

¹Kaikai Liang

Exploration of Emotion Analysis and Personalized Emotion Adjustment Algorithm in English Learning Process



Abstract: - An important component of human existence and behaviour is emotion. The automatic identification of emotions has gained significant importance in the domains of affective computing and human-machine interaction in recent times. Getting pictures of people's facial expressions is one of the simplest and least expensive methods for identifying emotions among the several physiological and kinematic data that may be utilised. Because of variations in anatomy, culture, and environment, developing a generalised, cross-subject model for emotion identification from facial expression remains difficult. This study presents an intelligent classroom that uses e-learning to help persons with autism spectrum condition strengthen their emotional abilities. Monitoring autism child engagement and participation in class while looking for signs of attention is necessary for effective classroom education. Teaching quality is increasingly being determined by the teachers' capacity to assess and evaluate the behaviour of their pupils in the classroom. Modern classroom systems are enhanced with the most recent technologies in educational institutions to make them more interactive, student-centered, and personalised. Even with these tools, it can be challenging for teachers to gauge their pupils' levels of interest and focus. This study makes use of contemporary technology to establish a real-time, intelligent vision-based classroom that can track students' moods, attendance, and levels of focus even when they are wearing face masks. In order to evaluate students' attention or lack thereof in a classroom, we trained AlexNet models that recognise student behaviour, including recognising facial expressions. Another designed to manage the irregular relationship between pupils and activities is the Interaction Students Education Neural Network (Int_Edu_NN). The Q-matrix is used to firstly determine the pupil variable and the assignment indicator. This network is analyzed using two datasets in terms of various parameters and hence the proposed Int_Edu_NN achieves 99.% and 98% of accuracy.

Keywords: student education, Emotion detection, neural network, action detection, interaction model, class room monitor.

I. INTRODUCTION

For social interactions in society, representations gleaned from a person's facial expressions are crucial [1]. Nonverbal communication relies heavily on emotional expression and eye gaze direction [2]. Facial emotion identification is a fundamental idea in the study of social cognition, claim the authors of [3]. Autism spectrum disease (ASD) sufferers may exhibit repetitive behaviour patterns, difficulty interacting with others, poor verbal and nonverbal communication, and inadequate social-emotional reciprocity from an early age [4]. Throughout their lives, people with ASD may struggle to learn and develop critical social emotional skills [5]. Growing up requires the ability to recognise facial reactions, though as stated in [6], conventional techniques and conversations in groups can prove to be successful at these kinds of duties. Multiple avenues of study have been developed with the goal to propose strategies that utilise the application of technology for communication and information (ICT) for the individualised enhancement of emotional skills [7]. According to [10], these methods allow the user to insert themselves into the real world more comfortably, entertainingly, and safely. Additionally, these tools can be utilised on a variety of hardware and software, including tablets, desktop computers, and smartphones. According to data from the study shown in [8], our society is adopting strategies developed with ICT in digital game environments. E-learning (EL) is one of the strategies used in this situation. The authors of [9] claim that ELs are both entertaining and a way to advance facial emotion identification abilities. Additionally, computational methods used for face emotion identification can benefit devices like ELs. The development of new technologies that can incorporate ways for enhancing emotion-related skills with SGs and facial expression detection methods in a way that quantifies each patient's data throughout therapy sessions has hurdles.

One of the most significant platforms for video communication in today's culture is the video management system[10]. With the continued advancement of information technology, major video websites, businesses and organisations with pertinent requirements, elementary and secondary schools, and all types of training facilities have established their own video resource management systems, enabling networking and information management for video release and on-demand [11]. These days, network classrooms and micro classes are common, making a good system for managing instructional video resources all the more important. Currently, colleges and universities have essentially created their own websites with instructional resources[12]. Micro

¹ School of Foreign Languages, Zhengzhou University of Science and Technology, Zhengzhou, Henan, China, 450064

*Corresponding author's e-mail: liangkaikai521@126.com

classes have emerged as the primary network teaching strategy as a result of the "flipped classroom" teaching mode's popularity and the advancement of network technologies. Students can examine and review at any time and anyplace thanks to its special advantages of brief material and various class durations [13]. In advanced countries, many elementary and high school and academic facilities utilize a separate instructional video management program, through which pupils could indeed view and study online videos, as well as pose questions and join virtual discussions [14]. Higher education institutions also have their own comprehensive learning materials and multimedia management system. Currently, higher education institutions have essentially created their own websites with instructional resources[15]. These internet video processes for learning and instruction are primarily made out of two main categories: first, the skills training segments of the largest video websites; while these sections have a wealth of educational content, the classification is disorganised and the quality of the content is uneven, with some sections having good and others having bad content; second, the video management systems used by various educational institutions or online education platforms; these systems are very constricting for users, with the majority of them finding them to be very frustrating[16]. Excellent instructional materials cannot be supplied to students in a timely manner, which impedes the advancement of modernization and information technology in education. In order to effectively manage teaching video materials in higher education institutions, advanced innovation must be used [17].

1.1 Contributions of the paper

In this investigation, researchers provide an online course that helps people with autism spectrum disorders (ASD) develop their emotional intelligence. The individual's identity is initially captured by a video camera. Hence, Interaction Students Education Neural Network (Int_Edu_NN) that can adequately capture knowledge on how individuals engage with activities. The introduction of the monotonicity criterion aims to improve the interpretability of the identified child's conditions. If the factors that follow were tweaked, e-learning might prove beneficial and successful: increasing student and teacher liberty and convenience, by improving the contents of educational resources to enhance communication across all study elements. By efficiently arranging courses, preparing students, and expanding interactivity techniques for learning by integrating rich media and synchronous conversations, teachers can improve e-learning. To learn the distinguishing characteristics shared by face photos of various resolutions, the AlexNet is trained just once. It is built using recently suggested residential components. Our online course examines multimodal elements like body language, verbal intonation, and facial expressions. Following accurate portrayal of the mood, condition, and flaws, the user receives favourable feedback.

The present content is organised as follows: The second part of the presentation with tables offers a pertinent compilation of studies for autism student education utilising neural networks. Using superior emotion/action acknowledgment, a proposed Interaction Students Education Neural Network (Int_Edu_NN) based interaction model between autism students and teachers is provided in Section 3. The performance of the proposed model is displayed together with benchmarking techniques in part 4. The overall recommendation for the recommended approach is presented in the fifth subsection.

II. RELATED WORKS

The use of computer technologies to improve skills has increased in recent years. In practical settings, such as personalised student tutoring [18], advance detection of academic qualification [19], and student choice prediction [20], cognitive diagnostic of student status is a crucial and basic role. To address these issues, some traditional models of cognitive diagnostics have been suggested.

The work of [21] demonstrates, through a comprehensive study, how technology can be used to enhance people with ASD's conceptual skills, practical skills, social skills, and general skills. This study demonstrates that individuals with ASD are frequently inspired by new technology, and that using these tools to develop abilities can be enjoyable. For people with ASD, the usage of technological advances like artificial intelligence and augmented reality unquestionably creates a comfortable setting that encourages continuous learning. A comprehensive mapping investigation of the use of technologies for emotion recognition in kids with ASD is also demonstrated in the study in [22]. The study analysed the primary methods used and offered suggestions for equipment and user interface. These techniques can offer adaptability, accessibility, and simple real-world scenario adaptation. The primary studies that looked into SG tactics for people with ASD are presented in this section. The authors of [23] presented a method that would use 3D animations superimposed on the face to help people with ASD develop their facial expression skills. In this tool, a 3D face mask was created using photos

