IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 2, June 2024, pp. 1619~1624
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i2.pp1619-1624

Hand gesture-based automatic door security system using squeeze and excitation residual networks

Surya Prihanto, Nazrul Effendy, Nopriadi
Intelligent and Embedded System Research Group, Department of Nuclear Engineering and Engineering Physics,
Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

Article Info

Article history:

Received Jul 2, 2023
Revised Feb 1, 2024
Accepted Feb 15, 2024

Keywords:

Convolutional neural network
Hand gesture recognition
Residual network
Security
Squeeze excitation network

ABSTRACT

Viruses can be transmitted in various ways; one spreads through airborne droplets or the touch of multiple objects. This can occur in any area, including the entrance to the house or access to a room or deposit box. The spread of viruses that cause diseases like COVID-19 has caused many human casualties, and there is still the possibility of similar conditions appearing in the future. Several things need to be done to reduce the chances of spreading disease due to viruses, including developing contactless security support methods. This paper proposes a security system using hand gesture recognition using squeeze and excitation residual networks (SE-ResNet). This research offers a hand gesture recognition system for an automatic door system using SE-ResNet and the residual network (ResNet).

Corresponding Author:

Nazrul Effendy
Intelligent and Embedded System Research Group
Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering
Universitas Gadjah Mada
St. Grafika 2, Yogyakarta, Indonesia
Email: nazrul@ugm.ac.id

1. INTRODUCTION

The security system for accessing a room or opening a safety box still needs to be developed along with the continued development of the expertise of criminals trying to penetrate the security system. In a condition where there is a pandemic of an infectious disease such as COVID-19, the security system for access to a public space also needs special attention. In connection with efforts to prevent the spread of contagious diseases such as COVID-19 during the pandemic and to avoid the spread of infection due to the virus, using security systems in direct contact should be avoided [1].

Some security systems use biometric methods such as voice, face, fingerprint, and finger vein [2], [3]. Several security methods require touching sensors such as fingerprints and finger veins, which may increase the chances of spreading diseases such as COVID-19 [4]. In special situations such as the COVID-19 pandemic, gesture-based systems can play a significant role because they are without touching objects [5].

Several researchers have developed artificial intelligence for classification or recognition systems [6], [7]. Pratama *et al.* [8] conducted hand code classification using the convolutional neural network (CNN) for human sign language. Then, Fong *et al.* [9] also developed security system research using hand sign recognition. The method includes analyzing intensity profiling, color histogram, and dimensionality analysis combined with neural networks as the classifier. In addition to sign language and hand pattern gesture research, CNN can predict even poor-quality fingerprint classes. This paper proposes a hand gesture-based

automatic door security system using squeeze and excitation residual networks (SE-ResNet) as the classifier in the system. The SE-ResNet performance was investigated and compared with CNN performance.

The squeeze and excitation network (SE-Net) algorithm won first place in the classification field at the imageNet large-scale visual recognition challenge (ILSVRC) [10]. While residual network (ResNet) won the most advanced performance award at ILSVRC for classification, localization, and common objects in context (COCO) detection and segmentation tasks. Combining the two algorithms is often called SE-ResNet [11]–[13]. In addition, SE-Net were used in the health sector to detect arrhythmias classes based on 12-lead electrocardiogram (ECG) signals, resulting in superior classification performance [14].

The rest of this paper is described as follows. First, the research method is described in detail in section 2. Then, section 3 describes the experimental results, including the model accuracy and confusion matrix of the system using SE-ResNet and CNN. Finally, in section 4, the conclusion of this study is offered.

2. METHOD

2.1. Hand gesture-based automatic door security system

Figure 1 shows the proposed hand gesture-based automatic door security system consisting of a digital camera, preprocessing, classifier, and a mechanical system. We design the system using Python 3.0, Arduino IDE, and Tkinter [15]–[17]. The classifier will generate door security code recognition results.

