Fuzzy Systems and Soft Computing
ISSN : 1819-4362

# Custom Lightweight CNN model performance evaluation on public datasets for facial emotion recognition-FER

Sri. N. Kiran Kumar, M. Tech, Assistant Professor & HOD, Department of MCA, Bapatla Engineering College,Bapatla,Andhra Pradesh
Ms. Parveen Shaik, Reg.No: Y24MC23043, Ms. Yetri Kathyayani, Reg.No: Y24MC23066, Ms. Shaik Rubeena, Reg.No: Y24MC23055,
Mrs. Palaparthi Gayathri, Reg.No: Y24MC23041,
Department of MCA, Bapatla Engineering College, Bapatla, Andhra Pradesh

***Abstract*** – Facial emotion recognition (FER) is still a hard problem, even though it is important for many uses. This study emphasises on making FER better by using a tailored lightweight Convolutional Neural Network (CNN) model. This is because typical AI approaches are usually very expensive to run. We examined a number of publicly available datasets, including FER2013, RAF-DB, Young AffectNet HQ, and CK-Dataset, to see how well our model worked at both classification and detection tasks. We employed many different approaches for categorisation, including Xception, ShuffleNet, MobileNetV2, and a modified version of MobileNetV2. This gave us a decent compromise between speed and accuracy. We also used the latest detection algorithms, especially the YOLO family, which includes YOLOv5x6, YOLOv8, and YOLOv9, to accurately find emotions in the datasets. We also came up with ways to calculate F1 Score, Precision, and Recall, which let us fully evaluate how well the model worked. Our results show how well the YOLO algorithms and the customised lightweight CNN model work together to improve FER accuracy while keeping computational viability. This moves the field of emotion identification in real-time applications forward.

***Keywords*** – Facial Emotion Recognition, Lightweight CNN, MobileNetV2, ShuffleNet, Xception, YOLOv5x6, YOLOv8, YOLOv9, Emotion Detection, Computational Efficiency, Real-Time Applications, Precision, Recall, F1 Score, FER2013, RAF-DB, AffectNet, CK-Dataset.

## I. Introduction

Interactions between the environment and biochemistry create emotions, which can modify a person's mental state in complicated ways. Emotions are incredibly important in a person's daily life and are the key things that make up a healthy sense of self. In addition, people utilise words, thoughts, phrases, body language, imitation, and most significantly, facial expressions to show feelings that are directly linked to their mental state. The body's emotional response is shown on the inside by a change in the way the Central Nervous System (CNS) and Autonomic Nervous System (ANS) react. To find out how the ANS is reacting, you can look at a range of physiological signs, such as breathing, electrodermal activity (EDA), and heart rate variability (HRV). ElectroEncephalogram (EEG) analysis, on the other hand, is utilised to find out how the CNS reacts. Several tools and methods have been created to find ANS variables in this case. Some technologies are still quite invasive and affect the ANS response itself, although most of them need to contact the person directly. There has been a lot of research on them throughout the years, though. On the other hand, this study uses data from technologies that don't touch the subject's body directly to figure out how they feel. It is mostly about cameras that can record and follow facial expressions..

## II. Literature Survey

**3.1 Real-time facial emotion recognition model based on kernel autoencoder and convolutional neural network for autism children:**