taken from the user's frontal and lateral facial regions. The participant's face micromovements were represented in the 3D animations using the facial action coding system (FACS) hypothesis. The overlay of the 3D facial image made it possible to assess how the participants' expressed their emotions The authors of [24] created an application employing machine learning techniques to assist people with ASD in learning how to feel neutral, happy, sad, and angry. The representations were recorded on video, and the FKs of the users' faces were found using the Viola and Jones technique. These were categorised using the random forest technique. The authors created social circumstances that called for the portrayal of emotions in order to make the game appealing. In [25] suggests a revolutionary Student Academic Performance Predicting (SAPP) system. It has a design a proper and forecasts whether pupils will pass or fail using a 4-layer stacked Long Short Term Memory (LSTM) network, Random Forest (RF), and Gradient Boosting (GB) technique combo. Users' involvement with E-learning is modelled in [26] utilising the Quality of Interaction (QoI) indicator as part of a blended and collaborative learning strategy. The Temporal Convolutional Neural Networks (T-CNN) model evaluation offers insights to consumers and instructors to improve the educational environment. [27] suggests using homogeneous hierarchy divergence, also known as HHSKT, to trace the learner's short-term attentional information. We'll employ short-term memory improvement and hierarchical heterogeneous procedural knowledge to describe how diverse contact patterns affect learners. For individualised course study planning, the authors of [28] suggest the Complexity-based Attentive Interactive Student Performance prediction model (CAISP). The individual characteristics between pupils can be taken into account in CAISP by employing an attention network to assign dynamically different priorities and to enable individualised planning. In [29], the concept of performing automated analysis of a student's comprehension of the class taking part in active face-to-face classroom instruction is explored. It is essential in classroom instruction that every student can comprehend the lecture. The student's facial photos for the current study were taken throughout the lecture. Students assisted in labelling the facial photos based on how well they understood the lecture at the time. A three-dimensional DenseNet self-attention neural network (DenseAttNet) is trying to provide in [30] in order to recognise and assess student engagement in both contemporary and conventional educational programmes. Studies using the GA-BP neural network technique are found in [31]. It begins by providing a quick summary of the state of online education and the GA-BP neural network method. This article suggests a much more scientific virtual learning classroom attempting to teach quality assessment optimization technique as well as, in the end, validates the validity of the virtual learning instructional assessment method through the use of it in a university. Second, through to the inquiry of the internet schooling system in several aspects, it evaluates students' online learning classroom teaching quality from five aspects. Convolutional neural networks (CNNs), a fundamental deep learning model architecture, are presented in [32] as CNN Explainer, an interactive visualisation tool meant for non-experts.

The analysis of interactions data in the aforementioned works is not thorough enough, and some exercise elements (such as estimate and glide component) as well as the combination of different training characteristics are not modelled. Thus, the following structural structure is suggested.

Table-1 evaluation of existing methods

Author/year	Method	Advantage	disadvantage
Kukkar et al., (2023)	Student Academic Performance Predicting (SAPP) system	Simple and straightforward approach	For dealing with tags, a largelabeled information is required.
Awadet al.,(2023)	Temporal Convolutional Neural Networks	It is resistant to being overfit.	sluggish of forecasting engine
Ni et al., (2023)	homogeneous hierarchy divergence	outperforms a single classifier in performance	Users need large things to achieve greater results.
Wang et al., (2022)	attention network	It lowers variability.	Network training is challenging
Sethi et al., (2022)	DenseAttNet	More precise results	Number of iterations is more
Mehta et al., (2022)	GA-BP neural network	More accuracy	More cost
Xu et al., (2021)	Convolutional neural networks	Needs statistical factor	More computation cost
Wang et al., (2020)	Deep neural network	Access to multiple layers	Overfitting problem arises

III. SYSTEM MODEL

The three layers that make up the suggested framework for emotion detection are the (i) cloud layer, (ii) fog layer, and (iii) internet of things layer. IoT and fog are the two primary levels. There is a proposed mobile assistant application for IoT that can be used to take a picture of a youngster while using a smart device and send that picture of their face to the fog layer. According to the suggested DL approach, a controller called the fog server (FS) is in charge of obtaining this face image and identifying the emotion. The pretrained dataset is also stored in a database (DB) that is present in the fog layer. When anger, fear, sadness, or surprise are detected as emotions, the fog server sends an alert message to the parent device. The overall block diagram for the proposed emotion recognition based interaction improvement among students using AffectNet and ASSIST2009 dataset is shown in figure 1. Thus, Discrete Cosine Transforms are used to pre-process those datasets. After preprocessing, facial emotion and action recognition using Deep Coupled AlexNet, which generates multiple pre-training processes in its input space, offers a better way for the determiner to anticipate the micro affirmation encountering for the source images. This discriminator is also required to find generated and real distributions. Communication is enhanced with the Interaction Students Education Neural Network (Int_Edu_NN) based on expected emotionality.

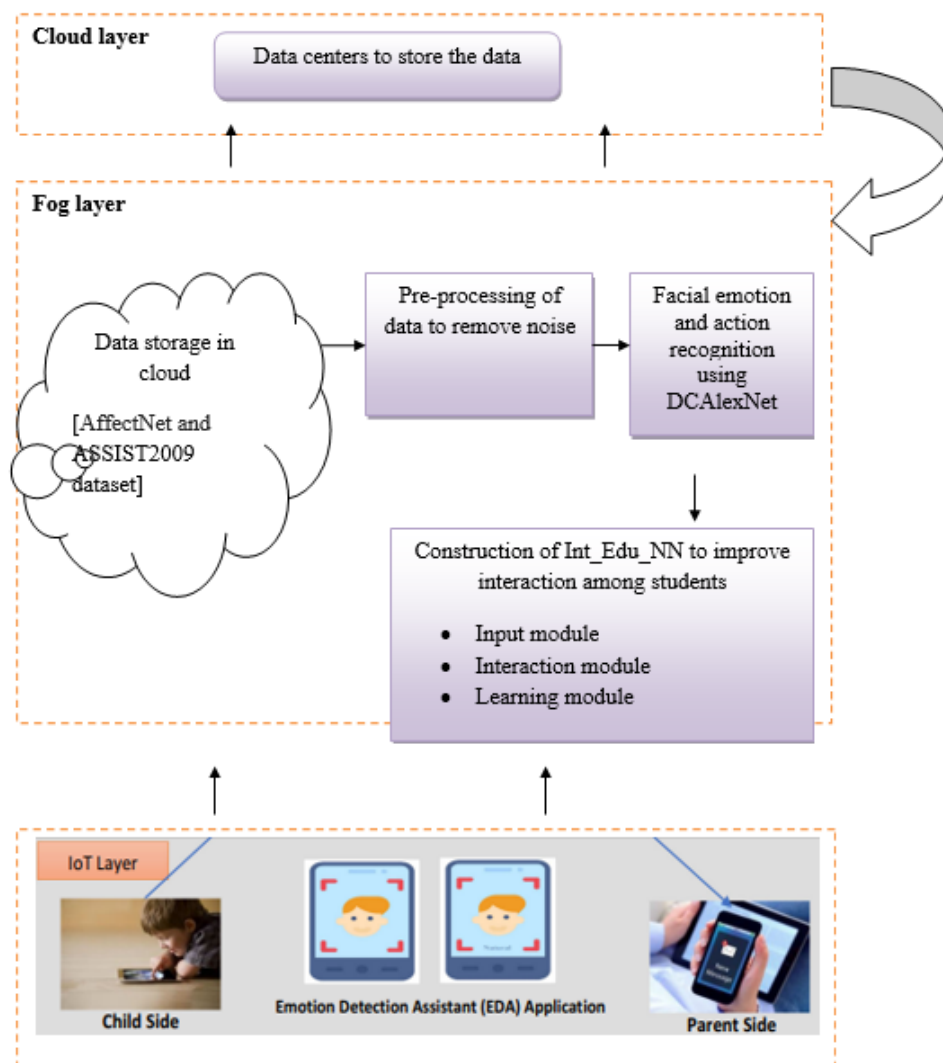


Figure-1 conceptual framework for interaction improvement for autistic children

- **Cloud Layer:** This cloud-based layer allows therapists and parents of autistic children to obtain statistics and data visualisations of the children's progress in learning to express their emotions. It is also helpful for parents to carry out efficient support work for the training, as well as for therapists to improve the rehabilitation training plan by making reference to the analysis results and pertinent medical data (such as the history of children). The cloud system enables system administrators to categorise, examine, and visualise a

sizable amount of abstract data and provide them to therapists and parents in an intuitive format, informing them of the children's development in understanding emotions.

- Fog Layer: Also known as a "fog node," this layer of technology consists of things like routers, gateways, access points, base stations, and certain fog servers. At the edge of a network are fog nodes. A hop can separate an edge from the end device. The Fog nodes are positioned between cloud data centres and endpoints.
- IoT Layer: While the youngster is using another app, the Emotion Detection Assistant (EDA) programme can run in the background. EDA is used to determine the child's emotional state.

