recognition technology to monitor classroom participation and motivate hearing-impaired students. The system precisely records students' in-class activities and real-time data, enabling the analysis of classroom dynamics and continuous refinement of instructional methods to enhance the quality of education.

Our study includes an investigation into methods for documenting classroom teaching activities, leading to the design of a teaching aid system specialized for this purpose. In the hardware component, we utilize the embedded ARM platform and cameras for image data acquisition. The server-side program is responsible for image format conversion, facial image acquisition, facial data transmission, and the compilation, transplantation, and debugging of the embedded experimental platform. On the software side, the client program is designed to acquire students' essential information, process facial image data, create a facial data repository, conduct facial recognition model training, display identity recognition results, and offer classroom data retrieval. These features facilitate monitoring of students' class participation, tracking of attendance, and an analysis of their learning progress within the teaching aid system's query functions.

LITERATURE REVIEW

Composition of Robot-assisted Teaching System

Compared with the traditional teaching system, in the robot-assisted teaching system, in addition to the four core elements of teacher, learner, teaching content, and teaching media, there is also a special element of the robot, which together constitute the robot-assisted teaching system.[16, 17]. Figure 1 shows the relationship between the elements of the robot-assisted teaching system.

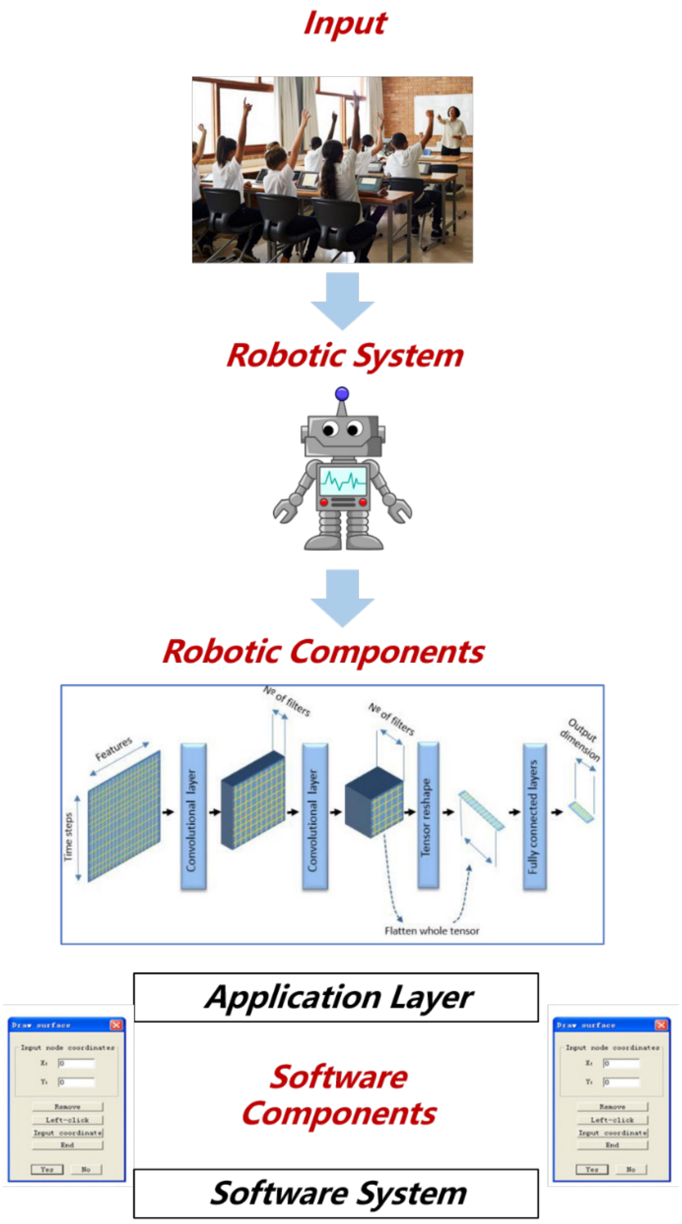

Fig. 1: Components and Interrelationships of Robot-assisted Teaching Systems

Teaching-assisted robots are equipped with a range of functions designed to support teaching, facilitate administrative tasks, and even lead instructional sessions. In the context of enhancing foreign language education, especially for hearing-impaired students in colleges and universities, this study emphasizes the essential features that teaching-assisted robots should incorporate. These fundamental functions encompass a big data analysis backend interface, network interface, backend programming control interface, machine translation, voice recognition, gesture recognition, front-end display, and interactive capabilities.

As depicted in Table 1, serial numbers 1 to 3 represent the core functions. These functions provide the foundational infrastructure for data connectivity and transmission, function configuration and control, as well as comprehensive big data analysis. In contrast, serial numbers 4 to 7 constitute the front-end interaction extension functions, primarily focused on resolving challenges related to listening, reading, and writing that often arise during human-robot interactions. These features are particularly vital when considering teaching assistance for hearing-impaired students, as they cater to their unique learning needs and communication requirements.

Table 1. Functional Requirements for Pedagogical Assistive Robots

Sl. No.	Functions	Content
1	Big data analytics backend interface	Robot data background control, including intelligent learning task analysis, learning ability analysis, evaluation and feedback, etc.
2	Network interfLace	High-speed networking to realize real-time data interoperability, including wired and wireless network ports, network port configuration management, etc.
3	Backend programming control interface	Contains context setting, behavior control, front-end display setting, human-machine interaction setting, etc.
4	Machine translation	Aiding human-computer dialogue, providing language selection, corpus management, text analysis feature matching, etc.
5	Speech recognition	Pronunciation characteristics selection, noise processing, text conversion, etc.
6	Gesture recognition	Gesture recognition, customization of specific gestures, etc., and expression and body recognition can be added at a later stage.
7	Front-end display and interaction	Visual display of network knowledge base resources, learning tasks and objectives, providing teaching resources customization, group matching, and other auxiliary management tools.

