Analysis of facial emotion recognition rate for real-time application using NVIDIA Jetson Nano in deep learning models

Usen Dudekula, Purnachand N
School of Electronics Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India.

Article Info

Article history:

Received May 14, 2022
Revised Dec 15, 2022
Accepted Dec 20, 2022

Keywords:

Deep learning
NVIDIA Jetson Nano
Transfer learning
Visual geometry group
Xception Architecture

ABSTRACT

Detecting facial emotion expression is a classic research problem in image processing. Face expression detection can be used to help human users monitor their stress levels. Perceiving an individual's failure to communicate specific looks might help analyze early psychological disorders. several issues like lighting changes, rotations, occlusions, and accessories persist. These are not simply traditional image processing issues, yet additionally, action units that make gathering activity of facial acknowledgment troublesome look information, and order of the demeanor. In this study, we use Xception taking into account Xception and convolution neural network (CNN), which is easy to focus on incredible parts like the face, and visual geometric group (VGG-19) used to extract the facial feature using the OpenCV framework classifying the image into any of the basic facial emotions. NVIDIA Jetson Nano has a high video handling outline rate. Accomplishing preferable precision over the recently evolved models on software. The average accuracies for standard data set CK+," on NVIDIA Jetson Nano, the accuracy rate is 97.1% in the Xception model in the convolutional neural network, 98.4% in VGG-19, and real-time environment accuracy using OpenCV, accuracy rate is 95.6%.

Corresponding Author:

Purnachand N
School of Electronics Engineering, VIT-AP University
G-30, Inavolu, Beside AP Secretariat Amaravati, Andhra Pradesh 522237, India
Email: chanduinece@gmail.com

1. INTRODUCTION

Acknowledgment of facial feelings is the strategy for recognizing human sentiments through looks. The human mind intuitively perceives feelings, and programming has now been made that can likewise perceive emotions we use Xception considering Xception and convolution neural network (CNN), which is easy to zero in on inconceivable parts like the face. Time, this technology is becoming more specific, and will finally be able to read emotions as our brains do. Using facial expressions and vocal tones, people usually interpret the profound emotional conditions of others, for example, excitement, sorrow, and rage. According to a survey [1]. The rapid development of artificial intelligence techniques for automatic facial expression processing (FER), including human-PC association (HCI), computer-generated reality (VR), increased reality (AR), the high-level driver helps frameworks (ADASs), and diversion. Consequently, various sensors, like electromyography (EMG), electrocardiogram (ECG), and camera is the most encouraging kind of sensor since it offers the most nitty gritty hints, for facial feeling acknowledgment (FER) inputs [2]–[7]. Due to a variety of challenges faced in the identification and recognition of lighting and accessories, partial occlusions, head deviation of the facial regions, and a low identification rate, FER remains difficult. Deep learning methodologies may offer a suboptimal solution to these problems [8], [9]. GPUs are also utilized in other applications that benefit from their parallel nature, including machine learning and deep learning applications.

The method executed on a negligible cost embedded structure thinking about the Raspberry Pi 3. The outcome was that the typical rate was 98.3% for face acknowledgment, with a 7.09 frame per second (FPS) [10]. Face acknowledgment calculation by incorporating a couple of face acknowledgment strategies that combine principal component analysis (PCA), linear discriminant analysis (LDA), based on Raspberry. This blend system has given precision processing units (GPUs) are electronic circuits originally designed for graphical and image processing. GPUs are also utilized in other applications that benefit from their parallel nature, including machine learning and deep learning applications.

Face acknowledgment calculation by incorporating a couple of face acknowledgment strategies that combine LDA, PCA based Raspberry pi 3. This blend technique has given a precision of 97% [11] Acknowledgment continuously outlines carried out Raspberry Pi in this work, the framework had the option to deal with only four 4FPS [12]. Friendly, stand-alone entry control framework utilizing facial acknowledgment, decreased the face include size by 40% of the Gabor LBP facial attributes. This framework accomplished facial acknowledgment exactness of 97.27%, with a 5.26 FPS handling speed in an indoor climate [13]. FPGA gives better productivity over CPU and GPU. It is likewise expressed that while CPUs and GPUs could offer high hypothetical maximized execution, they are not as proficiently used with BNNs, since binarized bit-level operations are more qualified for custom hardware equipment [14]. BNN design, which is tested on a Zed Board, containing a Zynq-7020 SoC, and then compared to an Intel Xeon server CPU, NVIDIA Tesla K40 GPU, and an NVIDIA Jetson TK1 embedded GPU. The paper shows significant improvement in performance and energy efficiency from the FPGA compared to the embedded GPU and the CPU. Compared to the GPU the FPGA is 8x slower but 6x more energy efficient. [15].

