Image classification-based transfer learning framework for image detection of IoT devices

Yuris Mulya Saputra, Hanung Addi Chandra Utomo, Ganjar Alfian
Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, Yogyakarta, Indonesia

Article Info

Article history:

Received Oct 23, 2023
Revised Dec 26, 2023
Accepted Jan 11, 2024

Keywords:

Artificial intelligence
Convolutional neural network
Image classification
Internet of things
Transfer learning

ABSTRACT

For artificial intelligence (AI) applications, centralized learning on a cloud server and local learning on an internet-of-things device may suffer from data privacy leakage due to data sharing and inaccurate prediction due to limited computing resources. Transfer learning has been proposed as one potential solution to the world's big data problems. Transfer learning eliminates the need for each internet-of-things device to share local data with the cloud server during the training process. Instead, it can go through the training process on its own, using a cloud server's pre-trained model with high accuracy. As a result, despite its limited computing resources, the internet ofthings device can still predict with high accuracy. This paper proposes a transfer learning model for improving image detection accuracy on IoT devices with restricted computation. To obtain accurate image classification, a deep learning approach based on convolutional neural networks is used. The proposed method with freeze and unfreeze approaches achieves a higher validation accuracy (up to 43.6%) and a lower validation loss (up to 6.5 times) than the non-transfer learning method, according to simulation results using three relevant internet-of-things datasets.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Ganjar Alfian
Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada
Bulaksumur, Sleman, Yogyakarta 55281, Indonesia
Email: ganjar.alfian@ugm.ac.id

1. INTRODUCTION

The growth of internet of things (IoT) based artificial intelligence applications following COVID-19 pandemic has influenced the global big data market size worldwide [1]. To provide meaningful information from the big data, centralized learning, in which all data from IoT devices is trained on a cloud server, can be utilized [2]. This is due to the unlimited computational capacity of the cloud server, which allows it to deal with massive amounts of data from numerous IoT devices. Nonetheless, when data from IoT devices is exchanged and processed on a cloud server, there is a substantial risk of information privacy breaches. To reduce the privacy risk, there is local learning, where each IoT device can train its own data locally [3]. However, this strategy may face inherent restrictions on the computation of the IoT devices, resulting in poor data training quality.

Transfer learning (TL) has emerged as a promising approach to overcome the limitations of the aforementioned methods [4]–[6]. Specifically, the cloud server has the capability to initially train a dataset containing a substantial amount of samples in order to generate a pre-trained model of superior quality. By leveraging transfer knowledge from the pre-trained model, each IoT device can independently perform the training process using its own local dataset. Consequently, the IoT device can maintain a high level of accuracy in its predictions while using less time and limited computing resources. Additionally, it can achieve this without the need of data sharing, thus preserving data privacy [7].

Multiple studies in [8]–[17], [18]–[26], [27]–[29] have examined the implementation of TL in diverse applications. The works in [8]–[11] proposes TL frameworks using pre-trained convolutional neural networks (CNN) models for medical images classifications including melanoma detection, anthracnose and red-rust leaf disease detection, diabetic retinopathy identification, and pneumonia classification. Meanwhile, some CNN-based TL approaches for brain and breast tumor detection using magnetic resonance images have been investigated in [12]–[14]. In [15]–[18], CNN-based TL methods are implemented for COVID-19 detection using chest X-ray images. The studies in [19]–[22] also apply CNN-based TL techniques to identify leukocytes as well as red blood cells for blood-related diseases and classify fundus for general retinal diseases diagnosis, respectively. The TL with pre-trained CNN models are also used for non-medical applications including Thai culture and Pitha traditional food images classification [23], [24], and other applications including land cover, fabric defect, birds' species, distracted driver classifications as well as age-invariant face recognition [25]–[29]. All the above literature studies show that the TL approach can achieve better accuracy performance than those of non-TL methods for various considered datasets and scenarios. Nonetheless, most of the above works are required to use high-resolution images with big file sizes to classify them correctly, which is difficult to be executed on IoT devices with limited computation.

The use of TL is then extended to IoT-based applications in the mobile edge networks for resource optimization, aiming to comply with the inherent challenge of IoT in terms of limited computing resource. For example, the work in [30] proposes a TL-based IoT framework for fire detection. Specifically, a pre-trained CNN model using MobileNet is executed for fine-tuning to improve the fire detection accuracy with fewer false positives. Meanwhile, a dynamic TL framework to improve the accuracy of IoT recognition ability is discussed in [31]. Through using the knowledge from additional sensing devices via the active learning process, the proposed framework can raise the recognition accuracy up to 13.05% compared to that of non-TL framework. Vu [32], a TL-based system IoT attack detection using deep learning and two auto-encoders is proposed. The experimental results show that the TL approach can remarkably improve the accuracy compared to the conventional deep learning method. Next, the work in [33] utilizes TL to enhance automated e-waste recycling classification accuracy in the smart cities. Meanwhile, a TL-based system for air pollution prediction in the city area is studied in [34]. Both works can improve the accuracy to 98% and 88%, respectively, compared to those of non-TL approaches. From the aforementioned studies, it is shown that TL approaches on IoT system help to improve the new model performance by utilizing the trained model from different datasets. However, the application of TL on IoT system still needs to be explored, especially for image classification with huge amounts of data.