In the multimedia teaching environment, the use of robots to assist teachers in teaching, teachers themselves can directly transmit teaching information,[18, 19] or through multimedia tools to present teaching information, but also targeted use of the robot's functional characteristics of the effective transmission of teaching information, there are usually the following three kinds of teaching methods.[20, 21]

Two-way Communication between Teachers and Students.[22]

Teachers pre-assemble and debug the robotics teaching system, and the teaching content is stored in the robot and teaching media, robot-assisted teachers control the teaching media for the learners to present, tell the teaching information, and according to the learner, learning situation to obtain feedback information. Students can seek help from the teacher when they encounter problems, and the teacher controls the whole teaching process and provides targeted counseling to the learners.

Student Independent Learning[23]

Teachers realize the debugging of the robot and teaching media, and indirectly deliver teaching information to learners by controlling the robot and teaching media, and do not get feedback directly based on the learning situation. By setting task-based learning objectives for learners to learn independently, learners can operate the teaching media or robot in real-time from the online resource base, and expert knowledge base to obtain knowledge information related to the teaching content.

Discovery Learning[24]

Teachers or robots do not directly teach teaching content, learners through the operation of the teaching media or robot demonstration obtain the basic teaching information, and then in turn for observation, assumptions, attempts, verification, adjustment, summarization, and a series of learning activities for discovery learning. The discovery learning model may also be accompanied by some kind of information dissemination process, allowing learners to communicate with the teacher through some kind of channel (robot or instructional media), requesting advice and obtaining guidance.

AI Teaching Assistance Robot Architecture

Building upon the aforementioned design scheme and harnessing the robot's speech and mobility capabilities, in conjunction with multimedia computers, this chapter introduces a novel model for robotic classroom lecturing, as illustrated in Figure 2.

Robot classroom lecturing represents an innovative educational approach specifically tailored to support teaching assistance for hearing-impaired children. This mode amalgamates the principles of robot-assisted teaching and traditional classroom lecturing, effectively integrating robotics and multimedia technology [25], [26]. On one front, the robot takes on the role of a teacher, performing educational tasks akin to an instructor. It aids teachers in explicating course content, which encompasses classroom introductions, posing questions, clarifying key concepts, and more.[27] Furthermore, it supports teachers in monitoring students' learning dynamics and classroom behavior, ensuring classroom discipline. The robot mimics the teacher's actions, mirroring their movements and following a predetermined path throughout the classroom.

Simultaneously, this approach seamlessly incorporates existing teaching media. The robot employs multimedia equipment to synchronize voice and physical movements with multimedia courseware presentations.[28] Through text, images, audio, video, and animations, the robot effectively conveys instructional information. This immersive approach encourages students, including hearing-impaired children, to actively engage with the learning materials. The ultimate aim is to create an educational environment conducive to independent exploration and learning, guided by the robot's interactive approach, fostering the development of problem-solving skills and encouraging a dynamic, participatory learning experience.

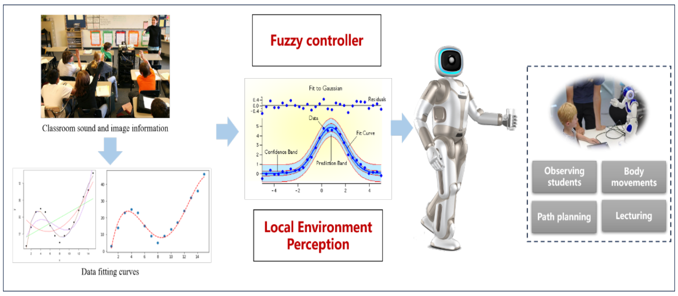

Fig. 2: Robot-assistedm Instruction Model Diagram

Throughout the teaching process, the human teacher changes from the original teaching leader to a teaching counselor, who observes the students' learning on the stage or in the laboratory, and gives timely guidance to the students when they encounter problems.[29, 30] This not only shares the teaching task of the teacher but also improves teaching efficiency.

According to the robot function setting above, robots with anthropomorphic and interactive features were selected for teaching practice in this study, and since there are not many commercially available robots on the market that meet the functional requirements, the results of screening and testing are shown in Table 2. Among them, ASIMO is a humanoid robot developed by Honda, Japan, with bipedal upright walking, programming control and voice recognition; Alter is a bionic humanoid robot jointly developed by Osaka University and the University of Tokyo, Japan, with voice recognition and gesture mimicry and other anthropomorphic interactive functions; Pepper is a humanoid robot jointly developed by SoftBank Group of Japan and Aldebaran Robotics of France, with anthropomorphic and interactive functions; Pepper is a humanoid robot jointly developed by SoftBank Group of Japan and Aldebaran Robotics of France, with anthropomorphic and interactive functions. Pepper is a humanoid robot jointly developed by Japan's Softbank Group and France's Aldebaran Robotics, with voice and expression recognition, touch screen interaction, and other functions. The first two robots lack intelligent front-end display and interaction functions, and can only carry out human-machine interaction through the background programming control interface and voice input, while Pepper not only has an interactive touch screen, but is also configured with visual sensors, infrared sensors, tactile sensors, etc., and at the same time, it can be used to develop more customized tools through the specific software development toolkit and is conducive to the collection of data on teaching activities, which is more in line with the requirements of the relevant functions of teaching The relevant functional requirements of assistive robots, so this study uses this robot to carry out teaching practice activities, see Figure 3.