Various examinations have as of late regarding deep learning, concentrated on FER issues in design acknowledgment, FER has had astounding achievement. Every one of the past investigations has gained significant headway in the field of emotion-recognizable proof when contrasted with past endeavors, yet they come up short on a clear procedure for distinguishing key facial regions for emotion recognition. The main goal of this work is to implement neural networks of different types to compare their performance when run on two different platforms. The comparison should highlight the differences between the tested hardware and real-time OpenCV in software using the same deep-learning models. The average accuracies for the standard data set"CK+," on NVIDIA Jetson Nano, the accuracy rate is 97.1% in the Xception model in CNN, 98.4% in VGG-19, and real-time environment accuracy using OpenCV, accuracy rate is 95.6%. Our examination is organized as follows. Area 2 is about the Comparison of CNN'S on Hardware. The proposed, Deep learning models are characterized in Section 3. The experimental outcomes and translation for Section 4 are reported. Area 5 gives the end explanation.

2. COMPARISON OF CONVOLUTIONAL NEURAL NETWORKS ON HARDWARE

Hardware can be used to accelerate computer processes or computing operations by identifying and utilizing the different strengths of hardware. This is also true for machine learning applications, where hardware is a necessity for achieving high-performance inference. Applications using neural networks use the fact that matrix multiplication and binary arithmetic can be drastically accelerated using parallel hardware. Harper in [16]. Comparison will be made using a few chosen example hardware platforms. This comparison will focus on relevant parameters such as performance, ease of use, power consumption, and flexibility in the design.

2.1. Field programmable gate arrays (FPGAs)

FPGAs are semiconductor devices that can be programmed and reprogrammed to the desired functionality or application after manufacturing. FPGA consists of two basic components: lookup tables (LUTs) and flip-flops (FFs). The LUTs are truth tables that handle combinatorial logic. Instead of having a set number of logical gates ready, each LUT can be customized to work as any logic gate [17]. Flipflops are binary registers that save the state between clock cycles by holding either a 1 or 0 until the next clock edge comes. there are two other components of an FPGA that should be discussed. Firstly, block RAM (BRAM) is memory located within the FPGA. While memory can also be located outside of the FPGA, like with EPROM, SRAM, or SD card, BRAM is used for data that needs to be accessed without going outside of the FPGA through the I/O blocks [18]. Secondly, there are DSP slices. These are usually prebuilt multiplier-accumulate circuitry used when certain common implementations are too resource-intensive and complex. ArtyZ7-20 board by Diligent. It is based around the Zynq-7020 System on chip (SoC) and has several functional peripherals, such as USB ports, an Ethernet port, buttons, switches, LEDs, and Arduino shield style connectors are shown in Figure 1.

