Cyberspace: Jurnal Pendidikan Teknologi informasi
Volume 10, Nomor 2, Oktober 2026, hal. 69 – 80
ISSN 2598-2079 (print) | ISSN 2597-9671 (online)

PERFORMANCE EVALUATION OF MOBILENETV2 AND EFFICIENTNET-B0 FOR EUROSAT LAND COVER CLASSIFICATION

Andriani Putri[1], Sri Azizah Nazhifah*[2], Usfita Kiftiyani[3], Maulyanda[4]

[1,2,4]Department of Informatics, Universitas Syiah Kuala, Banda Aceh, Indonesia
[3]Department of Informatics, Universitas Islam Negeri Sunan Kalijaga, Yogyakarta, Indonesia
E-mail: [1]andrianiputri@usk.ac.id, [2]sriazizah07@unsyiah.ac.id, [3]usfita.kiftiyani@gmail.com, [4]maulyanda@usk.ac.id

Abstract

Classifying land cover through satellite imagery is an essential remote-sensing task; nevertheless, the performance of pretrained convolutional neural networks (CNNs) depends on the selected architecture and the extent to which the pretrained features are fine-tuned for the specific domain. This study evaluates MobileNetV2 and EfficientNet-B0 for land cover classification using the RGB version of the EuroSAT dataset. Both architectures made use of a common experimental setup that included resizing images to 224 × 224 pixels, transforming them into tensors, applying a 70:15:15 division for training, validation, and testing, utilizing pretrained weights, and evaluating performance via accuracy, precision, recall, F1-score, and confusion matrices. Three configurations were examined for each architecture: feature extraction, Fine-Tuning 10, and Fine-Tuning 20. The results suggest that feature extraction creates a strong baseline, whereas fine-tuning greatly improves classification accuracy for both models. EfficientNet-B0 reached 93.73% accuracy through feature extraction, rising to 98.35% and 98.40% in both fine-tuning trials. MobileNetV2 enhanced from 92.91% with feature extraction to 97.80% and 97.98% post fine-tuning. The results at the class level indicate that fine-tuning is especially advantageous for visually challenging categories like River and Highway. In general, EfficientNet-B0 delivers the best classification results, whereas MobileNetV2 continues to be a strong lightweight option. The findings likewise suggest that further parameter adjustment yields decreasing benefits after the pretrained representation has become adequately tailored to the target dataset.

Keywords: *EuroSAT, land cover classification, transfer learning, MobileNetV2, EfficientNet-B0, fine-tuning*

1. Introduction

Environmental monitoring, urban planning, agricultural evaluation, disaster management, and natural resource management are all supported by the classification of land use and land cover (LULC) using satellite imagery. The demand for automated techniques that can reliably and accurately identify spatial patterns has resulted from the growing amount of Earth-observation pictures. Because convolutional neural networks (CNNs) can learn hierarchical visual representations directly from picture input, decreasing reliance on manually created features, deep learning has emerged as a key strategy in this setting [1], [2].

Classifiers like Support Vector Machines, Random Forests, and k-Nearest Neighbors

are frequently combined with manually created spectral, textural, or spatial data in traditional remote sensing classification techniques. These pipelines may not accurately capture the intricate spatial relationships found in satellite imagery, despite being successful in many applications. They also heavily rely on the quality of the manufactured features. By learning feature representations in tandem with the classification goal, CNNs overcome this constraint [1], [4].

However, it takes a lot of labeled data and processing power to train a deep CNN from random initialization. By repurposing representations acquired from big datasets like ImageNet and tailoring them to a specific remote-sensing issue, transfer learning offers a workable substitute. Feature extraction, which keeps the pretrained feature extractor frozen, and fine-tuning, which lets the pretrained parameters change based on the target dataset, are two popular approaches. The distinction is crucial because fine-tuning allows domain adaptation, while a frozen network evaluates the transferability of generic properties [1], [6].

Because it offers annotated Sentinel-2 picture patches across ten land-cover categories, EuroSAT is a commonly utilized benchmark for this purpose. It is appropriate for controlled comparisons across deep-learning architectures and transfer-learning configurations due to its standardized data and class definitions [5]. However, many researchers report a single transfer-learning configuration or focus on architecture-level comparisons, paying less attention to the behavior of the same architecture when the pretrained representation is frozen as opposed to adapted. The impact of architectural selection and feature adaptation can thus be better understood by conducting a controlled comparison between MobileNetV2 and EfficientNet-B0 using the identical preprocessing, data split, and evaluation protocol [1], [4], [7].

This study evaluates MobileNetV2 and EfficientNet-B0 for EuroSAT land-cover classification, with particular emphasis on the effects of CNN architecture and transfer-learning strategy. The experiments compare feature extraction and fine-tuning configurations using the same preprocessing, data split, and evaluation protocol.

2. Research Method

The experimental workflow consists of dataset preparation, image preprocessing, model configuration, transfer-learning, training, and performance evaluation as shown in Fig. 1. Both CNN architectures were trained and evaluated under the same general protocol so that performance differences could be interpreted primarily in terms of architecture and parameter adaptation.

