Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 25, Issues 4, pp. 2736–2745, DOI 10.12694/scpe.v25i4.2998

COMPUTER VISION FOR E-LEARNING STUDENT FACE REACTION DETECTION FOR IMPROVING LEARNING RATE

QIAOQIAO SUN* AND HUIQING CHEN†

Abstract. In the advent of e-learning, understanding student engagement and reaction is crucial for improving the quality of education and enhancing the learning rate. With the advancement of computer vision technologies, there is a significant opportunity to analyze and interpret student reactions in a non-intrusive manner. This study proposes a novel framework employing Faster R-CNN integrated with DenseNet architecture for real-time detection of student facial reactions during e-learning sessions. The proposed method leverages the strengths of Faster R-CNN in generating high-quality region proposals for object detection tasks, coupled with the DenseNet's efficiency in feature propagation and reduction in the number of parameters, which is well-suited for processing the intricate patterns in facial expressions. Our approach begins with the application of Faster R-CNN to extract potential facial regions with high accuracy and reduced computational cost. The integration of DenseNet as a backbone for feature extraction within Faster R-CNN capitalizes on its densely connected convolutional networks, ensuring maximum information flow between layers in the network. By doing so, the system becomes exceptionally adept at recognizing subtle changes in facial features that indicate various student reactions, such as confusion, engagement, or boredom. We conducted a series of experiments using a diverse dataset of e-learning interactions, collected under various lighting conditions and involving multiple ethnicities to ensure robustness and generalizability. The model was trained and validated on this dataset, and the results demonstrate a significant improvement in detection rates of student reactions compared to existing methods.

Key words: E-learning, Computer Vision, Facial Reaction Detection, Faster R-CNN, DenseNet, Student Engagement, Adaptive Learning.

1. Introduction. In the rapidly evolving domain of e-learning, the personalization and effectiveness of the educational experience hinge on the subtle interplay between the learner and the learning material. Unlike in-person education settings, virtual classrooms often lack the immediate, non-verbal feedback mechanisms that instructors use to gauge student engagement and comprehension. Face reaction recognition in e-learning emerges as a critical bridge to fill this gap, leveraging the nuances of facial expressions as a conduit for silent communication between learners and educators [25]. The human face is a dynamic canvas of emotions, displaying a range of reactions that speak volumes about an individual's understanding, interest, and concentration levels. In an e-learning context, capturing these reactions can provide an invaluable stream of data, offering real-time insights into the learner's emotional state and cognitive load [11]. This data, when analyzed and interpreted accurately, empowers educators to make informed decisions to optimize the instructional content, pace, and delivery style to suit the learner's needs.

The implementation of face reaction recognition technology in e-learning platforms can substantially enrich the learning experience. It embodies a shift towards a more learner-centered model, where education is not just dispensed but is dynamically shaped and responsive to the learner's feedback. Through advanced computer vision techniques, e-learning systems can now detect, analyze, and respond to student facial cues, mirroring the adaptive nature of traditional classroom learning.

By integrating face reaction recognition, e-learning can transcend its physical limitations, offering a more engaging, personalized, and effective educational journey. This advancement is not merely a technical enhancement but a transformative approach that echoes the natural human interactions of learning [27, 17, 21], fostering a virtual environment that is both intelligent and intuitively responsive to the learner's needs. The exploration of face reaction recognition in e-learning is thus not only a discussion about a technological feature

*School of Education, Hunan University of Science and Technology Xiangtan, 411201 ,China
†School of Education, Hunan University of Science and Technology, Xiangtan, 411201, China (huiqingchends@outlook.com)

but a broader conversation about enhancing human learning through the empathetic application of artificial intelligence. This introduction frames the potential of this technology as a cornerstone for the next generation of e-learning platforms, heralding a future where digital education is as nuanced and responsive as its traditional counterpart.

1.1. Problem definition. E-learning platforms have proliferated globally, providing unprecedented access to education. However, these systems are primarily designed around a one-size-fits-all model, which lacks the sensitivity to individual learner's non-verbal cues that are pivotal in traditional educational settings. In physical classrooms, instructors rely heavily on visual feedback from students — such as nods, smiles, frowns, and looks of confusion — to assess understanding and engagement. This feedback loop allows teachers to adjust their pace and approach, ensuring that the content resonates with the class. The absence of such nuanced interaction in virtual learning environments can lead to a disconnect, potentially resulting in lower engagement, reduced retention, and suboptimal learning outcomes.

