MITOR: Jurnal Teknik Elektro

Analysis of SVD-Based Image Compression for Efficient Deep Learning in Face Mask Classification

Eric Sean Kesuma*, Ali Zainal Abidin, Bhakti Yudho Suprapto, Suci Dwijayanti, Baginda Oloan Siregar
Department of Electrical Engineering – Sriwijaya University
Palembang, Indonesia
*ericseankesuma@unsri.ac.id

Abstract – *This study investigates the use of Singular Value Decomposition (SVD) as an image compression technique to improve the efficiency of deep learning models for face mask classification. The proposed approach applies SVD-based image compression using different values of k ($k = 10$, 30, and 50) prior to training a MobileNetV2 model. Experimental results demonstrate that SVD-based compression can significantly reduce computational costs while maintaining high classification performance. The model trained using $k = 50$ achieves the highest accuracy of 99.43%, slightly outperforming the model trained using the original dataset. Furthermore, compressed datasets require shorter training times, with the fastest configuration ($k = 30$) achieving a substantial reduction in training duration. The findings indicate that moderate compression levels provide an optimal balance between computational efficiency and classification accuracy, whereas excessive compression may lead to performance degradation due to the loss of important image features. In addition, the training process exhibits faster convergence when compressed datasets are utilized, indicating improved learning efficiency. Overall, the results confirm that SVD-based image compression is an effective preprocessing technique for enhancing deep learning efficiency without significantly compromising classification accuracy.*

Keywords – *Singular Value Decomposition; Image Compression; MobileNetV2; Face Mask Classification; Deep Learning Efficiency.*

I. Introduction

RECENT advances in deep learning have significantly improved the performance of image classification systems across various applications [1]. Convolutional Neural Networks (CNNs) and their modern variants have demonstrated remarkable capability in extracting complex features from visual data with high accuracy [2, 3]. However, the effectiveness of deep learning models is often dependent on large-scale datasets, which require substantial storage capacity and computational resources. This challenge becomes more critical in resource-constrained environments, where limitations in hardware and memory can hinder the training and deployment of deep learning models.

To address these challenges, efficient data representation and compression techniques have become increasingly important. One widely used method in numerical analysis and signal processing is Singular Value Decomposition (SVD), a matrix factorization technique that enables low-rank approximation and efficient data representation while preserving essential structural information [4]. By retaining only the most significant singular values, SVD enables data size reduction without significant loss of information.

Singular Value Decomposition (SVD) has been widely applied in image compression and signal processing because of its ability to generate low-rank approximations while preserving important structural information [4, 5]. However, limited research has investigated the impact of SVD-based dataset compression on deep learning performance, particularly in balancing data efficiency and classification accuracy. Most existing studies focus on conventional image compression techniques or model compression methods, such as pruning and quantization, rather than dataset-level compression using SVD [6, 7]. Recent studies have demonstrated that low-rank and matrix decomposition approaches can effectively reduce data redundancy and improve computational efficiency in deep learning systems [8–11]. Nevertheless, the relationship between SVD-based image compression and classification performance remains insufficiently explored.

The manuscript was received on June 11, 2026, revised on June 15, 2026, and published online on July 31, 2026. Emitor is a Journal of Electrical Engineering at Universitas Muhammadiyah Surakarta with ISSN (Print) 1411 – 8890 and ISSN (Online) 2541 – 4518, holding Sinta 3 accreditation. It is accessible at https://journals2.ums.ac.id/index.php/emitor/index.

In this study, SVD is applied as an image compression technique to reduce dataset size prior to model training. A MobileNetV2-based deep learning model is selected due to its lightweight architecture and computational efficiency, making it suitable for resource-constrained applications [12–14]. Face mask classification is used as a case study to evaluate the effectiveness of the proposed approach [15–20].

The main contribution of this research is an analysis of the trade-off between image compression levels and deep learning performance. Specifically, the study investigates how different levels of SVD-based compression affect classification accuracy while reducing dataset size. The findings provide valuable insights into the potential of SVD-based data reduction for improving computational efficiency without significantly degrading model performance. This approach is particularly relevant for the development of efficient deep learning systems in practical and resource-constrained environments.

The remainder of this paper is organized as follows. Section II describes the proposed methodology. Section III presents the results and discussion. Finally, Section IV concludes the paper and outlines directions for future work.