2.1 EuroSAT RGB Dataset

The RGB version of EuroSAT was used as the target dataset. The dataset contains 27,000 Sentinel-2 image patches distributed among ten land-cover classes: AnnualCrop, Forest, HerbaceousVegetation, Highway, Industrial, Pasture, PermanentCrop, Residential, River, and SeaLake [8]. The dataset was loaded using PyTorch torchvision ImageFolder, with directory names used as class labels. Table 1 shows the distribution of EuroSAT RGB images.

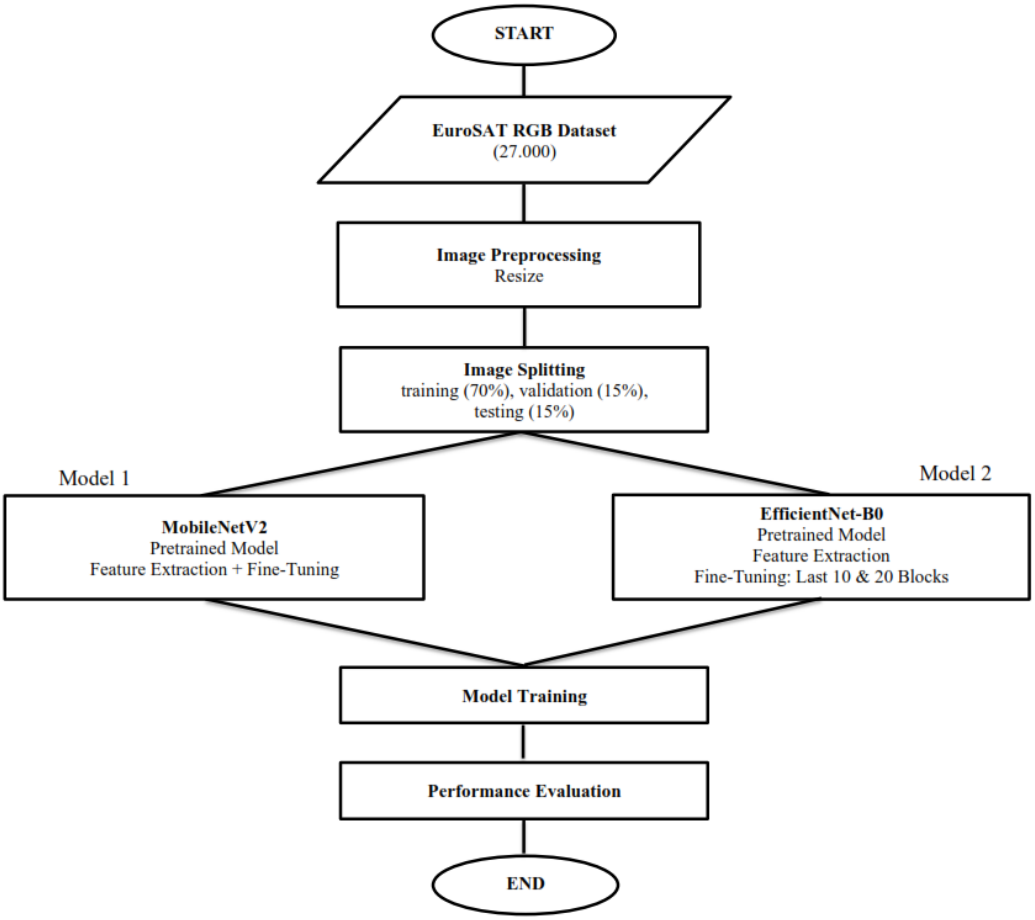

Figure 1. The overall workflow

TABLE 1. DISTRIBUTION OF EUROSAT RGB IMAGES ACROSS THE TEN LAND-COVER CLASSES

Class ID	Class	Images
0	AnnualCrop	3000
1	Forest	3000
2	HerbaceousVegetation	3000
3	Highway	2500
4	Industrial	2500
5	Pasture	2000
6	PermanentCrop	2500
7	Residential	3000
8	River	2500
9	SeaLake	3000

The same dataset was used for all six experimental configurations. Each directory corresponds to a land-cover class, allowing the directory structure to provide the target class labels automatically. The use of EuroSAT is appropriate for this study because it provides a standardized benchmark for evaluating deep-learning approaches to satellite-image classification. Fig 2 shows the RGB images from all the ten classes.



Figure 2. Representative RGB images from the ten EuroSAT land-cover classes

2.2 Image Processing and Image Splitting

Each RGB image was resized to 224 × 224 pixels and converted into a PyTorch tensor using ToTensor(). The resulting input representation contains three channels in the order used by the RGB images, with tensor values normalized to the [0, 1] range by the transformation. The same preprocessing procedure was applied to both CNN architectures to avoid introducing differences that could confound the model comparison.