1.1 Model of EDA

When the kid uses another application, the Emotion Detection Assistant (EDA) programme may operate in the background as well. EDA is employed to determine the child's emotional state. There's no issue, as demonstrated in figure 2, if the observed feeling is normal or pleasure. A associated programme running on the parent's smartphone gets a warning signal if the feeling that's being identified is one of anger, fear, sadness, or surprise. Autism affects a child's capacity to convey emotion, which is lower than average. Therefore, EDA is very helpful and necessary to let parents and carers to be informed if their child is not feeling well.

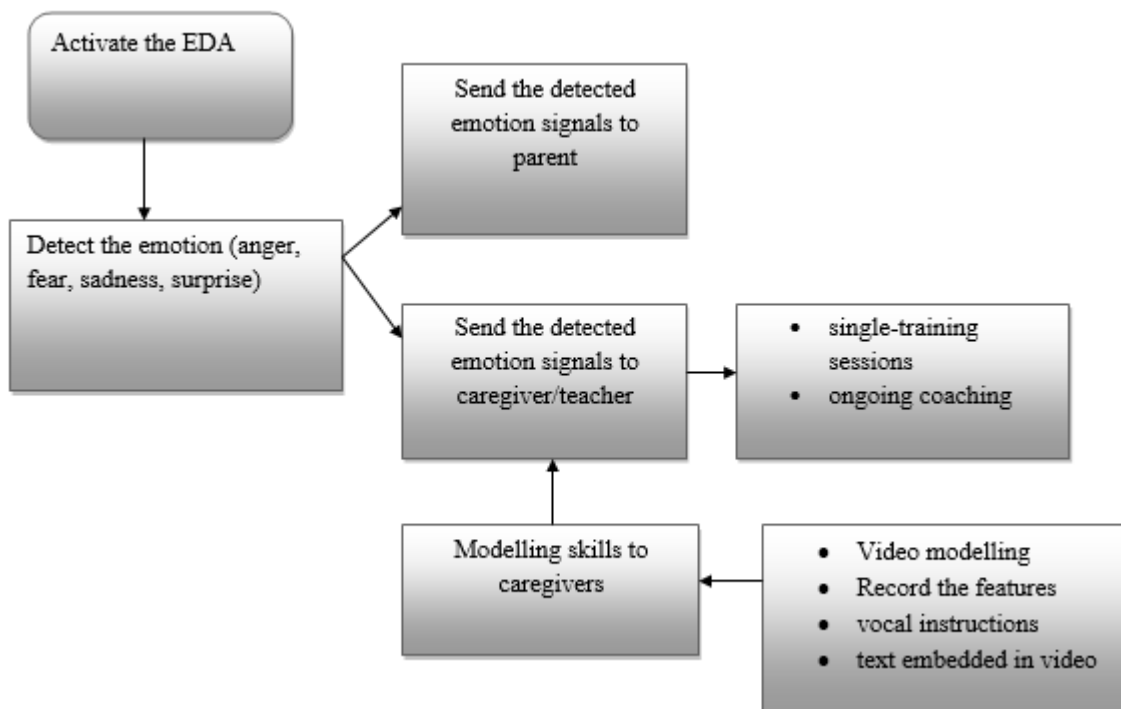


Figure -2 overall process of EDA

- Continuum of instructional assistance for carers: The kind of assistance for carers will probably fall into one of the following three groups: (a) giving directions, (b) offering one-time instruction, or (c) giving continuing mentoring. Additionally, educators can use these many stages across a variety of subject areas. For instance, the instructor could provide parents written instructions on how to set up their child's EDA devices whilst continuing to guide and encourage parents in putting an effective communication education regimen into practise.
- Demonstrating skills to carers: While demonstrating support programmes, a teacher may do so alongside a peer, by way of a video, or by having the carer role-play as the student. In order for the carer to apply the essential support tools throughout the educational sessions, the teacher must give clear directions during the modelling. Modelling might be done in-person or online. If the instructor employs video modelling, they must capture the crucial elements that must take place when the carer provides teaching, either verbally or through text that is incorporated in the video.

3.2 DCT based preprocessing

DCT converts the image into a spectral domain, and attributes in to spectral factor where the upper-left quadrant of the image contains the vast majority of the energy. In other words, the upper-left quadrant of the low-frequency

elements, which are focused, include the most significant visual data. DCT has indeed been extensively employed in image synthesis, including image encoding and image image retrieval, as a result of this capability. a common manufactured element in the spectral domain known as upper-left curvelet values, that is a one-dimensional vector. The most popular DCT-II variation, which itself is characterized as follows: is one of the eight standard DCT variants. These kernels are so considered in order to rejoin maximally and can be applied separately to input picture(image) in order to produce separate measurements of gradient component in each orientation i.e., V(i) and H(j). These can then be combined together to find an absolute gradient magnitude at each point

$$|\Delta X| = |V^2(i)| + |H^2(j)| \tag{1}$$

Angle of edge orientation with spatial gradient is denoted as,

$$\text{Angle}(\Delta X) = \tan^{-1} \frac{V(i)}{H(j)} \tag{2}$$

Computation of image gradient is dependent on obtaining partial derivatives at each pixel and hence grey levels are obtained from images. To select visual features at a finer level of granularity, the system utilizes deep models with 11 to 18 weight layers and only 7*7 filter layers. The number of Activation Maps in the third, fourth, and fifth Convolutional layers was increased from (385,384,256)(385,384,256) to (512,1024,512)(512,1024,512), which improved the amount of features the network could identify. First, 3*3 MaxPooling, and a dropout layer with a parameter of 0.8 was added; second, the local response normalisation was utilised in the output layer; and third, the batch size was adjusted to 1000 during neural network training. The pooling layer came after the convolutional layer, which decreased the dimension of the feature and prevented overfitting. The convolutional layer employed a 3× 3 filter and exponential linear unit activation function. The dropout layer employed a value of 0.8, the maxpool layer a 4× 4 filter, and the upsampling layer a 7× 7 filter. The network structure was flattened using convolutional layer consolidation. Ultimately, it was output using the sigmoid function and convolutional layer. Each learnable parameter is represented mathematically by a partial derivative of the loss, and a single update of a parameter is expressed as follows:

$$x(l) = x(l) - \beta \frac{\partial J(x,a)}{\partial x(l)} = x(l) - \beta \sum_{i=1}^n \frac{\partial J(x,a;r(i),s(i))}{\partial x(l)} - \Omega x \tag{3}$$

$$y(l) = y(l) - \beta \frac{\partial J(x,a)}{\partial x(l)} = y(l) - \beta \sum_{i=1}^n \frac{\partial J(y,a;r(i),s(i))}{\partial y(l)} \tag{4}$$

where,

x and y are the weight matrix and offset vector for each layer, correspondingly, and $(a(i), b(i)), 1 \leq i \leq N$ is a predetermined set of samples. β is the update rate of the parameters. The activation values generated by the inception model are adjusted using the feed-forward layer. It generates values that are precisely below 1 & nearer to 0. Such value product in less computational load and are simpler to work with in the model. The length of the visual output generated by the inception model is decreased through pooling. Assume that sample $X^i \in R^n$ represents the raw input data and $y^i \in \{1,2, \dots k\}$ represents the matching ground truth label for sample X^i . Assuming the pre-trained AlexNet architecture has M layers overall, the pre-trained DCAlexNet architecture's weight combinations are $W = (W(1), \dots, W(M))$. The associated weights are $w = (w^{(1)}, w^{(2)}, \dots, w^{(m-1)})$ for each classifier in each hidden layer of the pre-trained AlexNet architecture. The relationships in between parameterization and the classifiers in the pre-trained AlexNet architecture are depicted in Calculations (5) and (6), respectively:

$$Z^{(m)} = f(Q^{(m)}) \text{ and } Z^{(0)} = X \tag{5}$$

$$Q^{(m)} = W^{(m)} - Z^{(m-1)} \tag{6}$$

In Equations (5) and (6), $Q^{(m)}$ denotes the convolved answers on the prior convolution layer, M is the total amount of layers in the pre-trained AlexNet architecture, m signifies a specific layer in the pre-trained AlexNet

architecture, and $f()$ is the pooling function on Q . Equation displays the overall objective function for the pre-trained AlexNet architecture (7).

$$F(W) = P(W) + Q(W) \tag{7}$$

where $P(W)$ and $Q(W)$ refer to the output objective and the summed companion objectives, which are defined in Equations (8) and (9), respectively.

$$P(W) = \log_x w(out) + W(w(out)) \tag{8}$$

$$Q(W) = \sum_{m=1}^{M-1} \log_x w(out) + W(w(out)) \tag{9}$$

where $w(out)$ refers to the classifier weight of the output layer. The final combined objective function of the pre-trained DCAlexNet architecture is defined in Equation (10).

$$w(out) + L(W, w(out)) * 2 + \sum_{m=1}^{M-1} a[\log_x w(out) + W(w(out)) - r] \tag{10}$$

3.3 Interaction students Education Neural network (Int_Edu_NN)

The Q-matrix, which is acquired via professional labelling, serves as the input section to retrieve the learning ideas that the activities possess and to initialise the pupils' knowledge of each learning idea. The interface module employs a neural network to model the contact matrices and ultimately predicts pupil achievement to assess whether or not students have mastered the knowledge notion. This is built around the three-parameter exponential theoretical model as shown in flow chart figure-3.