In this paper, a CNN-based TL framework to relieve and speed up the image detection for IoT devices is presented. First, an image classification-based universal dataset with a large number of samples is trained on a cloud server to produce a high-quality pre-trained model. Then, transfer knowledge extracted from the pre-trained CNN model is sent to the participating IoT devices. Using this pre-trained CNN model, each IoT device can freeze or unfreeze the base convolutional model for a local training on the IoT device. For the freeze-model approach, the IoT device does not need to retrain the entire model. Instead, the IoT device can extract meaningful features from local data at the device through the representations from the pre-trained model. Meanwhile, for the unfreeze-model method, the IoT device can first unfreeze the base convolutional model and then fine-tune that model with the additional new model from the device. Via experimental results using three IoT-based image datasets, i.e., vehicle detection dataset, human activity recognition dataset, and aerial scene dataset, the proposed TL framework can achieve better validation accuracy (up to 43.6%) and less validation loss (by 6.5 times) compared with those of non-TL method. This is the first work that compares both freeze and unfreeze model approaches with three usable IoT datasets for the TL process. In the following sections, the details of TL training processes on the cloud server and IoT devices using freeze and unfreeze models are explored. Then, complete comparisons in terms of validation accuracy, validation loss, and learning performances are discussed.

2. METHOD

Commonly, Figure 1 shows the proposed TL model. Assume that there exist a cloud server and a set of IoT devices, i.e., $\mathcal{N} = \{1, \dots, \mathrm{n}, \dots, \mathrm{N}\}$. In particular, the cloud server is connected to N participating IoT devices through Wi-Fi or cellular networks, in a specified IoT network for specific time. The cloud server orchestrates unlimited computation resources to conduct the learning process with many samples in a dataset. Meanwhile, IoT devices have limited computing resources to execute the training processes for the TL. Here, there are two main steps for the proposed CNN-based TL framework in IoT services as follows: i) cloud server-based training process and ii) IoT-based training process.

2.1. Cloud server-based training process

This training process requires a large number of image samples in a dataset to provide a high-quality pre-trained model. Additionally, the training process is only conducted once at a cloud server. In particular, huge number of learning participants, e.g., smartphone and camera users, can first collect image data from smartphones and cameras. Then, the participants can share the collected image data to the cloud server for the centralized learning process. Prior to the training process, the cloud server can set the hyperparameter settings and control the number of training rounds. Through the training process, the cloud server can extract the important features of the accumulated image data as a pre-trained model.

This pre-trained model contains million number of parameters that can be extracted as a transfer knowledge. This may include many convolutional 1D layers, batch normalization layers, convolutional 2D layers, max pooling layers, rectified linear unit (ReLU) activation functions, and fully connected layers [35]. Then, the transfer knowledge for the IoT-based training processes is generated by freezing/unfreezing the base convolutional model of the pre-trained model via feature extraction/fine tuning, respectively.

2.2. IoT-based training process

After completing the cloud server-based training process, the training process using transfer knowledge at each participating IoT device in $\mathcal{N}$ can be seen in Figure 1. Specifically, the IoT device can first gather limited number of image samples via various smart cameras for a short time period. For example, cameras embedded at the traffic lights can capture different types of vehicles to monitor and analyze the road traffic conditions. Then, cameras at public places can take some pictures of people activities to provide crowd monitoring. These cameras can also be embedded on flying unmanned aerial vehicles for aerial scene detection. To this end, the IoT device can separate between training and validation image samples, and then perform image data augmentation to provide data variety and avoid overfitting problem through training image data transformation, e.g., horizontal/vertical flipping and rotation.

Next, the IoT device can build a local CNN model using feature extraction by freezing the base convolutional model such that only few parameters can be trained from the pre-trained model. Then, the IoT device can add a global average pooling 2D layer to downsample the feature map and a dropout layer with a specific rate for overfitting reduction. Finally, a fully connected layer using a softmax function is applied for the image classification with specific number of classes. By utilizing a certain learning rate and optimizer, the local model can be compiled and used to train the training image samples *X-train*, i.e., image features, and *y-train*, i.e., image labels. In this case, the training processes are repeated until the training time reaches the pre-defined TL time threshold.