The dataset was divided into 70% training, 15% validation, and 15% testing subsets, corresponding to 18,900, 4,050, and 4,050 images, respectively. The training set was used to update trainable parameters, the validation set was used to monitor learning and select the best model according to validation loss, and the independent test set was reserved for the final evaluation.

2.3 CNN Architectures and Transfer-Learning Configurations

MobileNetV2 and EfficientNet-B0, two pretrained CNN architectures, were assessed as feature-learning backbones for EuroSAT land-cover categorization. Both models were modified to fit the 10 EuroSAT classes by substituting the original classification layer with ImageNet-pretrained weights. The experiments investigate how the pretrained representations' adaptation to the target dataset is impacted by feature extraction and fine-tuning.

1) MobileNetV2

A lightweight CNN called MobileNetV2 was created for computationally effective image processing. It makes use of depthwise convolutions, linear bottlenecks, and inverted residual blocks [9]. In this study, a ten-class output layer is used in place of the original classifier to adapt the pretrained model to EuroSAT.

Three configurations were assessed: Feature Extraction, where all feature layers are frozen; Fine-Tuning 10, where the final ten feature modules are trainable; and Fine-Tuning 20, where all feature modules are trainable. These settings evaluate the impact of feature adaptation on classification performance. MobileNetV2's appropriateness for computationally limited remote-sensing applications makes it noteworthy as well [10].

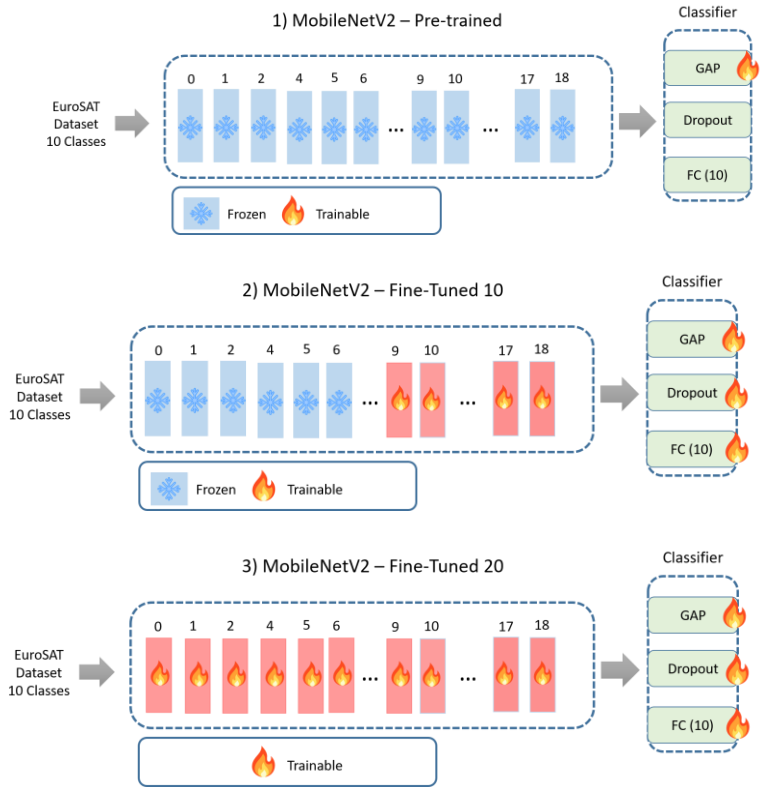

Figure 3. MobileNetV2 transfer-learning configurations for EuroSAT land-cover classification: (1) Feature Extraction, (2) Fine-Tuning 10, and (3) Fine-Tuning 20

2) EfficientNet-B0

EfficientNet is a CNN family based on compound scaling of network depth, width, and input resolution [11]. EfficientNet-B0 serves as the baseline architecture, offering a balance between classification performance and computational efficiency. Its applicability to remote-sensing land-cover classification has been investigated using the EuroSAT dataset and other land-cover benchmarks [12], [13].

In this study, EfficientNet-B0 is adapted to the ten EuroSAT classes by replacing its original classifier. Feature Extraction keeps the pretrained feature extractor frozen, whereas Fine-Tuning allows the feature extractor to update together with the classifier. The two fine-tuning configurations are implemented using different slices of the model.features module.

In the implemented PyTorch model, the EfficientNet-B0 feature sequence contains fewer than 20 top-level elements. Therefore, both model.features[-10:] and model.features[-20:] include the complete feature extractor. The two experiments are consequently interpreted as full-feature-extractor fine-tuning configurations rather than a strict 10-versus-20-block comparison.



Figure 4. EfficientNet-B0 transfer-learning configurations for EuroSAT land-cover classification: (1) pretrained feature extraction and (2) fine-tuning with the feature extractor unfrozen

2.4 Model Training and Evaluation

All experiments were trained for 20 epochs using Cross-Entropy Loss and the Adam optimizer with an initial learning rate of 0.001. These choices are consistent with training strategies adopted in previous deep transfer-learning studies for land-use and land-cover classification [14], [15]. A ReduceLROnPlateau scheduler was employed to reduce the learning rate when the validation loss stopped improving. The model checkpoint corresponding to the lowest validation loss was retained for final testing. Classification performance was evaluated on the 4,050-image test set using accuracy, macro-averaged precision, recall, and F1-score. Confusion matrices were used to examine class-specific classification errors and the effect of fine-tuning on class separation.