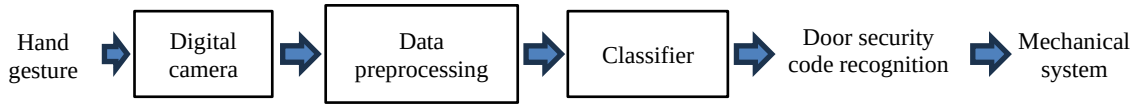

Figure 1. The proposed hand gesture-based automatic door security system

2.2. Dataset

This study collects hand gesture images from five subjects using a digital camera. The dataset comprises seven hand gesture types: type 0, type 1, type 2, type 3, type 4, type 5, and empty. The dataset is acquired using a digital camera and divided into 67% as training data and 33% as test data.

2.3. Data pre-processing

Before the training process of the machine learning model, it is necessary to convert the image from red, green and blue (RGB) to binary scale. The conversion is commonly called thresholding and is the most straightforward image segmentation method, which converts the image from grayscale to binary form. A binary image is an image that has only two colors, namely black and white. This conversion stage needs to be done because it provides several advantages in the data training process on the model. The binary thresholding technique converts the background image into white and the hand object into black, often called skin color detection [18], [19]. In addition to the data thresholding process, constant illuminance control is performed in the environmental system. Lighting monitoring is one of the essential things in some experiments, such as the need for illumination studies and the measurement of lux in a particular environment [20], [21].

2.4. Classifier

This paper, investigated CNN and SE-ResNet as classifier of hand gesture images. SE-ResNet implements a mechanism that connects the output layer with the input from the previous layer as the input of a residual block activation function [22]. The mechanism is called skip connection, which aims in the training process to pass through many layers without changing and keeping the norm of the error gradient stable.

Meanwhile, SE-Net aims to improve the network representation performance in the inter-channel's interdependence model on convolution features. The performance mechanism of SE-Net allows the network to perform feature recalibration so the network can adaptively modify the weight of each feature map. This is useful for selectively acquiring global information and emphasizing fewer valuable features. The hyperparameters of CNN and SE-ResNet are designed in the same way. Both use optimizers, regulation, and data augmentation, differing only in early stopping because early stopping will stop adjusting the training process of each scenario model [23]–[25]. Figure 2 shows a diagram of CNN, while Figure 3 shows that of SE-ResNet.

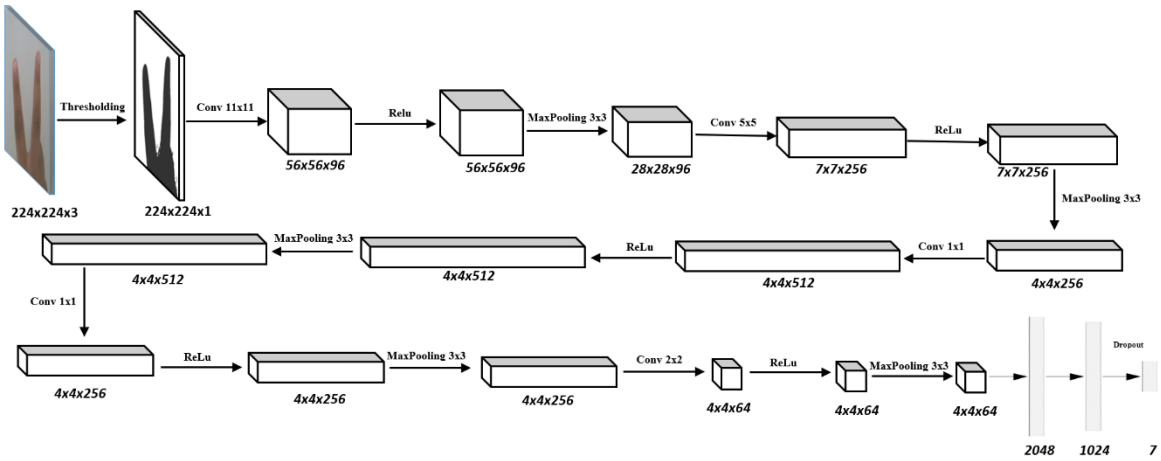

Figure 2. A diagram of CNN [26]