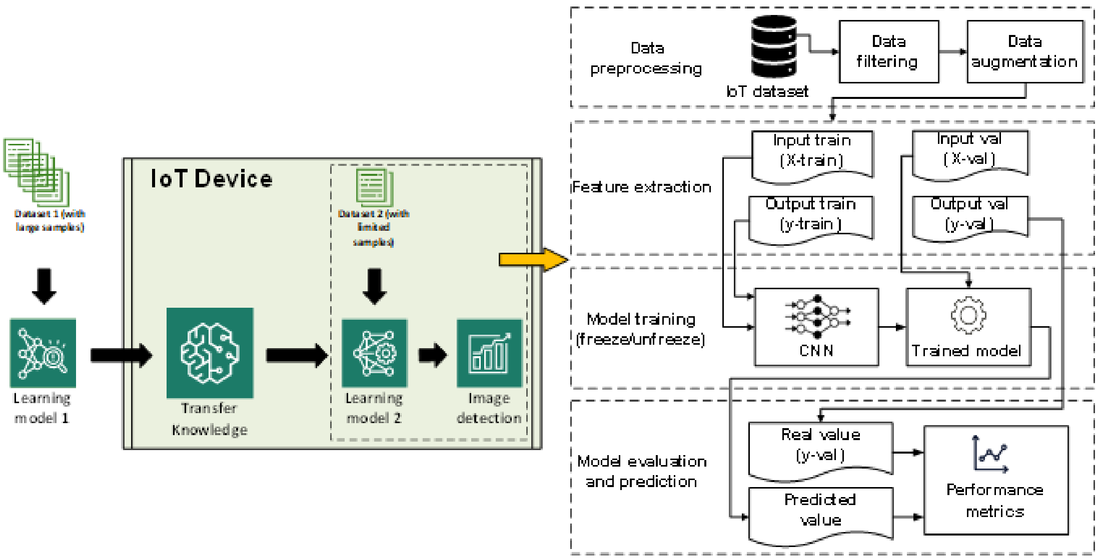


Figure 1. The CNN-based TL framework for IoT devices

To provide a higher prediction accuracy, the IoT device can perform a fine-tuning process. Particularly, all parameters from the base convolutional model are unfreezed such that the million number of parameters now can be trained. Then, the same processes apply until the local model converges, or the additional training time reaches the new TL time threshold. Upon completing CNN model training, the final local trained model is produced. This trained model is then used to predict the validation input image data *X-val* from the IoT dataset. To this end, through comparing the predicted value with the real value, i.e., validation output image data *y-val*, the performance evaluation including confusion matrix can be generated and analysed.

3. RESULTS AND DISCUSSION

To investigate the performance comparison between the proposed CNN-based TL methods and conventional CNN-based non-TL approach, a pre-trained model MobileNetV2 [35] that contain 3.5M parameters and is suitable for image classification problems is utilized. Furthermore, three relevant IoT image datasets with various number of samples are used. Particularly, vehicle detection dataset with 2K images and 6 labels [36], human activity recognition dataset with 12K samples and 15 labels [37], and aerial scene detection dataset with 24K samples and 6 labels [38] are applied.

To implement the training processes using the CNN, *TensorFlow with NVIDIA T4 Tensor Core GPU* [36] is used. For the proposed TL methods, the TL-CNN with freeze and unfreeze mechanisms are adopted. Meanwhile for the non-TL method, the conventional CNN with three sequential convolutional 2D and max pooling layers are used. To this end, the accuracy, loss, and learning performances for both proposed CNN-based TL methods and conventional CNN-based non-TL approach are investigated in the following.

3.1. Accuracy performance

In this section, the validation accuracy perfomance comparison is first investigated. From Table 1, the TL with freeze and unfreeze methods can achieve high validation accuracy by 93.4%. Particularly, TL-CNN unfreeze method can reach the accuracy of 43.6% and 43.5% higher than that of baseline CNN for vehicle detection dataset and human activity recognition dataset, respectively. This is due to additional transfer knowledge with all trainable parameters coming from the cloud server to help the training process at the IoT device. Although, the validation accuracy of TL-CNN freeze is slightly lower than that of TL-CNN unfreeze method for both datasets due to some untrainable parameters, the TL-CNN freeze method can achieve the accuracy of 0.3% higher than that of the TL-CNN unfreeze method for aerial scene dataset. Meanwhile, the baseline CNN suffers from bad validation accuracy for all utilized datasets as typical CNN layers without TL approach such as convolutional 2D layers, max pooling layers, and dense layers are used. This indicates that the transfer knowledge from freeze and unfreeze models can help IoT device to enhance its accuracy without using huge computation resource and leaking its data privacy.

Table 1. The validation accuracy performances for all datasets

Method	Vehicle	Human activity	Aerial scene
CNN	49.8%	30.1%	73.3%
TL-CNN Freeze	92.5%	70.7%	93.1%
TL-CNN Unfreeze	93.4%	73.6%	92.8%

3.2. Loss performance

Aligned with the validation accuracy performance, the same trend of loss validation performance is explained in Table 2. Specifically, a lower loss indicates that a better learning quality is performed. When vehicle detection and human activity recognition datasets are applied as shown in Table 2, the TL-CNN unfreeze method can achieve the validation loss 6.5 times and 2.5 times lower than that of baseline CNN without TL method. Meanwhile, the TL-CNN freeze method can improve the validation loss 3.6 times lower than that of the baseline CNN for aerial scene dataset. This implies that the use of TL for CNN can reduce the loss which aligns with the learning quality to produce better model accuracy. Additionally, the lower loss of CNN using the proposed TL approaches means that better prediction for image classification can be achieved.

3.3. Learning performance

To evaluate the accuracy and loss performances in more detail for both the training and validation processes, the learning performances for various scenarios with 20 training rounds are conducted. From Figure 2 when vehicle detection dataset is used, generally, the accuracy gets improved when the training round increases. To this end, it is observed that the TL-CNN unfreeze method can act as the upper bound performance of the training and validation processes. Meanwhile, the baseline CNN can perform as the lower bound performance of the training and validation processes. All methods can reach stable accuracy performance after 12 epochs have been executed. For the dynamic performance of loss, it is clearly observed that the loss performance gets reduced when the epoch increases. However, due to bad trained model of the baseline CNN, it suffers from unstable validation loss performance. This provides an insight that the use of TL with transfer knowledge can reduce the error loss more when pre-trained model is sent to the IoT device.

