Accredited SINTA 2 Ranking
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online at: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 9 No. 1 (2025) 1 - 9 e-ISSN: 2580-0760

Comparison of Transfer Learning Architecture Performance for Indonesian Auction Object Classification

Hanif Noer Rofiq[1]*
[1]Directorate Transformation and Information Systems, Directorate General of State Assets Management, Ministry of Finance, Jakarta, Indonesia
[1]hanif.noer94@gmail.com

Abstract

The Indonesian auction, one of the sources of Indonesia's income for Non-Tax State Revenue (PNBP), faces challenges in accurately classifying auction objects, limiting revenue optimisation. This research aims to compare the performance of several transfer learning architectures on the Indonesian Auction Object Dataset, which includes categories such as Buildings, Cars, Motorbikes, and Salvage Materials. Seven pre-trained transfer learning models—MobileNetV2, NASNetMobile, EfficientNetV2B0, DenseNet121, Xception, InceptionV3, and ResNet50V2—were evaluated against a baseline model, focusing on validation accuracy, model size, and computational efficiency. MobileNetV2, NASNetMobile, DenseNet121, Xception, InceptionV3, and ResNet50V2 all achieved 100% validation accuracy, outperforming the baseline model's 96.5% accuracy. MobileNetV2 stands out for its efficiency, reaching 100% accuracy in just eight epochs with a compact model size of 11.1 MB. In contrast, EfficientNetV2B0 performed poorly on this dataset, achieving only 25% validation accuracy. These findings confirm that transfer learning architectures can significantly improve auction object classification accuracy while reducing the model size and training time, highlighting the benefit of transfer learning for optimising Indonesian auction systems.

Keywords: *image classification; indonesian auction; PNBP; transfer learning*

How to Cite: H. N. Rofiq, "Comparison of Transfer Learning Architecture Performance for Indonesian Auction Object Classification", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 9, no. 1, pp. 1 - 9, Jan. 2025.
DOI: https://doi.org/10.29207/resti.v9i1.6082

1. Introduction

The Ministry of Finance Regulation [1] defines an auction as a public sale of goods aimed at achieving the highest possible price, with auction fees charged to the seller and/or buyer for each auction conducted. These fees contribute significantly to Non-Tax State Revenue (PNBP). which reached IDR 974.24 billion in 2023, more than double the IDR 467.68 billion revenue generated in 2018 [2]. This substantial growth underscores the increasing importance of auctions as a revenue collection source and highlights the need for continuous improvement in auction management practices, as the current business processes still have room for enhancement to maximize revenue collection [3], [4].

The increase in non-tax revenue is closely tied to the frequency of auctions, which has risen significantly since the introduction of the new Indonesian auction domain (lelang.go.id) in 2018. Sujak and Rofiq [5] highlight that as the frequency of auctions increases, the challenge of efficiently classifying diverse auction objects becomes more complex, mainly due to the limited number of auctioneers available to oversee the process. The research also found that inefficiencies in auction classification often arise from differing interpretations between sellers and regulators. Saputri et al. [6] further note that auctioneers focus primarily on verifying the formal legal requirements of auction requests without assessing the actual condition or authenticity of the items. As a result, the responsibility for ensuring accurate auction object descriptions falls entirely on the seller. These issues underscore the need for more robust, automated solutions to enhance classification accuracy, optimising the auction system's non-tax revenue.

Building on these challenges, this research extends Sujak and Rofiq's [5] research, which explored AI's potential for optimising PNBP through auction fee classification in the Indonesian auction system. which explored the potential of artificial intelligence (AI) to

Received: 25-09-2024 | Accepted: 23-12-2024 | Published Online: 19-01-2025

optimise PNBP through auction object classification to determine its fees in the Indonesian auction system. Their research demonstrated the effectiveness of AI in improving auction processes, specifically by leveraging a custom Convolutional Neural Network (CNN) model. This CNN achieved a validation accuracy of 96.5%, with a relatively large model size of 692.4 MB. While this performance was promising, there remain several areas where further improvements are necessary, particularly in achieving higher accuracy and reducing the model size.

The CNNs approach has proven effective in similar applications, such as Jareño et al. [7], which utilised CNNs to classify fish size and species in Spanish fish auctions, and Bobulski and Szymoniak [8] applied CNNs to categorise items in Dutch auctions. The Indonesian auction system presents unique challenges. The dataset used in these auctions includes an object that can't be found in another country or auction process, that is, salvage materials. A more refined approach is needed to enhance the classification performance, especially considering that accurate classification directly impacts the collection of PNBP, one of the revenue sources for the Indonesian government. Misclassification of auction objects can lead to inaccurate fees imposed, potentially resulting in revenue loss for the government. Therefore, optimising accuracy is essential for ensuring that the auction process contributes effectively to state revenue. Additionally, a smaller model size is equally important, as it allows for seamless integration into existing auction systems, which often have computational and storage constraints. A more compact model ensures faster processing times, reduces the need for extensive hardware resources, and supports deployment in environments with limited computational power, such as mobile devices or smaller servers.

