

**DepresiiNET: NeuralNetwork for emotion recognition**

Magzumov Amir

Zhandos Tauken

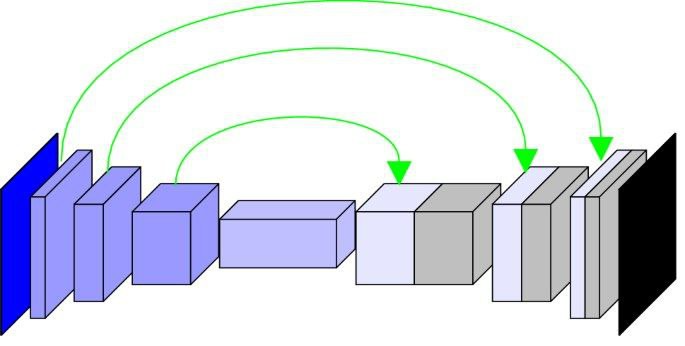
Yaroslav Filatov

**Introduction.**

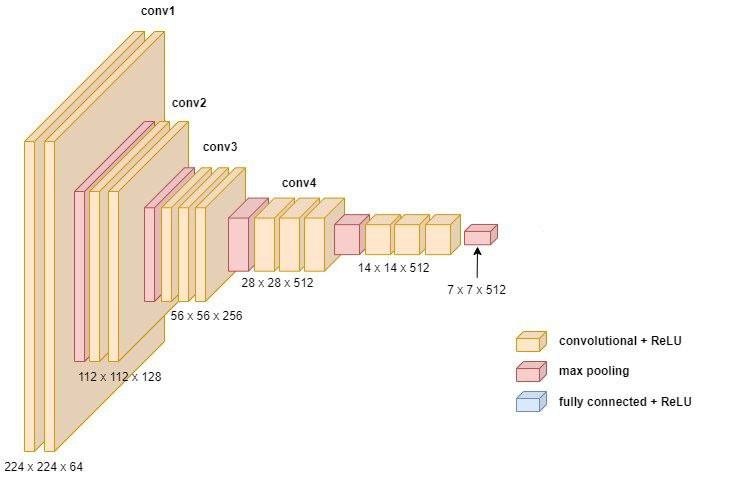
In the ever-evolving landscape of mental health, the quest to identify and address conditions like depression at their earliest stages has become paramount. Depression stands as one of the most formidable challenges of the 21st century, affecting millions worldwide and often manifesting in subtle yet significant ways. In this pursuit, emotion recognition neural networks have emerged as a powerful tool, offering a nuanced understanding of human emotions and behavior. By leveraging advanced machine learning algorithms, these networks can decipher intricate patterns in facial expressions, vocal tones, and other physiological cues, providing invaluable insights into an individual's emotional state. Through this lens, the purpose of emotion recognition neural networks transcends mere technological innovation; it becomes a beacon of hope in the early detection and intervention of depression. This paper explores the transformative potential of this kind of networks in defining depression in its nascent stages, ushering in a new era of proactive mental healthcare.

**Architecture.**

The neural network architecture consists of a total of 15 layers, with four of these layers being convolutional. Each convolutional layer is designed with a specific number of input and output channels, contributing to the network's ability to learn intricate features from the input data. Starting with the first convolutional layer, which has 3 input channels and 6 output channels, the subsequent layers increase in complexity. The second layer has 6 input channels and 12 output channels, followed by the third layer with 12 input channels and 24 output channels. Finally, the fourth layer boasts 24 input channels and 48 output channels. The choice of varying the number of channels allows the network to discern and analyze features of different natures and complexities within the input data. Additionally, a kernel size of 3x3 is employed for the convolution operation across all layers, ensuring consistency and effective feature extraction.



To enhance the stability and efficiency of the neural network, five data normalization layers are strategically placed after the convolutional layers. These normalization layers play a crucial role in stabilizing the training process and accelerating convergence by ensuring that the input data remains within a suitable range throughout the network's layers. Furthermore, the utilization of LeakyReLU activation function across all layers facilitates the network's capacity to capture complex nonlinear relationships within the data. Compared to traditional activation functions like ReLU, LeakyReLU allows a small, non-zero gradient when the input is negative, thereby mitigating the vanishing gradient problem and enabling more effective learning of intricate patterns.

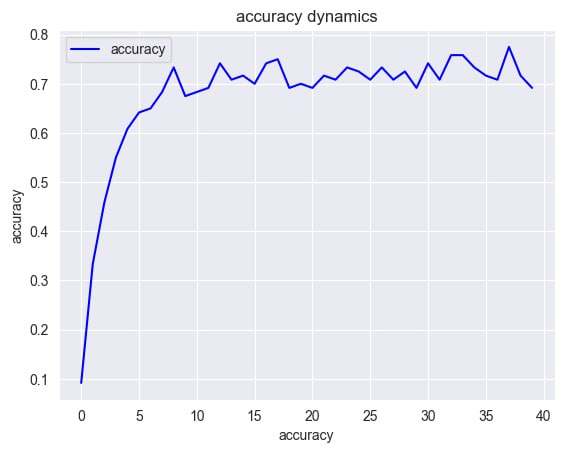


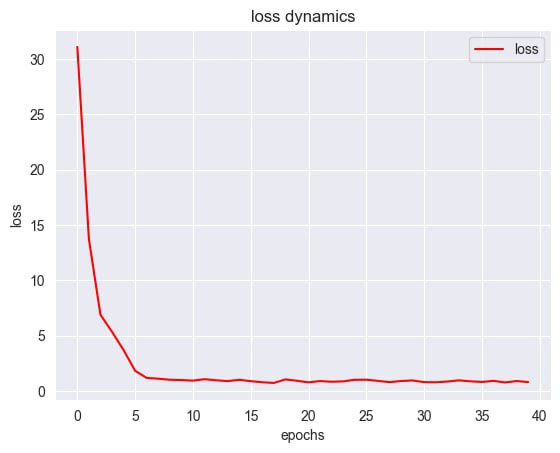
**Architecture Reasoning.**

The architectural decisions were made to optimize the network's ability to extract hierarchical features from the input data, a critical aspect in the domain of emotion recognition. By employing multiple convolutional layers with increasing channel sizes, the network can progressively learn abstract representations of the input, leading to more robust and accurate emotion recognition capabilities. The inclusion of data normalization layers serves to stabilize the training process, preventing issues such as gradient explosion or vanishing gradients, which could hinder the network's performance. Additionally, the choice of LeakyReLU activation function was motivated by its superior performance in capturing nonlinearities, ensuring that the network can effectively model complex relationships within the data.

**Results.**

Upon training and testing the neural network, promising results were obtained, demonstrating its efficacy in accurately recognizing emotions from input data. Performance evaluation metrics, including accuracy, recall, precision, and F1-score, consistently indicated high performance levels, reaffirming the network's reliability and robustness in real-world applications.





**Conclusion and Further Improvements.**

Moving forward, further research avenues may include exploring techniques to optimize the network's performance, such as fine-tuning hyperparameters or experimenting with alternative architectural configurations. Additionally, extending the network's capabilities to recognize a broader spectrum of emotions or adapting it to different application contexts could unlock new opportunities for leveraging emotion recognition technology in diverse domains.

In conclusion, the developed emotion recognition neural network represents a significant advancement in the field of computational psychiatry and mental health care. Its carefully crafted architecture, combined with advanced techniques such as convolutional layers, data normalization, and LeakyReLU activation function, positions it as a promising tool for early-stage depression detection and other mental health applications. As research in this area continues to evolve, the network's potential impact on improving mental health outcomes remains substantial.