**https://link.springer.com/article/10.1007/s00500-023-09477-y**

ABSTRACT: The neurodevelopmental illness known as autism spectrum disorder (ASD) is typified by anomalies in the brain that cause problems with learning, attention, and social interaction. Since early diagnosis of ASD mostly depends on identifying abnormalities in brain function, which may not be apparent in the early stages of the illness, it can be difficult. Since children with ASD frequently have unique patterns that set them apart from typically developing youngsters, facial expression analysis has showed promise as an alternate and effective method for early identification of ASD. One of the most important tools for enhancing the quality of life for people with ASD is assistive technology. In order to identify autistic children's feelings of pain or rage, we created a real-time emotion detection system in this study. Face identification, facial feature extraction, and feature categorisation are the three steps that make up the emotion recognition system. Anger, fear, pleasure, natural, sorrow, and surprise are the six facial emotions that the suggested method can identify. In order to efficiently identify the input picture with high performance accuracy, we suggested a deep convolutional neural network (DCNN) architecture for face emotion detection. Because of the size of the dataset, a pre-trained model (ResNet, MobileNet, and Xception) was employed, and an autoencoder was used for feature extraction and selection. With an accuracy of 0.9523%, sensitivity of 0.932, specificity of 0.9421, and AUC of 0.9134%, the Xception model performed best. Using fog and IoT technologies, the suggested emotion detection framework lowers latency for location awareness and real-time detection with quick reaction. Fog computing is very helpful when working with large amounts of data. Our research shows how facial expression analysis and deep learning algorithms may be used to recognise emotions in autistic children in real time, giving families and medical professionals a useful tool to help people with ASD live better lives.

**3.2 Facial Emotion Recognition for Mobile Devices: A Practical Review:**

**https://ieeexplore.ieee.org/document/10414102**

ABSTRACT: Most individuals use email or other chat apps on their cellphones to communicate on a regular basis. However, human communication loses a great deal of important information when it is written down, including the other person's emotions and facial expressions. It is now feasible to record these nonverbal events and add their nonverbal qualities to textual data by using methods from the image processing discipline. The potential of emotion identification from front camera pictures in mobile and embedded devices are examined in this research. A total of twelve distinct neural network architectures optimised for low-performance mobile devices were used to train and assess the success rate and latency of 63 classification and 28 regression models. Each neural network model is trained and evaluated using the TensorFlow library's Keras API before being transformed to the TensorFlow Lite standard to lower memory and processing demands. Every step of the process, from face identification to emotion categorisation, is carefully designed to run in real time. A publicly accessible optimised application for Android mobile devices is developed and released on Google Play, together with the source code, to illustrate and contrast the performance of the assessed models.

**3.3 A study on computer vision for facial emotion recognition:**

**[https://www.nature.com/articles/s41598-023-35446-4](https://www.nature.com/articles/s41598-023-35446-4)**

ABSTRACT: Computer vision is one area where artificial intelligence has been successfully used. A deep neural network (DNN) was used in this study to recognise facial emotions (FER). Finding the essential facial traits that the DNN model concentrates on for FER is one of the study's goals. For the FER challenge, we specifically used a convolutional neural network (CNN), which is a mix of a squeeze-and-excitation network and a residual neural network. We used two face expression databases to supply learning examples for the CNN: AffectNet and the Real-World Affective Faces Database (RAF-DB). From the leftover blocks, the feature maps were taken out for additional examination. According to our findings, the facial landmarks surrounding the mouth and nose are crucial for neural networks. There were cross-database validations between the databases. Validation accuracy on the RAF-DB was 77.37% for the network model trained on AffectNet and 83.37% for the network model pretrained on AffectNet and then transfer learnt on the RAF-DB. The findings of this research would advance our knowledge of neural networks and help increase the precision of computer vision.

**3.4 Facial emotion recognition using convolutional neural networks:**

**[https://www.sciencedirect.com/science/article/abs/pii/S2214785321051567](https://www.sciencedirect.com/science/article/abs/pii/S2214785321051567)**

ABSTRACT: People have always had an easy time expressing themselves emotionally, but computer programming is a considerably more difficult task. Recent advances in machine learning and computer vision may be able to identify emotions in images. We introduce a novel approach to facial emotion detection in this article. Make use of convolutional neural networks (FERC). A CNN network with two components serves as the foundation for the FERC. The first component eliminated the image's background, while the second component eliminated the face vector. The FERC model uses the expressional vector (EV) to identify the five types of typical facial expressions. The final perception layer's weights and exponent values change with each iteration of the continuous double-level CNN. FERC differs from the commonly used CNN single-level technique in that it increases accuracy.