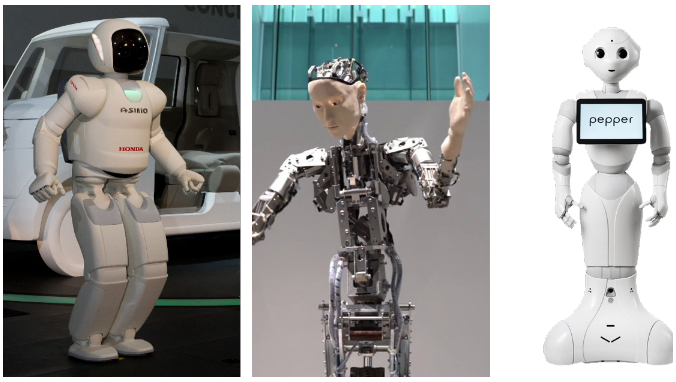

Fig. 3: Robot Appearance
(from left to right, ASIMO, Alter, Pepper)

In conducting the cost-benefit analysis, we meticulously examined initial setup costs, ongoing maintenance costs, and potential benefits in terms of instructional quality, student engagement, and learning outcomes.

Initial setup costs: Hardware investment: Includes the acquisition of AI-assisted instructional robotics equipment and associated sensor technology, and is estimated to cover the cost of purchasing and configuring hardware such as the robot body, camera, and sound sensors. Software development and customization: Costs for system development, software customization, and user interface design are considered to ensure that the system meets educational needs and is compatible with the school environment.

Human Resources: Salary and training costs for project management, technical support, and training teams were estimated to ensure that the educational institution has adequate professional support.

Ongoing maintenance costs: Hardware Maintenance and Updates: Includes regular maintenance of robotics equipment and sensors as well as updates and upgrades as necessary to ensure continued stable system performance. Software Updates and Technical Support: Regular updates, maintenance, and technical support for the system software are considered to maintain the functionality and safety of the system.

Training Costs: Consideration has been given to the cost of training faculty and staff to fully utilize and operate the AI-assisted instruction system.

Potential Benefits: Increased quality of instruction: Expected to improve the quality of instruction and promote deeper student understanding through real-time feedback, personalized learning paths, and resource adaptability.

Increased Student Engagement: It is expected that the introduction of AI-assisted instructional robots will stimulate student interest, increase engagement, and help create a more interactive learning environment.

Improved Learning Outcomes: By monitoring students' progress, strengths, and weaknesses in real-time, the system is able to personalize support and is expected to improve students' academic performance.

Combining the costs and potential benefits of these aspects, we provide a detailed and comprehensive economic model to help decision-makers weigh the costs and benefits when making investment decisions and to ensure that the AI-assisted instruction system invested in is both economically sound and able to provide substantial educational benefits to students.

METHODOLOGY

Teaching Aid System Software Design

The structure of the client software part is shown in Figure 4, and the software functions are mainly realized through the Qt Creator program. Through the way of multiple threads to realize each function module, which is conducive to the smooth operation of the program. The system software functions include system login, image acquisition, data transmission, image recognition, data saving, and data query.[31]

In the first part, the service program receives the login request from the client, and after passing the authentication, the service end transmits the collected image data to the client through a wired network. The second part is the data acquisition module of the embedded experiment platform, this module uses an ARM ARM-embedded experiment platform, connects with a Logitech camera through a USB interface, and realizes image acquisition by calling the service program. In the third part, after receiving the image data, the client needs to process the received image data and display the recognition results. In the fourth part, while displaying, we save the received image data and recognized data. In the fifth part, we save the recognized results as the basis of classroom performance in the database and can display the classroom information through the query function.

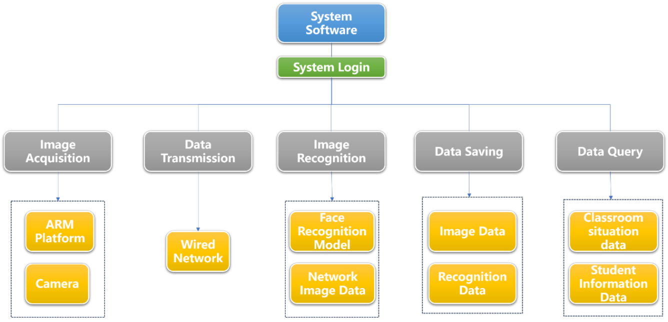

Fig. 4: Software Architecture

In this system, the image acquisition platform uses the embedded experimental platform, and the embedded system is based on the Linux kernel system, we need to choose the appropriate image acquisition framework, and the V4L2 framework is one of the commonly used frameworks, which can be applied to our image acquisition program. the V4L2 framework mainly supports the acquisition method of direct data reading and the acquisition method of memory mapping.

The direct data reading method is mainly used to capture static images. The memory-mapped way is suitable for continuous video data acquisition. As we need to save the video information and continuous image data, we need to use the memory-mapped acquisition method. And use multi-threading technology to realize that the thread for capturing images and the thread for sending data are separated, which can improve the processing speed and is conducive to the stable operation of the server-side program.

Adaboost-based Face Detection Algorithm

Face detection is an important node in face recognition technology, its role is to determine whether there is a human face in an image. When we detect an image in which there is a human face, we will detect the face information recorded, the information content mainly includes the characteristics of the human body's five senses and the geometric position of each other, that is, what we call the size of the face and location information. AdaBoost is an iterative algorithm, the basic idea is to train different classifiers for the same training set, which we call weak classifiers, by putting all these weak classifiers together to form a final classifier, we will use the same training set to train different classifiers. The basic idea is to train different classifiers, which we call weak classifiers, for the same training set, and then we assemble all these weak classifiers to form a final classifier, which we call a strong classifier, so as to realize face detection.