II. Research Methods

This study adopts an experimental approach to evaluate the effectiveness of Singular Value Decomposition (SVD)-based image compression in improving the efficiency of deep learning models for face mask classification. The proposed methodology consists of several stages, including dataset preparation, image preprocessing, SVD-based compression, model training, and performance evaluation. Different compression levels are investigated to analyze the trade-off between computational efficiency and classification accuracy. The overall framework is designed to compare the performance of MobileNetV2 trained on the original dataset and on datasets compressed using different values of the rank parameter k. Performance is assessed using classification metrics, including accuracy, precision, recall, and F1-score, as well as training time to evaluate computational efficiency.

i. System Overview

The overall workflow of the proposed system is illustrated in Figure 1. The process begins with dataset preparation and preprocessing, during which the input images are resized and normalized. Subsequently, the dataset is divided into two processing paths, namely the original dataset and the SVD-compressed dataset.

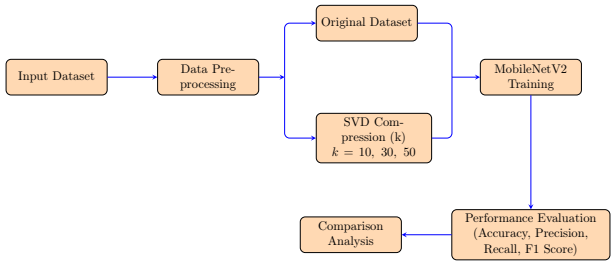

Figure 1: Overall workflow of the proposed SVD-based image compression and classification system.

For the compressed dataset, Singular Value Decomposition (SVD) is applied to reduce the dimensionality of the image data while preserving its essential features. Given a matrix A, the SVD can be expressed as Eq. (1):

$$A = U\Sigma V^T \tag{1}$$

where U and V are orthogonal matrices, and Σ is a diagonal matrix containing singular values. SVD is widely used in numerical analysis, signal processing, and image compression [5].

To achieve data compression, a low-rank approximation of the original matrix can be constructed as shown in Eq. (2):

$$A_k = U_k \Sigma_k V_k^T \tag{2}$$

where k represents the number of retained singular values, U_k and V_k contain the first k columns of U and V, respectively, while Σ_k contains the k largest singular values. A smaller value of k results in higher compression but may lead to greater information loss, whereas a larger value of k preserves more image details at the expense of reduced compression.

In image processing, each image can be represented as a matrix. By applying SVD using Eq. (1), the image matrix can be approximated using only the largest singular values while discarding smaller ones. This process reduces data dimensionality and enables efficient compression while preserving the most important image features.

To investigate the effect of compression on classification performance, SVD is applied using different values of the rank parameter ($k = 10$, 30, and 50), resulting in images with different compression levels. The original and compressed datasets are subsequently used for model training and evaluation.

For the classification task, MobileNetV2 is employed as the deep learning model. MobileNetV2 is a lightweight convolutional neural network designed for efficient computation in resource-constrained environments. It utilizes depthwise separable convolutions and inverted residual structures to reduce computational

complexity while maintaining competitive classification performance.

Due to its efficiency and low computational requirements, MobileNetV2 is well suited for applications requiring fast inference and reduced memory usage. In this study, it is used to evaluate the effectiveness of SVD-based image compression for image classification. Together, SVD-based compression and MobileNetV2 form the basis of the proposed approach, where image data are compressed prior to model training.

The trained models are evaluated using several performance metrics, including accuracy, precision, recall, and F1-score. Finally, the results obtained from the different compression configurations are compared to assess the impact of SVD-based image compression on deep learning classification performance.

TensorFlow and Keras are used as the primary frameworks for implementing and training the deep learning model. These frameworks provide efficient tools for building, training, and optimizing neural network architectures for image-based applications.

ii. Dataset Preparation

The dataset used in this study consists of facial images categorized into two classes: *with mask* and *without mask*. The dataset was obtained from the Face-Mask-Detection dataset published by Debargha Mitra Roy on Kaggle [21]. It contains 19,345 images in total, consisting of 9,608 images in the *with-mask* class and 9,737 images in the *without-mask* class.

All images are resized to 224×224 pixels to match the input size required by the MobileNetV2 architecture. Furthermore, image normalization is applied to scale pixel values into a suitable range for deep learning processing.