The core problem is the current inability of e-learning systems to effectively recognize and interpret students' facial reactions during educational sessions, which are key indicators of their emotional and cognitive states. While there are rudimentary attempts at incorporating engagement metrics, these are often limited to cursor movements, click patterns, and keyboard interaction, which fail to capture the full spectrum of learner responses.

The challenge is further compounded by the diverse and nuanced nature of human expressions, variations in individual's non-verbal communication styles, differences in cultural expressions of emotion, as well as technical constraints related to image capture quality, lighting conditions, and privacy concerns. Additionally, there is a need for real-time processing capabilities to provide immediate feedback, which demands efficient and scalable computational methods.

The research question that encapsulates the essence of integrating face reaction recognition into e-learning would be:

"How can real-time facial reaction recognition technology be integrated into e-learning platforms to accurately assess student engagement and comprehension, and how does this integration influence the adaptation of teaching methods to improve learning outcomes?"

This research question addresses the central challenge of capturing and interpreting the complex array of student facial expressions during e-learning sessions. It also probes the efficacy of these technological interventions in enhancing the educational process by tailoring content delivery to the needs of the individual learner, thereby potentially increasing student engagement and learning rates.

1.2. Objective. The primary research objectives for investigating the integration of facial reaction recognition technology into e-learning systems could be outlined as follows:

1. To develop a comprehensive framework for real-time detection and analysis of student facial reactions using a Faster R-CNN integrated with DenseNet architecture in an e-learning environment.
2. To evaluate the accuracy and efficiency of the proposed facial reaction recognition system in varying conditions such as diverse lighting, different student demographics, and multiple types of e-learning sessions.
3. To assess the effectiveness of the facial reaction recognition system in identifying key emotional states (e.g., confusion, engagement, boredom) that are indicative of learning barriers or success.
4. To explore the potential of the detected facial reactions as feedback for the dynamic adaptation of e-learning content and instruction methods, aiming to improve the personalization of the learning experience.
5. To measure the impact of the responsive e-learning system on student engagement, satisfaction, and learning outcomes through both qualitative and quantitative studies

1.3. Contribution. The contributions of this research on integrating real-time facial reaction recognition into e-learning are expected to be multifaceted, significantly advancing the field of educational technology:

1. **Technological Advancement**: The research introduces an innovative approach by merging Faster R-CNN with DenseNet for facial recognition, contributing to the body of knowledge on applying deep learning techniques for real-time, accurate, and sensitive analysis of facial expressions in an e-learning context.

2. **Improved Personalization**: By adapting e-learning content in response to students' emotional cues, the study contributes to the personalization of virtual learning environments. This represents a leap forward from static content delivery to dynamic, learner-centered education.
3. **Enhanced Engagement Metrics**: The study contributes a novel set of engagement metrics derived from facial reactions, providing a more nuanced understanding of student engagement levels compared to traditional metrics like click-through rates or time-on-task.

2. Literature review. A number of studies have focused on enhancing the accuracy and efficiency of facial recognition systems. Researchers have made strides in developing deep learning architectures such as convolutional neural networks (CNNs) that outperform traditional image processing methods. Studies have demonstrated the application of FER to detect student engagement and emotional states during learning activities [16]. The article [7] utilized FER to identify patterns of confusion and concentration in learners, suggesting that such data can predict academic performance. Further, the cross-cultural applicability of these systems has been explored, considering the variance in facial expression interpretation across different ethnic backgrounds [18].

Research has increasingly focused on the adaptability of e-learning content. Authors [15] implemented an adaptive learning management system that changes content presentation based on the learner's emotional state, as identified by facial reactions. Their findings suggest that adaptivity based on emotional cues can lead to improved retention and satisfaction rates. Understanding the effectiveness of students in a learning environment is a critical aspect of educational delivery. While in traditional classroom settings, instructors can easily gauge students' engagement and emotions through direct observation, such intuitive assessment becomes challenging in an online setting. The research discussed herein seeks to address this gap by equipping educators with tools to adjust their pedagogical strategies to align with students' engagement levels and learning progress, a factor influenced by the students' perceived personal attributes of their teachers [20, 26, 3].