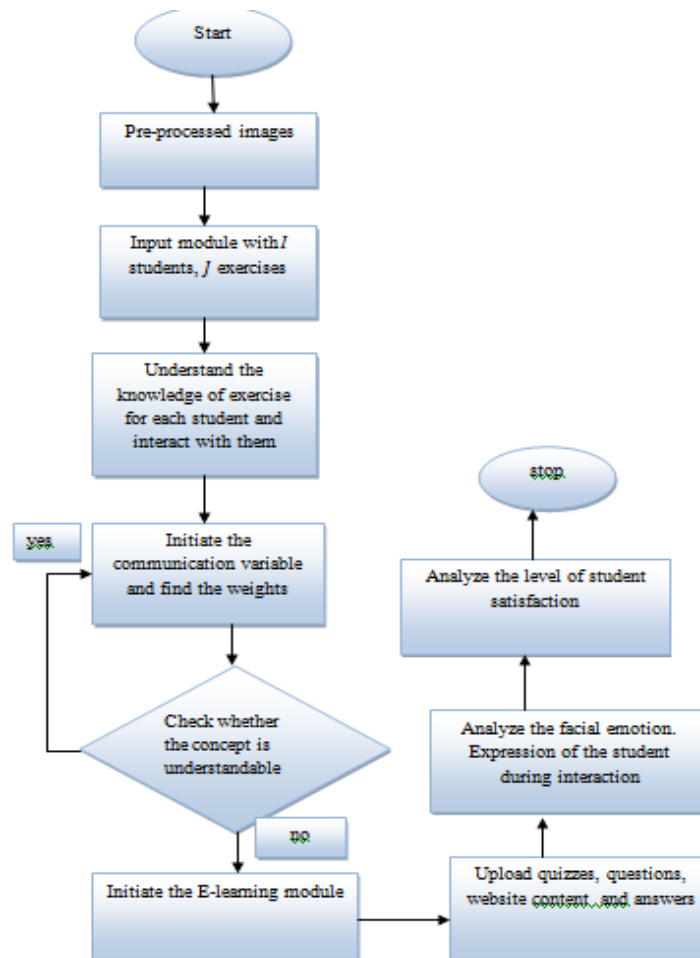


Figure-3 Flowchart for Int_Edu_NN

3.3.1 Input module

Suppose that the virtual learning system has are I students, J exercises, and $Know$ knowledge concepts. Establish the students' development is a significant proficiency for students $stud_no$ first, and then use hm to represent the students' one-hot vector. The student variable δ_{vec} is then determined by Equation (11)

$$\delta_{vec} = sigmoid(hm * A) \tag{11}$$

Where $\delta_{vec} = (\delta^1, \delta^2, \dots, \delta^K), \delta^i \in [0,1]$ indicates the student's range of knowledge concept, and $know_i, hm > \epsilon [0,1]^{1*N}; A \in R^{N*K}$ represents a trainable matrix. Also, because the model's goal is to help students gauge their level of understanding, the student vector is a comprehensible vector akin to pupil competence evaluation.

3.3.2 Interaction module

In order to standardise the input data, a characterised layer is placed even before kernel function of the fully linked layers and the output units in this component, which has one interface level, two layers that are completely connected, one and output units of the neural network. Equations (12) and provide the solution for the interaction layer (13).

$$k = (1.2 * exer * \delta - b) * a \tag{12}$$

$$k = f * (1 - k) + (1 - status) * k \tag{13}$$

where $f * (1 - exer * \delta)$ indicates the shift in student status when students pass the test the workout with the correlating understanding level, and $((1 - status) * (exer * \delta))$ reflects the alteration in pupil condition once respondents answered the exercise regularly without obtaining the corresponding understanding but by trying to guess. k denotes the communication variable after learners interact with the workout characteristics. Calculations below demonstrate how the fully linked layer and the output units fit the contact variable x after getting it.

$$D_1 = \partial(BN(wei_1 * k^T)) \tag{14}$$

$$D_2 = \partial(BN(wei_2 * D_1)) \tag{15}$$

$$y = \partial(BN(wei_3 * D_2)) \tag{16}$$

where wei_1, wei_2, wei_3 are the weight parameters for the neural network, and ∂ is the sigmoid of the activation function. In the component, we employ the perceptron gaussian to fulfill the stationarity condition and improve the comprehensibility of the data available in order to reduce the black-box characteristic of the neuron. In need for the neural network's outputs to grow steadily over time, each component of wei_1, wei_2, wei_3 is constrained to be progressive.

3.3.3 E-learning module

The true grade tr_gr and the projected grade $pred_gr$ are cross-entropy functions that make up the loss function of the model. In more detail, r represents the actual grade for exercise E whereas y represents the student hm grade for exercise E as predicted by the Int_Edu_NN model. As a result, the equation below can be used to determine the model's loss.

$$loss = -\sum_{i=1}^M (r_i \log y_i + (1 - r_i) \log(1 - y_i)) \tag{17}$$

Due to the possibility of student cheating, evaluation in e-learning systems is regarded as a challenging task. Concurrent evaluations, that consist of tests, inquiries, and responses, make up the majority of online tests. Homework, projects, and case studies are the main components of asynchronous evaluation. In e-learning, constant evaluation approaches need to be emphasised. Continuous e-assessment might be created to enhance student learning outcomes and can be aligned with the expectations of self-reflection assessments.

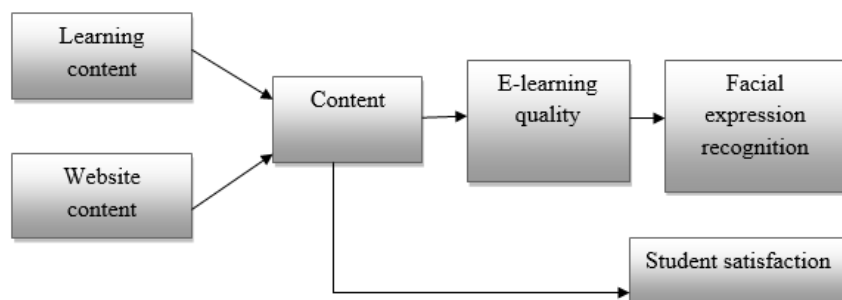


Figure-4 conceptual framework for E-learning module

For better comprehension of the link between online instruction material, student satisfaction, and the intermediary function of e-learning quality, an analytical approach is used. Figure 4 illustrates how the perceived negative effects of being on campus are employed as a mediating factor between the effectiveness of e-learning and satisfaction among students. The questionnaire was developed in English, and the construct items—learning content, website content, e-learning quality, and student satisfaction—were borrowed from prior studies as shown in table 1

Table-1 denotation of type of content and statements

Type of content	statements
Learning content	<ul style="list-style-type: none"> • gives us enough material for learning. • provides updated information • It offers the educational materials We require.
Web site content	<ul style="list-style-type: none"> • It makes good use of audio and video components. • It effectively makes use of multimedia capabilities and sequences/graphics. • It offers accurate details.
E-Learning Quality	<ul style="list-style-type: none"> • I receive teaching from online courses that ranges in quality from "poor" to "excellent." • The educational website appears to be current. • Instructions on the educational webpage are clear.
Student Satisfaction	<ul style="list-style-type: none"> • I'm happy I decided to sign up for the online classes. • My decision to join in online classes was intelligent, and I believe that paying for an educational programme were the right decision.
Student facial expression	<ul style="list-style-type: none"> • Am perplexed by his idea • bored during this session • I'm unwinding and learning more with this session.