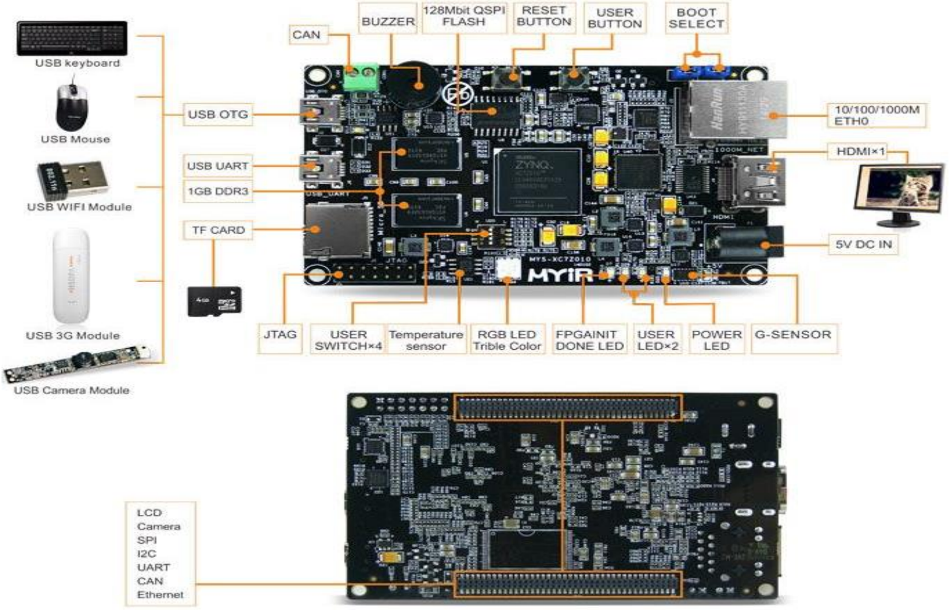


Figure 1. Zynq-7020 system on chip (SoC)

2.2. Graphical processing units (GPUs)

GPUs are electronic circuits originally designed for graphical and image processing. They are highly parallel, often containing several hundred cores that can perform many simpler computations at once. GPUs are often used today to relieve a CPU when required and they are found in all computers, mobile phones, and other graphical rendering devices. GPUs are also utilized in other applications that benefit from their parallel nature, including machine learning. NVIDIA has further developed the use of GPUs in other fields with their CUDA parallel platform for GPUs [19]. The Jetson Nano runs full Ubuntu and can consequently be used as a stand-alone computer. As such, it can take advantage of frameworks, tools, and libraries available for Ubuntu to increase efficiency during development. It same as a Raspberry Pi style single-board computer with a 128-core Maxwell GPU. It also features a 1.43 GHz ARM A57 CPU, and 4 GB of RAM and uses a MicroSD card for storage as shown in Figure 2.

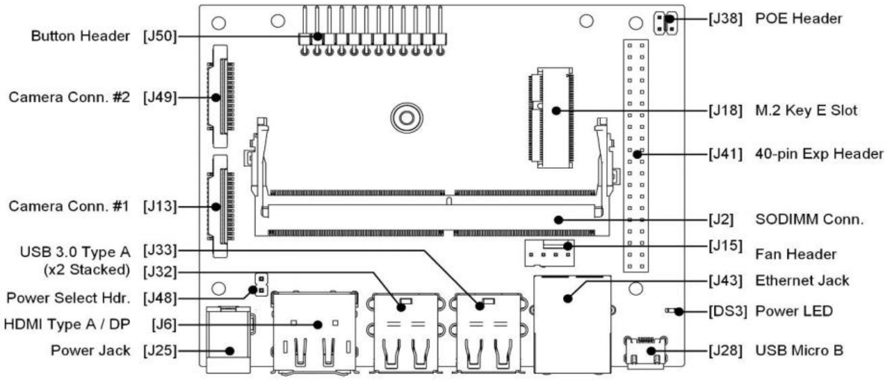


Figure 2. Jetson Nano board used for the GPU implementation

3. PROPOSED MODEL

Modern methodology for recognition of emotion acknowledgment utilizing equipment and constant implementation strategy. The proposed strategy comprises the accompanying advances: facial features extraction from Xception and vgg19. Perceiving facial emotion states with the assistance of the suggested model. With regards to confront discovery, the current examination utilizes the broadly used Viola-Jones system [20]. The whole procedure of the GPU and FPGA work process of FER is portrayed in Figure 3.

For the test examination, we utilized CK+dataset [21]. Outcomes manifest that the proposed models can proficiently play out every one of the assignments like recognition furthermore, arrangement with seven distinct feelings utilizing Xception calculation and VGG-19. We approved our models by making a real-time acknowledgment framework that achieves the errands of face recognition, and emotion state classification simultaneously, our proposed architecture's identification of faces is used for facial emotion recognition using the new proposed approach.

Insights from the Xception estimate have inspired the model we propose. This strategy integrates residual and depth-wise separable convolutions. With the goal of the learnt components becoming the qualification of the primary component map and the best features, the optimum preparation is modified by additional modules between two arising levels. Separating spatial cross-correlations from channel cross-correlations is the major goal of these layers. To do this, they apply a DxD filter to each of the M input channels before combining the resulting signals using N 1x1xM convolution filters to produce N output channels. When compared to standard convolutions, depth-wise separable convolutions speed up processing by a factor of 1 N + 1 D2. Each of the 4 residuals deep depth-wise convolutions in the proposed architectural model is followed by a gathering standardisation movement and a ReLUs enactment job. Finally, a soft-max activation function is used to average out the input classes that are expected, and a pooling layer is used to reduce the size of the network.