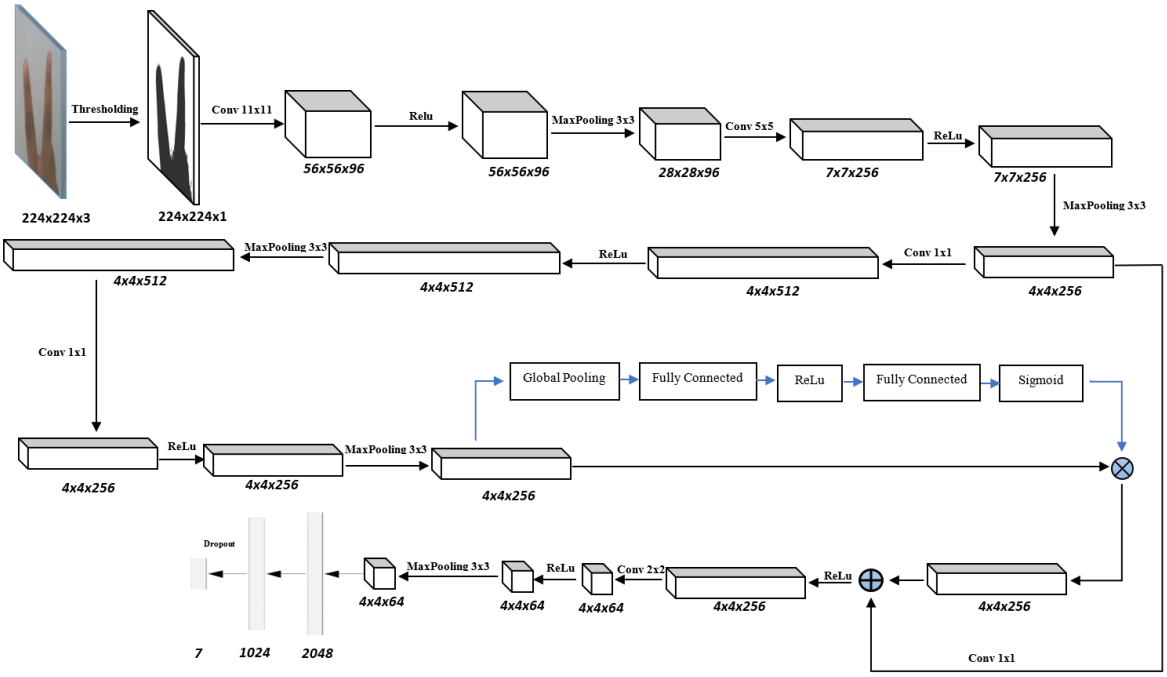

Figure 3. A diagram of SE-ResNet [26]

2.5. Mechanical system

After we get a good machine learning model, the next step is to integrate it with the mechanical system of opening or closing the door. The mechanical system utilizes an Arduino microcontroller programmed using the C++ programming language, while the language used for hand gesture classification uses the Python programming language [17], [27], [28]. This paper adds a connecting library between the Python and Arduino C++ programming languages. The library connecting the C++ and Python programming languages is PyFirmata. The Firmata library implements the Firmata protocol to communicate with software on the host computer [29]–[31]. PyFirmata can connect several programming languages with the Arduino IDE language.

3. RESULTS AND DISCUSSION

The results of this research consist of two main parts: model accuracy and confusion matrix. The accuracies of the training process of CNN and SE-ResNet models are shown in Figures 4 and 5. CNN's

training and validation accuracy is 92.69% and 91.84%, respectively. SE-ResNet achieved a training accuracy of 95.35% and a validation accuracy of 95.24%. With the training process running, the accuracy of the two models tends to increase. Accuracy SE-ResNet is more stable, reaching a maximum value than CNN accuracy.

Regarding data prediction, in predicting 42 new images, SE-ResNet is better than CNN. SE-ResNet correctly predicts 100% of the test images. While CNN correctly predicts 92.86% of the test images. The results indicate that SE-ResNet performs better than CNN. Figures 6 and 7 show the confusion matrix of CNN and SE-ResNet models [32]–[34]. SE-ResNet correctly recognized 100% of the hand gestures, while CNN correctly recognized 92.86%. This indicates that SE-ResNet is more effective in detecting hand gestures. Both models use the same optimizer, learning rate, augmentation, and dropout. The purpose of the dropouts to the SE-Net block is to overcome overfitting [35], [36]. Early stops are algorithms widely used to determine the better time to stop the training process before overfitting occurs. This technique monitors the validation error and stops the training when the error starts to increase [37].