The present research, which examined the e-learning quality's mediation function among e-content and student happiness, has important ramifications for the discipline of online education. The study's findings show that the learning materials and websites made available as part of the online study environment are both crucial components of the quality of e-learning, positively affecting both the e-learning experience and student happiness. As a way to foster a sense of involvement amongst participants and assure the quality of the online educational system and satisfaction among students, administrators and educators need to devote close attention to the creation of materials and structuring of the program's structure. As the only interface between the learner and teacher in the present pandemic situation, the website content offered in the e-learning platform should be simple to use and deliver important information. In addition, the possibility of contracting the disease on campus during the pandemic had no impact on the relationship between the effectiveness of online instruction and learners' pleasure. Therefore, it is crucial to maintain the greatest degree of quality in e-learning in order to please students. The abrupt shift towards e-learning was anticipated to result in students being pleased with whatever level of e-learning they are receiving due to the risk of contracting a virus in offline classes. The possibility of contracting the virus on campus during the pandemic had no impact on the relationship between the quality of e-learning and learners' pleasure. To capture and hold students' interest and encourage creative problem-solving, effective websites and instructional materials featuring infographics, video clips, forums, and quizzes are required. Along with offering high-quality e-learning, the instructors and web designers should include these technologies to create an effective online learning programme. Knowing the students' perceived level of satisfaction with the present material and e-learning quality offered by the various online courses is important for the content designers and instructors. Effective website content can result in visitors having a favourable attitude and learners being satisfied. Additionally, the fact that content is directly unrelated to student satisfaction and that e-learning quality

mediates this relationship further demonstrates the need for instructors to pay attention to e-learning quality in addition to content creation in order to improve student satisfaction. As a result, e-learning, which acts as a mediator between the material and learner satisfaction, has a significant impact on learner satisfaction.

3.4 Emotion classification

A *Int_Edu_NN* method built on the MobileNet architecture uses the ROIs as input. The applications' latency and size constraints were addressed by this model's use of depth-separable convolutional (DSC) layers with width and resolution multipliers as hyperparameters. These features try to reduce the number of parameters in the convolutional layers by splitting the convolution phase into two operations: depth-wise convolution and point-wise convolution (11 size). After the convolutional layers, the batch normalisation and activation processes are used. Batch normalisation is a method for standardising inputs that helps with the model's training phase. The rectified linear unit (ReLU) technique was used for the activation stage. The purpose of the pooling layer was to simplify the information in the convolutional layer output. The max-pooling operation, which only extracts the greatest value for the output, was utilised in this phase. By summarising the data, it is possible to avoid overfitting and decrease the amount of weights that must be learned. The final stage employed the decreasing slope and soft maximum layers. Three fully connected (FC) layers have been altered and introduced to the recommended technique. All the convolutional layer characteristics and the manually created features were utilised in the categorization. The algorithm used by CNN may learn characteristics that are unique to a given issue by using handmade features, that can enhance outcomes. The 70% ROIs from every set of images were used to train the machine learning algorithm across 10000 periods with an optimised Adam algorithm and a batch volume of one photo patch. The Adam optimisation tool was configured to use the default rates of learning of the Keras system, which is 0.0001.

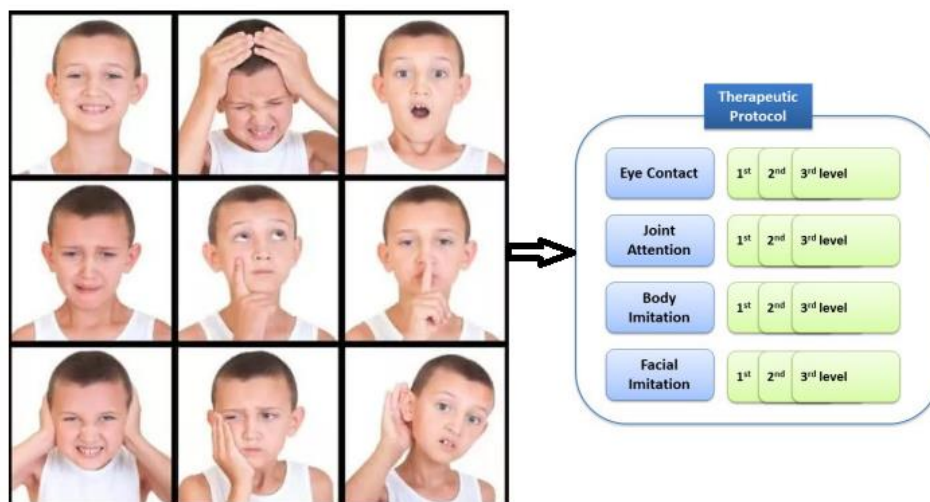


Figure-5 Facial Expression classification in different levels

First-Level Imitation of Facial Expression The machine addressed the kid by name. The robot exclaims, "I am (angry, happy, sad, scared), you do it!" when it makes eye contact and displays the emotions (fear, anger, happiness, and sorrow). (stimulus). The robot says "Good!" if the youngster accurately mimics the behaviour (intended behavioural reaction). The child's name is then repeated, and music is played (reinforcement). **Second-Level Facial Expression Imitation** With the exception of the reinforcement, this level is identical to the preceding one. There is no music offered. **Third-Level Facial Expression Imitation** Additionally, the reinforcement changes in this level. "How do you feel?" the robot asks the child as they mimic the emotion on the face. For each of the four fundamental emotions—anger, fear, happiness, and sadness—each level will be repeated.

IV. SIMULATION ANALYSIS

The proposed *Int_Edu_NN* model evaluated the performance with consideration of the face recognition dataset presented as follows:

- **Emotion dataset**-The AffectNet dataset, an accessible dataset, was utilized to build the emotion system. It is a massive dataset with 120 terabytes of information and about 1 million photos that have been carefully sorted into following categories (neutral, angry, sad, fear, happy, surprise, disgust, and contempt). Due to the

COVID-19 contagion, the majority of students donned masks during lectures. Researchers used facemask training to build the emotional network to overcome this problem. The features in the photographs were hidden using a machine vision program called MaskTheFace [33] that is hosted on Scribd. The facial tilting and eight basic facial traits are identified through a dlib-based face detection approach, enabling the application of mask. Based on the facial tilt, a match is created between the mask template and the face. Mask dataset collection is difficult in a number of situations. MaskTheFace offers a variety of masks, including surgical, textile, gas, N95, and KN95 masks.

- **Interaction dataset**-Each sample in the ASSISTment web - based learning platform's open dataset ASSIST2009 has two distinct iterations of the model theory. Throughout this research, the skill-builder edition is utilised, that offers a log of pupil replies to activities and assigns a learning idea identifier to each encounter. Following screening, 4163 responses of the respondents for 17,746 tasks with a total of 123 learning categories was retrieved for ASSIST2009. [34].



Figure 6 sample images of dataset

4.1 Experimental analysis

The experimental result is carried out by using the parameters such as accuracy, precision, Root Mean Square Error (RMSE), Root Absolute Error (RAE). These parameters are compared with three state of art methods such as Student Academic Performance Predicting (SAPP) system [25], Temporal Convolutional Neural Networks (T-CNN) [26], Attentive Interactive Student Performance prediction model (CAISP) [28]with the proposed Interaction students Education Neural network (Int_Edu_NN) . These are analyzed for two datasets such as AffectNet and ASSIST2009.