To ensure a fair comparison, all six model configurations were trained using the same dataset split (70% training, 15% validation, and 15% testing), input size of 224 × 224 pixels, ImageNet-pretrained weights, Adam optimizer with an initial learning rate of 0.001, and a maximum of 20 epochs. The model checkpoint with the lowest validation loss was selected for evaluation on the test set.

3. Results and Discussion

The six experiments reveal a consistent trend where both architectures gain advantages by tailoring their pretrained feature representations for the EuroSAT domain. The curves for training and validation accuracy are displayed in Fig. 5(a)–(f), whereas the associated training and validation loss curves are illustrated in Fig. 6(a)–(f). Panels (a)–(c) illustrate EfficientNet-B0 with Feature Extraction, Fine-Tuning 10, and Fine-Tuning 20, respectively, while panels (d)–(f) show the related MobileNetV2 setups.

As shown in Fig. 5, the Feature Extraction configurations exhibit a gradual improvement in accuracy during training. At epoch 20, EfficientNet-B0 reaches approximately 91.3% training accuracy and 93.9% validation accuracy, while MobileNetV2 reaches approximately 91.6% and 93.1%, respectively. In contrast, the fine-tuned configurations show a faster increase in training accuracy, approaching 100%, while validation accuracy stabilizes at around 98%. This indicates that the pretrained features already provide a useful basis for EuroSAT classification, but updating the

feature extractor allows the models to learn representations that are better suited to the target land-cover classes.

In the later epochs, training accuracy continues to increase while validation accuracy changes only slightly. This difference suggests that the models continue to fit the training data after most of the improvement in generalization has already been achieved. The pattern indicates a limited degree of overfitting in the later stages; however, the relatively stable validation accuracy shows that the models maintain good generalization to unseen samples.