Table 2. The validation loss performances for all datasets

Method	Vehicle	Human activity	Aerial scene
CNN	1.4723	2.3447	0.7196
TL-CNN Freeze	0.237	0.9438	0.1973
TL-CNN Unfreeze	0.2248	0.9237	0.2588

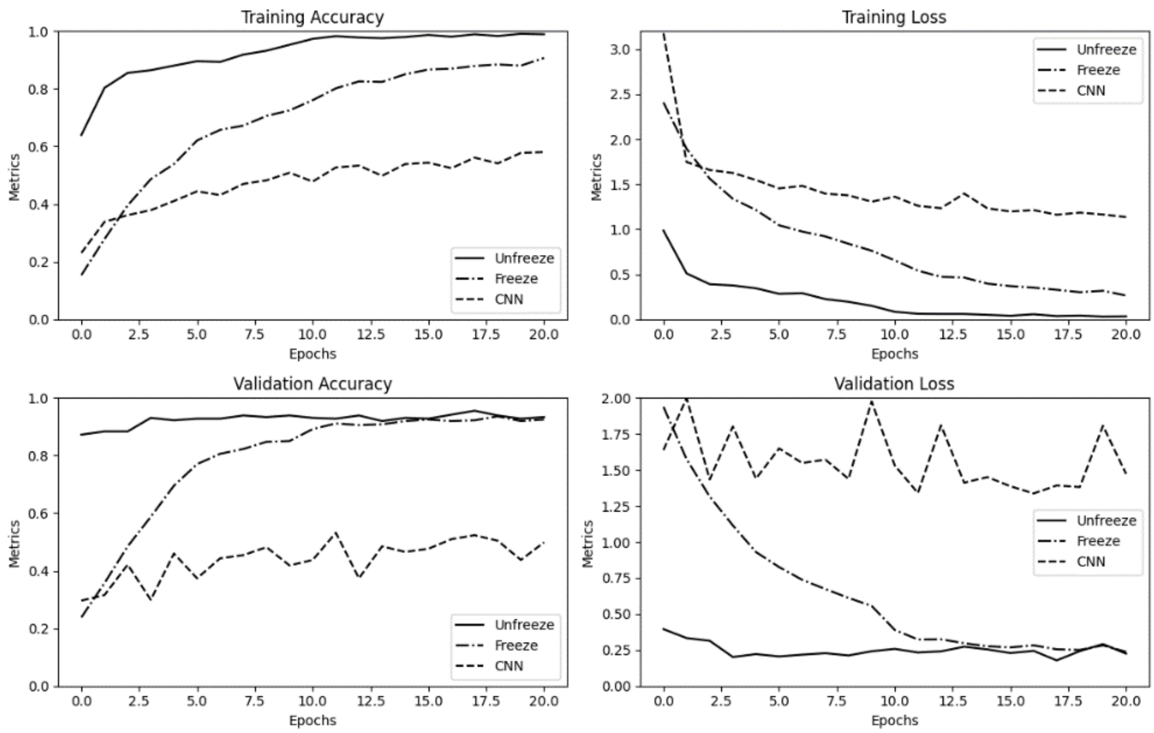


Figure 2. The accuracy and loss performances for different learning approaches when vehicle detection dataset is used

Similar to the vehicle detection dataset, the same trend of training and validation performances for human activity recognition dataset can be seen in Figure 3 in which the TL-CNN unfreeze and baseline CNN methods perform as the upper and lower bounds of the learning performance, respectively. Nevertheless, a slightly different trend can be observed in Figure 4 when the aerial scene dataset is used. Particularly, at the of the training rounds, the TL-CNN freeze method can improve both its validation accuracy and loss a little and achieve the highest validation accuracy and loss compared to other methods. This is because the base trained model that is freezed in the TL-CNN freeze method may provide bad trainable parameters when it is unfreezed in the TL-CNN unfreeze method. The results reveal that the performance of TL method is also influenced by the quality of trainable parameters. Here, the parameters may provide a good performance when they are untrained or vice versa depending on the trained dataset.

To show prediction efficacy of the proposed TL approach, the confusion matrices for all datasets and all methods are discussed in Figure 5. As shown in Figures 5(a) and 5(b), the confusion matrix of TL-CNN with unfreeze method for vehicle detection and human activity recognition dataset can provide the best performance because it can achieve the best validation accuracy 93.4% and 73.6%, respectively. The TL-CNN with freeze method can provide the best confusion matrix performance with validation accuracy 93.1% when aerial scene dataset is used in Figure 5(c). From those figures, the diagonal cells can achieve the highest values which infer that both methods can predict new IoT images well. All the confusion matrices in Figure 5 also show that the additional TL approach consistently provides the best prediction performance for image detection on IoT devices.