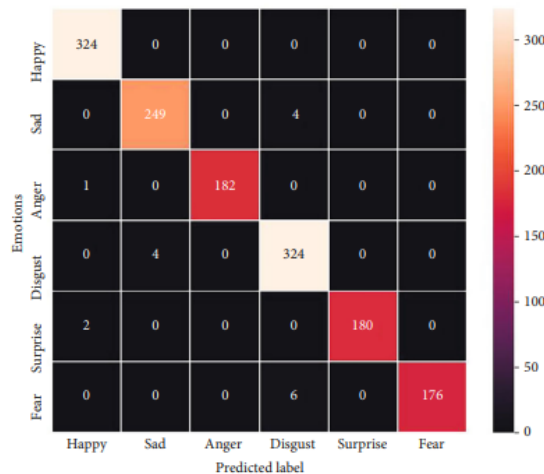


Figure 7 confusion matrix

The atop Figure 7 illustrates the confusion matrix for features employing classifier testing and training paradigm for Int_Edu_NN classifier where the rows portray the predicted class and columns portray the actual class of data relevant to predict emotion. The crosswise colors portray the tested and trained networks that are rightly and wrongly classified. The column on the right portray each anticipated class whereas the row below portrays the execution of each real class.

Table-2 analysis of accuracy

AffectNet					ASSIST2009				
Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN	Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN
100	93	86	94	99	100	95	87	93.6	98
200	94.8	87.9	93.5	99.4	200	94.6	87.9	92	98.6
300	95	87.5	94.8	99.6	300	95	88	92.5	97
400	94	88	94.3	99	400	95.9	88.5	93	97.9
500	93.5	88.9	94.2	98.9	500	96	86	93.6	98.3

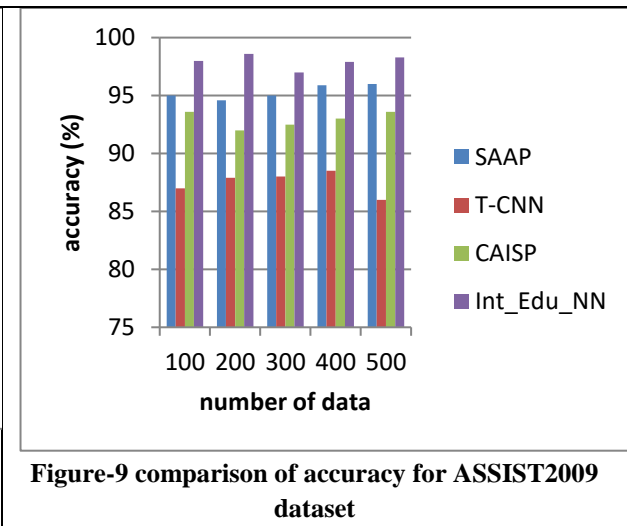
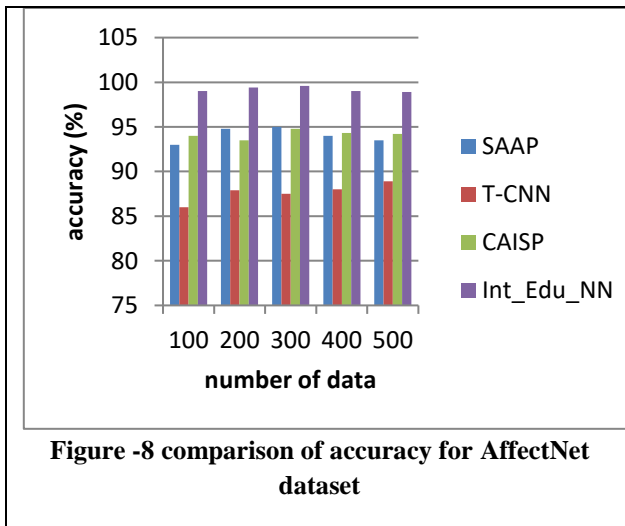


Figure 8 and 9 depicts the accuracy evaluation. When analyzing AffectNet dataset, the existing SAAP, T-CNN and CAISP achieves 94.23%, 86% and 94% of accuracy, whereas the proposed Int_Edu_NN achieves 99.6% which is 2.43%, 13.6% and 5.1% better than forementioned existing methods, When analyzing ASSIST2009 dataset we can see that the existing method achieves 95%, 87% and 93.8% of accuracy, whereas the proposed Int_Edu_NN achieves 98%, which is 3%, 11% and 5.8% better results.

Table-3 analysis of precision

AffectNet					ASSIST2009				
Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN	Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN
100	89	84	79	91	100	87	89	84	92
200	89.5	84.6	78.9	90.8	200	87.9	89.6	83.6	91.3
300	88.9	85	78.6	90.9	300	86	89.4	84.2	92.5
400	88.6	85	79.3	91.3	400	86.5	88	83	92
500	89	84	79	91	500	87	88.4	84	92.6

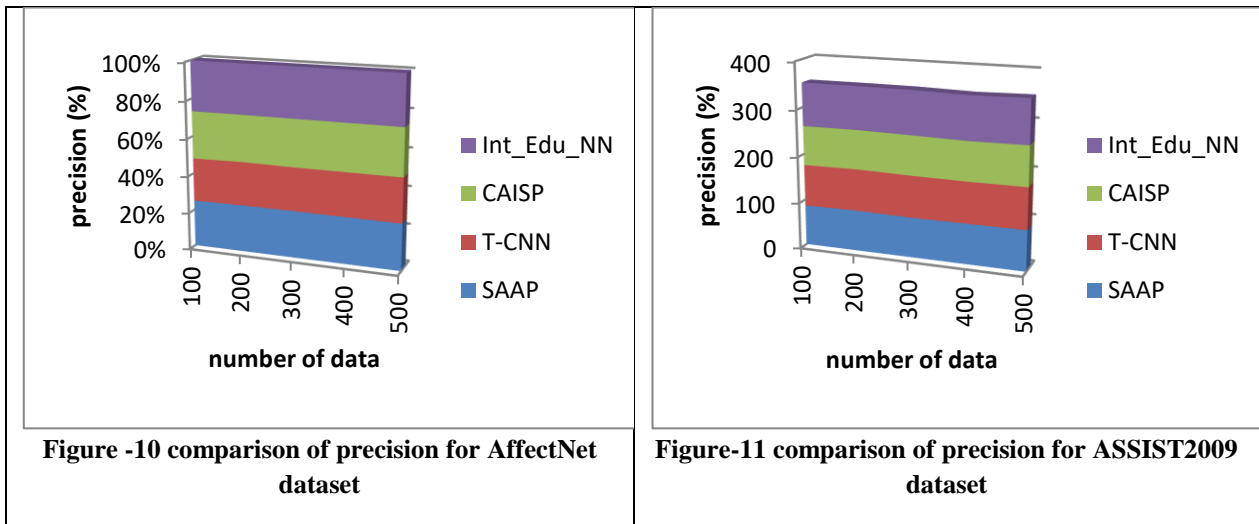


Figure 10 and 11 depicts the precision evaluation. When analyzing AffectNet dataset, the existing SAAP, T-CNN and CAISP achieves 29%, 20% and 20% of precision, whereas the proposed Int_Edu_NN achieves 21% which is 2%, 7% and 12% better than forementioned existing methods, When analyzing ASSIST2009 dataset we can see that the existing method achieves 47.3%, 39% and 49.7% of precision, whereas the proposed Int_Edu_NN achieves 52.7%, which is 5%, 3% and 8% better results.

Table-4 analysis of RMSE

AffectNet					ASSIST2009				
Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN	Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN
100	89	57.5	79	34	100	87	47.3	84	24
200	89.5	56	78.9	34.3	200	87.9	47	83.6	24.5
300	88.9	55.9	78.6	33	300	86	46	84.2	25
400	88.6	57	79.3	35.3	400	86.5	46.3	83	24.1
500	89	57.2	79	34	500	87	47	84	24