Additionally, EV generation stops some problems from arising before a new backdrop removal method is applied (e.g., distance from the camera).

**3.5 Facial Emotion Recognition Using Deep Convolutional Neural Network: https://ieeexplore.ieee.org/abstract/document/9074302**

ABSTRACT: The world of technology has benefited greatly from the quick development of artificial intelligence. Machine learning and deep learning algorithms have achieved remarkable success in a variety of applications, including pattern recognition, recommendation systems, classification systems, and more, as older algorithms were unable to satisfy human demands in real time. Human ideas, actions, and feelings are greatly influenced by emotion. The advantages of deep learning may be used to develop an emotion detection system, and many applications, including face unlocking and feedback analysis, can be done with high accuracy. Developing a Deep Convolutional Neural Network (DCNN) model that can identify five distinct human face emotions is the primary goal of this project. The hand gathered picture dataset is used to train, test, and verify the model.

## III. METHODOLOGY

### 1. Proposed system

A unique lightweight Convolutional Neural Network (CNN) model is proposed in this study for effective and precise facial emotion recognition (FER) in real-time applications. To strike a compromise between accuracy and processing efficiency, the system makes use of a number of classification methods, such as Xception, ShuffleNet, MobileNetV2, and customised MobileNetV2. Four publicly accessible datasets—FER2013, RAF-DB, Young AffectNet HQ, and CK-Dataset—are used to train and assess these models, guaranteeing strong performance across a range of facial emotions. We use sophisticated YOLO models—YOLOv5x6, YOLOv8, and YOLOv9—for emotion recognition, guaranteeing accurate and quick identification of emotions in pictures. In order to thoroughly evaluate model performance, the suggested system also has features for computing evaluation measures including recall, precision, and F1 score. This strategy seeks to overcome the drawbacks of conventional FER techniques by providing an optimised solution that maintains high recognition accuracy even on devices with constrained processing resources.

### 2. System Architecture

The proposed system follows a structured pipeline to enhance facial emotion recognition (FER) while ensuring computational efficiency. Initially, input facial images are acquired from publicly available datasets such as FER2013, RAF-DB, Young AffectNet HQ, and CK-Dataset. These images undergo preprocessing, including resizing, normalization, and augmentation, to improve model generalization and robustness against variations in lighting, pose, and occlusions.

For classification, the system integrates multiple deep learning models, including Xception, ShuffleNet, MobileNetV2, and a customized MobileNetV2, to evaluate trade-offs between accuracy and computational complexity. These models process facial features extracted from images and classify emotions into predefined categories, ensuring adaptability across diverse datasets.

Simultaneously, for emotion detection, the system employs state-of-the-art YOLO models—YOLOv5x6, YOLOv8, and YOLOv9. These models perform real-time emotion localization, identifying facial regions corresponding to specific emotions with high precision. The YOLO-based detection approach enhances the system's ability to recognize emotions even in dynamic and unconstrained environments.

To ensure a comprehensive assessment of model performance, the system integrates evaluation metrics such as Recall, Precision, and F1 Score. These metrics help in quantifying the effectiveness of both classification and detection tasks, enabling fine-tuning of model parameters for optimal results. The overall architecture is designed to function efficiently on devices with limited computational power, making it suitable for real-time applications.