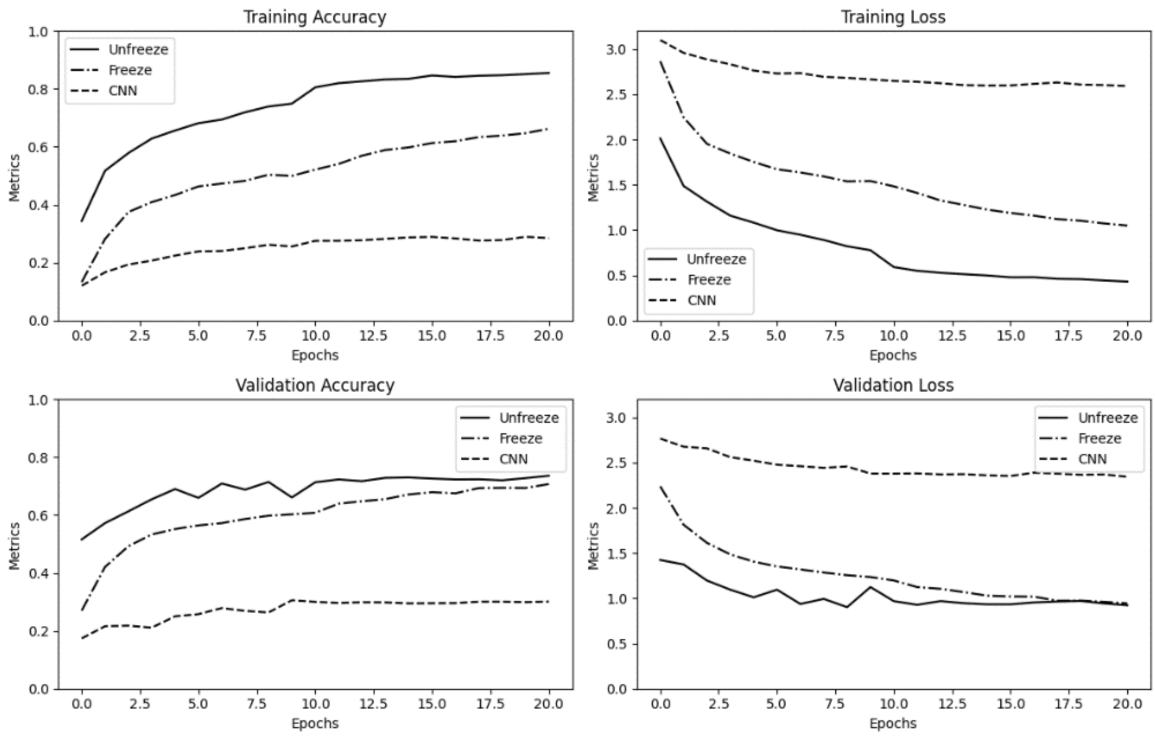


Figure 3. The accuracy and loss performances for different learning approaches when human activity recognition dataset is used