In response to these challenges, this research explores the application of transfer learning techniques. Transfer learning offers significant advantages by leveraging knowledge from previously learned tasks [9]. Transfer learning also addresses the challenge of insufficient training data [10], [11] and significantly reduces the time and computational resources required for model training [12]. This approach is particularly suitable for this case, as the dataset is limited to 250 images per category.

Transfer learning has been widely adopted in various image classification tasks, offering the advantage of achieving high accuracy even with smaller datasets and more efficient model architectures. This technique leverages pre-trained models, which have been trained on large datasets, and fine-tunes them on specific tasks, allowing for rapid and effective adaptation to new image classification tasks.

For instance, Sailaja and VenuGopal [13] demonstrated the usability of transfer learning by achieving 100% accuracy in Parkinson's Disease detection. Their research enhanced the ResNet50V2 model with additional dense and dropout layers, allowing it to capture more complex features while mitigating overfitting. This illustrates how transfer learning can be fine-tuned to deliver exceptional results in medical diagnostics.

Similarly, Chen et al. [14] achieved remarkable success in plant disease identification by employing a squeeze-and-excitation MobileNet model with twice transfer learning, achieving an accuracy of 99.78%. Kaur et al. [10] applied the InceptionV3 model to detect rice leaf diseases, achieving 96% accuracy. Their approach highlights how transfer learning can be adapted to various domains within agriculture, where precise identification of plant diseases is crucial for effective crop treatment and management. Additionally, Foong [15] utilised ResNet50 to detect rotten fruit, achieving 98.89% accuracy. This further gives another example of the effectiveness of transfer learning in food quality assessment, where accurate defect detection is essential for ensuring product quality.

Despite the widespread application of transfer learning across various domains, there remains limited research on its application to auction object classification, particularly within the context of the Indonesian auction system. This domain presents unique challenges, such as classifying various auctioned items with limited training data. Exploring transfer learning in this context could offer significant benefits, improving the accuracy and efficiency of auction object classification while addressing the specific needs of the Indonesian government to automate the classification process in the Indonesian auction system, ensuring that non-tax state revenue imposed is accurate and reliable.

This research aims to fill this gap by applying transfer learning techniques to enhance the accuracy of auction object classification and reduce model size for seamless integration into existing systems. The findings will provide practical insights to enhance the efficiency of the Indonesian auction system and further optimise its contribution to non-tax state revenue.

2. Research Methods

This research employed a comparative experimental design to evaluate the performance of various transfer learning architectures on the Indonesian Auction Object Dataset, as used by Sujak and Rofiq [5]. The dataset consists of four categories: Buildings, Cars, Motorbikes, and Salvage Materials. It has 1000 images split into 200 training images and 50 validation images per category, each 512x512 pixels.

2.1 Model Selection and Benchmarking

To select which transfer learning models to choose from the available models, the Keras Application benchmark [16] was referenced, considering model size, Top-1, and Top-5 accuracy. Keras, a Python-based neural network API, is used to run TensorFlow [17].

The chosen models were required to be smaller than the base model size of 692.4 MB while achieving high accuracy. To evaluate these models, a Size-Accuracy Weight (SW) metric was introduced as a reference metric, considering both the model size and classification performance. The equation for SW is shown in Formula 1.

$$SW = \frac{1}{3} x (\frac{BM1}{BM2}) + T1 + T5 \quad (1)$$

BM1 refers to the base model size, which is 692.4 MB, while BM2 represents the benchmark model size, which is available in the Keras Application Benchmark [16] and varies depending on the selected transfer learning architecture. T1 is Top-1 accuracy, and T5 is Top-5 accuracy, which, in the Keras Application benchmark, Top-1 and Top-5 accuracy refer to a model's performance on the ImageNet validation dataset. The widely accepted hypothesis is that models with higher accuracy gained from the ImageNet dataset show superior performance on a broad range of applications in another task [18]. ImageNet, a large-scale, human-annotated dataset with 1.28 million images for training, 50,000 for validation, and 100,000 for testing, covers 1,000 different classes [19], making it a robust benchmark for evaluating model performance. The reasoning behind this hypothesis is that models trained on such a diverse and comprehensive dataset as ImageNet will develop a strong ability to generalise, which translates into superior performance when these models are fine-tuned for other specific tasks.