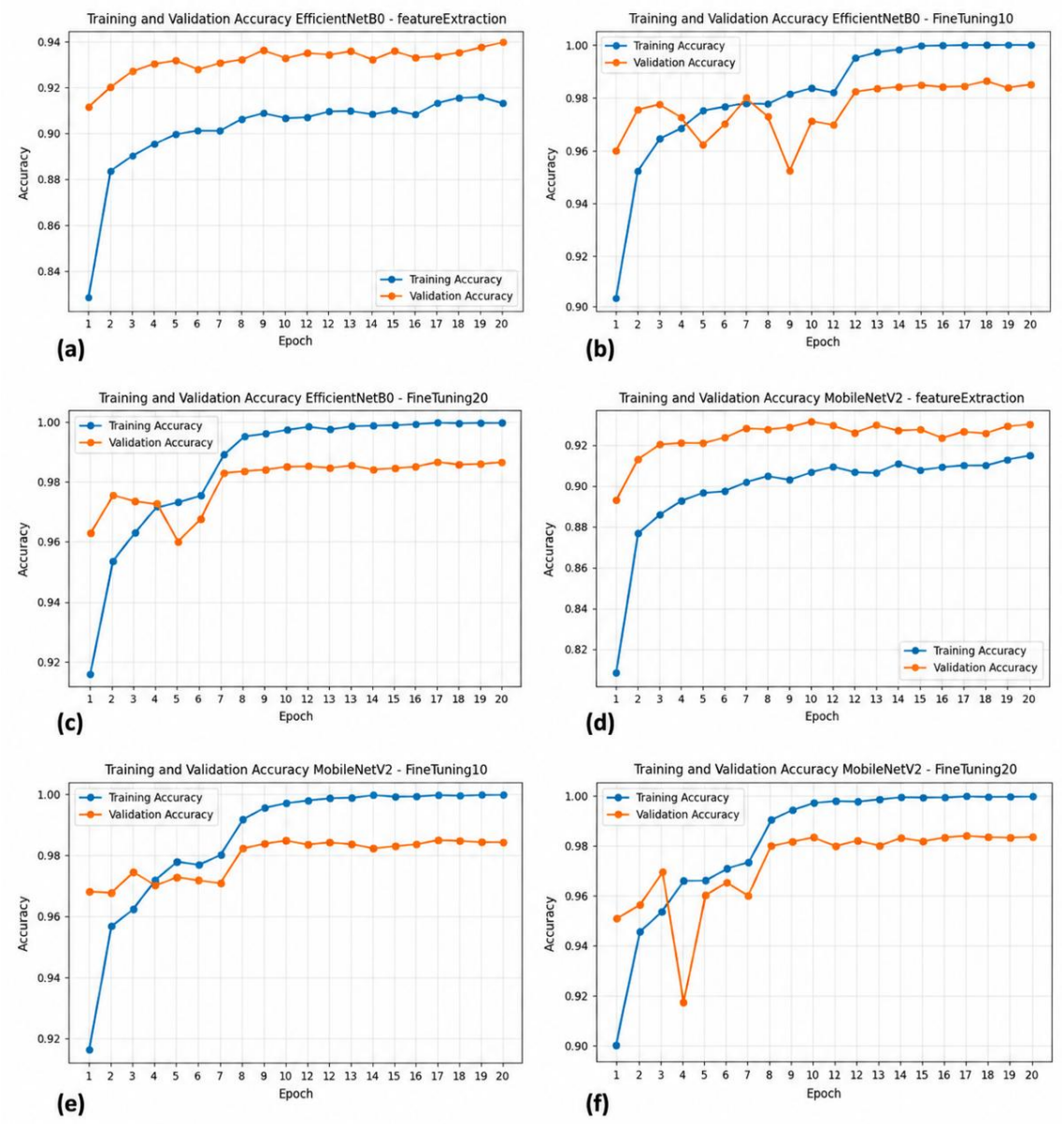

Figure 5. Training and validation accuracy for (a) EfficientNet-B0 Feature Extraction, (b) EfficientNet-B0 Fine-Tuning 10, (c) EfficientNet-B0 Fine-Tuning 20, (d) MobileNetV2 Feature Extraction, (e) MobileNetV2 Fine-Tuning 10, and (f) MobileNetV2 Fine-Tuning 20

The corresponding loss curves in Fig. 6 support this observation. The Feature Extraction configurations show a gradual decrease in training and validation loss, consistent with their slower improvement in accuracy. The fine-tuned models, on the

other hand, show a more pronounced reduction in training loss, reaching very low values in the later epochs. At the same time, validation loss remains higher and tends to stabilize, which is consistent with the gap observed between training and validation accuracy. Thus, the loss curves confirm that fine-tuning enables more effective adaptation of the pretrained representations to the EuroSAT classification task.