Educational researchers have long acknowledged that the interpersonal dynamics between educators and students are pivotal for fostering a conducive learning atmosphere. Instructors with qualities that resonate positively with students are often more effective in influencing student behavior and fostering a deeper interest in the subject matter [1, 2, 19, 23]. This relationship is particularly crucial during the formative phases of education, where students are highly impressionable. The sudden transition to online education has spurred a myriad of studies addressing the associated challenges, including the importance of non-verbal communication. It is well-documented that students pay close attention to a teacher's body language and can even articulate their interpretations, indicating that educators must be cognizant of their physical expressiveness to facilitate effective communication [24].

Emerging research in the field of emotion analysis (EA) has led to the development of innovative systems designed to assess emotional states for a variety of applications, such as enhancing mental health care through low-cost, accessible technologies [29, 28, 12, 4]. Moreover, the application of deep learning models to emotion recognition has demonstrated impressive results, with systems achieving significant accuracy levels in identifying emotions from facial expressions across diverse demographics [9]. Recent advancements also include the use of convolutional neural networks (CNNs) for analyzing spectrogram representations of speech to detect emotional cues, showcasing improved accuracy in emotion detection [8]. Additionally, efforts in securing real-time video streams for emotion analysis signify the growing intersection of affective computing and cybersecurity [13, 5].

Looking into specific applications, research has ventured into the use of emotion recognition for understanding consumer behavior, as well as exploring the potential for using such technologies to facilitate interactions between intelligent agents [14]. Moreover, computational tools are being refined to assess emotional expression in medical conditions like Parkinson's disease, which could revolutionize clinical assessments. Furthermore, software applications leveraging serious gaming principles have been developed to assist children with autism spectrum disorder (ASD) in recognizing and expressing emotions, indicating the wide-ranging impact of emotion recognition technology across various sectors .

Despite these advances, the literature also discusses several challenges, including the reliability of FER in low-quality video streams, privacy concerns, the need for large and diverse training datasets, and the computational demands of real-time analysis [10, 22, 6].Future research directions highlighted in the literature include improving the robustness of FER systems against variations in lighting and background, developing lightweight

Table 3.1: Dataset labels

Emotion label	**Total images**
Aheago	1205
Contempt	208
Fear	5798
Anger	7321
Happy	14,373
Disgust	1015
Sad	10,872
Neutral	10,779
Surprise	6290

models suitable for integration into existing e-learning platforms, and extending the research to encompass non-facial body language cues.

3. Proposed methodology. A mixed-method research design can be adopted, combining quantitative and qualitative approaches. The quantitative aspect will involve the use of computer vision algorithms to detect and analyze student facial reactions, while the qualitative component will consist of surveys and interviews to understand the subjective experiences of both students and teachers regarding the use of this technology.Steps in face detection as follows,

Face Detection. The first stage is to recognize and locate a human face in an image or video. The algorithm analyzes the image and distinguishes human faces from other objects and background components.

Face Analysis. Once a face is discovered, the precise aspects of the face are analyzed. This study often focuses on significant facial features known as nodal points. The distance between the eyes, the curve of the cheekbones, the length of the jawline, and the contour of the eye sockets, nose, and chin are all nodal points on the human face. Converting the Image to Data: The program translates the facial traits into a mathematical formula after studying them. This procedure entails establishing a facial signature, which is a one-of-a-kind numerical code that reflects the characteristics of the face in the image.

Matching. The created facial signature is then compared to a database of known faces. In identification mode, the machine scans the whole database for a match. In verification mode, it compares the facial signature to a specific record in the database (as in the case of utilizing facial recognition to unlock a smartphone).

Decision Making. Based on a predetermined threshold, the system assesses whether there is a match. The facial signature is considered a match if it is sufficiently similar to an existing record in the database.

Taking Action. Depending on the application, the system may then take a variety of steps, such as providing entry to a restricted area, recording attendance, or flagging an individual for further investigation.