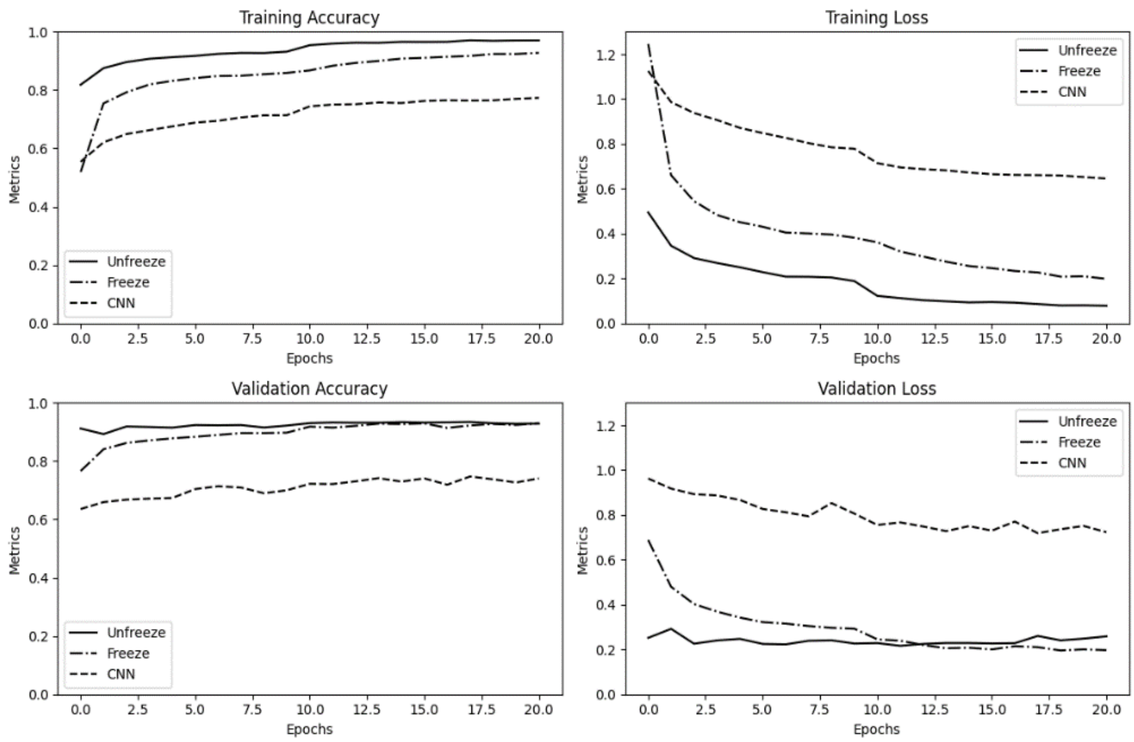
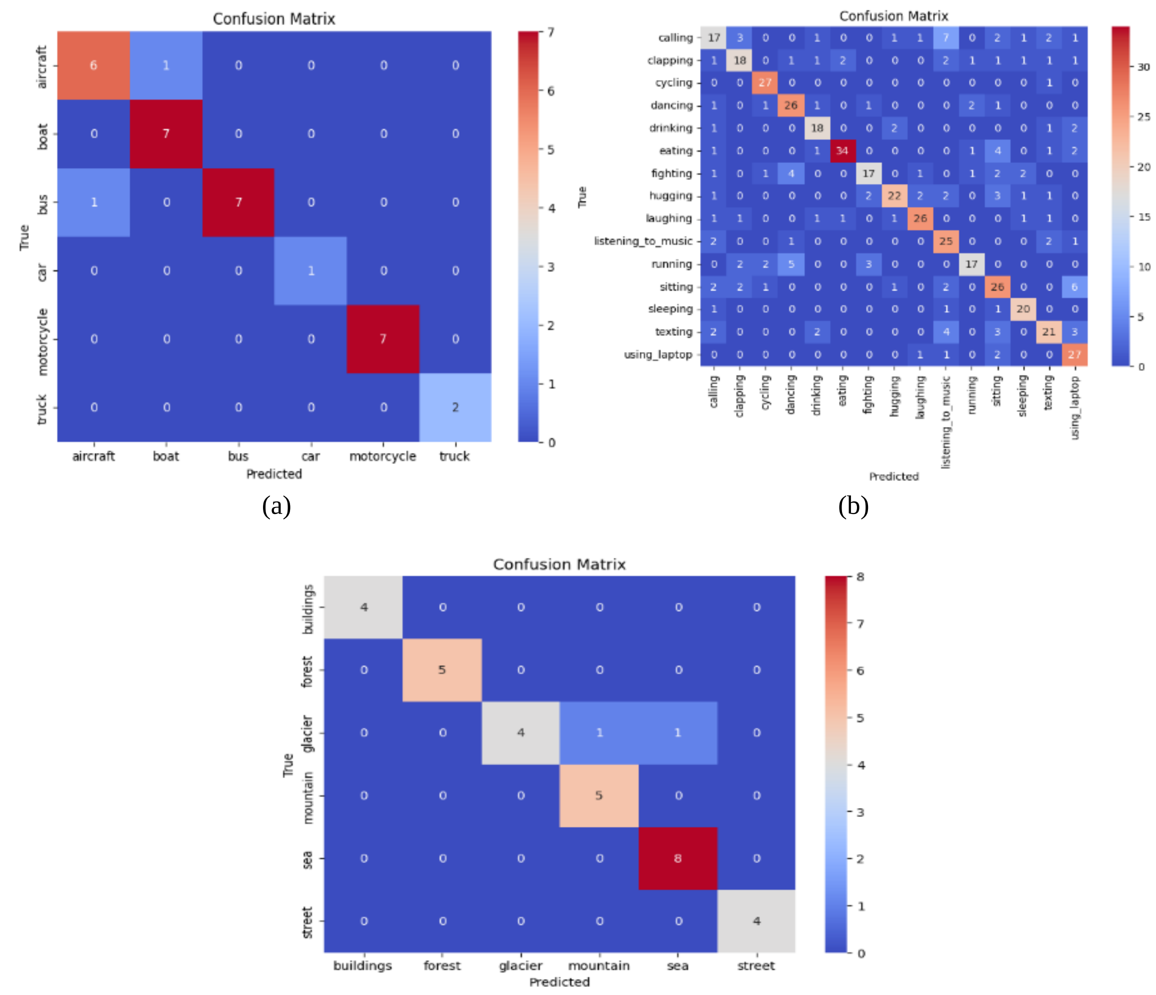


Figure 4. The accuracy and loss performances for different learning approaches when aerial scene dataset is used



(a) (b)

(c)

Figure 5. The best confusion matrix: (a) TL-CNN unfreeze for vehicle detection dataset, (b) TL-CNN unfreeze for human activity recognition dataset, and (c) TL-CNN freeze for aerial scene dataset

4. CONCLUSION

This paper presents a transfer learning framework based on image classification, which aims to accelerate and alleviate image detection for IoT devices. Specifically, the cloud server transmits knowledge extracted from a pre-trained CNN model to the participating IoT devices. Through utilizing the pre-trained CNN model, each IoT device has the capability to either freeze or unfreeze the base model for local training on the IoT device. The experimental results. which utilized three IoT-based image datasets, demonstrate that the suggested TL framework can attain a validation accuracy that is higher up to 43.6% and a validation loss that is 6.5 times lower compared to the non-TL method. This suggests that the proposed TL approach has the potential to serve as an alternative solution for IoT devices that have limited data and computing resources through generating a prediction model with high accuracy.

REFERENCES

[1] "Big data and business analytics market statistics 2030." https://www.alliedmarketresearch.com/big-data-and-business-analytics-market (accessed Aug. 07, 2023).