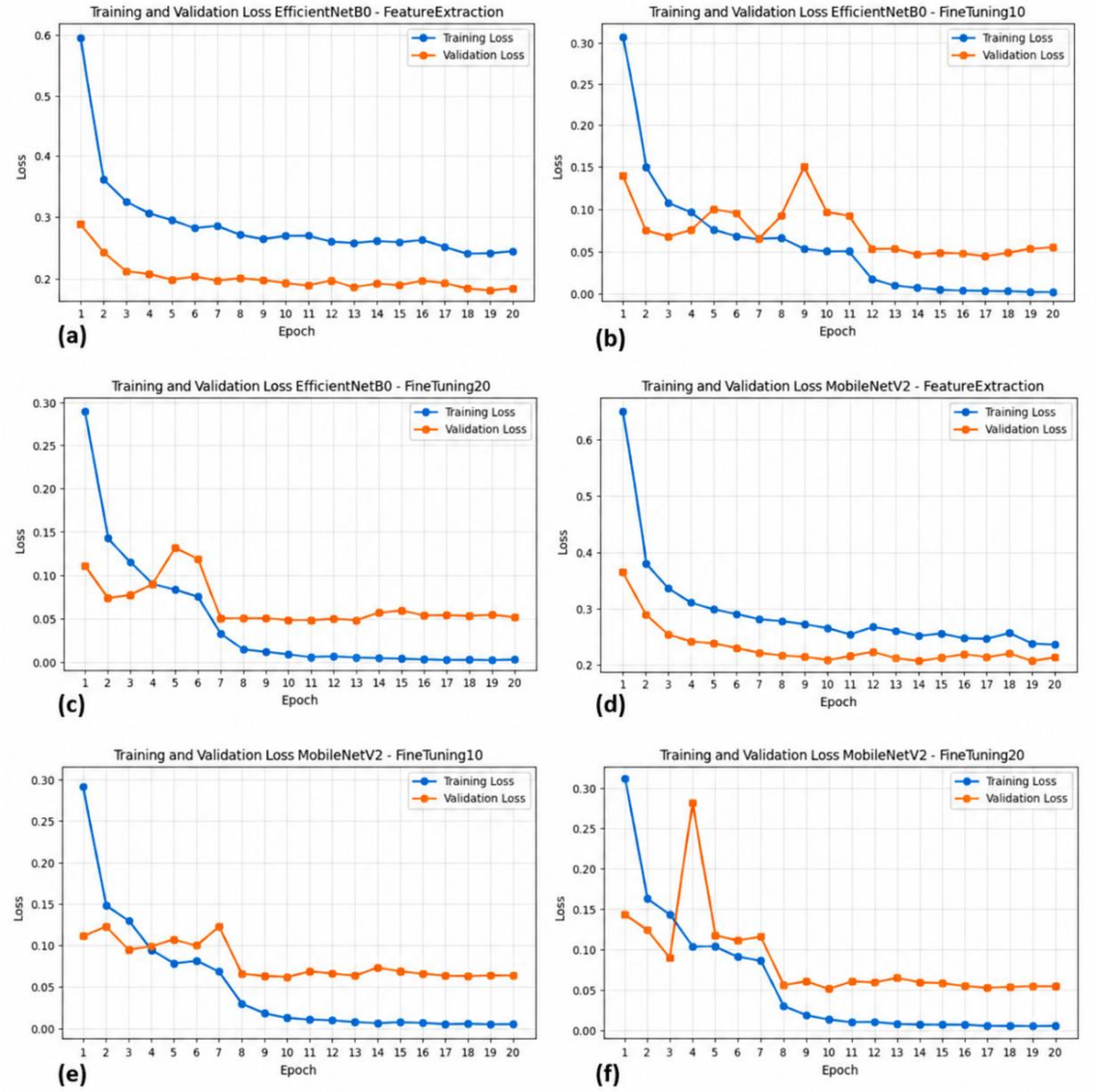

Figure 6. Training and validation loss for (a) EfficientNet-B0 Feature Extraction, (b) EfficientNet-B0 Fine-Tuning 10, (c) EfficientNet-B0 Fine-Tuning 20, (d) MobileNetV2 Feature Extraction, (e) MobileNetV2 Fine-Tuning 10, and (f) MobileNetV2 Fine-Tuning 20

Overall, Figs. 5 and 6 indicate that the main performance improvement is obtained by allowing the pretrained feature representations to adapt to the target dataset. Although further optimization continues to improve training performance, the relatively stable validation results indicate that the benefit to generalization becomes smaller in the later epochs. This provides a clear basis for the subsequent comparison of the final test-set performance.

TABLE 2. OVERALL TEST-SET PERFORMANCE OF THE SIX EXPERIMENTAL CONFIGURATIONS

Architecture	Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
EfficientNet-B0	Feature Extraction	**93.73**	93.52	93.46	**93.47**
EfficientNet-B0	Fine-Tuning 10	98.35	98.36	98.29	98.32
EfficientNet-B0	Fine-Tuning 20	98.40	98.36	98.32	98.34
MobileNetV2	Feature Extraction	**92.91**	92.69	92.81	**92.74**
MobileNetV2	Fine-Tuning 10	97.80	97.72	97.74	97.72
MobileNetV2	Fine-Tuning 20	97.98	97.94	97.92	97.92

Table 2 indicates that EfficientNet-B0 reaches a test accuracy of 93.73% and a macro F1-score of 93.47% when using Feature Extraction. Following fine-tuning, the accuracy rises to 98.35% and 98.40%, while the macro F1-score improves to 98.32% and 98.34%, respectively. MobileNetV2 shows a similar pattern, rising from 92.91% accuracy and 92.74% macro F1-score with Feature Extraction to 97.80% and 97.98% post fine-tuning. These findings indicate that the main improvement in performance is linked to modifying the pretrained feature representation instead of merely substituting the final classifier.

EfficientNet-B0 has a slight but steady edge over MobileNetV2. The difference in accuracy between the related architectures is 0.82 percentage points for Feature Extraction, 0.55 percentage points for Fine-Tuning 10, and 0.42 percentage points for Fine-Tuning 20. The reduced difference post fine-tuning suggests that MobileNetV2 significantly gains from domain-specific adjustment and can nearly match the performance of EfficientNet-B0 once its more profound representations are modified.

Despite this, EfficientNet-B0 continues to be the top-performing architecture in the current experiments.The distinction between Fine-Tuning 10 and Fine-Tuning 20 is relatively minor: 0.05 percentage points for EfficientNet-B0 and 0.18 percentage points for MobileNetV2. These minor improvements suggest decreasing returns after the pretrained representation has been adequately adjusted to the target dataset.

TABLE 3. CLASS-LEVEL F1-SCORE (%) FOR THE SIX EXPERIMENTAL CONFIGURATIONS

Class	EffNet FE	EffNet FT10	EffNet FT20	Mobile FE	Mobile FT10	Mobile FT20
AnnualCrop	94.67	97.13	97.66	93.75	96.79	96.38
Forest	97.85	99.66	99.55	97.19	99.32	98.98
HerbaceousVegetation	93.26	97.62	97.82	91.92	97.30	97.29
Highway	88.20	98.36	98.34	86.92	98.34	98.73
Industrial	96.33	98.66	98.80	94.69	98.53	98.94
Pasture	91.12	97.14	96.46	92.96	95.19	95.81
PermanentCrop	91.40	97.10	97.12	90.98	95.54	96.30
Residential	98.00	99.33	99.45	97.03	99.22	99.56
River	86.73	98.79	99.06	85.98	97.98	98.12
SeaLake	97.15	99.39	99.15	95.96	99.02	99.14

The results at the class level in Table 3 offer a clearer perspective on the variations in performance. In Feature Extraction, both architectures find River and Highway to be some of the most challenging categories. EfficientNet-B0 achieves F1-scores of 86.73% for River and 88.20% for Highway, whereas MobileNetV2 achieves 85.98% and 86.92%, respectively. These reduced scores signify that the standard pretrained representations are not as effective at distinguishing visually similar or spatially diverse classes.