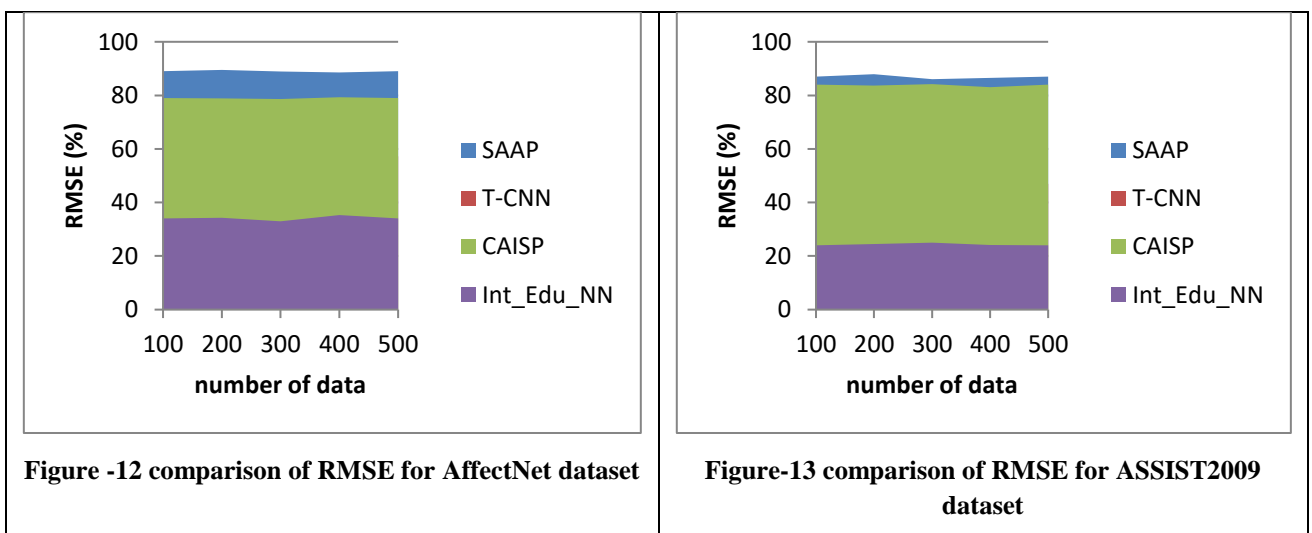


Figure 12 and 13 depicts the RMSE evaluation. When analyzing AffectNet dataset, the existing SAAP, T-CNN and CAISP achieves 89%, 55.8% and 78% of RMSE, whereas the proposed Int_Edu_NN achieves 34.2% which is 55%, 21.8% and 24.2% lesser than forementioned existing methods, When analyzing ASSIST2009 dataset we can

see that the existing method achieves 87%, 44.5% and 83.2% of RMSE , whereas the proposed Int_Edu_NN achieves 24%, which is 66%,31% and 74% better results.

Table-5 analysis of RAE

AffectNet					ASSIST2009				
Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN	Number of data	SAAP	T-CNN	CAISP	Int_Edu_NN
100	77	56	67	34.2	100	45.3	56.3	34.9	19.3
200	76	54.3	67.4	34	200	44	54	35	19
300	76.9	53	66.3	35.2	300	44.9	59.4	35.8	19.5
400	78	52.1	65	34	400	44.2	54	35.9	18.9
500	77.4	51	68	34.2	500	45	54.8	35	19

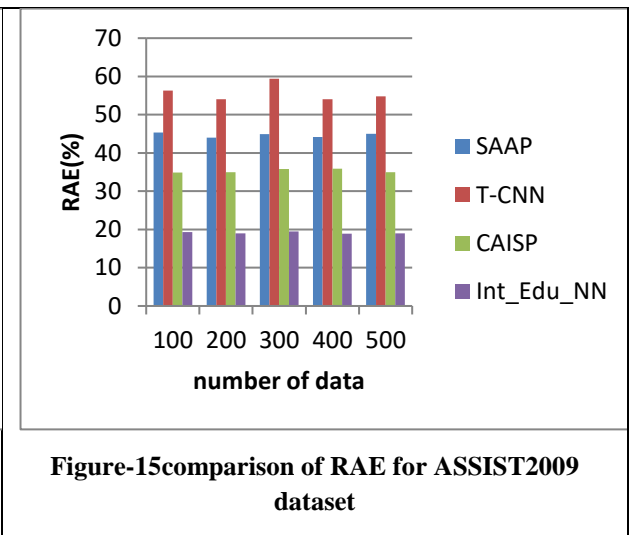
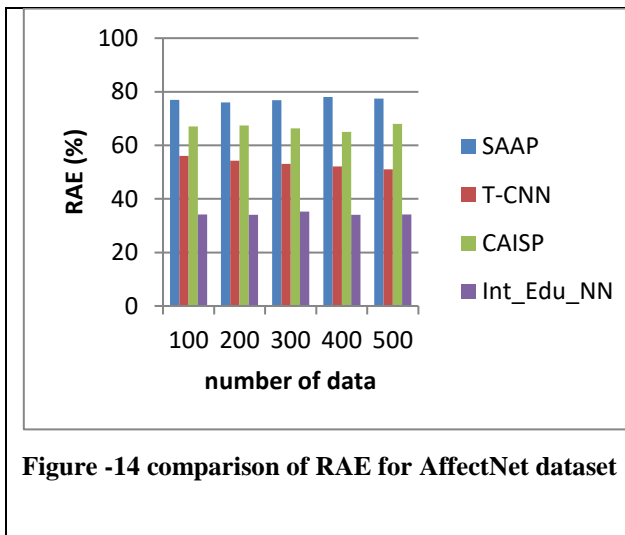


Figure 14 and 15 depicts the RAE evaluation. When analyzing AffectNet dataset, the existing SAAP, T-CNN and CAISP achieves 77%,54.8% and 65.9% of RAE, whereas the proposed Int_Edu_NN achieves 34% which is 55%, 21.8% and 24.2% lesser than foremenioned existing methods, When analyzing ASSIST2009 dataset we can see that the existing method achieves 45%, 54.9% and 35.9% of RAE , whereas the proposed Int_Edu_NN achieves 19%, which is 66%,31% and 74% better results.

Table- 6 comparative analysis for AffectNet dataset

Parameters	SAAP	T-CNN	CAISP	Int_Edu_NN
Accuracy (%)	94.23	86	94	99.6
Precision (%)	89	84	79	91
RMSE (%)	89	55.8	78	34.2
RAE (%)	77	54.3	65.9	34

Table-7 comparative analysis for ASSIST2009 dataset

Parameters	SAAP	T-CNN	CAISP	Int_Edu_NN
Accuracy (%)	95	87	93.8	98
Precision (%)	87	89	84	92
RMSE (%)	87	46.5	83.2	24
RAE (%)	45	54.9	35.9	19

4.2 Implications

Numerous studies have been done in attempt to understand ASD, however there is still room for detecting and efficiently treating this illness. In this study, we suggested a system that uses multiple machine learning, deep learning, and pre-trained transfer learning models to analyse facial photographs of children (such as autism/normal photos). However, this study diagnoses autism/non-autism from static photos; it is not fundamentally built for emotional/behavioral face identification. Additionally, we have used deep learning models for categorization that can automatically perform robust feature extraction to a degree that is nearly impossible to identify by simple observation due to their subtlety. Because we concentrated on classifying autism using static images, the experiment did not include video sequence analysis. Autism must be identified early in order to ensure that the appropriate procedures are taken. Our framework can help the medical field identify these neurological disorders more quickly than current systems do. Additionally, transfer learning models work better with mobile devices like smartphones, tablets, etc. As a result, they can be readily included into a mobile app, which will help health professionals or doctors recognise autism at an autistic resource centre. Additionally, parents can use this software at home to more precisely and promptly identify autism. With this method, we could alter and detect these cases faster than with any other method by using a smartphone camera to capture facial photos of youngsters. Once more, this paradigm is incredibly easy to use, economical, and requires few resources to integrate into devices.

V. CONCLUSION

Without relying on the well-known game engines and hardware, we were able to develop a tool for people with ASD to improve their skills in emotion detection and recognition thanks to the computational methods used in the EL. The principles used in the EL enabled the creation of a dynamic, interactive application capable of stimulating the users involved in the interventions' interest in a casual, engaging setting during concept evaluation phases. This paper suggests the Interaction Students Education Neural network (Int_Edu_NN) as a solution to the issues with conventional interaction improvement frameworks because of inadequate consideration of various exercise features and challenges in managing the non-linear relationship between learners and activities. Integrating data preparation and homogeneity constraint procedures in the neural network enhances the diagnostics findings' precision and readability while also speeding up the model's completion. E-learning quality and content have become popular as a way to minimise disruptions to education while also meeting the needs of students. The risk of contracting a virus on campus has forced students to rely on the online learning environment, but this risk does not necessarily indicate that students are happy with their online education, even though e-learning quality is a big factor in student satisfaction. The study's findings thus imply that in order to please students, the highest standard of e-learning quality must be upheld by institutions, administrators, and teachers alike. The actual findings demonstrate that the Int_Edu_NN outperforms comparable traditional models both in real data sets, validating the utility of the numerous workout features and the neural network with data preprocessing. The next study will concentrate on incorporating the idea of online learning with pupil attention quantitative determination.