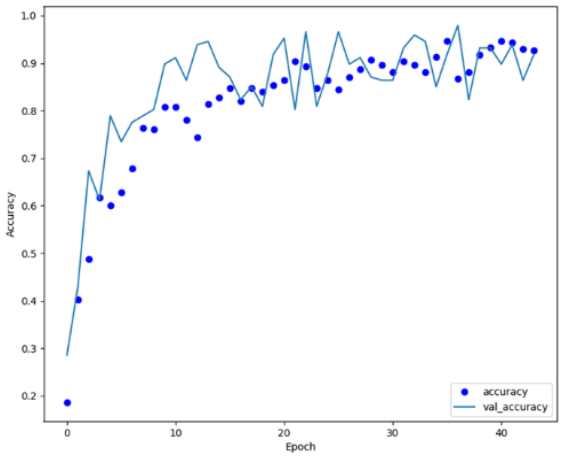

Figure 4. The accuracy of CNN [26]

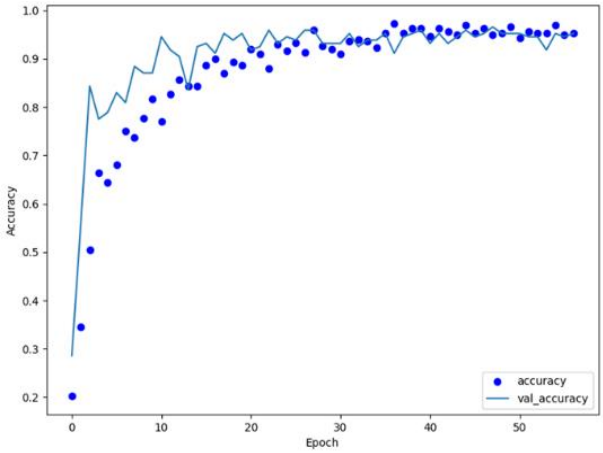

Figure 5. The accuracy of SE-ResNet [26]

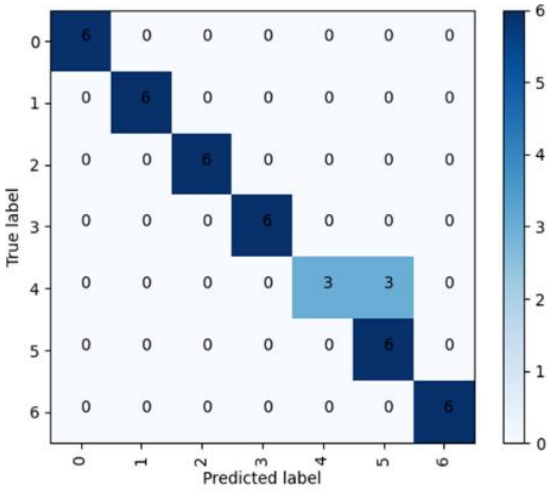

Figure 6. Confusion matrix of CNN [26]

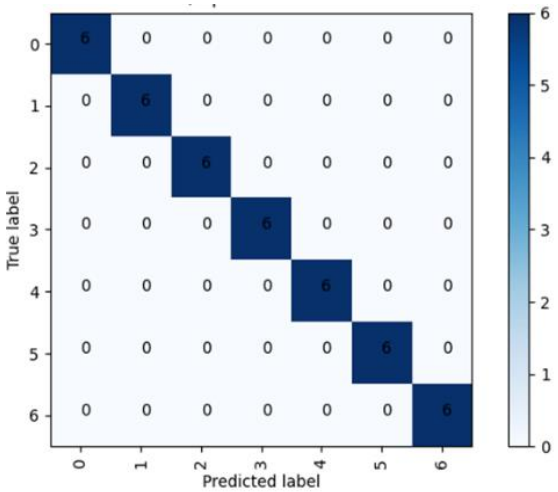

Figure 7. Confusion matrix of the SE-ResNet [26]

4. CONCLUSION

This paper proposed a hand gesture automatic home door security system based on SE-ResNet. Based on the experimental results, the model on the hand gesture dataset with binary-type segmentation reached a training accuracy of 95.85%, and the validation accuracy reached 95.31%. The SE-ResNet can classify 100% of the hand gesture images, while CNN correctly recognized 92.86%.