Face detection is an important node in face recognition technology, its role is to determine whether there is a human face in an image.[32, 33] When we detect an image in which there is a human face, we will detect the face information recorded, the information content mainly includes the characteristics of the human body's five senses and the geometric position of each other, that is, what we call the size of the face and location information. AdaBoost is an iterative algorithm, the basic idea is to train different classifiers for the same training set, which we call weak classifiers, by putting all these weak classifiers together to form a final classifier, we will use the same training set to train different classifiers .[34, 35] The basic idea is to train different classifiers, which we call weak classifiers, for the same training set, and then we assemble all these weak classifiers to form a final classifier, which we call a strong classifier, to realize face detection.[36]

Only a portion of the features obtained from our computation using the integral map can be used to construct a classifier. To solve this problem, we use the AdaBoost algorithm to obtain a more discriminative subset of features to construct more efficient classifiers. This is achieved by continuously changing the weight distribution of the training data to construct weak classifiers, by linearly combining the weak classifiers to form strong classifiers, and by continuously iterating the error rate of the strong classifiers to a certain desired value.

If we have a training dataset of T, as shown in Equation (1), where $X_i \in R_n$, i represents the feature vector of the sample, n represents the dimension, and $Y_i \in \{+1,-1\}$ represents the category label of sample i. To ensure that all the samples have the same weights during the learning process of the classifier, we need to initialize the weight distribution of the training samples. As shown in Equation (2), it can be used to iteratively learn the basic classifier, and the classification error rate in the kth round is computed, where $G_k(X)$ is the weak classifier of the current iteration, i is the trained sample, and w_{ki} is the weights of i.

$$T=\{(X_1,Y_1),(X_2,Y_2),\cdots,(X_M,Y_M)\} \tag{1}$$

$$R_{error\_k}=\frac{\sum G_k(x_i)\neq w_{ki}}{\sum_{i=1}^{m} w_{ki}} \tag{2}$$

Generally, the process of detecting faces consists of feature extraction and classification, it is very easy to extract features using an integral map and build a classifier on top of feature extraction. AdaBoost improves face detection's speed and performance by cascading multiple weak classifiers to form a cascading classifier. This cascaded classifier is similar to a degenerate decision tree, as shown in Figure 5. The classifiers in each separate stage are trained using AdaBoost and tuned to minimize the leakage rate, and for each position, the n+1th classifier is entered only if the nth classifier does not exclude it.

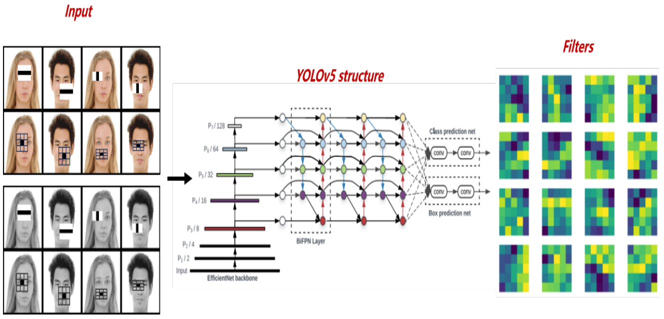

Fig. 5: Cascading Classifiers

SVM-based Face Recognition Algorithm and Implementation

Support Vector Machine abbreviated as SVM, also known as Support Vector Machine, is a class of classifiers that uses a loss function to compute empirical risk using supervised learning and optimize the structural risk, it is a robust classifier that can be used to achieve binary classification of data, Support Vector Machines can be applied to small samples as well as high-dimensional spaces, so it is more appropriate to use it for face recognition.

The classification process of the support vector machine is to solve for the maximum margin hyperplane of the learning samples, that is, to separate the two categories in the hyperplane that can be found in the two categories of data points in the longest distance, and the number of such hyperplanes is not unique. The data points we are looking for are, in each category, the data points with the shortest distance to the hyperplane.

After we use PCA for face image dimensionality reduction, we need to select a suitable classifier to realize face classification and recognition, for the multi-classification problem, we use SVM classification using the one-to-one voting method. To avoid the phenomenon of a person corresponding to multiple identity information, the classifier is trained using one-to-one voting, if the maximum number of votes has M categories, we can derive the sample with the smallest distance by calculating the distance between the sample and the mean of the M training samples, and the sample is the recognition sample, the recognition process is shown in Figure 6.

In this paper, we mainly compare different kernel functions and different c-values, and the experimental platform we chose is Matlab software, which has the advantages of many kinds of simulation, good compatibility, and beautiful graphical interface, etc. Before the experimental simulation, we need to prepare the face sample library to be used. Before the experimental simulation, we need to prepare the face sample libraries to be used in the experiment, the face libraries used in this experiment are mainly the ORL face sample library, the YALE face sample library, and the homemade face library. In the ORL face library, there are 400 face sample images, of which 200 face sample images are used for training and the remaining face sample images are used for prediction. In the same way, in the YALE face sample library, we select 75 of the face sample images for training and the remaining 75 for prediction, and in the homemade face sample library, we select 675 of the face sample images for training and the remaining 675 for testing. The specific face sample information is shown in Figure 7.

Data Analysis

In the SVM face recognition process, we need to pay attention to the selection of kernel function. The experiment uses the image information after PCA dimensionality reduction and the SVM method for face