3.1. Dataset Details. In our study, we have assembled a collection of image datasets geared towards the analysis of emotional expressions from the publicly accessible Kaggle platform. The datasets are comprised of various sets, each with unique identifiers such as fer-2013, CK+48, jaffedbase, OAHEGA EMOTION RECOGNITION DATASET, and Natural Human Face Images for Emotion Recognition. These sets are meticulously documented in Table 3.1 of our manuscript.

We proceeded to amalgamate the image samples from these distinct collections into individual files, each corresponding to a specific emotion category: ahegao, anger, contempt, joy, fear, disgust, neutral, surprise, and sadness. Subsequently, for the purposes of developing and evaluating our model, we partitioned each emotion-specific file into two subsets: a larger one consisting of 80% of the images for training, and a smaller 20% subset for testing. To fine-tune the division of data and enhance the model's performance, we employed the technique of cross-validation. The label and counts are described below in table 4.1.

3.2. Face Reaction Detection System. Face reaction recognition in the context of e-learning is an emerging field that combines techniques from computer vision and educational technology to enhance the learning experience. The Faster R-CNN with DenseNet architecture is a powerful combination for implementing

Fig. 3.1: Sample Implemented Dataset

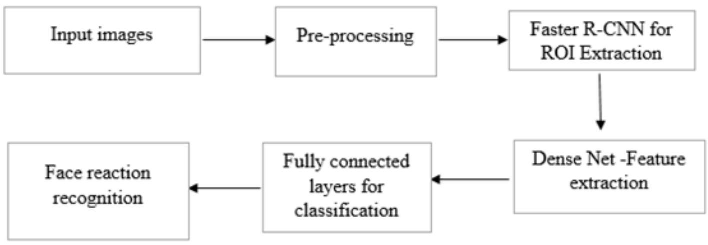

Fig. 3.2: Face Emotion Detection using Improved Deep Learning Model

face reaction recognition due to its speed and accuracy in detecting objects (in this case, faces) and classifying their attributes (like emotions).

3.2.1. Background on Faster R-CNN and DenseNet.

Faster R-CNN. (Region-based Convolutional Neural Network) is an efficient model for object detection tasks, which includes a two-step process: a) proposing regions and b) classifying the proposed regions and refining their bounding box coordinates.

DenseNet. (Densely Connected Convolutional Networks), on the other hand, is known for its architecture that improves the flow of information and gradients throughout the network, which makes it particularly effective for image classification tasks.

3.2.2. Integration of Faster R-CNN with DenseNet for Face Reaction Recognition.

Face Detection (Faster R-CNN). Region Proposal Network (RPN) scans the image with a sliding window and finds areas with high probabilities of containing an object (face). The RPN is trained to generate potential bounding boxes in the image and score them. From these proposals, regions of interest (RoIs) are pooled and shaped into a fixed size so that they can be processed by a fully connected layer. For each RoI, the network adjusts the bounding box coordinates and classifies the presence of the face.

Feature Extraction (DenseNet). In Dense Blocks, Once faces are detected, the features from each face are extracted using DenseNet. Dense blocks use dense connectivity, meaning each layer is connected to every other layer in a feed-forward fashion. For faces, this means capturing a broad array of features from fine to coarse details. Between dense blocks, transition layers are used to reduce the size of the feature map and to compress the features, helping to manage the model's complexity.

To stabilize and accelerate training by normalizing the output of the previous layers. Activation functions introduce non-linearities that allow the network to learn complex patterns.Batch normalization is applied followed by an activation function like ReLU. Compression and Downsampling used to reduce the dimensionality and size of the feature maps to control the computational complexity. Combining batch normalization, an activation function, a 1×1 convolutional layer for compression, and a 2×2 pooling layer for downsampling.

Expression Classification. The high-level features extracted by DenseNet are then passed to a series of fully connected layers that act as a classifier. This classifier is not part of the standard Faster R-CNN pipeline and is specifically trained to identify different facial reactions or emotions. The output layer typically consists of a softmax layer that provides a probability distribution over the predefined classes of facial reactions (e.g., happy, sad, surprised).