ACKNOWLEDGEMENTS

We thank Universitas Gadjah Mada for providing facilities and financial support for this research. This research is funded by the Final Project Recognition Grant Universitas Gadjah Mada Number 5075/UN1.P.II/Dit-Lit/PT.01.01/2023.

REFERENCES

[1] Q. Chen, "Can we migrate COVID-19 spreading risk?," *Frontiers of Environmental Science and Engineering*, vol. 15, no. 3, pp. 1–4, 2021, doi: 10.1007/s11783-020-1328-8.

[2] S. Nafisah and N. Effendy, "Voice biometric system: The identification of the severity of Cerebral Palsy using mel-frequencies stochastics approach," *International Journal of Integrated Engineering*, vol. 11, no. 3, pp. 194–206, 2019, doi: 10.30880/ijie.2019.11.03.020.

[3] T. H. Obaida, A. S. Jamil, and N. F. Hassan, "Real-time face detection in digital video-based on Viola-Jones supported by convolutional neural networks," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 3, pp. 3083–3091, 2022, doi: 10.11591/ijece.v12i3.pp3083-3091.

[4] V. A. Yulianto, N. Effendy, and A. Arif, "Finger vein identification system using capsule networks with hyperparameter tuning," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 4, pp. 1636–1643, 2023, doi: 10.11591/ijai.v12.i4.pp1636-1643.

[5] M. S. Alam, K. C. Kwon, and N. Kim, "Implementation of a Character Recognition System Based on Finger-Joint Tracking Using a Depth Camera," *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 3, pp. 229–241, 2021, doi: 10.1109/THMS.2021.3066854.

[6] S. N. Sembodo, N. Effendy, K. Dwiantoro, and N. Muddin, "Radial basis network estimator of oxygen content in the flue gas of debutanizer reboiler," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 3, pp. 3044–3050, 2022, doi: 10.11591/ijece.v12i3.pp3044-3050.

[7] N. Effendy, N. C. Wachidah, B. Achmad, P. Jiwandono, and M. Subekti, "Power estimation of G.A. Siwabessy Multi-Purpose Reactor at start-up condition using artificial neural network with input variation," in *2016 2nd International Conference on Science and Technology-Computer (ICST)*, IEEE, 2016, pp. 133–138. doi: 10.1109/ICSTC.2016.7877362.

[8] Y. Pratama, E. Marbun, Y. Parapat, and A. Manullang, "Deep convolutional neural network for hand sign language recognition using model e," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 5, pp. 1873–1881, 2020, doi: 10.11591/eei.v9i5.2027.

[9] S. Fong, Y. Zhuang, I. Fister, and I. Fister, "A biometric authentication model using hand gesture images," *BioMedical Engineering Online*, vol. 12, no. 1, pp. 1–18, 2013, doi: 10.1186/1475-925X-12-111.

[10] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 2011–2023, 2020, doi: 10.1109/TPAMI.2019.2913372.

[11] X. Zhang *et al.*, "SERNet: Squeeze and Excitation Residual Network for Semantic Segmentation of High-Resolution Remote Sensing Images," *Remote Sensing*, vol. 14, no. 19, pp. 1–19, 2022, doi: 10.3390/rs14194770.

[12] X. Li, F. Zhang, Z. Sun, D. Li, X. Kong, and Y. Zhang, "Automatic heartbeat classification using S-shaped reconstruction and a squeeze-and-excitation residual network," *Computers in Biology and Medicine*, vol. 140, 2022, doi: 10.1016/j.compbiomed.2021.105108.