Fine-tuning greatly enhances these challenging categories. Through Fine-Tuning 20, EfficientNet-B0 achieves an F1-score of 99.06% for River and 98.34% for Highway, whereas MobileNetV2 attains scores of 98.12% and 98.73%. The significant

enhancement for River is especially revealing as it indicates that fine-tuning alters not just the overall accuracy but also the model's capacity to differentiate classes that were inadequately separated by the static pretrained representation. The lower performance of River and Highway under Feature Extraction may reflect greater visual variability and overlap with other land-cover categories, making their discrimination more dependent on target-domain feature adaptation

The confusion matrices displayed in Fig. 7 support the results at the class level. The Feature Extraction settings reveal more noticeable off-diagonal values, suggesting that a greater number of samples are being allocated to rival classes. Following fine-tuning, the matrices exhibit a pronounced diagonal structure, indicating enhanced differentiation among the ten land-cover classes. The alignment between the confusion matrices and the class-level F1-scores reinforces the understanding that the performance improvement is widespread instead of being focused on a few classes.

Collectively, the six experiments demonstrate that architecture and transfer-learning approaches serve different functions. EfficientNet-B0 consistently reaches the top absolute performance, while MobileNetV2 stays very competitive post fine-tuning. The main impact is the shift from frozen to adjusted characteristics, whereas the extra benefit from more comprehensive fine-tuning is fairly minor. These results suggest that adapting pretrained CNN features to different domains is crucial for achieving precise EuroSAT land-cover classification, while MobileNetV2 provides a strong lightweight option.

4. Conclusion

This research assessed MobileNetV2 and EfficientNet-B0 for EuroSAT land-cover classification within a unified transfer-learning framework. Both architectures gained significantly from fine-tuning in contrast to feature extraction, showing that customizing pretrained visual representations to the properties of satellite imagery is crucial for attaining excellent classification results. EfficientNet-B0 achieved the highest overall performance, attaining 98.40% accuracy and 98.34% macro F1-score, whereas MobileNetV2 achieved 97.98% accuracy and 97.92% macro F1-score in its optimal setup. The analyses of both class-level and the confusion matrix indicated that fine-tuning significantly lowers mistakes for tougher categories such as River and Highway. The slight difference between the two fine-tuning setups suggests diminishing benefits from further parameter adjustment. These results endorse the use of EfficientNet-B0 when prioritizing classification performance, whereas MobileNetV2 continues to be a strong choice for scenarios that prefer a lightweight CNN design.

Future work could extend this study by evaluating the models on other remote sensing datasets and testing alternative transfer-learning and fine-tuning strategies. Further investigation of lightweight architectures could also be conducted to improve computational efficiency while maintaining high classification accuracy, particularly for resource-constrained remote sensing applications