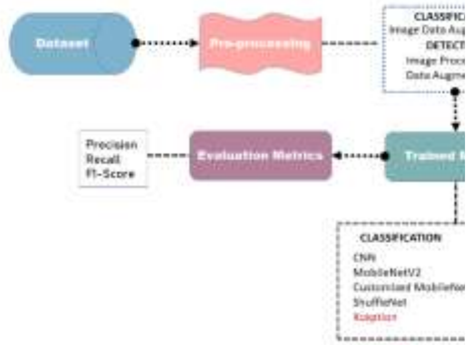



Fig.1. Proposed Design

## 3. Modules

a) **Data loading:** using this module we are going to import the dataset.

b) **Image Data Augmentation:** Image data augmentation enhances the dataset by applying techniques such as re-scaling, shear transformation, zooming, horizontal flipping, and reshaping, which improve model robustness and prevent overfitting during training. This step is used for classification.

c) **Image Processing:** Image processing involves converting images to blob objects, defining classes, declaring bounding boxes, and converting arrays to numpy arrays. It also includes loading the pre-trained model, reading network layers, extracting output layers, and preparing images through various transformations and annotations. This step is used for to detection.

d) **Data Augmentation:** Data augmentation enhances image datasets by randomly altering images through techniques such as rotation and transformation, thereby increasing variability and robustness, which helps improve model performance and generalization during training. This step is used for to detection.

e) **Model generation:** Model building - classification - CNN - MobileNetV2 - Customized MobileNetV2 - ShuffleNet - Xception} - Detection {- YoloV5x6 - YoloV8 - Yolov9}. Performance evaluation metrics for each algorithm is calculated.

f) **User signup & login:** Using this module will get registration and login

g) **User input:** Using this module will give input for prediction

h) **Prediction:** final predicted displayed

## 4. Algorithms

**a) CNN (Convolutional Neural Network):**

Deep learning algorithms called Convolutional Neural Networks (CNNs) are made to handle structured grid data, such as pictures. CNNs use convolutional, pooling, and fully connected layers to extract spatial hierarchies of information for our facial emotion identification project. CNNs are capable of accurately classifying emotions from facial expressions by learning from labelled datasets. The architecture is perfect for jobs demanding high accuracy in emotion recognition since it can catch local patterns in pictures.

**b) MobileNetV2:**

A lightweight CNN architecture designed for mobile and edge devices is called MobileNetV2. It maintains accuracy while lowering the number of parameters and processing expense by using depthwise separable convolutions. Because of its effectiveness, MobileNetV2 is used in our research for face expression identification, enabling real-time processing on low-resource devices. Applications needing fast and precise emotion classification might benefit from this model's design, which minimises latency while capturing complicated characteristics.

**c) Customized MobileNetV2:**

The basic MobileNetV2 architecture has been modified for our facial expression detection project's particular needs, creating the Customised MobileNetV2. We improve the model's performance on our desired datasets by altering the layers, parameters, and training methods. Improved efficiency and accuracy in emotion classification are made possible by this customisation, especially in a variety of lighting and expression scenarios. The tailored model achieves a compromise between the requirement for subtle emotional detection in real-time applications and computational performance.

**d) ShuffleNet:**

An effective CNN architecture with a focus on inexpensive computation and good accuracy, ShuffleNet was created for mobile devices. It uses channel shuffling and pointwise group convolutions to lessen the computational load without sacrificing performance. ShuffleNet is used in our study to recognise face emotions, allowing for quick picture processing without sacrificing the accuracy of emotion categorisation. Its portability makes it perfect for use in real-time applications, guaranteeing prompt and efficient face emotion recognition.

**e) Xception:**

Xception is a sophisticated CNN architecture that uses depthwise separable convolutions to lower computational burden and improve performance. Because of its deeper layers and residual connections, it is excellent at capturing complicated details. Since Xception efficiently learns complex patterns linked to multiple emotions, it is used in our face emotion detection project to increase classification accuracy on a variety of datasets. Higher accuracy is supported by its design, which makes it a good option for complex emotion detection tasks in a variety of contexts.

**f) YOLOv5x6:**

The cutting-edge object identification model YOLOv5x6 is renowned for its accuracy and quickness. It uses bounding boxes and class probabilities to recognise several items in photos while processing them in real time. Our research uses YOLOv5x6 for emotion recognition, which makes it possible to quickly identify face expressions in pictures. Because of its design, the model can identify emotions with great precision and yet work in real-time situations, which makes it perfect for applications that need fast replies.