[13] H. juan Zhu, W. Gu, L. min Wang, Z. cheng Xu, and V. S. Sheng, "Android malware detection based on multi-head squeeze-and-excitation residual network," *Expert Systems with Applications*, vol. 212, 2023, doi: 10.1016/j.eswa.2022.118705.

[14] R. Ge *et al.*, "Convolutional squeeze-and-excitation network for ECG arrhythmia detection," *Artificial Intelligence in Medicine*, vol. 121, 2021, doi: 10.1016/j.artmed.2021.102181.

[15] R. Yadav and H. Raheman, "Development of an artificial neural network model with graphical user interface for predicting contact area of bias-ply tractor tyres on firm surface," *Journal of Terramechanics*, vol. 107, pp. 1–11, 2023, doi: 10.1016/j.jterra.2023.01.004.

[16] H. Bhamre, S. Deshmukh, H. Shintre, P. Ghodke, and S. Shinde, "Design and development of bottle sorting machine using Arduino," *Materials Today: Proceedings*, vol. 77, pp. 1023–1027, 2023, doi: 10.1016/j.matpr.2023.02.252.

[17] E. J. Davis and B. H. Clowers, "Low-cost Arduino controlled dual-polarity high voltage power supply," *HardwareX*, vol. 13, pp. 1–14, 2023, doi: 10.1016/j.ohx.2022.e00382.

[18] Y. Luo, G. Cui, and D. Li, "An Improved Gesture Segmentation Method for Gesture Recognition Based on CNN and YCbCr," *Journal of Electrical and Computer Engineering*, vol. 2021, pp. 1–9, 2021, doi: 10.1155/2021/1783246.

[19] E. K. Brunsgaard, J. Jensen, and D. Grossman, "Melanoma in skin of color: Part II. Racial disparities, role of UV, and interventions for earlier detection," *Journal of the American Academy of Dermatology*, vol. 89, no. 3, pp. 459–468, 2023, doi: 10.1016/j.jaad.2022.04.057.

[20] X. Yang *et al.*, "Comprehensive LiDAR simulation with efficient physically-based DART-Lux model (I): Theory, novelty, and consistency validation," *Remote Sensing of Environment*, vol. 272, 2022, doi: 10.1016/j.rse.2022.112952.

[21] M. Rastegari, S. Pournaseri, and H. Sanaieian, "Analysis of daylight metrics based on the daylight autonomy (DLA) and lux illuminance in a real office building atrium in Tehran," *Energy*, vol. 263, 2023, doi: 10.1016/j.energy.2022.125707.

[22] A. Zaeemzadeh, N. Rahnavard, and M. Shah, "Norm-Preservation: Why Residual Networks Can Become Extremely Deep?," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 11, pp. 3980–3990, 2021, doi: 10.1109/TPAMI.2020.2990339.

[23] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019, doi: 10.1186/s40537-019-0197-0.

[24] H. Salem, A. E. Kabeel, E. M. S. El-Said, and O. M. Elzeki, "Predictive modelling for solar power-driven hybrid desalination system using artificial neural network regression with Adam optimization," *Desalination*, vol. 522, 2022, doi: 10.1016/j.desal.2021.115411.

[25] R. K. Yadav and Anubhav, "PSO-GA based hybrid with Adam Optimization for ANN training with application in Medical Diagnosis," *Cognitive Systems Research*, vol. 64, pp. 191–199, 2020, doi: 10.1016/j.cogsys.2020.08.011.

[26] S. Prihanto, "The hand movement-based automatic door security system using a squeeze and excitation residual network algorithm (In Indonesia: Sistem keamanan pintu otomatis berbasis gerakan tangan menggunakan algoritma squeeze and excitation residual network)," Universitas Gadjah Mada, 2023.

[27] J. Droujko, F. Kunz, and P. Molnar, "Ötz-T: 3D-printed open-source turbidity sensor with Arduino shield for suspended sediment monitoring," *HardwareX*, vol. 13, pp. 1–15, 2023, doi: 10.1016/j.ohx.2023.e00395.