[2] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019, doi: 10.1109/COMST.2019.2904897.

[3] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019, doi: 10.1109/COMST.2019.2924243.

[4] F. Zhuang *et al.*, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.

[5] C. T. Nguyen *et al.*, "Transfer learning for wireless networks: A comprehensive survey," *Proceedings of the IEEE*, vol. 110, no. 8, pp. 1073–1115, Aug. 2022, doi: 10.1109/JPROC.2022.3175942.

[6] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *Artificial Neural Networks and Machine Learning – ICANN 2018*, 2018, pp. 270–279. doi: 10.1007/978-3-030-01424-7_27.

[7] H. Lin, J. Hu, X. Wang, M. F. Alhamid, and M. J. Piran, "Toward secure data fusion in industrial IoT using transfer learning," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 7114–7122, Oct. 2021, doi: 10.1109/TII.2020.3038780.

[8] S. S. Reddy, V. V. S. R. Raju, C. R. Swaroop, and N. Pilli, "Evaluation of deep learning models for melanoma image classification," *International Journal of Public Health Science (IJPHS)*, vol. 12, no. 3, pp. 1189–1199, Sep. 2023, doi: 10.11591/ijphs.v12i3.22983.

[9] R. Y. Patil, S. Gulvani, V. B. Waghmare, and I. K. Mujawar, "Image based anthracnose and red-rust leaf disease detection using deep learning," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 6, pp. 1256–1263, Dec. 2022, doi: 10.12928/telkomnika.v20i6.24262.

[10] Y. S. Devi and S. P. Kumar, "A deep transfer learning approach for identification of diabetic retinopathy using data augmentation," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 4, pp. 1287–1296, Dec. 2022, doi: 10.11591/ijai.v11.i4.pp1287-1296.

[11] T. H. Mandeel, S. M. Awad, and S. Naji, "Pneumonia binary classification using multi-scale feature classification network on chest x-ray images," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 4, pp. 1469–1477, Dec. 2022, doi: 10.11591/ijai.v11.i4.pp1469-1477.

[12] N. Q. Al-Ani and O. Al-Shamma, "A review on detecting brain tumors using deep learning and magnetic resonance images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 4, pp. 4582–4593, Aug. 2023, doi: 10.11591/ijece.v13i4.pp4582-4593.

[13] A. A. Pravitasari *et al.*, "UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 3, pp. 1310–1318, Jun. 2020, doi: 10.12928/telkomnika.v18i3.14753.

[14] A. M. Hassan, M. B. El-Mashade, and A. Aboshosha, "Deep learning for cancer tumor classification using transfer learning and feature concatenation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 6, pp. 6736–6743, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6736-6743.

[15] D. E. Cahyani, A. D. Hariadi, F. F. Setyawan, L. Gumilar, and S. Setumin, "COVID-19 classification using CNN-BiLSTM based on chest X-ray images," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 3, pp. 1773–1782, Jun. 2023, doi: 10.11591/eei.v12i3.4848.

[16] H. Imaduddin, F. Y. Ala, A. Fatmawati, and B. A. Hermansyah, "Comparison of transfer learning method for COVID-19 detection using convolution neural network," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 1091–1099, Apr. 2022, doi: 10.11591/eei.v11i2.3525.

[17] M. Al-Smadi, M. Hammad, Q. B. Baker, S. K. Tawalbeh, and S. A. Al-Zboon, "Transfer deep learning approach for detecting coronavirus disease in X-ray images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 6, pp. 4999–5008, Dec. 2021, doi: 10.11591/ijece.v11i6.pp4999-5008.

[18] N. Alrefai and O. Ibrahim, "Deep learning for COVID-19 diagnosis based on chest X-ray images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 5, pp. 4531–4541, Oct. 2021, doi: 10.11591/ijece.v11i5.pp4531-4541.

[19] M. S. Jarjees, S. S. M. Sheet, and B. T. Ahmed, "Leukocytes identification using augmentation and transfer learning based convolution neural network," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 2, pp. 314–320, Apr. 2022, doi: 10.12928/telkomnika.v20i2.23163.

[20] H. A. Abdulkarim, M. A. A. Razak, R. Sudirman, and N. Ramli, "A deep learning AlexNet model for classification of red blood cells in sickle cell anemia," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 2, pp. 221–228, Jun. 2020, doi: 10.11591/ijai.v9.i2.pp221-228.

[21] W. Setiawan, M. I. Utoyo, and R. Rulaningtyas, "Transfer learning with multiple pre-trained network for fundus classification," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 3, pp. 1382–1388, Jun. 2020, doi: 10.12928/telkomnika.v18i3.14868.

[22] B. Bhatkalkar, A. Joshi, S. Prabhu, and S. Bhandary, "Automated fundus image quality assessment and segmentation of optic disc using convolutional neural networks," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 816–827, Feb. 2020, doi: 10.11591/ijece.v10i1.pp816-827.

[23] M. Rattaphun and K. Songsri-in, "Thai culture image classification with transfer learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 6, pp. 6259–6267, Dec. 2023, doi: 10.11591/ijece.v13i6.pp6259-6267.