REFERENCE

- [1] W. C. D. Souza, M. A. G. Feitosa, S. Eifuku, R. Tamura, and T. Ono, "Face perception in its neurobiological and social context," *Psychology & Neuroscience*, vol. 1, no. 1, pp. 15–20, 2008.
- [2] C. C. de Santana, W. C. D. Souza, and M. A. G. Feitosa, "Recognition of facial emotional expressions and its correlation with cognitive abilities in children with down syndrome," *Psychology & Neuroscience*, vol. 7, no. 2, pp. 73–81, 2014.
- [3] N. J. Sasson, A. E. Pinkham, K. L. Carpenter, and A. Belger, "The benefit of directly comparing autism and schizophrenia for revealing mechanisms of social cognitive impairment," *Journal of Neurodevelopmental Disorders*, vol. 3, no. 2, pp. 87–100, 2011.
- [4] A. M. O. D. Lima, M. R. D. A. Medeiros, P. D. P. Costa, and C. A. S. Azoni, "Analysis of softwares for emotion recognition in children and teenagers with autism spectrum disorder," *Revista CEFAC*, vol. 21, no. 1, 2019.
- [5] E. J. Jones, T. Gliga, R. Bedford, T. Charman, and M. H. Johnson, "Developmental pathways to autism: a review of prospective studies of infants at risk," *Neuroscience & Biobehavioral Reviews*, vol. 39, pp. 1–33, 2014.
- [6] R. C. Pennington and M. Carpenter, "Teaching written expression to students with autism spectrum disorder and complex communication needs," *Topics in Language Disorders*, vol. 39, no. 2, pp. 191–207, 2019.
- [7] K. Sagayaraj, C. R. Gopal, and S. Karthikeyan, "The efficacy of technology and non-technology based intervention for children with autism spectrum disorder: a meta-analysis," *International journal of Innovative Science and Research Technology*, vol. 5, pp. 863–868, 2020.
- [8] J. W. Tan and N. Zary, "Diagnostic markers of user experience, play, and learning for digital serious games: a conceptual framework study," *JMIR serious games*, vol. 7, no. 3, article e14620, 2019.

- [9] A. Solinska-Nowak, P. Magnuszewski, M. Curl et al., "An overview of serious games for disaster risk management - prospects and limitations for informing actions to arrest increasing risk," *International Journal of Disaster Risk Reduction*, vol. 31, pp. 1013–1029, 2018.
- [10] S. Wang, H. Zhang, and X. Meng, "Design of video teaching system based on virtual reality technology," *Electronics Science Technology and Application*, vol. 7, no. 4, p. 72, 2021.
- [11] H. G. Hernandez Palma, R. Pitre Redondo, and A. Sierra Parodi, "Methodology for the design of an integrated management system based on the iso 9001: 2015 and gtc 180:2008 standards for higher education institutions," *Indian Journal of Science and Technology*, vol. 11, no. 25, pp. 1–5, 2018.
- [12] Y. Choi, S. Pak, S. Pak, and H. Chang, "A study on design of security management system based on the type of display industrial technology leakage," *Korean Journal of Industry Security*, vol. 10, no. 3, pp. 109–140, 2020.
- [13] A. S. Abdulraheem, S. R. M. Zeebaree, and A. M. Abdulazeez, "Design and implementation of electronic human resource management system for duhok polytechnic university," *Technology Reports of Kansai University*, vol. 62, no. 4, pp. 1407–1420, 2020.
- [14] C.-H. Chen, F. Song, F.-J. Hwang, and L. Wu, "A probability density function generator based on neural networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 541, Article ID 123344, 2020.
- [15] A. North-Samardzic and M. De Witt, "Designing a human resource management simulation to engage students," *Journal of Management Education*, vol. 43, no. 4, pp. 359–395, 2019.
- [16] M. Y. N. F. Antonius Kadaranta, "Developing and improving competences of educators/teaching staffs of Indonesian national police school," *Psychology*, vol. 58, no. 2, pp. 120–125, 2021.
- [17] L. G. Jin and C. H. Zhao, "Water detection video management system of coal mine Industry and Mine," *Automation*, vol. 44, no. 9, pp. 102–104, 2018.
- [18] Wan, S.; Niu, Z. A hybrid e-learning recommendation approach based on learners' influence propagation. *IEEE Trans. Knowl. Data Eng.* 2019, 32, 827–840.
- [19] . Bernacki, M.L.; Chavez, M.M.; Uesbeck, P.M. Predicting achievement and providing support before STEM majors begin to fail. *Comput. Educ.* 2020, 158, 103999.
- [20] Jiang, P.; Wang, X. Preference cognitive diagnosis for student performance prediction. *IEEE Access* 2020, 8, 219775–219787. [CrossRef]
- [21] K. Valencia, C. Rusu, D. Quiñones, and E. Jamet, "The impact of technology on people with autism spectrum disorder: a systematic literature review," *Sensors*, vol. 19, no. 20, p. 4485, 2019.
- [22] R. Pavez, J. Díaz, and D. Vega, "Emotion recognition in children with ASD using technologies: a systematic mapping study," in 2019 38th International Conference of the Chilean Computer Science Society (SCCC), pp. 1–8, IEEE, 2019.
- [23] C.-H. Chen, I.-J. Lee, and L.-Y. Lin, "Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders," *Research in Developmental Disabilities*, vol. 36, pp. 396–403, 2015.
- [24] A. Dapogny, C. Grossard, S. Hun et al., "Jemime: a serious game to teach children with ASD how to adequately produce facial expressions," in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), pp. 723–730, IEEE, 2018.
- [25] Kukkar, A., Mohana, R., Sharma, A., & Nayyar, A. (2023). Prediction of student academic performance based on their emotional wellbeing and interaction on various e-learning platforms. *Education and Information Technologies*, 1-30
- [26] Awad, A., AlShehhi, A., Dias, S. B., Hadjileontiadis, S. J., & Hadjileontiadis, L. J. (2023, January). On Modeling LMS Users' Quality of Interaction Using Temporal Convolutional Neural Networks. In *Technology and Innovation in Learning, Teaching and Education: Third International Conference, TECH-EDU 2022, Lisbon, Portugal, August 31–September 2, 2022, Revised Selected Papers* (pp. 145-154). Cham: Springer Nature Switzerland.
- [27] Ni, Q., Wei, T., Zhao, J., He, L., & Zheng, C. (2023). HHSKT: A learner–question interactions based heterogeneous graph neural network model for knowledge tracing. *Expert Systems with Applications*, 215, 119334.
- [28] Wang, P., Li, L., Wang, R., Xie, Y., & Zhang, J. (2022). Complexity-based attentive interactive student performance prediction for personalized course study planning. *Education and Information Technologies*, 1-23.
- [29] Sethi, K., & Jaiswal, V. (2022). PSU-CNN: prediction of student understanding in the classroom through student facial images using convolutional neural network. *Materials Today: Proceedings*, 62, 4957-4964.

- [30] Mehta, N. K., Prasad, S. S., Saurav, S., Saini, R., & Singh, S. (2022). Three-dimensional DenseNet self-attention neural network for automatic detection of student's engagement. *Applied Intelligence*, 52(12), 13803-13823.
- [31] Xu, X., & Liu, F. (2021). Optimization of online education and teaching evaluation system based on GA-BP neural network. *Computational Intelligence and Neuroscience*, 2021, 1-9.
- [32] Wang, Z. J., Turko, R., Shaikh, O., Park, H., Das, N., Hohman, F., ... & Chau, D. H. P. (2020). CNN explainer: learning convolutional neural networks with interactive visualization. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 1396-1406
- [33] Anwar, A.; Raychowdhury, A. Masked Face Recognition for Secure Authentication. 2020. Available online: <https://arxiv.org/abs/2008.11104> (accessed on 19 January 2023)
- [34] Wu, M.; Tan, L.; Xiong, N. A structure fidelity approach for big data collection in wireless sensor networks. *Sensors* 2014, 15, 248–273