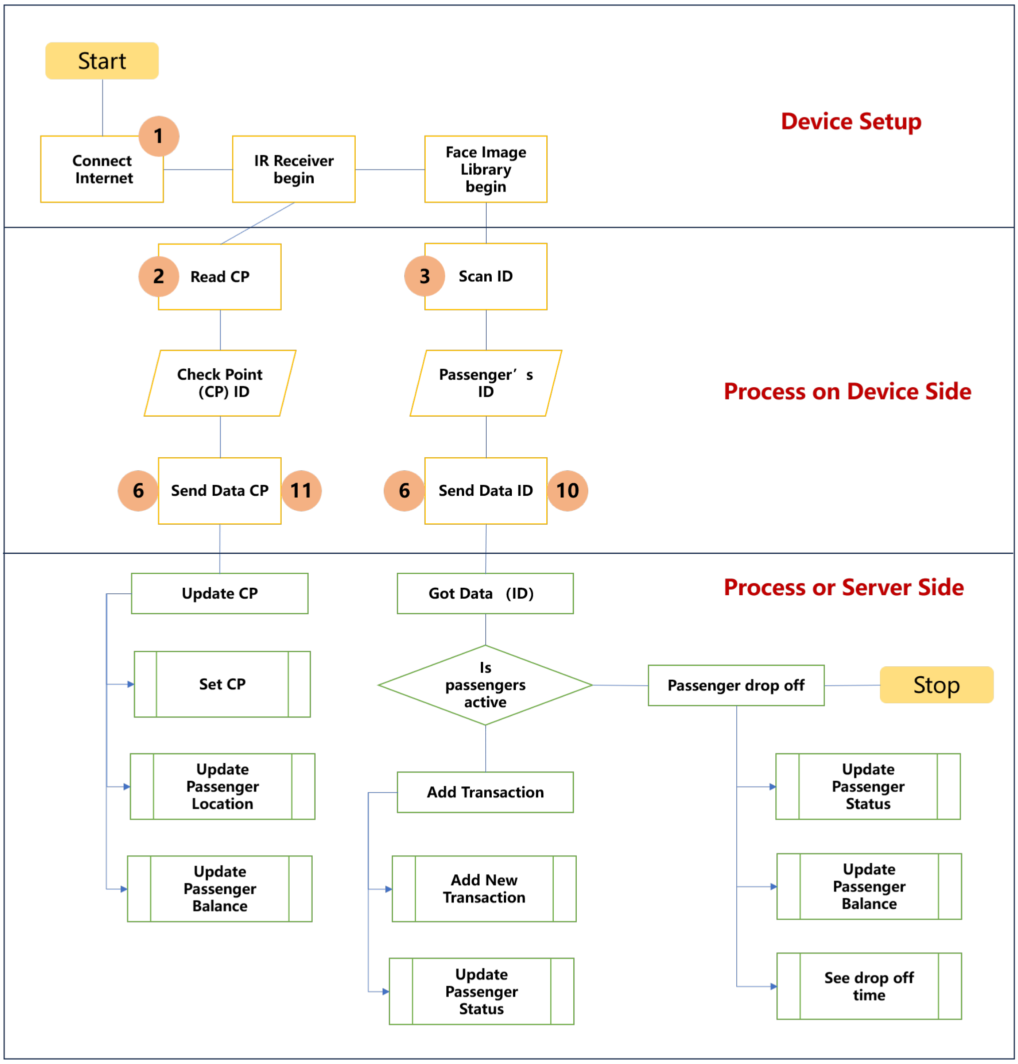

Fig. 6: Identification Process

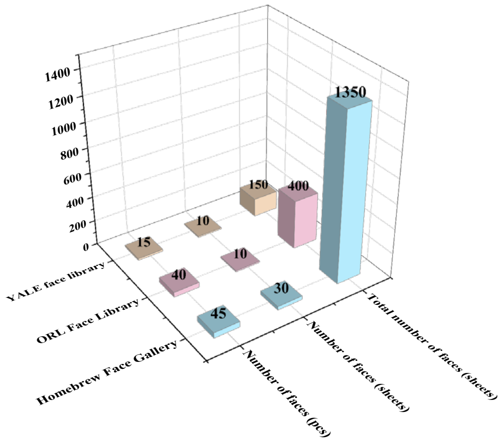

Fig. 7: Face Sample Data Details

classification and recognition. In this paper, we mainly focus on four kinds of kernel functions, namely, linear kernel function, polynomial kernel function, RBF Gaussian kernel function, and sigmoid kernel function, and compare the recognition results, to come up with the recognition of different kernel functions, as shown in Figure 8.

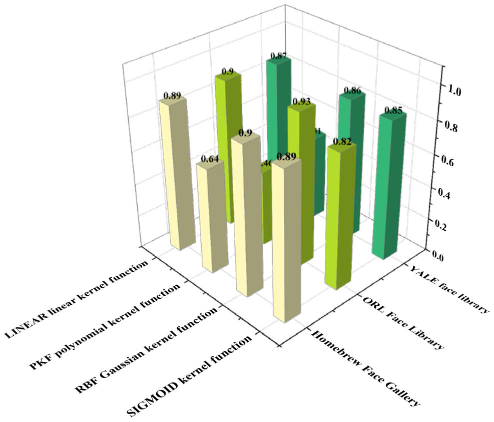

Fig. 8: Recognition of Different Kernel Functions

From Figure 8, we can see that we use the RBF Gaussian kernel function to achieve face recognition, and its recognition rate is higher than other kernel functions; the linear kernel function and the sigmoid kernel function have similar recognition rates for the two kernel functions when recognizing the YALE sample library against the homemade face library and the ORL sample library; the recognition rate of the two kernel functions is the same when recognizing the homemade face library; the polynomial kernel function has the lowest recognition rate, which increases as the number of faces in the face sample library increases, but the difference in recognition rate is large compared to the other kernel functions.

RESULTS AND DISCUSSION

Characterization of Student Classroom Scenario Data

In various applications, the data's characteristics can vary due to specific educational scenarios and requirements. For instance, in a teaching assistance system for hearing-impaired children, data characteristics predominantly revolve around visual and text-based information. The focus of this paper is within a classroom setting, where a camera is positioned directly in front of the robot to capture the video content of students during lessons. The objective is to detect and record students' gestures and visual interactions.

As this classroom student behavior association method proposed in this paper lacks established evaluation standards and corresponding test sets for hearing-impaired children, the precision of behavior association is evaluated through researcher observation and statistical analysis, primarily addressing two crucial aspects: the accuracy of behavior association and the processing time for a single frame image.