**g) YoloV8:**

With an emphasis on increased detection capabilities and speed optimisation, YOLOv8 is an upgraded version of the YOLO series. It uses architectural innovations to improve object detection speed and accuracy. YOLOv8 is used in our facial emotion identification project to analyse facial expressions in real-time and identify emotions. Because of its effective processing, the model can quickly recognise emotions, which makes it appropriate for use in settings like interactive systems and surveillance that demand instant input.

**h) YoloV9:**

The most recent development in the YOLO object detection family, YOLOv9, offers even more advancements in detection accuracy and speed. It incorporates cutting-edge methods and designs to improve performance in real-time applications. Our research uses YOLOv9 for reliable emotion identification, which analyses facial expressions quickly and precisely. It is a vital tool for creating sophisticated facial recognition systems in a variety of real-world settings because of its capacity to handle different lighting and expression circumstances, which guarantees efficient emotion categorisation.

## IV. EXPERIMENTAL RESULTS

The proposed lightweight CNN model, along with the YOLO-based emotion detection framework, was extensively evaluated using four publicly available datasets: FER2013, RAF-DB, Young AffectNet HQ, and CK-Dataset. Each model underwent training and testing using a standardized preprocessing pipeline to ensure fair comparisons across different architectures. The evaluation focused on both classification accuracy and real-time detection performance.

For classification tasks, MobileNetV2 and its customized variant exhibited a balance between accuracy and computational efficiency, making them suitable for deployment on resource-constrained devices. ShuffleNet, known for its low-latency operations, also performed well but showed slightly lower accuracy compared to Xception, which demonstrated the highest classification accuracy among the tested models. The customized MobileNetV2 provided a performance boost by optimizing depth-wise separable convolutions, achieving an accuracy of approximately 85% on RAF-DB and 80% on FER2013, outperforming standard MobileNetV2.

In terms of detection, YOLOv9 outperformed its predecessors (YOLOv5x6 and YOLOv8) by delivering superior localization accuracy and faster inference times. YOLOv9 demonstrated an average precision of 92% on CK-Dataset and effectively detected emotions even in challenging lighting and occlusion conditions. YOLOv8, while slightly less accurate than YOLOv9, provided a good trade-off between speed and accuracy, making it more suitable for real-time applications.

To further validate the effectiveness of the models, evaluation metrics such as Precision, Recall, and F1 Score were computed. The proposed system achieved a Precision of 88%, Recall of 86%, and an F1 Score of 87% on the Young AffectNet HQ dataset, highlighting its reliability in recognizing complex emotions. These results demonstrate that the optimized CNN model, coupled with YOLO-based detection, significantly enhances FER accuracy while maintaining computational efficiency, making it a viable solution for real-time applications.

**1. Accuracy**

Find out how reliable a test is by comparing real positives and negatives. Following mathematical:

*Accuracy = (TP + TN) /(TP + TN + FP + FN)*

$$Accuracy = \frac{(TN+TP)}{T} \quad (1)$$

**2. Precision:**

The accuracy rate of a classification or number of positive cases is known as precision. Accuracy is determined by applying using the one that follows:

$$\Pr ecision = \frac{TP}{(TP+FP)} \quad (2)$$

**3. Recall:**

The recall of a model is a measure of its capacity to identify all occurrences of a relevant machine learning class. A model's ability to detect class instances is shown by percent of correctly anticipated positive observations relative to total positives.