The dataset is divided into training and testing sets using an 80:20 ratio. Based on this split, approximately 15,476 images are used for training and 3,869 images are used for testing. The training set consists of approximately 7,686 *with-mask* images and 7,790 *without-mask* images, while the testing set contains approximately 1,922 *with-mask* images and 1,947 *without-mask* images. The dataset includes diverse variations in lighting conditions, facial orientations, and background settings to improve model robustness.

iii. SVD-Based Image Compression

In this study, Singular Value Decomposition is applied as an image compression technique to reduce dataset size prior to deep learning training. Each input image is represented as a matrix and decomposed using Eq. (1). To perform compression, only the top-k singular values are retained while the remaining values are discarded.

This process produces a low-rank approximation of the original image using Eq. (2), significantly reducing data size while preserving essential features. Three compression levels are evaluated to analyze the trade-off between data reduction and model performance:

1. $k = 10$ (high compression)
2. $k = 30$ (medium compression)
3. $k = 50$ (low compression)

Each image in the dataset is processed using this approach to generate multiple compressed datasets corresponding to different values of k. The compressed images are then reconstructed from the decomposed matrices and used as input for model training. This enables a comparative analysis between the original dataset and compressed datasets in terms of classification performance and computational efficiency.

iv. Deep Learning Model Architecture

In this study, a MobileNetV2-based deep learning model is employed for image classification. MobileNetV2 is a lightweight convolutional neural network designed for efficient computation [12, 13], making it suitable for applications involving resource-constrained environments.

The model architecture consists of several key components. The input layer accepts images with a size of $224 \times 224 \times 3$. The feature extraction process is performed using depthwise separable convolutions, which significantly reduce the number of parameters and computational cost compared to standard convolutional layers.

MobileNetV2 introduces inverted residual blocks with linear bottlenecks, which improve feature representation while maintaining efficiency. Each block consists of an expansion layer, depthwise convolution, and projection layer. This structure enables the model to learn meaningful features with reduced computational overhead. The architecture is illustrated in Figure 2.

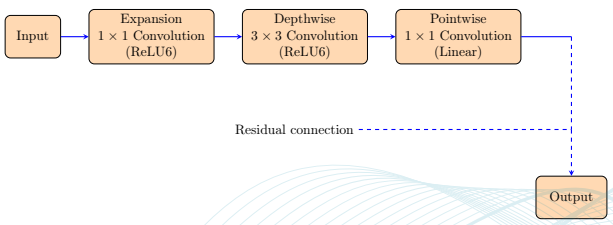

Figure 2: MobileNetV2 architecture showing inverted residual blocks and depthwise separable convolutions.

After feature extraction, a global average pooling layer is applied to reduce the spatial dimensions of

the feature maps. The output is then passed to a fully connected layer for classification. Finally, a softmax activation function is used to produce probability outputs for two classes: *with mask* and *without mask*.

MobileNetV2 is selected in this study due to its balance between classification accuracy and computational efficiency. This makes it an appropriate choice for evaluating the impact of SVD-based image compression on deep learning performance.

v. Training Configuration

The deep learning model is trained using the same configuration across all datasets to ensure a fair comparison between the original dataset and SVD-compressed datasets. The training process is performed using the Adam optimizer with a learning rate of 0.001.

The batch size is set to 32, allowing efficient memory utilization while maintaining stable gradient updates. The model is trained for a maximum of 50 epochs, while early stopping is employed to terminate training automatically when validation performance no longer improves.

All input images are resized to 224×224 pixels and normalized before being fed into the model. The training process is conducted separately for each dataset configuration, including the original dataset and compressed datasets with $k = 10$, $k = 30$, and $k = 50$.

vi. Experimental Setup

The experiment was conducted using Python with TensorFlow/Keras on an Apple MacBook Pro equipped with an Apple M3 chip (base model). The dataset consisted of 19,345 images and was divided into training and testing sets using an 80:20 ratio.

All configurations were kept identical for the original and SVD-compressed datasets to ensure a fair comparison. The model employed MobileNetV2 as the backbone architecture, with an input size of $224 \times 224 \times 3$, the Adam optimizer, a learning rate of 0.001, a batch size of 32, and a maximum of 50 epochs with early stopping.

For reproducibility, future implementations should additionally report the exact software versions, random seed initialization, validation strategy, and whether ImageNet pre-trained weights were utilized during training.

