Kashihara, A. et al. (Eds.) (2024). Proceedings of the 32nd International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Exploring Student Emotion via Facial Expressions Using Transfer Learning

Tita HERRADURA[a,b]*, Macario CORDEL II[b] & Merlin Teodosia SUAREZ[b]
[a]*De La Salle University-Dasmarinas, Philippines*
[b]*De La Salle University, Philippines*
*trherradura@dlsud.edu.ph

Abstract: Recognizing student emotions can significantly enhance the learning process. This study investigates the effectiveness of transfer learning with VGG-16 and ResNet-18 models for classifying student emotions based on facial expressions. Leveraging pre-trained models and employing cross-validation, we achieved a robust valence classification accuracy of 92% on FER2013 dataset. However, when applied to MAHNOB-HCI and ACADEMO datasets characterized by limited and subtle emotional cues, performance declined to approximately 82% with overfitting. To enhance model generalization and mitigate overfitting, strategies such as data augmentation, regularization techniques, and hyperparameter optimization are proposed. Our findings demonstrate the effectiveness of transfer learning in recognizing student emotions, which may significantly impact education through personalized learning, improved student engagement, and early intervention.

Keywords: academic emotion, facial expression recognition (FER), transfer learning, VGG-16, ResNet-18

1. Introduction

Human emotion is very important and should be managed to ensure quality of life (Ekman, 2007). For example, a student, when faced with a difficult lesson such as a mathematical or computer programming problem, may tend to experience confusion or frustration when he does not know the solution. Hence, emotion plays a significant role in learning.

There are four emerging emotions of a student during learning – engagement, confusion, frustration, and boredom, these are called academic emotions (Craig, et al., 2004). Engagement is a positive affect that is beneficial to learning, while the other three emotions are negative emotions that could be detrimental to learning and, thus, need to be identified.

Emotion has unique signals; one of the most identifiable signals is shown in the face (Ekman, 2003). This implies that the human face is an accessible and obvious means of detecting emotions. Emerging technologies such as computer vision and neural networks have been valuable tools for facial expression recognition (FER). Various research on FER has been done using diverse neural network architectures such as Long Short-Term Memory (LSTM), hybrid convolutional neural network-recurrent neural network (RNN), and other pre-trained models (Li & Lima, 2021; Febrian et al., 2023). These works delved into detecting basic emotion and mainly utilized publicly available datasets.

In this study, we intend to utilize students' facial expressions while solving a computer programming task to recognize their emotions to provide necessary intervention to maximize learning. Recognizing the emotion of the student using facial expressions may pose challenges because of limited and subtle facial expressions during learning. With this premise, transfer learning (TL) technique was utilized on pre-trained models such as VGG-16 and ResNet-18 for emotion classification.

2. Methodology

Figure 1 presents the diagram of the research methodology conducted in the study. We utilized three different datasets: FER2013, MAHNOB-HCI (Soleymani, et al., 2011) and ACADEMO, an academic emotion dataset (Herradura, et al., 2014). ACADEMO was collected from first-year college students, who were enrolled in an introductory course in computer programming.

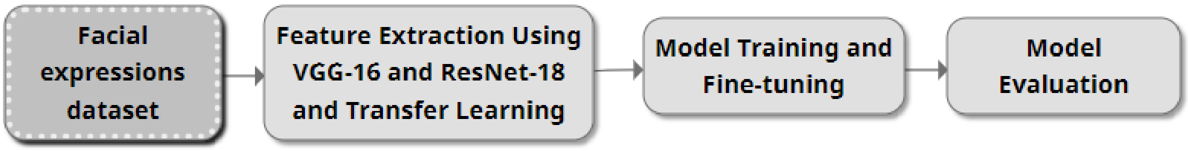

Figure 1. The diagram depicts the process of emotion recognition using facial expressions.

The valence emotion model (positive and negative) was adopted since facial expressions often more clearly convey valence than arousal. From the three datasets, we extracted the face images with positive and negative emotion labels. Only the happy and sad images were selected from FER2013 database. Face images with positive and negative valence emotion labels were also selected from MAHNOB-HCI database. The ACADEMO dataset has a limited number of images. The four academic emotions are divided into valence – positive and negative. Those with *engagement* labels were tagged as positive while *confusion, frustration and boredom* labels were tagged as negative. Table 1 summarizes the distribution of the three datasets to training/validation and testing sets.