[28] U. Schmitt, B. Moser, C. S. Lorenz, and A. Réfrégier, "sympy2c: From symbolic expressions to fast C/C++ functions and ODE solvers in Python," *Astronomy and Computing*, vol. 42, pp. 1–11, 2023, doi: 10.1016/j.ascom.2022.100666.
[29] P. P. Ray, D. Dash, and D. De, "Real-time event-driven sensor data analytics at the edge-Internet of Things for smart personal healthcare," *Journal of Supercomputing*, vol. 76, no. 9, pp. 6648–6668, 2020, doi: 10.1007/s11227-019-03089-w.
[30] W. Kim, H. Ko, H. Yun, J. Sung, S. Kim, and J. Nam, "A generic Internet of things (IoT) platform supporting plug-and-play device management based on the semantic web," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–11, Sep. 2019, doi: 10.1007/s12652-019-01464-2.
[31] M. S. Mohammed *et al.*, "Low-cost autonomous car level 2: Design and implementation for conventional vehicles," *Results in Engineering*, vol. 17, pp. 1–12, 2023, doi: 10.1016/j.rineng.2023.100969.
[32] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-Label Confusion Matrix," *IEEE Access*, vol. 10, pp. 19083–19095, 2022, doi: 10.1109/ACCESS.2022.3151048.
[33] Y. Wang, Y. Jia, Y. Tian, and J. Xiao, "Deep reinforcement learning with the confusion-matrix-based dynamic reward function for customer credit scoring," *Expert Systems with Applications*, vol. 200, 2022, doi: 10.1016/j.eswa.2022.117013.
[34] D. Valero-Carreras, J. Alcaraz, and M. Landete, "Comparing two SVM models through different metrics based on the confusion matrix," *Computers and Operations Research*, vol. 152, pp. 1–12, 2023, doi: 10.1016/j.cor.2022.106131.
[35] L. Cai, H. Li, W. Dong, and H. Fang, "Micro-expression recognition using 3D DenseNet fused Squeeze-and-Excitation Networks," *Applied Soft Computing*, vol. 119, 2022, doi: 10.1016/j.asoc.2022.108594.
[36] S. Tyagi and S. N. Talbar, "CSE-GAN: A 3D conditional generative adversarial network with concurrent squeeze-and-excitation blocks for lung nodule segmentation," *Computers in Biology and Medicine*, vol. 147, 2022, doi: 10.1016/j.compbiomed.2022.105781.
[37] M. Al-Qurishi, T. Khalid, and R. Souissi, "Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues," *IEEE Access*, vol. 9, pp. 126917–126951, 2021, doi: 10.1109/ACCESS.2021.3110912.

BIOGRAPHIES OF AUTHORS

Surya Prihanto received a B.Eng. degree in Physics from Universitas Negeri Yogyakarta in 2019 and a M.Eng. degree in Engineering Physics from Universitas Gadjah Mada in 2023. He is a research assistant at the Intelligent and Embedded System Research Group, Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada. His research interests are machine learning and its application in engineering. He can be contacted at email: surya.prihanto@mail.ugm.ac.id.

Nazrul Effendy received a B.Eng. degree in Instrumentation Technology of Nuclear Engineering and a M.Eng. degree in Electrical Engineering from Universitas Gadjah Mada in 1998 and 2001. He received a Ph.D. in Electrical Engineering from Chulalongkorn University, Thailand, in 2009. He was a research fellow at the Department of Control and Computer Engineering, Polytechnic University of Turin, Italy, in 2010 and 2011 and a visiting researcher in Shinoda Lab (Pattern Recognition & its Applications to Real World), Tokyo Institute of Technology, Japan, in 2009. He is an associate professor and the Intelligent and Embedded System Research Group coordinator at the Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada. He is a member of the Indonesian Association of Pattern Recognition, the Indonesian Society for Soft Computing, the Indonesian Artificial Intelligence Society, and the International Association for Pattern Recognition. He can be contacted at email: nazrul@ugm.ac.id.

Nopriadi received his Ph.D. from Tokyo Institute of Technology, Japan. He is an assistant professor at the Department of Nuclear Engineering and Engineering Physics, Universitas Gadjah Mada. His research interests include biomedical engineering, instrumentation, and sensor technology. He can be contacted at email: nopriadi@ugm.ac.id.