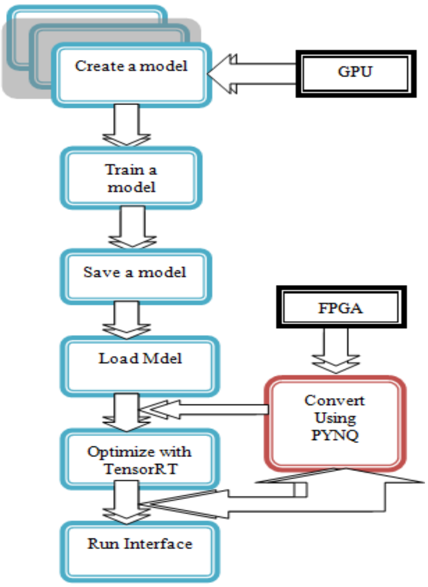


Figure 3. Intended workflow for the implementation of the FER in NVIDIA Jetson Nano and FPGA

VGG16 and VGG19 are visual geometry group models. Both employ face emotion detection features derived from pre-trained networks built on ImageNet. During the training stage, the information greyscale pictures are preprocessed by directing strength standardization and pixel scaling on the pixel values. In any case, VGG ought to be made with 224x224x3 sources of info, consequently pooling = 'avg' is indicated be-front the arrangement it deals ant input shape, and afterward, add own model head completely convolutional or thick. These pictures are given as a contribution to the VGG organization. TheVGG contains the most accurate five pooling layers. Adjusting empowers one to refresh the architectural design by disposing of the layer heads that were already completely associated, offering new, recently instated layers.

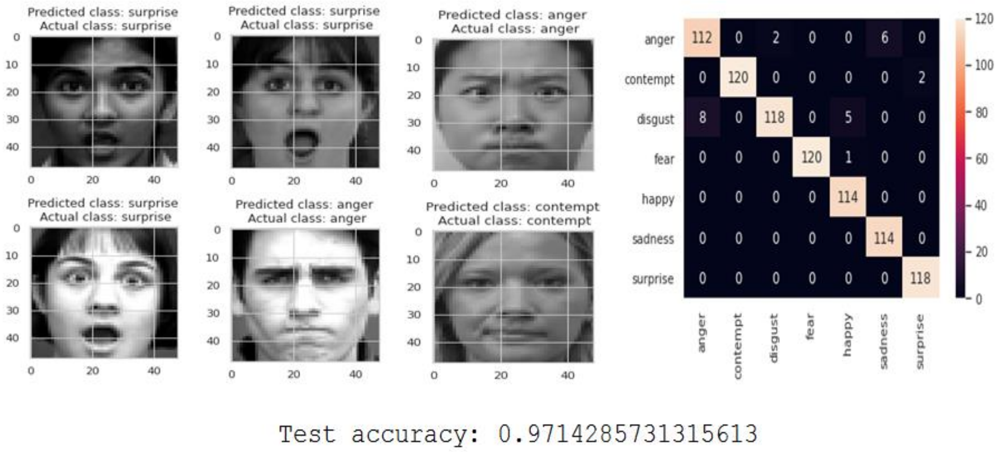
Algorithm Method: Emotion recognition using the Xception network and VGG-19 networks on NVIDIA Jetson Nano.

TensorRT was demonstrated using an example that comes with the TRT installation. These needed the usage of a converter to convert the frozen graph into UFF format for TensorRT. Although the convert-to-uff program is available with a Python 3x installation, it will not operate on the Jetson Nanos ARM core. A GitHub repository was utilized instead. This repository also includes Jupyter notebooks, however, because the Jetson Nano is a stand-alone computer, they have been converted to Python files that may be interacted with locally. Instructions for obtaining a pre-processed subset of the dataset for which the code was built were also given in the repository.