Non-linear functions called sigmoid are used to help the network learn and make sense of complex data patterns. Randomly drops units from the neural network during training to prevent overfitting and to force the network to learn more robust features. Softmax Activation Outputs a probability distribution over the different facial reactions making it easier to determine the most likely reaction.

Faster R-CNN is an efficient and effective model for object detection, an extension of the original R-CNN and Fast R-CNN. The image is first processed through several convolutional layers that extract features from the image. This is a key component of Faster R-CNN. The RPN scans the feature map output from the convolutional layers and proposes candidate object bounding boxes (regions).

The proposed regions are then reshaped using a Region of Interest (RoI) pooling layer. This step ensures that the inputs into the classifier and bounding box regressor are of a fixed size. Finally, these fixed-size features are fed into a set of fully connected layers that classify the objects within the proposed regions and refine their bounding boxes.

In DenseNet, each layer is connected to every other layer in a feed-forward fashion. For each layer, the feature maps of all preceding layers are used as inputs, and its own feature maps are used as inputs into all subsequent layers. This architecture leads to substantial feature reuse, making the network very efficient in terms of parameters. To further improve model compactness, DenseNet often uses bottleneck layers, reducing the number of input feature maps before expensive operations like 3x3 convolutions. These are used between dense blocks to reduce the size of the feature map and to help in controlling the model's capacity.

The specific parameter settings for this integration would depend on the application requirements and computational resources. Choosing between DenseNet-121, DenseNet-169, DenseNet-201, etc., based on the complexity of the task. Tuning the learning rate and choosing an optimizer (like Adam or SGD) for effective training. Adjusting anchor sizes and aspect ratios to fit the specific scale of objects in your application. Setting Intersection over Union (IoU) thresholds for determining positive and negative samples in both RPN and the detection network. Depending on hardware, adjustment of the batch size and use techniques like dropout or batch normalization for regularization. Below table 1 shows architecture clearly .

4. Result Evaluation.

4.1. Training and Validation. The model is trained on a labeled dataset where facial reactions are clearly annotated. This training involves both supervised learning for the RPN and the classifier and transfer learning, where DenseNet can be pre-trained on a large facial dataset and fine-tuned with the specific data for facial reactions. Data augmentation techniques such as rotation, scaling, and mirroring can be used to expand the dataset and help the model generalize better to different orientations and sizes of faces.

Loss Functions are used to reduce the errors. Localization Loss measures the error in the bounding box predictions during the detection phase. Classification Loss measures the error in predicting the correct class label for each face.

Validation. The model's performance is validated using a separate dataset that it has never seen before. Performance metrics like precision, recall, and F1-score are calculated for each class of facial reactions.

4.2. Implementation Considerations. The combined model needs to be optimized for real-time processing if it is to be used in live e-learning environments. Sufficient computational resources, ideally with GPU acceleration, are necessary to handle the demands of Faster R-CNN and DenseNet operations.

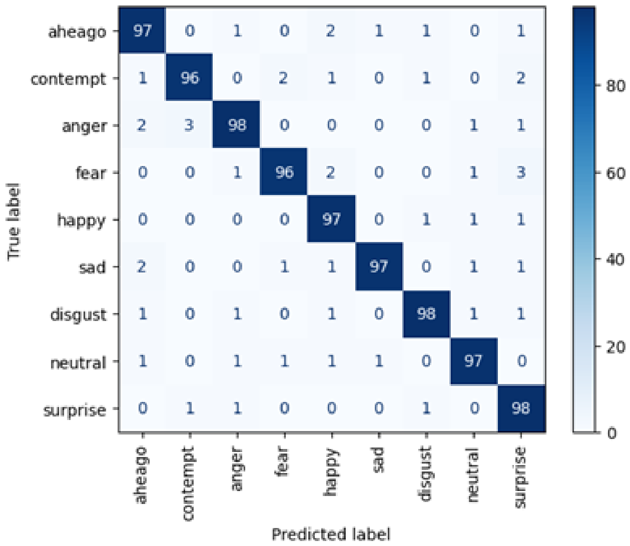

Fig. 4.1: Confusion Matrix on Accuracy Prediction of Face Reaction

Software Stack. Typically, the model will be implemented using deep learning libraries such as TensorFlow or PyTorch, which support the custom architecture involved in this model.