Additionally, we have established strict access controls to ensure that authorized personnel can access student data only when necessary. We record and monitor access to the data, as well as any changes made to the data, to safeguard the integrity and traceability of the data.

Regarding data storage, we choose secure and reliable cloud storage service providers and ensure that they comply with international standards and industry best practices. These service providers not only provide a high level of data security, but are also equipped with disaster recovery mechanisms to ensure that data is effectively protected under any circumstances.

Finally, we assess the security of our systems through regular security reviews and vulnerability scans, and promptly update and patch any potential security vulnerabilities. This comprehensive set of measures aims to ensure that student data is protected with maximum security while complying with relevant regulations and ethical guidelines.

To ensure diversity in the test data, the study constructs a test set by recording classroom videos in different educational contexts. This ensures that the students represented exhibit various communication styles and gestures, which are especially pertinent in assisting hearing-impaired children. The recorded videos are referred to as Video 1 and Video 2. Images are extracted from these videos at regular intervals to compile a dataset, including 579 images from Video 1 and 571 images from Video 2. These images are subsequently employed for testing, enabling an analysis of the accuracy of behavior association and the processing time for a single frame.

During testing, the images are sequentially input into the network, and each image is annotated with associated student IDs, facilitating observation and statistical analysis by researchers, particularly when dealing with hearing-impaired children. The evaluation process involves comparing the IDs assigned to students in the first frame with those of students in subsequent frames, assessing the accuracy of individual student associations. This approach is particularly crucial for understanding and enhancing the learning experiences of hearing-impaired students within the classroom. The accuracy of student behavior association is calculated using a specific formula, where “cori” represents the number of individuals correctly associated in the i-th image, “all” indicates the total number of associated students in the i-th image, “n” stands for the number of test images, and “ActLinkrate” signifies the precision of student behavior association in assisting hearing-impaired children in their learning journey.

$$ActLink_{rate} = \frac{\sum_{i=2}^{n} cor_i}{\sum_{i=2}^{n} all_i} \tag{2}$$

In this paper, the proposed behavioral association method consists of two steps: firstly filtering candidate targets by position distance, then comparing the feature similarity of the candidate targets, and confirming whether the candidate targets are the same body or not by the position distance and feature similarity between the targets. In this paper, experiments are carried out on positional distance, positional distance combined with high-level features, and positional distance combined with low-level features, and the experimental results are shown in Table 2, which are analyzed in terms of the accuracy of behavioral association and single-frame processing time.

Statistical Analysis of Behavior

To ensure that the assessment of the behavioral associations of children with hearing impairment is accurate and effective, we have developed a comprehensive assessment framework that takes into account their special needs and learning environment. The framework is based on seven key areas:

First, we pursued children’s active participation in learning and interaction with the AI robot by setting assessment goals and corresponding indicators, including visualization rate, response speed, and conversation frequency.

Secondly, concentration and learning effectiveness were focused on ensuring that they could effectively absorb the content by assessing gaze time, task performance,

Table 2. Table of Results of Behavioral Association Experiments with Different Modalities

Method	Video 1		Video 2	
	Accuracy	Time/ms	Accuracy	Time/ms
Position distance	72.69%	30	69.97%	32
Position distance + P1 advanced features	75.36%	46	73.25%	48
Position distance + P2 advanced features	76.21%	48	76.26%	36
Position distance + P3 advanced features	75.32%	48	77.31%	38
Position distance + P4 advanced feature	78.54%	48	78.11%	38
Position distance + color feature	89.21%	42	90.31%	36
Position distance + edge feature	82.31%	45	84.65%	41
Position distance + (color, edge)	86.54%	44	80.23%	40

learning content comprehension, and classroom engagement.

Quality of emotional expression and engagement is the third area of focus, where we assess students' ability to express emotions in learning and the quality of their interactions with the AI robot, including expression diversity, emotional responses, and active participation.

Social skills and cooperation form the fourth part of the assessment framework, where we assess the degree of student's social skills and cooperation with peers and the robot, including interaction, social skills, and shared communication.

Task Completion and Feedback Acceptance and Adjustment are the fifth areas of focus, ensuring that students are effectively engaged in learning by assessing the accuracy of task completion, feedback response, and behavioral adjustment.

Finally, to ensure the cultural adaptability of the assessment system, we specifically considered the cultural sensitivity of engagement. We verified the validity of the above indicators in different cultural contexts to ensure the universality and comparability of the assessment results.

This comprehensive assessment framework aims to provide an in-depth understanding of how children with hearing impairment interact and learn in the teaching and learning process, respecting their special needs to ensure the most appropriate and effective educational support.

Students' behavior reflects the level of interest in the class and the motivation to listen to the class. By counting students' behaviors in the classroom and forming effective feedback to help lecturers analyze after class, the quality of teaching can be effectively improved. In this paper, we analyze the students' interest level in the class by counting the students' behaviors in the whole classroom, including the percentage of each behavior in the classroom and the changes in the behaviors, and then analyze the students' interest level in the class.

$$Actcls_{rate} = \frac{\sum_{i=2}^{n} Aact_i}{\sum_{i=2}^{n} Allact_i} \tag{3}$$

In this section, using the example of the video presented in the previous section, we elucidate the method employed in this paper for statistically analyzing classroom student behaviors. By calculating the proportions of all student behaviors during a class session and providing the results to the instructing teacher, the teacher can gain a visual understanding of the overall performance of students in the class. The method for calculating the overall proportion of student behaviors in this paper is shown in Equation (4), where "Aacti" represents the quantity of a certain type of behavior in the i-th image, "Aact" comprises five categories, including raising the head, lowering head, standing, leaning on the desk, and turning head. "Allacti" denotes the total number of behaviors detected in the i-th image, "n" represents the number of images detected, and "Actclrate" represents the proportion of a certain type of behavior in the entire class.