III. Results and Discussion

This section presents the experimental results obtained from training the MobileNetV2 model using both the original dataset and SVD-compressed datasets. The objective is to analyze the impact of SVD-based image compression on classification performance and computational efficiency.

i. Training Results

The model is trained with a maximum limit of 50 epochs across all configurations to ensure a consistent experimental setup. Early stopping is applied so that the training process can stop automatically when the validation performance becomes stable and no longer improves.

In practice, the training process stops before reaching the maximum epoch limit in several configurations. This indicates that the model converges earlier than the predefined maximum number of epochs. In particular, SVD-compressed datasets tend to converge faster compared to the original dataset, suggesting that compression reduces data redundancy and simplifies the learning process.

The model trained on the original dataset converges at 30 epochs, while SVD-based datasets converge significantly faster. Specifically, the model trained with $k = 50$ converges at 16 epochs, $k = 30$ at 12 epochs, and $k = 10$ at 19 epochs. These results indicate that SVD-based compression can accelerate the learning process by reducing data redundancy and simplifying feature representation.

ii. Performance Comparison

Table 1 presents the comparison of classification performance and training efficiency obtained from the original dataset and the SVD-compressed datasets using different values of the rank parameter k.

Table 1: Performance Comparison of Original and SVD-Compressed Datasets

Dataset	Accuracy (%)	Precision	Recall	F1-score	Training Time (s)
Direct	99.12	0.9912	0.9912	0.9912	3043
SVD ($k = 50$)	99.43	0.9943	0.9943	0.9943	1615
SVD ($k = 30$)	99.33	0.9933	0.9933	0.9933	1243
SVD ($k = 10$)	95.66	0.9566	0.9566	0.9566	1893

The results presented in Table 1 demonstrate that the proposed SVD-based image compression approach is capable of maintaining excellent classification performance while significantly reducing computational requirements. Across all configurations, the MobileNetV2 model achieves high classification accuracy, indicating that the essential discriminative features required for face mask classification are largely preserved even after image compression.

Among all evaluated configurations, the dataset compressed using SVD with $k = 50$ achieves the highest classification accuracy of 99.43%, slightly outperforming the model trained on the original dataset,

which achieves an accuracy of 99.12%. This finding suggests that moderate image compression may act as an implicit noise-reduction mechanism by removing redundant or less informative image components while preserving the dominant structural features required for classification. As a result, the compressed dataset may improve the model's ability to generalize and reduce the influence of irrelevant information during the training process.

Similarly, the configuration with $k = 30$ achieves an accuracy of 99.33%, which remains comparable to that of the original dataset. The small difference in performance between the original dataset and the moderately compressed datasets indicates that substantial data reduction can be achieved without significantly degrading classification capability. This observation highlights the effectiveness of SVD in preserving the most informative image characteristics through low-rank approximation.

In contrast, the configuration with $k = 10$ exhibits a noticeable decline in performance, achieving an accuracy of 95.66%. Although this result still indicates acceptable classification capability, the reduction in accuracy suggests that excessive compression removes important image information that contributes to feature extraction and decision making. Consequently, the model encounters greater difficulty in distinguishing between masked and unmasked faces when the compression level becomes too aggressive.

The precision, recall, and F1-score values within each configuration are nearly identical due to the balanced nature of the dataset and the symmetric distribution of classification errors. Since the numbers of false positives and false negatives are relatively similar, the evaluation metrics converge to comparable values. This consistency indicates that the model maintains balanced predictive performance across both classes and does not exhibit a significant bias toward either the masked or unmasked category.

From the computational perspective, the impact of SVD-based compression is even more evident. The original dataset requires 3043 seconds of training time, whereas all compressed datasets achieve substantially shorter training durations. The most efficient configuration is obtained with $k = 30$, requiring only 1243 seconds, corresponding to a reduction of approximately 59.15% in training time compared with the original dataset. Meanwhile, the configuration with $k = 50$ requires 1615 seconds, representing a reduction of approximately 46.93%, while maintaining the highest classification accuracy among all tested configurations.

These findings reveal an important trade-off between compression level, classification performance, and computational efficiency. Moderate compression levels ($k = 30$ and $k = 50$) provide a favorable balance by preserving classification accuracy while significantly reducing training time. Therefore, these configurations are particularly suitable for resource-constrained environments where computational resources, memory capacity, and energy consumption are critical considerations.