[24] S. Shakil, A. A. K. Akash, N. Nabi, M. Hassan, and A. Haque, "PithaNet: A transfer learning-based approach for traditional pitha classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 5, pp. 5431–5443, Oct. 2023, doi: 10.11591/ijece.v13i5.pp5431-5443.

[25] A. Beljadid, A. Tannouche, and A. Balouki, "Fabric defect classification using transfer learning and deep learning," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 3, pp. 1378–1385, Sep. 2023, doi: 10.11591/ijai.v12.i3.pp1378-1385.

[26] M. Alswaitti, L. Zihao, W. Alomoush, A. Alrosan, and K. Alissa, "Effective classification of birds' species based on transfer learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 4, pp. 4172–4184, Aug. 2022, doi: 10.11591/ijece.v12i4.pp4172-4184.

[27] I. Nasri, M. Karrouchi, H. Snoussi, K. Kassmi, and A. Messaoudi, "DistractNet: A deep convolutional neural network architecture for distracted driver classification," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 2, pp. 494–503, Jun. 2022, doi: 10.11591/ijai.v11.i2.pp494-503.

[28] A. A. Moustafa, A. Elnakib, and N. F. F. Areed, "Optimization of deep learning features for age-invariant face recognition," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 2, pp. 1833–1841, Apr. 2020, doi: 10.11591/ijece.v10i2.pp1833-1841.

[29] M. Hasanat, W. Khan, N. Minallah, N. Aziz, and A.-U.-R. Durrani, "Performance evaluation of transfer learning based deep convolutional neural network with limited fused spectro-temporal data for land cover classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 6, pp. 6882–6890, Dec. 2023, doi: 10.11591/ijece.v13i6.pp6882-6890.

[30] A. Gaur, A. Singh, A. Verma, and A. Kumar, "Artificial intelligence and multi-sensor fusion based universal fire detection system for smart buildings using IoT techniques," *IETE Journal of Research*, pp. 1–13, Jun. 2022, doi: 10.1080/03772063.2022.2088626.

[31] N. Wang, T. Li, Z. Zhang, and L. Cui, "TLTL: An active transfer learning method for Internet of Things applications," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6. doi: 10.1109/ICC.2019.8761694.

[32] L. Vu, Q. U. Nguyen, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, "Deep transfer learning for IoT attack detection," *IEEE Access*, vol. 8, pp. 107335–107344, 2020, doi: 10.1109/ACCESS.2020.3000476.

[33] N. Baker, P. Szabo-Müller, and U. Handmann, "Transfer learning-based method for automated e-waste recycling in smart cities," *EAI Endorsed Transactions on Smart Cities*, Jul. 2018, doi: 10.4108/eai.16-4-2021.169337.
[34] J. Kalajdjieski *et al.*, "Air pollution prediction with multi-modal data and deep neural networks," *Remote Sensing*, vol. 12, no. 24, Dec. 2020, doi: 10.3390/rs12244142.
[35] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.
[36] "NVIDIA T4." Accessed: Oct. 12, 2023. [Online]. Available: https://www.nvidia.com/en-us/data-center/tesla-t4/
[37] "Human action recognition (HAR) dataset." https://www.kaggle.com/datasets/meetnagadia/human-action-recognition-har-dataset (accessed Oct. 12, 2023).
[38] "Image scene classification of multiclass." https://www.kaggle.com/datasets/puneet6060/intel-image-classification (accessed Oct. 12, 2023).

BIOGRAPHIES OF AUTHORS

Yuris Mulya Saputra is currently an Assistant Professor as well as the Deputy Head at Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, Indonesia and an Adjunct Fellow at School of Electrical and Data Engineering, University of Technology Sydney. He obtained Ph.D. in Electrical and Data Engineering from the University of Technology Sydney (Australia) in 2022. His research interests include mobile computing, energy and economic efficiency, machine learning, IoT, and optimization problems for wireless communication networks. He is affiliated with IEEE as a Senior Member. He is currently an active Reviewer for various Q1 journals including IEEE TMC, IEEE WCM, IEEE IoT Journal, IEEE TNNLS, IEEE TBD, IEEE TCCN, and IEEE TCE. He can be contacted at email: ym.saputra@ugm.ac.id.

Hanung Addi Chandra Utomo is currently working as a Data Engineer in PT Global Data Inspirasi, Indonesia. He graduated from Universitas Gadjah Mada in July 2023 as an Applied Bachelor of Engineering majoring in Internet Engineering Technology, Department of Electrical Engineering and Informatics, Vocational College. His research interests include machine learning, deep learning, computer vision, natural language processing, data engineering, and backend development. He can be contacted at email: hanungutomo45@mail.ugm.ac.id.

Ganjar Alfian currently works as Assistant Professor at Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, Indonesia since 2022. Previously, he worked as Assistant Professor at Dongguk University Seoul, Republic of Korea for five years. In July 2017, he was a short-term Visiting Researcher with the VSB-Technical University of Ostrava, Czech Republic. He received Dr.Eng degrees from the Department of Industrial and Systems Engineering, Dongguk University, Seoul, South Korea, in 2016. His research interests include artificial intelligence, RFID, IoT, machine learning, carsharing, and simulation. He is listed in the World's Top 2% Scientist 2022 by Stanford University. He can be contacted at email: ganjar.alfian@ugm.ac.id.