As shown in Figure 9, the rate of raising heads in this class session reaches 62.57%, indicating a generally positive overall behavior. However, the rate of lowering heads is relatively high at 33.06%, suggesting that students often lower their heads. In addition to the overall proportion of student behaviors, the variations in student behaviors during the class provide insights into changes in student interest and engagement. In this paper, student behaviors are quantified on a per-minute basis, with the proportion of behaviors calculated every minute. The changes in the overall proportion of student behaviors during the class are presented to the instructing teacher in the form of a line chart.

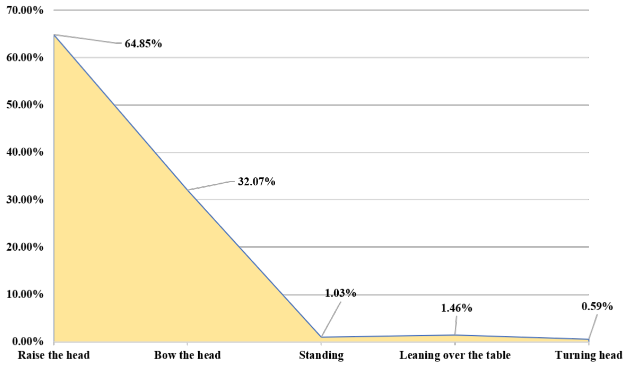

Fig. 9: Overall Classroom Student Behavior Percentage Chart

CONCLUSION

This paper outlines the system architecture of an AI-assisted teaching robot, specifically tailored to provide support in teaching assistance for hearing-impaired children. The robot leverages deep neural networks to record classroom teaching situations, with the primary goal of continuously refining teaching methods and elevating the quality of education through the analysis of classroom data. The face detection phase is facilitated by the AdaBoost algorithm, while facial classification and recognition rely on SVM. The paper conducts dedicated investigations into the AdaBoost and SVM algorithms, validating their results through experimental simulations, and focusing on their application for hearing-impaired students.

The software client, implemented with multithreading techniques, assumes responsibility for student facial image capture, classroom situation recording, and data retrieval. The image capture module utilizes the V4L2 framework, and the server transmits the captured image data to the client through a wired network, employing the socket network communication mechanism under the Linux system. To facilitate facial image processing, the paper establishes an Open CV development environment. This environment enables the training of models for recognizing student identities, tracking student behavior, and storing image data, all designed to cater to the needs of hearing-impaired children.

Expanding on the foundation of classroom student behavior detection, this paper takes strides in correlating individual student behaviors over continuous periods. This achievement hinges on the analysis of positional characteristics and feature similarities of students across video frames. The accuracy of behavior correlation exceeds 90%. To provide educators with a more comprehensive understanding of student's classroom engagement, the paper incorporates statistical analysis to quantify classroom behavior. This includes calculating overall behavior proportions, monitoring behavior changes, and assigning scores to behaviors to gauge students' attentiveness in class. These endeavors are dedicated to supporting teaching instructors in their quest to enhance teaching quality, particularly in the context of hearing-impaired students.

REFERENCES

[1] Tyagi, A. K., Fernandez, T. F., Mishra, S., & Kumari, S. (2020, December). Intelligent automation systems at the core of industry 4.0. In *International conference on intelligent systems design and applications* (pp. 1-18). Cham: Springer International Publishing.

[2] Ahmad, T., Zhang, D., Huang, C., Zhang, H., Dai, N., Song, Y., & Chen, H. (2021). Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities. *Journal of Cleaner Production*, *289*, 125834..

[3] Alam, A. (2021, December). Should robots replace teachers? Mobilisation of AI and learning analytics in education. In *2021 International Conference on Advances in Computing, Communication, and Control (ICAC3)* (pp. 1-12). IEEE.

[4] Urban, R., Haluzová, S., Strunga, M., Surovková, J., Lifková, M., Tomášik, J., & Thurzo, A. (2023). AI-assisted CBCT data manag ement in modern dental practice: benefits, limitations and innovations. *Electronics*, *12*(7), 1710..

[5] Chao, C. Y., Chen, Y. T., & Chuang, K. Y. (2015). Exploring students' learning attitude and achievement in flipped learning supported computer aided design curriculum: A study in high school engineering education. *Computer Applications in Engineering Education*, *23*(4), 514-526..

[6] Liu, F., & Yang, K. (2021). Exploration on the teaching mode of contemporary art computer aided design centered on creativity. *Computer-Aided Design and Applications*, *19*(S1), 105-116..

[7] Li, H., Cui, C., & Jiang, S. (2024). Strategy for improving the football teaching quality by AI and metaverse-empowered in mobile internet environment. *Wireless Networks*, *30*(5), 4343-4352..

[8] Chen, Y. (2020). Optimization of music teaching methods based on multimedia computer-aided technology. *Computer-Aided Design and Applications*, *18*(1)..

[9] Cao, H. (2022). Entrepreneurship education-infiltrated computer-aided instruction system for college Music Majors using convolutional neural network. *Frontiers in psychology*, *13*, 900195..

[10] Wang, Q. (2023). Computer Aided Teaching System for Football Teaching and Training Based on Video Image..

[11] Hennessy, S., Harrison, D., & Wamakote, L. (2010). Teacher factors influencing classroom use of ICT in Sub-Saharan Africa. *Itupale online journal of African studies*, *2*(1), 39-54..

[12] Mahajan, G. (2012). Multimedia in teacher education: Perceptions & uses. *Journal of Education and Practice*, *3*(1), 5-13..

[13] DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *37*(6), 582-601..