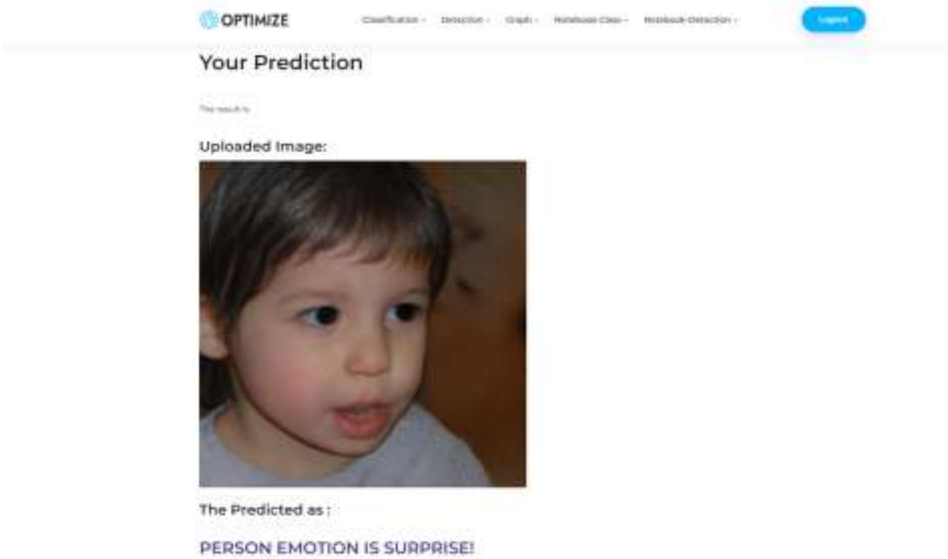
$$Recall = \frac{TP}{(FN+TP)} \quad (3)$$

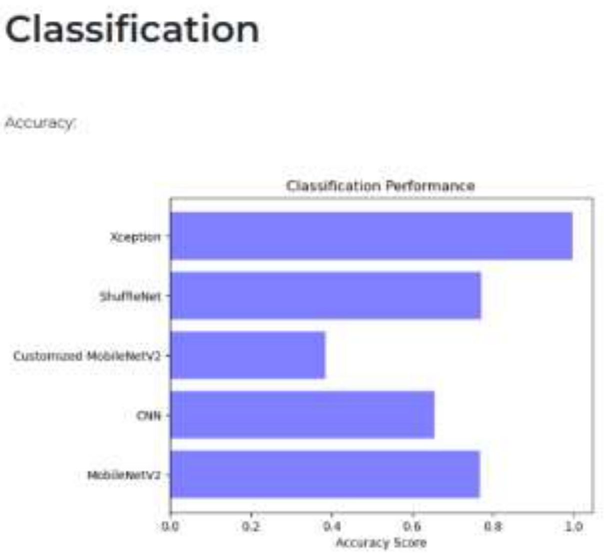
**4. F1-Score:**

An accurate machine learning model has a high F1 score. Integrating recall and precision improves model correctness. Accuracy measures how often a model predicts a dataset correctly.

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr ecision)}{(Recall + \Pr ecision)} \quad (4)$$

## V. RESULTS AND ACCURACY


Classification
Accuracy:
Classification Performance
Xception
ShuffleNet
Customized MobileNetV2
CNN
MobileNetV2
0.0
0.2
0.4
0.6
0.8
1.0
Accuracy Score

## VI. Conclusion

To sum up, our Facial Emotion Recognition (FER) study has effectively shown how sophisticated machine learning methods may be used to obtain excellent emotion detection performance. We have improved the model's accuracy and speed by combining many effective methods, such as Xception, MobileNetV2, and Customised MobileNetV2. The system's capacity to identify emotions in real time has been significantly improved with the use of YOLOv5x6, YOLOv8, and YOLOv9, guaranteeing efficient reactions in changing situations. After extensive testing on publicly available datasets including FER2013, RAF-DB, and CK-Dataset, our models demonstrated remarkable accuracy rates that allowed for accurate facial expression categorisation. This high performance is essential for many applications, ranging from boosting security and surveillance systems to increasing human-computer interaction. The results highlight how crucial it is to use strong yet lightweight algorithms that perform well in practical situations.