- Step 1: Load the ck+dataset
- Step 2: Split the dataset tanning and testing
- Step 3: Apply the preprocessing technique
- Step 4: Create deep learning models like Xception and VGG-19
- Step 5: Simulate the models using hardware NVIDIA Jetson Nano
- Step 6: Create the real-time recognition environment using OpenCV
- Step 7: Calculate the accuracy

4. EXPERIMENTAL RESULTS AND ANALYSIS

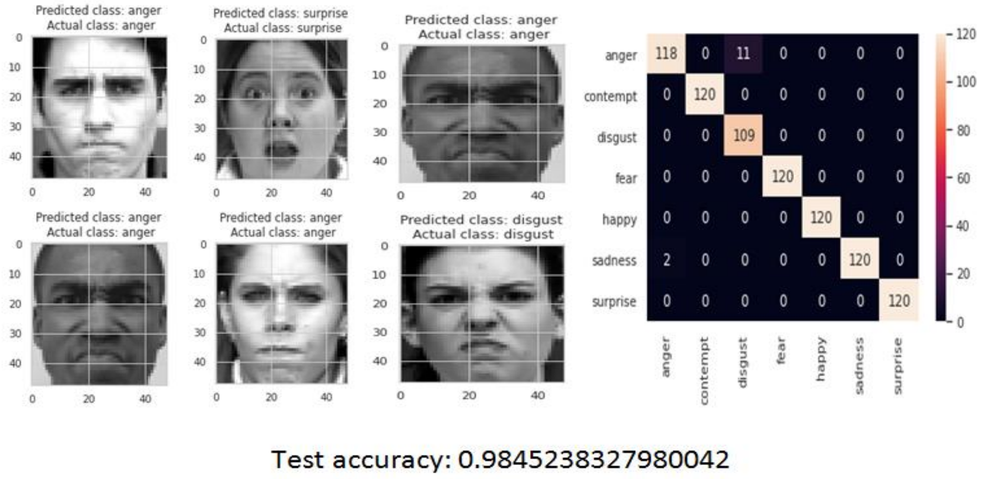
VGG19 is 19convolution layers, three completely associated layers, five MaxPooling layers, and a SoftMax activation layer. Activation layers work as an outlet layer with nodes of seven classes [22]. The Xception model is a 71-layered design, a lengthy rendition of the Inception model, yet with an excellent presentation ability. VGG-19 model is supplemented by a global average pooling layer, and a ReLU function [23]. Here the dataset was parted into two sub-exhibits by the Train-Test split methodology. The Training part included 80% of the first information, while the testing part represented 20% of the first dataset. After handling the subtleties of the multitude of models, an examination utilizing execution was directed to distinguish the best arrangement model. In this correlation investigation, the boundary measurements used for recognizing models were accuracy. Every one of the models' yields is portrayed through the probability of having a place with a particular class displayed in Figure 4 and Figure 5. comparison and contrast, it is empirical that the proposed classification techniques in Table 1.



	anger	contempt	disgust	fear	happy	sadness	surprise
anger	112	0	2	0	0	6	0
contempt	0	120	0	0	0	0	2
disgust	8	0	118	0	5	0	0
fear	0	0	0	120	1	0	0
happy	0	0	0	0	114	0	0
sadness	0	0	0	0	0	114	0
surprise	0	0	0	0	0	0	118

```
Test accuracy: 0.9714285731315613
```

Figure 4. Xception model test accuracy and their confusion matrix



	anger	contempt	disgust	fear	happy	sadness	surprise
anger	118	0	11	0	0	0	0
contempt	0	120	0	0	0	0	0
disgust	0	0	109	0	0	0	0
fear	0	0	0	120	0	0	0
happy	0	0	0	0	120	0	0
sadness	2	0	0	0	0	120	0
surprise	0	0	0	0	0	0	120

```
Test accuracy: 0.9845238327980042
```

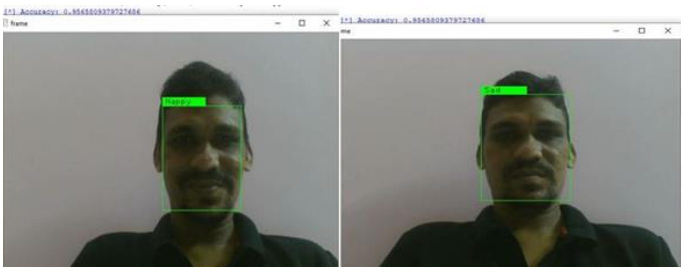
Figure 5. VGG-19 model test accuracy and their confusion matrix