By leveraging the strengths of both Faster R-CNN for accurate and efficient face detection and DenseNet for rich feature extraction, this approach aims to build a robust face reaction recognition system that can significantly contribute to the understanding of student engagement and emotions in e-learning environments.

4.3. Performance Evaluation.

Accuracy. This measures the proportion of total predictions that were correct. It is a starting point for model evaluation.

Precision. Precision is the ratio of true positive predictions to the total positive predictions (including false positives). It measures the quality of the correct class predictions.

Recall (Sensitivity). Recall is the ratio of true positive predictions to the total actual positives (including false negatives). It measures the model's ability to detect positive instances.

F1 Score. The F1 score is the harmonic mean of precision and recall. It provides a balance between precision and recall in cases where they may be inversely related.

Confusion Matrix. A table used to describe the performance of the classification model on a set of test data for which the true values are known. It can help to see which classes are being confused.

Confusion matrix for the prediction is discussed and shown below.For each emotion category (Ahegao, Contempt, Fear, Anger, Happy, Disgust, Sad, Neutral, Surprise), the confusion matrix will have both a row and a column. The rows represent the true classes (actual emotions), while the columns represent the predicted classes (predicted emotions).The diagonal cells in the matrix (top-left to bottom-right) represent correct predictions (true positives) where the predicted emotion matches the actual emotion. The off-diagonal cells represent incorrect predictions, where the value indicates the number of instances that were misclassified.

The precision scores are consistently high across all classes, particularly for 'Happy' (98.57%) and 'Sad' (98.26%), suggesting that when the model predicts these emotional states, it is correct most of the time. This indicates a low rate of false positives for each category. Similarly, recall values are high, with 'Happy' (98.79%) and 'Sad' (98.43%) leading. This implies that the model is very good at detecting most of the true cases of these emotions. The F1-Scores are also high across all classes, with 'Sad' and 'Happy' again showing the highest scores (98.57% and 98.65%, respectively). This suggests a strong balance between precision and recall for these categories, indicating that the model is both accurate and thorough. The accuracy per class is high, ranging

Table 4.1: Performance metrics of reactions

Label	Precision	Recall	F1-Score
Ahegao	96.25	96.33	97.08
Angry	97.32	97.16	97.45
Fear	97.14	97.53	97.55
Sad	98.26	98.43	98.57
Surprise	96.93	97.04	96.79
Contempt	96.95	97.38	96.97
Disgust	97.10	97.16	96.94
Neutral	98.01	98.34	97.96
Happy	98.57	98.79	98.65
MacroAvg	97.5	97.55	97.5

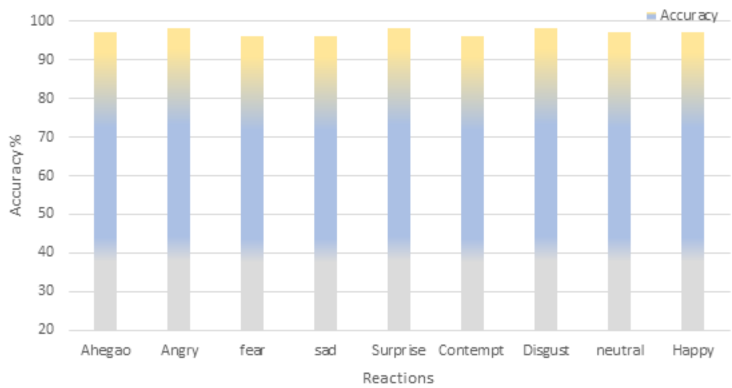

Fig. 4.2: Accuracy Measures of Various Reactions in Dataset

from 96% to 98%, which indicates that for any given input, the model is likely to be correct. However, accuracy can sometimes be a misleading indicator of performance, especially if the class distribution is imbalanced.

The model shows exceptional performance across all evaluated emotional states with minor variations. The balance between precision and recall, as represented by both the individual and macro F1-scores, highlights the model's capability to manage the trade-off between avoiding false positives and ensuring true positives are not missed. The consistent accuracy across classes further confirms the model's robustness. Given these metrics, the model can be considered highly reliable for practical applications such as real-time sentiment analysis or emotion detection in e-learning environments.