[14] Bennell, P. (2004). Teacher motivation and incentives in sub-Saharan Africa and Asia. *Knowledge and Skills for development, Brighton*, *1*(1), 1-52..

[15] Noah, H. J., & Eckstein, M. A. (2001). *Fraud and education: The worm in the apple*. Rowman & Littlefield..

[16] Zhexenova, Z., Amirova, A., Abdikarimova, M., Kudaibergenov, K., Baimakhan, N., Tleubayev, B., ... & Sandygulova, A. (2020). A comparison of social robot to tablet and teacher in a new script learning context. *Frontiers in Robotics and AI*, *7*, 99..

[17] Randall, N. (2019). A survey of robot-assisted language learning (RALL). *ACM Transactions on Human-Robot Interaction (THRI)*, *9*(1), 1-36..

[18] Chin, K. Y., Hong, Z. W., & Chen, Y. L. (2014). Impact of using an educational robot-based learning system on students' motivation in elementary education. *IEEE Transactions on learning technologies*, *7*(4), 333-345..

[19] Alimisis, D., & Kynigos, C. (2009). Constructionism and robotics in education. *Teacher education on robotic-enhanced constructivist pedagogical methods*, 11-26..

[20] Altin, H., & Pedaste, M. (2013). Learning approaches to applying robotics in science education. *Journal of baltic science education*, *12*(3), 365..

[21] Alam, A. (2021, December). Should robots replace teachers? Mobilisation of AI and learning analytics in education. In *2021 International Conference on Advances in Computing, Communication, and Control (ICAC3)* (pp. 1-12). IEEE..

[22] Dowden, T., Pittaway, S., Yost, H., & McCarthy, R. (2013). Students' perceptions of written feedback in teacher education: Ideally feedback is a continuing two-way communication that encourages progress. *Assessment & Evaluation in Higher Education*, *38*(3), 349-362..

[23] Meyer, B., Haywood, N., Sachdev, D., & Faraday, S. (2008). What is independent learning and what are the benefits for students. *Department for Children, Schools and Families Research Report*, *51*, 1-6..

[24] Svinicki, M. D. (1998). A theoretical foundation for discovery learning. *Advances in physiology education*, *275*(6), S4..

[25] Liang, J. C., & Hwang, G. J. (2023). A robot-based digital storytelling approach to enhancing EFL learners' multimodal storytelling ability and narrative engagement. *Computers & Education*, *201*, 104827..

[26] Selwyn, N. (2019). *Should robots replace teachers?: AI and the future of education*. John Wiley & Sons..

[27] Kagan, D. M. (1992). Implication of research on teacher belief. *Educational psychologist*, *27*(1), 65-90..

[28] Chang, C. W., Lee, J. H., Chao, P. Y., Wang, C. Y., & Chen, G. D. (2010). Exploring the possibility of using humanoid robots as instructional tools for teaching a second language in primary school. *Journal of Educational Technology & Society*, *13*(2), 13-24..

[29] Blase, J., & Blase, J. (2003). *Handbook of instructional leadership: How successful principals promote teaching and learning*. Corwin Press..

[30] Brophy, J. (1988). Educating teachers about managing classrooms and students. *Teaching and teacher Education*, *4*(1), 1-18..

[31] Inoue, T. (2018). TI Workbench, an integrated software package for electrophysiology and imaging. *Microscopy*, *67*(3), 129-143..

[32] Yow, K. C., & Cipolla, R. (1997). Feature-based human face detection. *Image and vision computing*, *15*(9), 713-735..

[33] Wiskott, L., Fellous, J. M., Krüger, N., & Von Der Malsburg, C. (2022). Face recognition by elastic bunch graph matching. In *Intelligent biometric techniques in fingerprint and face recognition* (pp. 355-396). Routledge..

[34] Collins, M., Schapire, R. E., & Singer, Y. (2002). Logistic regression, AdaBoost and Bregman distances. *Machine Learning*, *48*, 253-285..

[35] Wu, P., & Zhao, H. (2011, January). Some analysis and research of the AdaBoost algorithm. In *International Conference on Intelligent Computing and Information Science* (pp. 1-5). Berlin, Heidelberg: Springer Berlin Heidelberg..

[36] Tuzel, O., Porikli, F., & Meer, P. (2006). Region covariance: A fast descriptor for detection and classification. In *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part II 9* (pp. 589-600). Springer Berlin Heidelberg..

[37] Kavitha, M. (2024). Advances in wireless sensor networks: From theory to practical applications. *Progress in Electronics and Communication Engineering, 1*(1), 32–37. https://doi.org/10.31838/PECE/01.01.06

[38] Barhoumi, E. M., Charabi, Y., & Farhani, S. (2024). Detailed guide to machine learning techniques in signal processing. *Progress in Electronics and Communication Engineering, 2*(1), 39–47. https://doi.org/10.31838/PECE/02.01.04

[39] Chakma, K. S. (2025). Flexible and wearable electronics: Innovations, challenges, and future prospects. Progress in Electronics and Communication Engineering, 2(2), 41–46. https://doi.org/10.31838/PECE/02.02.05

[40] Muralidharan, J. (2024). Machine learning techniques for anomaly detection in smart IoT sensor networks. *Journal of Wireless Sensor Networks and IoT, 1*(1), 15-22. https://doi.org/10.31838/WSNIOT/01.01.03

[41] Kumar, T. M. S. (2024). Integrative approaches in bioinformatics: Enhancing data analysis and interpretation. *Innovative Reviews in Engineering and Science, 1*(1), 30-33. https://doi.org/10.31838/INES/01.01.07

[42] Rahim, R. (2024). Review of modern robotics: From industrial automation to service applications. *Innovative Reviews in Engineering and Science, 1*(1), 34-37. https://doi.org/10.31838/INES/01.01.08