Table 1. The proposed classification methods are empirically validated through comparison and contrast

Method	Dataset	Accuracy (%) in Xception architecture	Accuracy (%) in VGG-19 architecture	Accuracy (%) in OpenCV
Deep convolutional neural network [24]	CK+	92.81		
Deep neural networks [25]	CK+	91.9		
Convolutional neural networks [26]	CK+	60		
Support vector machine (SVM) [27]	CK+	95.71		
Proposed model	Ck+	97.1		
Deep CNN based features, (VGG16) [28]	CK+		86.04	
Google Net and Alex Net [29]	CK+		83.0	
Inception [30]	CK+		93.2	
Dynamic cascade classifier [31]	CK+		97.8	
Proposed model	Ck+		98.4	
The real-time recognition rate for Proposed models on OpenCV	Ck+			95.7

4.1. Results from quantitative studies on the suggested model

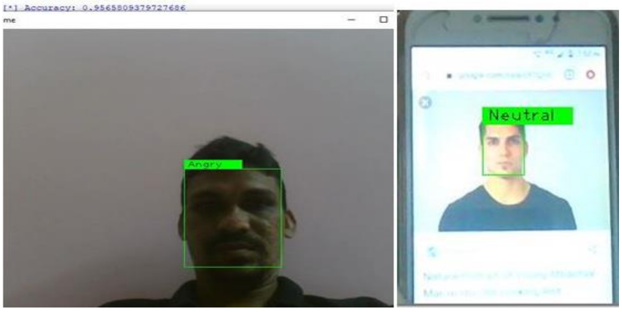
Testing facial appearance via webcam. A green rectangular area represents the face in each of the figures. Figure 6 demonstrates s good identification cases: happy and sad extremely precise facial gestures. Figure 7 demonstrates s good identification cases: angry and neutral. Figure 8 demonstrates good identification cases: fear and surprise.



[*] Accuracy: 0.9565809379727686

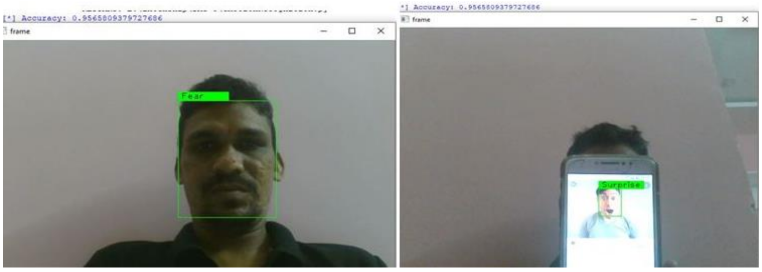
[*] Accuracy: 0.9565809379727686

Figure 6. Facial image samples for good identification cases like happy and sad



[*] Accuracy: 0.9565809379727686

Figure 7. Facial image samples for good identification cases like angry and neutral



[*] Accuracy: 0.9565809379727686

*] Accuracy: 0.9565809379727686

Figure 8. Facial image samples for good identification cases like fear and surprise

The novel aspect of the proposed method Standard implementations of depthwise separable convolutions (like TensorFlow) typically perform 1X1 convolution after channel-wise spatial convolution, but the modified depthwise separable convolution flips this order. This is disregarded as unnecessary because, when used in a stacked configuration, only minute variations appear at the beginning and end of all connected inception modules. After the initial operation, the Inception Module's original non-linearity becomes apparent. The intermediate ReLU non-linearity has been removed in the modernised depthwise separable convolution, Xception. To reiterate, this is how the proposed method provides a highly efficient means of raising precision levels.

5. CONCLUSION

Various calculations propose the most ideal model to enslave the issues of existing techniques. The exploration used the dataset and extricated highlights from 3 shading channels and the grayscale-separated pictures. The element extraction execution was assisted and contrasted with the advanced picture handling procedures like VGG-19, Xception, and real-time environment. The live, real-time testing of the system was a huge success. Because of the low-power requirements of the Jetson Nano, the decisions that were made were satisfying. There is a similar set of features on the board included with the NVIDIA Jetson Nano Developer Kit. Compared to the Raspberry Pi3, the Jetson Nano's processor is considerably more powerful, and the Nano's GPU is even more powerful than that of the Pi4. In terms of runtime and real-world performance, the Jetson Nano is far ahead of the competition, allowing for lower dormancy and higher throughput in serious learning applications.