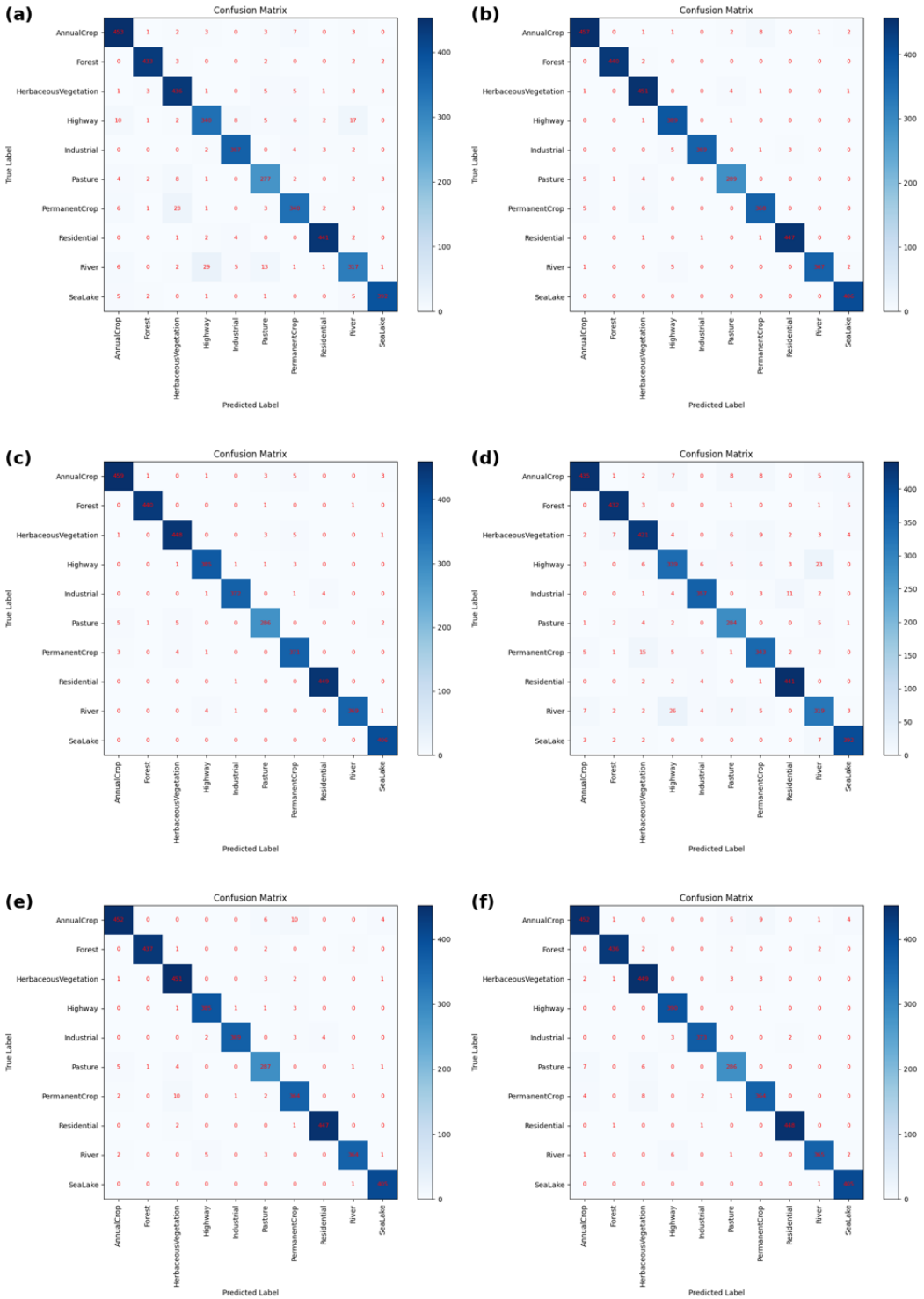

Figure 7. Confusion matrices for (a) EfficientNet-B0 Feature Extraction, (b) EfficientNet-B0 Fine-Tuning 10, (c) EfficientNet-B0 Fine-Tuning 20, (d) MobileNetV2 Feature Extraction, (e) MobileNetV2 Fine-Tuning 10, and (f) MobileNetV2 Fine-Tuning 20.

References

[1] R. Naushad, T. Kaur, and E. Ghaderpour, "Deep Transfer Learning for Land Use and Land Cover Classification: A Comparative Study," Sensors, vol. 21, no. 23, Art. no. 8083, 2021.
[2] A. Sharma et al., "A Classification Method for Satellite Imagery Utilizing Transfer Learning Techniques," Proc. 2025 7th Int. Conf. Signal Processing, Computing and Control (ISPCC), 2025.
[3] J. Plested, M. Phiri, and T. Gedeon, "Deep transfer learning for image classification: A survey," Artificial Intelligence Review, vol. 59, no. 3, p. 100, 2026.
[4] M. Rouba and M. E. A. Larabi, "Improving remote sensing classification with transfer learning: Exploring the impact of heterogeneous transfer learning," *Engineering Proceedings*, vol. 56, no. 1, p. 316, 2023.
[5] A. S. S. et al., "Multi-label remote sensing image classification using MobileNetV2," in Proc. IEEE ICCCNT, 2024.
[6] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019.
[7] A. Alem and S. Kumar, "Transfer learning models for land cover and land use classification in remote sensing image," Applied Artificial Intelligence, vol. 36, no. 1, Art. no. 2014192, 2022, doi: 10.1080/08839514.2021.2014192.
[8] I. Papoutsis, N.-I. Bountos, A. Zavras, D. Michail, and C. Tryfonopoulos, "Benchmarking and scaling of deep learning models for land cover image classification," ISPRS Journal of Photogrammetry and Remote Sensing, 2022.
[9] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.
[10] L. Pham, C. Le, D. Ngo, A. Nguyen, J. Lampert, A. Schindler, and I. McLoughlin, "A light-weight deep learning model for remote sensing image classification," in Proc. 13th Int. Symp. Image Signal Process. Anal. (ISPA), 2023, doi: 10.1109/ISPA58351.2023.10279679.
[11] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. 36th Int. Conf. Mach. Learn. (ICML), vol. 97, 2019, pp. 6105–6114.
[12] B. Sarıçayır and C. Özcan, "EfficientNet Deep Learning Model for Satellite Image Classification Using the EuroSAT Dataset," in Proceedings of the Eighth International Workshop on Computer Modeling and Intelligent Systems, 2025.
[13] I. Papoutsis, N. I. Bountos, A. Zavras, D. Michail, and C. Tryfonopoulos, "Benchmarking and scaling of deep learning models for land cover image classification," ISPRS J. Photogramm. Remote Sens., vol. 195, pp. 250–268, Jan. 2023.
[14] R. Naushad, T. Kaur, and E. Ghaderpour, "Deep transfer learning for land use and land cover classification: A comparative study," Sensors, vol. 21, no. 23, Art. no. 8083, 2021.
[15] H. Dastour and Q. K. Hassan, "A comparison of deep transfer learning methods for land use and land cover classification," Sustainability, vol. 15, no. 10, Art. no. 7854, 2023.