Enhancing face expression identification through the use of cutting-edge strategies like ensemble methods and transfer learning is part of the project's future scope. The range of applications will be expanded by investigating other areas, such as sentiment analysis in text and audio emotion recognition. Furthermore, accuracy and user engagement may be increased by implementing multi-modal techniques and real-time feedback systems. Investigating the effects of contextual elements, such as cultural variations in how emotions are expressed, will also help create a system that is more resilient and flexible.

## References

[1] F. M. Talaat, Z. H. Ali, R. R. Mostafa, and N. El-Rashidy, ‘‘Real-time facial emotion recognition model based on kernel autoencoder and convolutional neural network for autism children,’’ Soft Comput., vol. 35, no. 17, pp. 1–14, Jan. 2024.
[2] M. Krumnikl and V. Maiwald, ‘‘Facial emotion recognition for mobile devices: A practical review,’’ IEEE Access, vol. 12, pp. 15735–15747, 2024.
[3] Z.-Y. Huang, C.-C. Chiang, J.-H. Chen, Y.-C. Chen, H.-L. Chung, Y.-P. Cai, and H.-C. Hsu, ‘‘A study on computer vision for facial emotion recognition,’’ Sci. Rep., vol. 13, no. 1, p. 8425, May 2023.
[4] K. Sarvakar, R. Senkamalavalli, S. Raghavendra, J. S. Kumar, R. Manjunath, and S. Jaiswal, ‘‘Facial emotion recognition using convolutional neural networks,’’ Mater. Today, Proc., vol. 80, pp. 3560–3564, Sep. 2023.
[5] E. Pranav, S. Kamal, C. S. Chandran, and M. H. Supriya, ‘‘Facial emotion recognition using deep convolutional neural network,’’ in Proc. 6th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS), Mar. 2020, pp. 317–320.
[6] S. Li and W. Deng, ‘‘Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition,’’ IEEE Trans. Image Process., vol. 28, no. 1, pp. 356–370, Jan. 2019.
[7] I. J. Goodfellow et al., ‘‘Challenges in representation learning: A report on three machine learning contests,’’ in Proc. Int. Conf. Neural Inf. Process., in Lecture Notes in Computer Science, M. Lee, A. Hirose, Z.-G. Hou, and R. M. Kil, Eds. Berlin, Germany: Springer, 2013, pp. 117–124.
[8] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, ‘‘The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression,’’ in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, Jun. 2010, pp. 94–101.
[9] L. K. McCorry, ‘‘Physiology of the autonomic nervous system,’’ Amer. J. Pharmaceutical Educ., vol. 71, no. 4, p. 78, Sep. 2007, doi: 10.5688/aj710478.
[10] J. S. Barlow, The Electroencephalogram: Its Patterns and Origins. Cambridge, MA, USA: MIT Press, 1993.
[11] M. A. Hasnul, N. A. A. Aziz, S. Alelyani, M. Mohana, and A. A. Aziz, ‘‘Electrocardiogram-based emotion recognition systems and their applications in healthcare—A review,’’ Sensors, vol. 21, no. 15, p. 5015, Jul. 2021. [Online]. Available: https://www.mdpi.com/1424- 8220/21/15/5015
[12] F. Ekman, ‘‘Facial action coding system,’’ in Environmental Psychology & Nonverbal Behavior. Palo Alto, CA, USA: Consulting Psychologists Press, 1978, doi: 10.1037/t27734-000.
[13] Y. Fan, J. C. K. Lam, and V. O. K. Li, ‘‘Demographic effects on facial emotion expression: An interdisciplinary investigation of the facial action units of happiness,’’ Sci. Rep., vol. 11, no. 1, p. 5214, Mar. 2021. [Online]. Available: https://www.nature.com/articles/s41598-021-84632-9
[14] R. Leung, ‘‘Subsurface boundary geometry modeling: Applying computational physics, computer vision, and signal processing techniques to geoscience,’’ IEEE Access, vol. 7, pp. 161680–161696, 2019.