Overall, the results confirm that SVD-based image compression can effectively reduce data redundancy and computational complexity while maintaining competitive classification performance. The combination of high accuracy and reduced training time demonstrates the potential of SVD as an efficient preprocessing technique for deep learning applications involving large-scale image datasets.

iii. Confusion Matrix Analysis

Figure 3 presents the confusion matrix of the model using SVD compression with $k = 30$.

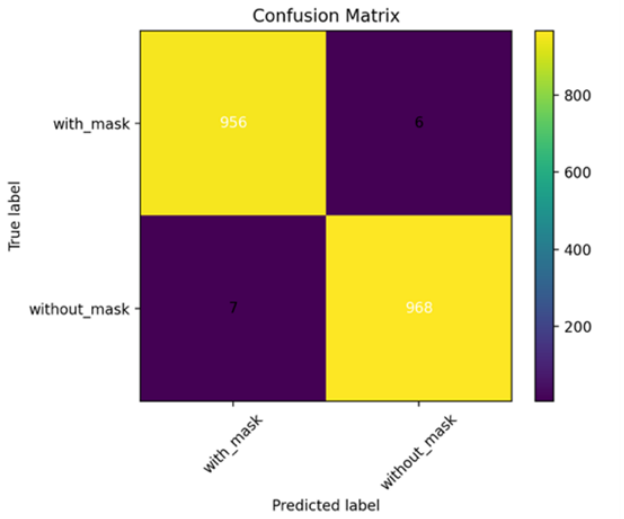

Figure 3: Confusion matrix of the model using SVD compression with $k = 30$.

The confusion matrix shown in Figure 3 provides a detailed visualization of the classification performance achieved by the proposed MobileNetV2 model when trained using the SVD-compressed dataset with $k = 30$. The matrix indicates that the majority of samples from both classes are correctly classified, while only a small number of samples are assigned to incorrect categories. This observation demonstrates the effectiveness of the proposed compression approach in preserving the essential image features required for accurate face mask classification.

A closer examination of Figure 3 reveals that the numbers of true positive and true negative predictions are substantially higher than the numbers of false pos-

itive and false negative predictions. This indicates that the model is capable of accurately distinguishing between images containing masks and those without masks. The relatively small number of classification errors further suggests that the compressed dataset retains sufficient discriminative information for reliable feature extraction and decision making.

The balanced distribution of correctly classified samples across both classes also confirms that the model does not exhibit a significant bias toward either category. This observation is consistent with the high precision, recall, and F1-score values reported in Table 1. Since both classes achieve similarly strong recognition performance, the resulting evaluation metrics remain highly consistent and indicate robust classification capability.

Furthermore, the confusion matrix provides additional evidence that moderate SVD compression does not significantly degrade classification performance. Although image compression reduces the dimensionality of the input data, the dominant structural and visual characteristics of facial images are still preserved. Consequently, the deep learning model can continue to identify relevant patterns associated with mask usage without experiencing a substantial loss of predictive accuracy.

Another important observation is that the misclassified samples are likely associated with challenging image conditions, such as variations in illumination, facial pose, partial occlusions, image blur, or masks that do not fully cover facial regions. Such conditions can make feature extraction more difficult and may lead to ambiguity during the classification process. Nevertheless, the relatively small number of errors indicates that the proposed model remains robust under a wide range of real-world scenarios.

iv. Classification Results

Examples of classification results obtained using the trained model are shown in Figure 4.

As illustrated in Figure 4, the system is able to correctly classify faces with and without masks under various conditions. The classification results demonstrate that the model can accurately identify the presence or absence of masks, even under variations in lighting conditions, facial orientations, image quality, and background complexity. These results indicate that the learned feature representations are sufficiently robust to distinguish masked and unmasked faces despite environmental variations that commonly occur in real-world scenarios.

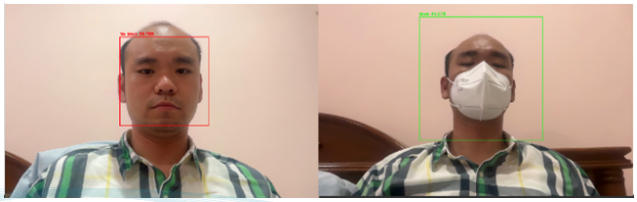

Figure 4: Classification results using the trained model.

The predicted labels are displayed together with confidence values, indicating the reliability of the model predictions. In most cases, the confidence scores are close to one, suggesting that the model has a high degree of certainty when assigning class labels. High confidence values are generally associated with well-defined facial features and clear mask visibility, whereas slightly lower confidence values may occur in images containing challenging conditions such as partial occlusion, non-frontal face poses, shadows, or variations in illumination.