REFERENCES

[1] K. Kaulard, D. W. Cunningham, H. H. Bülthoff, and C. Wallraven, "The MPI facial expression database — a validated database of emotional and conversational facial expressions," *PLoS One*, vol. 7, no. 3, p. e32321, Mar. 2012, doi: 10.1371/journal.pone.0032321.

[2] F. Dornaika and B. Raducanu, "Efficient facial expression recognition for human robot interaction," in *Computational and Ambient Intelligence*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 700–708. doi: 10.1007/978-3-540-73007-1_84.

[3] C. Bartneck and M. J. Lyons, "HCI and the face: Towards an art of the soluble," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2007, vol. 4550 LNCS, no. PART 1, pp. 20–29. doi: 10.1007/978-3-540-73105-4_3.

[4] S. Hickson, N. Dufour, A. Sud, V. Kwatra, and I. Essa, "Eyemotion: classifying facial expressions in VR using eye-tracking cameras," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Jan. 2019, pp. 1626–1635. doi: 10.1109/WACV.2019.00178.

[5] M. A. Assari and M. Rahmati, "Driver drowsiness detection using face expression recognition," in *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Nov. 2011, pp. 337–341. doi: 10.1109/ICSIPA.2011.6144162.

[6] C. Zhan, W. Li, P. Ogunbona, and F. Safaei, "A real-time facial expression recognition system for online games," *Int. J. Comput. Games Technol.*, vol. 2008, pp. 1–7, 2008, doi: 10.1155/2008/542918.

[7] A. Mourão and J. Magalhães, "Competitive affective gaming," in *Proceedings of the 21st ACM international conference on Multimedia*, Oct. 2013, pp. 83–92. doi: 10.1145/2502081.2502115.

[8] R. Samad and H. Sawada, "Extraction of the minimum number of Gabor wavelet parameters for the recognition of natural facial expressions," *Artif. Life Robot.*, vol. 16, no. 1, pp. 21–31, Jun. 2011, doi: 10.1007/s10015-011-0871-6.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, vol. 2, pp. 1097–1105.

[10] A. Soetedjo and I. K. Somawirata, "Implementation of face detection and tracking on a low cost embedded system using fusion technique," in *2016 11th International Conference on Computer Science & Education (ICCSE)*, Aug. 2016, pp. 209–213. doi: 10.1109/ICCSE.2016.7581582.

[11] N. R. Borkar and S. Kuwelkar, "Real-time implementation of face recognition system," in *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, Jul. 2017, pp. 249–255. doi: 10.1109/ICCMC.2017.8282685.

[12] N. Mustakim, N. Hossain, M. M. Rahman, N. Islam, Z. H. Sayem, and M. A. Z. Mamun, "Face Recognition system based on raspberry Pi platform," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, May 2019, pp. 1–4. doi: 10.1109/ICASERT.2019.8934485.

[13] H. Lee, S.-H. Park, J.-H. Yoo, S.-H. Jung, and J.-H. Huh, "Face recognition at a distance for a stand-alone access control system," *Sensors*, vol. 20, no. 3, p. 785, Jan. 2020, doi: 10.3390/s20030785.

[14] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, May 2009, doi: 10.1016/j.imavis.2008.08.005.

[15] R. Zhao *et al.*, "Accelerating binarized convolutional neural networks with software-programmable FPGAs," in *Proceedings of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, Feb. 2017, pp. 15–24. doi: 10.1145/3020078.3021741.

[16] L. Deng *et al.*, "Model compression and hardware acceleration for neural networks: A comprehensive survey," in *Proceedings of the IEEE,* 2020, 485-532, doi: 10.15439/2021F127.

[17] I. Kuon, R. Tessier, and J. Rose, "FPGA architecture: Survey and challenges," *Foundations and Trends® in Electronic Design Automation*, vol. 2, no. 2, pp: 135-253, 2008, doi: 10.1561/1000000005.