In support of this, studies have consistently shown that supervised pre-trained models that perform well on ImageNet tend to achieve better results when applied to different tasks, including those that differ significantly from the original ImageNet classification task. This strong correlation between ImageNet accuracy and downstream task performance has made it a standard practice to use ImageNet accuracy as a key criterion for selecting pre-trained models in transfer learning [18].

To simplify the process of selecting the most suitable transfer learning architecture, this research not only relies on SW metrics but also incorporates the widely accepted practice of selecting transfer learning architecture based on their performance on ImageNet accuracy. Such as calculating the mean of Top-1 and Top-5 accuracy to deliver a more balanced assessment of model performance. This approach ensures that the evaluation captures not just the model's ability to predict the most likely class (Top-1) but also its capability to rank the correct class within its top five predictions (Top-5). The mathematical formula for this calculation is presented in Formula 2.

$$accuracy\ mean = \frac{Top\ 1\ Accuracy + Top\ 5\ Accuracy}{2} \quad (2)$$

The remaining models were then ranked using the Size-Accuracy Weight, a metric that balances model accuracy with its model's size. This approach ensures that the selected model's architecture not only performs well in terms of accuracy but is also manageable in terms of storage and deployment. Models that exceeded 100 MB in size and had an accuracy mean below 80% were excluded from the transfer learning architecture selection. This threshold was set to prioritise models that offer a strong balance between high performance and practical usability, ensuring they are both effective and resource-efficient for real-world applications. The results from this transfer learning architecture selection are shown in Table 1.

As seen in Table 1, several transfer learning architectures, such as EfficientNetV2B0 and EfficientNetB0, DenseNet121 and DenseNet201, as well as ResNet50 and ResNet50V2, belong to the same architectural family [20], [21], [22]. Therefore, only the highest-ranked model from each architecture family was selected. Seven models, MobileNetV2, NASNetMobile, EfficientNetV2B0, DenseNet121, Xception, InceptionV3, and ResNet50V2, were chosen for further evaluation, as they met the specified criteria.

Table 1. Keras benchmark calculation result

No	Model	Size (MB)	Top-1 Accuracy	Top-5 Accuracy	SW	Accuracy Mean
1.	**MobileNetV2**	14	71.30%	90.10%	17.024	80.70%
2.	**NASNetMobile**	23	74.40%	91.90%	10.589	83.15%
3.	**EfficientNetV2B0**	29	78.70%	94.30%	8.535	86.50%
4.	EfficientNetB0	29	77.10%	93.30%	8.527	85.20%
5.	EfficientNetB1	31	79.10%	94.40%	8.023	86.75%
6.	**DenseNet121**	33	75.00%	92.30%	7.552	83.65%
7.	EfficientNetV2B1	34	79.80%	95.00%	7.371	87.40%
8.	EfficientNetB2	36	80.10%	94.90%	6.994	87.50%
9.	EfficientNetV2B2	42	80.50%	95.10%	6.081	87.80%
10.	EfficientNetB3	48	81.60%	95.70%	5.399	88.65%
11.	DenseNet169	57	76.20%	93.20%	4.614	84.70%
12.	EfficientNetV2B3	59	82.00%	95.80%	4.505	88.90%
13.	EfficientNetB4	75	82.90%	96.40%	3.675	89.65%
14.	DenseNet201	80	77.30%	93.60%	3.455	85.45%
15.	EfficientNetV2S	88	83.90%	96.70%	3.225	90.30%
16.	**Xception**	88	79.00%	94.50%	3.201	86.75%
17.	**InceptionV3**	92	77.90%	93.70%	3.081	85.80%
18.	**ResNet50V2**	98	76.00%	93.00%	2.918	84.50%
19.	ResNet50	98	74.90%	92.10%	2.912	83.50%