Furthermore, the qualitative results presented in Figure 4 are consistent with the quantitative evaluation metrics reported in Table 1. The high accuracy, precision, recall, and F1-score values indicate that the model is capable of maintaining reliable classification performance across different dataset configurations. This consistency between qualitative observations and quantitative measurements further validates the effectiveness of the proposed approach.

Another important observation is that the use of SVD-based image compression does not visibly degrade the classification capability of the model at moderate compression levels. Even though the compressed images contain fewer data components, the essential facial characteristics required for classification are still preserved. Consequently, the model remains capable of accurately distinguishing between masked and unmasked faces while benefiting from reduced computational requirements.

The experimental results also demonstrate the potential applicability of the proposed system in practical deployment scenarios, such as public facility monitoring, access control systems, healthcare environments, transportation hubs, and other smart surveillance applications. The combination of SVD-based image compression and MobileNetV2 provides an effective balance between computational efficiency and classification accuracy, making the proposed approach suitable for implementation on devices with limited processing power and memory resources.

Overall, the classification results confirm that the proposed framework successfully maintains high recognition performance while reducing computational complexity. This finding highlights the feasibility of integrating image compression techniques into deep learn-

ing pipelines to achieve more efficient and scalable image classification systems.

v. *Efficiency Analysis*

In addition to classification performance, training time is analyzed to evaluate computational efficiency. The results summarized in Table 1 show that SVD-based compression significantly reduces training time compared to the original dataset.

The original dataset requires 3043 seconds for training, whereas compressed datasets substantially reduce training duration. The fastest training process is achieved at $k = 30$, requiring only 1243 seconds. This reduction demonstrates that SVD effectively decreases computational cost and accelerates model convergence.

vi. *Trade-Off Analysis*

The trade-off between compression level and classification performance is clearly observed from the results presented in Table 1. Moderate compression levels ($k = 30$ and $k = 50$) maintain high accuracy while improving computational efficiency.

Interestingly, the configuration with $k = 50$ achieves the highest accuracy, indicating that moderate compression may help remove redundant information and improve model generalization. However, excessive compression at $k = 10$ leads to a noticeable decrease in performance because important image features are lost during the compression process.

The experimental results demonstrate that SVD-based image compression is an effective preprocessing technique for improving the efficiency of deep learning models. The precision, recall, and F1-score values remain nearly identical across all configurations due to the balanced dataset and weighted averaging strategy.

Furthermore, the faster convergence observed in SVD-compressed datasets highlights the advantage of compression in simplifying the learning process. Overall, moderate compression levels provide an optimal balance between efficiency and classification performance, making SVD a promising approach for efficient deep learning applications.

IV. Conclusion

This study investigated the use of Singular Value Decomposition (SVD) as an image compression technique to improve the efficiency of deep learning models for face mask classification. The experimental results demonstrated that SVD-based compression can significantly reduce computational cost while maintaining high classification performance.

The model trained using SVD with $k = 50$ achieved the highest accuracy of 99.43%, slightly outperforming the model trained on the original dataset. In addition, SVD-compressed datasets required substantially less training time, with the fastest configuration ($k = 30$) achieving a significant reduction in computational time compared to the original dataset.

The results also indicated that moderate compression levels ($k = 30$ and $k = 50$) provide an optimal balance between efficiency and accuracy. Furthermore, the training process demonstrated faster convergence for compressed datasets, suggesting that SVD helps reduce data redundancy and simplify feature learning. However, excessive compression ($k = 10$) led to a noticeable decrease in performance, indicating the loss of important image features.

Overall, this study confirms that SVD-based image compression is an effective preprocessing technique for improving deep learning efficiency without significantly compromising classification accuracy. Future work may explore the integration of SVD with other deep learning architectures and its application to more complex image classification tasks.

Acknowledgment

The authors would like to express their gratitude to Sriwijaya University for providing academic support and research facilities. The authors also appreciate all individuals who contributed directly or indirectly to the completion of this study.