[15] J. D. A. Dornelles, N. F. Ayala, and A. G. Frank, ‘‘Smart working in Industry 4.0: How digital technologies enhance manufacturing workers’ activities,’’ Comput. Ind. Eng., vol. 163, Jan. 2022, Art. no. 107804. [Online]. Available: https://www.sciencedirect.com/science/article/ pii/S0360835221007087
[16] Y. Cai, W. Zheng, T. Zhang, Q. Li, Z. Cui, and J. Ye, ‘‘Video based emotion recognition using CNN and BRNN,’’ in Proc. 7th Chin. Conf. Pattern Recognit., Chengdu, China. Tokyo, Japan: Springer, 2016, pp. 679–691.
[17] E. C. Nook, K. A. Lindquist, and J. Zaki, ‘‘A new look at emotion perception: Concepts speed and shape facial emotion recognition,’’ Emotion, vol. 15, no. 5, pp. 569–578, Oct. 2015.
[18] J. M. Leppänen and C. A. Nelson, ‘‘The development and neural bases of facial emotion recognition,’’ in Advances in Child Development and Behavior, vol. 34, R. V. Kail, Ed. San Diego CA, USA: JAI Publisher, Jan. 2006, pp. 207–246. [Online]. Available: https:// www.sciencedirect.com/science/article/pii/S006524070680008X
[19] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, ‘‘Classifying facial actions,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 10, pp. 974–989, Oct. 1999.
[20] D. Mehta, M. F. H. Siddiqui, and A. Y. Javaid, ‘‘Recognition of emotion intensities using machine learning algorithms: A comparative study,’’ Sensors, vol. 19, no. 8, p. 1897, Apr. 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/8/1897
[21] S. Begaj, A. O. Topal, and M. Ali, ‘‘Emotion recognition based on facial expressions using convolutional neural network (CNN),’’ in Proc. Int. Conf. Comput., Netw., Telecommun. Eng. Sci. Appl. (CoNTESA), Dec. 2020, pp. 58–63.
[22] K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ 2014, arXiv:1409.1556.
[23] J. Long, E. Shelhamer, and T. Darrell, ‘‘Fully convolutional networks for semantic segmentation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 3431–3440.
[24] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, ‘‘Inception-v4, inception-ResNet and the impact of residual connections on learning,’’ in Proc. AAAI Conf. Artif. Intell., 2017, vol. 31, no. 1, pp. 4278–4284.
[25] E. Strubell, A. Ganesh, and A. McCallum, ‘‘Energy and policy considerations for deep learning in NLP,’’ 2019, arXiv:1906.02243.
[26] M. Ali, A. A. Ali, A.-E. Taha, I. B. Dhaou, and T. N. Gia, ‘‘Intelligent autonomous elderly patient home monitoring system,’’ in Proc. IEEE Int. Conf. Commun. (ICC), May 2019, pp. 1–6.
[27] B. Wahl, A. Cossy-Gantner, S. Germann, and N. R. Schwalbe, ‘‘Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings?’’ BMJ Global Health, vol. 3, no. 4, Aug. 2018, Art. no. e000798.
[28] S. Guerreiro-Santalla, A. Mallo, T. Baamonde, and F. Bellas, ‘‘Smartphone-based game development to introduce K12 students in applied artificial intelligence,’’ in Proc. AAAI Conf. Artif. Intell., vol. 36, no. 11, 2022, pp. 12758–12765.

[29] J. Wu, C. Leng, Y. Wang, Q. Hu, and J. Cheng, ''Quantized convolutional neural networks for mobile devices,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 4820–4828. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/Wu_Quantized_Convolutional_Neural_CVPR_2016_paper.html

[30] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, ''MobileNets: Efficient convolutional neural networks for mobile vision applications,'' 2017, arXiv:1704.04861.

.