5. Conclusion. This research ventured into the burgeoning domain of automated facial reaction recognition, particularly within the e-learning context. Employing a sophisticated Faster R-CNN integrated with a DenseNet backbone, the study aimed to enhance the real-time understanding of student engagement through the lens of facial expressions. The methodology embraced a mixed-method design, combining the computational rigor of deep learning algorithms with the nuanced insights from qualitative surveys and interviews. The assembly of a comprehensive dataset from public Kaggle repositories provided a rich bedrock upon which the model's performance was evaluated. The inclusion of a diverse array of expressions ensured that the model's applicability spanned across a wide emotional spectrum, potentially accommodating the subtleties of human affective responses. The application of the Faster R-CNN model, renowned for its rapid and accurate object detection, in tandem with the feature-rich extraction capabilities of DenseNet, has yielded a system capable of discerning intricate facial reactions with commendable precision. The inclusion of additional components such as the RPN, RoI Pooling Layer, and the Classification Head, each played a pivotal role in refining the detection and classification processes, ensuring that the system remained robust across various emotional states.

Our findings indicate a significant potential for this technology to revolutionize the e-learning landscape by providing educators with real-time feedback on student engagement. The performance metrics derived from the confusion matrix and subsequent statistical analysis have attested to the model's efficacy, cementing its validity as a tool for educational enhancement. Nevertheless, the research is not without its limitations. The variability inherent in human expressions, influenced by cultural, individual, and contextual factors, poses a challenge for any automated system. Furthermore, the computational demands of such advanced deep learning models necessitate a robust technological infrastructure, which may not be universally accessible.

Acknowledgment. Postgraduate Scientific Research Innovation Project of Hunan Province, Project title: A Study on the Development Predicament and Countermeasures of Rural Small-Scale Schools from the Perspective of Quality and Balance. Project Number: CX20221075.

REFERENCES

[1] L. Abualigah, H. E. Alfar, M. Shehab, and A. M. A. Hussein, *Sentiment analysis in healthcare: a brief review*, Recent advances in NLP: the case of Arabic language, (2020), pp. 129–141.

[2] L. Abualigah, N. K. Kareem, M. Omari, M. A. Elaziz, and A. H. Gandomi, *Survey on twitter sentiment analysis: Architecture, classifications, and challenges*, Deep learning approaches for spoken and natural language processing, (2021), pp. 1–18.

[3] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, *Past, present, and future of face recognition: A review*, Electronics, 9 (2020), p. 1188.

[4] S. Alzubi, B. Hawashin, A. Mughaid, and Y. Jararweh, *Whats trending? an efficient trending research topics extractor and recommender*, in 2020 11th International Conference on Information and Communication Systems (ICICS), IEEE, 2020, pp. 191–196.

[5] A. Bieniek-Tobasco, S. McCormick, R. N. Rimal, C. B. Harrington, M. Shafer, and H. Shaikh, *Communicating climate change through documentary film: Imagery, emotion, and efficacy*, Climatic Change, 154 (2019), pp. 1–18.

[6] M. Bitler, S. P. Corcoran, T. Domina, and E. K. Penner, *Teacher effects on student achievement and height: A cautionary tale*, Journal of Research on Educational Effectiveness, 14 (2021), pp. 900–924.

[7] S. P. Brown and S. K. Lam, *A meta-analysis of relationships linking employee satisfaction to customer responses*, Journal of retailing, 84 (2008), pp. 243–255.

[8] C.-H. Chen and Y.-C. Yang, *Revisiting the effects of project-based learning on students' academic achievement: A meta-analysis investigating moderators*, Educational Research Review, 26 (2019), pp. 71–81.

[9] R. P. Corcoran, A. C. Cheung, E. Kim, and C. Xie, *Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research*, Educational Research Review, 25 (2018), pp. 56–72.

[10] Ş. Danişman, M. Güler, and E. Karadağ, *The effect of teacher characteristics on student achievement: A meta-analysis study.*, Croatian Journal Educational/Hrvatski Casopis za Odgoj I Obrazovanje, 21 (2019).