Table 1. *Distribution of FER2013, MAHNOB-HCI and ACADEMO datasets with 80% Train-Val and 20% Test ratio*

	FER2013			**MAHNOB-HCI**			**ACADEMO**		
LABEL	***Train-Val***	***Test***	***Total***	***Train-Val***	***Test***	***Total***	***Train-Val***	***Test***	***Total***
Positive	7,186	1,797	**8,983**	5,133	1,284	**6,417**	1,907	477	**2,384**
Negative	4,861	1,216	**6,077**	5,140	1,285	**6,425**	1,719	430	**2,149**
			15,060			**12,842**			**4,533**

VGG-16 and ResNet-18 were used as pre-trained models. These simple deep learning models were chosen because they require less computational power for training and inference, have faster training time and can be deployed easily (Simonyan & Zisserman, 2014; He, et al., 2016). We utilized transfer learning to recognize the subtle facial expressions of students. Transfer learning is a machine learning technique that leverages knowledge gained from one task or domain and applies it to a different but related task or domain (Zhuang et al., 2020). Since VGG-16 and ResNet-18 were pre-trained on large-scale datasets, they might learn generalizable facial feature representations that are beneficial for subtle expressions.

We calibrated various hyperparameters to fine-tune the pre-trained models. We also applied the K-fold cross-validation technique, where K=5. Validation metrics such as accuracy and F1-score were used in this study.

3. Results and Discussion

The results of the comparative experiments are presented in Table 2. The results shown are the train, validation and test accuracy for the three datasets using VGG-16 and ResNet-18. The test accuracy results of FER2013 for VGG-16 and ResNet-18 models reached 92.03% and 92.50% respectively. Since FER2013 has a substantial number of images, it provides the model with sufficient data to learn robust features which reduce overfitting and improve generalization performance.

Both MAHNOB-HCI and ACADEMO accuracy results are overfitting as observed in Table 2, there is a significant gap between the training and validation accuracy for VGG-16 and ResNet-18. This may be due to the limited number of images and subtle facial expressions of the two datasets. Overfitting can be addressed by increasing training data, applying regularization, tuning hyperparameters and simplifying the network model. The test F1-score results as seen in Table 2 for VGG-16 and ResNet-18 indicate that the model has a good balance between precision and recall for the test set.

These findings may suggest that pre-trained models and transfer learning technique can be valuable tools to recognize student emotions from facial expressions. Possible impacts on student learning may include improved engagement, early intervention and personalized learning.

Table 2. Training, *validation and test set classification accuracy and test F1-score results using 5-fold cross-validation.*

	VGG-16				**RESNET-18**			
	ACCURACY			**TEST F1-Score**	**ACCURACY**			**TEST F1-Score**
DATASET	**TRAIN**	**VAL**	**TEST**		**TRAIN**	**VAL**	**TEST**	
FER2013	95.17%	91.23%	**92.03%**	0.92	96.09%	90.86%	**92.50%**	0.93
MAHNOB-HCI	89.61%	85.70%	**92.33%**	0.92	94.99%	88.04%	**90.27%**	0.90
ACADEMO	89.29%	81.84%	**82.47%**	0.82	94.57%	82.00%	**82.80%**	0.83

4. Conclusion and Recommendations

Transfer learning with VGG-16 and ResNet-18 effectively classified emotion in FER2013 with 92% accuracy but struggled with MAHNOB-HCI and ACADEMO, due to limited and subtle facial expressions, achieving 82% accuracy with overfitting. To improve the performance, data augmentation, regularization, and hyperparameter tuning may be considered.

This study demonstrates the potential of using pre-trained models and transfer learning to effectively recognize student emotions from facial expressions. The findings suggest that this technology can significantly impact education by enabling personalized learning, improving student engagement, and facilitating early intervention. We recommend the integration of emotion recognition technology in our educational institutions.

References

Craig, S., Graesser, A., Sullins, J., & Gholson, B. (2004). Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of educational media*, 29(3), 241-250.

Ekman, P. (2007). Emotions revealed: Recognizing faces and feelings to improve communication and emotional life. Macmillan.

Ekman, P., & Friesen, W. V. (2003). Unmasking the face: A guide to recognizing emotions from facial clues (Vol. 10).

Febrian, R., Halim, B. M., Christina, M., Ramdhan, D., & Chowanda, A. (2023). Facial expression recognition using bidirectional LSTM-CNN. *Procedia Computer Science*, 216, 39-47.

Herradura, T. R., Ilao, J. P., & Suarez, M. T. C. (2014, June). Exploring the Behavior of Novice Programmers' EEG Signals for Affect-based Student Modeling. *In Workshop on Utilizing EEG Input in Intelligent Tutoring Systems (ITS2014 WSEEG)* (p. 19).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Li, B., & Lima, D. (2021). Facial expression recognition via ResNet-50. *International Journal of Cognitive Computing in Engineering*, 2, 57-64.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition.

Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2011). A multimodal database for affect recognition and implicit tagging. *IEEE transactions on affective computing*, 3(1), 42-55.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76.