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. https://doi.org/10.1038/nature14539

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097–1105. https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

[3] M. Tan and Q. V. Le, "Efficientnetv2: Smaller models and faster training," in *Proceedings of the 38th International Conference on Machine Learning*, 2021, pp. 10 096–10 106. https://proceedings.mlr.press/v139/tan21a.html

[4] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: Johns Hopkins University Press, 2013. https://jhupbooks.press.jhu.edu/title/matrix-computations

[5] H. C. Andrews and C. L. Patterson, "Singular value decomposition (svd) image coding," *IEEE Transactions on Communications*, vol. 24, no. 4, pp. 425–432, 1976. https://doi.org/10.1109/TCOM.1976.1093309

[6] L. Deng, G. Li, S. Han, L. Shi, and Y. Xie, "Model compression and hardware acceleration for neural

networks: A comprehensive survey," *Proceedings of the IEEE*, vol. 108, no. 4, pp. 485–532, 2020. https://doi.org/10.1109/JPROC.2020.2976475

[7] S. Liang, S. Yin, L. Liu, W. Luk, and S. Wei, "Deep learning model compression: A survey," *Neurocomputing*, vol. 421, pp. 96–136, 2021. https://doi.org/10.1016/j.neucom.2020.09.106

[8] X. Yu, T. Liu, X. Wang, and D. Tao, "On compressing deep models by low-rank and sparse decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[9] Z. Lyu, T. Yu, F. Pan, Y. Zhang, J. Luo, D. Zhang, Y. Chen, B. Zhang, and G. Li, "A survey of model compression strategies for object detection," *Multimedia Tools and Applications*, vol. 83, no. 16, pp. 48 165–48 236, 2024. https://doi.org/10.1007/s11042-023-17192-x

[10] X. Ou, Z. Chen, C. Zhu, and Y. Liu, "Low rank optimization for efficient deep learning: Making a balance between compact architecture and fast training," *arXiv preprint arXiv:2303.13635*, 2023. https://arxiv.org/abs/2303.13635

[11] Y. Idelbayev and M. A. Carreira-Perpinan, "Low-rank compression of neural nets: Learning the rank of each layer," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020. https://openaccess.thecvf.com/content_CVPR_2020/html/Idelbayev_Low-Rank_Compression_of_Neural_Nets_Learning_the_Rank_of_Each_Layer_CVPR_2020_paper.html

[12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520. https://doi.org/10.1109/CVPR.2018.00474

[13] 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," *arXiv preprint arXiv:1704.04861*, 2017. https://arxiv.org/abs/1704.04861

[14] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, and D. J. Hemanth, "Ssdmnv2: A real time dnn-based face mask detection system using single shot multibox detector and mobilenetv2," *Sustainable Cities and Society*, vol. 66, p. 102692, 2021. https://doi.org/10.1016/j.scs.2020.102692

[15] S. Sethi, M. Kathuria, and T. Kaushik, "Face mask detection using deep learning: An approach to reduce risk of coronavirus spread," *Journal of Biomedical Informatics*, vol. 120, p. 103848, 2021. https://doi.org/10.1016/j.jbi.2021.103848

[16] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, "Maskedface-net: A dataset of correctly/incorrectly masked face images in the context of covid-19," *Smart Health*, vol. 19, p. 100144, 2021. https://doi.org/10.1016/j.smhl.2020.100144

[17] W. Boulila, A. Alzahem, A. Almoudi, M. Afifi, I. Alturki, and M. Driss, "A deep learning-based approach for real-time facemask detection," *arXiv preprint arXiv:2110.08732*, 2021. https://arxiv.org/abs/2110.08732

[18] S. Habib, M. Alsanea, M. Aloraini, H. S. Al-Rawashdeh, M. Islam, and S. Khan, "An efficient and effective deep learning-based model for real-time face mask detection," *Sensors*, vol. 22, no. 7, p. 2602, 2022. https://doi.org/10.3390/s22072602

[19] A. H. I. Al-Rammahi, "Face mask recognition system using mobilenetv2 with optimization function," *Applied Artificial Intelligence*, vol. 36, no. 1, p. 2145638, 2022. https://doi.org/10.1080/08839514.2022.2145638

[20] F. Fadly, T. B. Kurniawan, D. A. Dewi, M. Z. Zakaria, and P. A. A. Hisham, "Deep learning based face mask detection system using mobilenetv2 for enhanced health protocol compliance," *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 2067–2078, 2024. https://doi.org/10.47738/jads.v5i4.476

[21] D. M. Roy, "Face-mask-detection," Kaggle Dataset, 2026, accessed: 2026-06-03. https://www.kaggle.com/datasets/debarghamitraroy/face-mask-detection