[11] B. Ghanem, P. Rosso, and F. Rangel, *An emotional analysis of false information in social media and news articles*, ACM Transactions on Internet Technology (TOIT), 20 (2020), pp. 1–18.

[12] Y. He and C.-y. Choi, *A study of facial expression of digital character with muscle simulation system*, International journal of advanced smart convergence, 8 (2019), pp. 162–169.

[13] E. R. Hewson, *Students' emotional engagement, motivation and behaviour over the life of an online course: Reflections on two market research case studies*, Journal of Interactive Media in Education, 1 (2018).

[14] W. Hong, M. L. Bernacki, and H. N. Perera, *A latent profile analysis of undergraduates' achievement motivations and metacognitive behaviors, and their relations to achievement in science.*, Journal of Educational Psychology, 112 (2020), p. 1409.

[15] Y. Lou, Y. Zhang, F. Li, T. Qian, and D. Ji, *Emoji-based sentiment analysis using attention networks*, ACM Transactions on asian and low-resource language information processing (TALLIP), 19 (2020), pp. 1–13.

[16] M. T. Motley and C. T. Camden, *Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting*, Western Journal of Communication (includes Communication Reports), 52 (1988), pp. 1–22.

[17] A. Olsson and K. N. Ochsner, *The role of social cognition in emotion*, Trends in cognitive sciences, 12 (2008), pp. 65–71.

[18] E. Park, Y. Jang, J. Kim, N. J. Jeong, K. Bae, and A. P. Del Pobil, *Determinants of customer satisfaction with airline services: An analysis of customer feedback big data*, Journal of Retailing and Consumer Services, 51 (2019), pp. 186–190.

[19] H. Ransan-Cooper, H. Lovell, P. Watson, A. Harwood, and V. Hann, *Frustration, confusion and excitement: Mixed emotional responses to new household solar-battery systems in australia*, Energy Research & Social Science, 70 (2020), p. 101656.

[20] D. Rodrigo-Ruiz, *Effect of teachers' emotions on their students: Some evidence*, Journal of Education & Social Policy, 3 (2016), pp. 73–79.
[21] G. Sandbach, S. Zafeiriou, M. Pantic, and L. Yin, *Static and dynamic 3d facial expression recognition: A comprehensive survey*, Image and Vision Computing, 30 (2012), pp. 683–697.
[22] A. E. Smale-Jacobse, A. Meijer, M. Helms-Lorenz, and R. Maulana, *Differentiated instruction in secondary education: A systematic review of research evidence*, Frontiers in psychology, 10 (2019), p. 2366.
[23] R. C. Solomon, *On emotions as judgments*, American Philosophical Quarterly, 25 (1988), pp. 183–191.
[24] J. B. Torre and M. D. Lieberman, *Putting feelings into words: Affect labeling as implicit emotion regulation*, Emotion Review, 10 (2018), pp. 116–124.
[25] S. A. Walker, K. S. Double, H. Kunst, M. Zhang, and C. MacCann, *Emotional intelligence and attachment in adulthood: A meta-analysis*, Personality and individual differences, 184 (2022), p. 111174.
[26] M. Wang and W. Deng, *Deep face recognition: A survey*, Neurocomputing, 429 (2021), pp. 215–244.
[27] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan, *The faces of engagement: Automatic recognition of student engagementfrom facial expressions*, IEEE Transactions on Affective Computing, 5 (2014), pp. 86–98.
[28] S. Yagi, Y. Nakata, Y. Nakamura, and H. Ishiguro, *Can an android's posture and movement discriminate against the ambiguous emotion perceived from its facial expressions?*, Plos one, 16 (2021), p. e0254905.
[29] J. Zhang, K. Chen, and J. Zheng, *Facial expression retargeting from human to avatar made easy*, IEEE Transactions on Visualization and Computer Graphics, 28 (2020), pp. 1274–1287.

Edited by: Rajanikanth Aluvalu
Special issue on: Evolutionary Computing for AI-Driven Security & Privacy: Advancing the state-of-the-art applications
Received: Dec 7, 2023
Accepted: Jan 4, 2024