[18] C. W. Huang, C. J. Chang, M. Y. Lin, and H. Y. Tai, "Compact FPGA implementation of 32-bits AES algorithm using Block RAM," in *TENCON 2007-2007 IEEE Region 10 Conference,* 2007, pp:1-4, doi: 10.1109/TENCON.2007.4429126.

[19] X. Liu *et al*, "A hybrid GPU-FPGA based design methodology for enhancing machine learning applications performance," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 6, pp. 2309-2323, 2020, doi: https://doi.org/10.1007/s12652-019-01357-4.

[20] Y.-Q. Wang, "An analysis of the Viola-Jones face detection algorithm," *Image Process. Line*, vol. 4, pp. 128–148, Jun. 2014, doi: 10.5201/ipol.2014.104.

[21] 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 *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, Jun. 2010, pp. 94–101. doi: 10.1109/CVPRW.2010.5543262.

[22] W. W. Lo, X. Yang, and Y. Wang, "An Xception convolutional neural network for malware classification with transfer learning," in *2019 10th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, Jun. 2019, pp. 1–5. doi: 10.1109/NTMS.2019.8763852.

[23] S. H. Kassani, P. H. Kassani, R. Khazaeinezhad, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Diabetic retinopathy classification using a modified Xception architecture," in *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Dec. 2019, pp. 1–6. doi: 10.1109/ISSPIT47144.2019.9001846.

[24] J. D. Bodapati and N. Veeranjaneyulu, "Facial emotion recognition using deep Cnn based features," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 7, pp. 1928–1931, 2019.

[25] J. Li and E. Y. Lam, "Facial expression recognition using deep neural networks," in *2015 IEEE International Conference on Imaging Systems and Techniques (IST)*, Sep. 2015, pp. 1–6. doi: 10.1109/IST.2015.7294547.

[26] A. Lopes, E. Aguiar, A. F. D. Souza, and T. Santos, "Facial expression recognition with convolutional neural networks: coping with few data and the training sample order," *Pattern recognition*, vol. 61, pp. 610-628, 2017, doi: 10.1016/j.patcog.2016.07.026.

[27] V. Neagoe, A. Bărar, N. Sebe, and P. Robitu, "A deep learning approach for subject independent emotion recognition from facial expressions," in *Recent Advances in Image, Audio and Signal Processing*, 2013, vol. 978, no. Dl, p. 960.

[28] D. Y. Liliana, "Emotion recognition from facial expression using deep convolutional neural network," *J. Phys. Conf. Ser.*, vol. 1193, no. 6 Special Issue, p. 012004, Apr. 2019, doi: 10.1088/1742-6596/1193/1/012004.

[29] P. Giannopoulos, I. Perikos, and I. Hatzilygeroudis, "Deep Learning Approaches for Facial Emotion Recognition: A Case Study on FER-2013," 2018, pp. 1–16. doi: 10.1007/978-3-319-66790-4_1.

[30] M. Li, H. Xu, X. Huang, Z. Song, X. Liu, and X. Li, "Facial Expression Recognition with Identity and Emotion Joint Learning," *IEEE Trans. Affect. Comput.*, vol. 12, no. 2, pp. 544–550, Apr. 2021, doi: 10.1109/TAFFC.2018.2880201.

[31] A. M. Ashir, A. Eleyan, and B. Akdemir, "Facial expression recognition with dynamic cascaded classifier," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 6295–6309, May 2020, doi: 10.1007/s00521-019-04138-4.

BIOGRAPHIES OF AUTHORS

Usen Dudekula received his B.Tech. degree in Electronics and Communication Engineering (ECE) from Jawaharlal Nehru Technological University, Anantapuramu, and M. Tech degree in Embedded Systems from Jawaharlal Nehru Technological University, Anantapuramu. Currently, he is a Research Scholar at the School of Electronics Engineering, VIT-AP University, Andhra Pradesh-522237, India. His main research areas include image processing, pattern recognition, machine learning, and deep learning. He can be contacted at email: basha.834@gmail.com.

Purnachand N received his M.Tech. degree from VIT University, Vellore, India, and Ph.D. degrees from the University of Aveiro, Aveiro Portugal. He is currently working as Associate Professor at VIT-AP University, India. His areas of research include image processing, video processing, and pattern recognition. He can be contacted at email: chanduinece@gmail.com.