Several previous studies have highlighted the good results of various transfer learning models on different datasets. For instance, Yong et al. [23] demonstrated that MobileNetV2 outperformed the CNN model in waste classification, achieving 15.42% higher accuracy. Naskinova [24] applied NASNetMobile for pneumonia X-ray classification, finding that it can improve accuracy by an average of 5% and reduce loss to 15%. Fayyad and Mustakim [25] compared several transfer learning architectures for glaucoma eye disease classification, with EfficientNetV2B0 emerging as the top performer, achieving 89.77% accuracy. Zebari et al. [26] used DenseNet121 to enhance brain tumour classification, reaching an accuracy of 94.83%. Salim et al. [27] conducted a comparative study using various transfer learning algorithms for fruit classification tasks, where Xception achieved 99.13% accuracy on the Fruits-360 dataset and 97.73% on the Fruit Recognition dataset. Jaware et al. [28] applied InceptionV3 for colon cancer classification, achieving 98.86% training accuracy and 99.74% validation accuracy after 100 epochs. Additionally, Sailaja and VenuGopal [29] used ResNet50V2 for Parkinson's disease detection, achieving a perfect accuracy score of 100%, outperforming other models like MobileNetV2, which achieved 95%, and InceptionV3, which reached 99% accuracy. These studies collectively demonstrate that transfer learning algorithms can be implemented in various domains and demonstrate their reliability in producing high-performance models.

2.1 Data Preparation

After selecting the transfer learning architectures, the next step involved data preparation, which followed the approach used by Sujak and Rofiq [5]. This included data augmentation techniques such as rotations up to 10 degrees, width and height shifts up to 10% and horizontal flipping. Reflect mode was used to fill empty pixels. These augmentations were applied during data preparation using an image generator, ensuring consistent settings across the seven selected transfer learning architectures for comparison in this research.

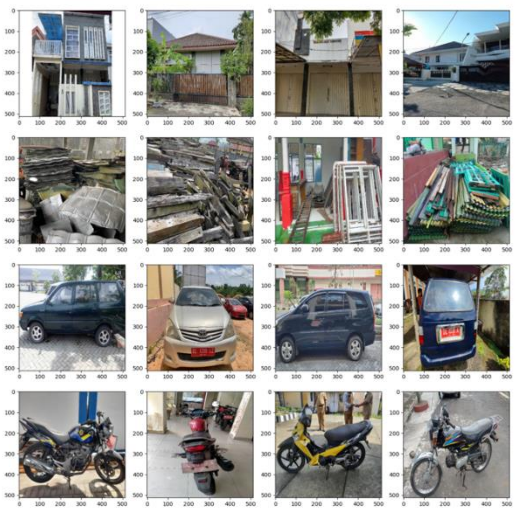

Figure 1. Sample Dataset

Figure 1 is the sample dataset of this research.

2.2 Modelling

Following data preparation, the next step is building the model. The first step is loading the pre-trained transfer learning model architectures with ImageNet weights. The original classification layers were removed by setting the arguments include_top to False, allowing custom layers to be added for the specific task. The global average pooling argument was applied to reduce each feature map by averaging all its elements to a single value. The trainable argument was set to False, freezing the pre-trained ImageNet weights to prevent it from being updated during training. This will preserve the knowledge and features learned from the extensive ImageNet dataset, allowing the model to focus on learning the specific features of the auction object classification task without losing the generalisation capabilities provided by the pre-trained weights.

To further improve the performance of the transfer learning models, fine-tuning was applied by adding a dense layer with 128 units. This dense layer acts as a fully connected layer, which integrates features extracted by the pre-trained layers and allows the model to learn more complex patterns specific to the Indonesian auction dataset. The ReLU (Rectified Linear Unit) activation function was used in this dense layer, followed by a dropout layer with a dropout rate of 0.5 to reduce the likelihood of overfitting. The final layer is configured with four output units using a softmax activation function to handle the multi-class classification task, classifying auction objects into four categories: Buildings, Cars, Motorbikes, and Salvage Materials.

The model was optimised using the Adam optimiser and a batch size of 32. The categorical cross-entropy loss function is used, which is appropriate for multi-class classification problems as it measures the model's performance by comparing the predicted class probabilities to the actual class labels. The training process was set to run for a maximum of 500 epochs to ensure that the model had enough training. An epoch is one complete pass through the entire training dataset, and allowing up to 500 epochs gives the model sufficient time to find its best model performance.

2.3 Performance Evaluation

To avoid overfitting and unnecessary computation, early stopping was implemented. This technique is set to monitor the validation loss during training. If the validation loss does not improve for ten consecutive epochs, it indicates that the model may no longer be learning or might start to overfit the training data, and the training process is automatically stopped. Validation loss is used to determine which epoch performs best and to identify the best model because it measures how well the model generalises to unseen data, so the model with the lowest validation loss was selected as the final model [30], [31].

In addition to early stopping, model checkpointing was used to save only the best weight. This means that the model's performance on the validation dataset was continuously monitored during training, and only the version with the lowest validation loss was saved. This approach helps to ensure that the final model saved retains its best performance while avoiding overfitting, saving both time and computational resources so it can perform better on the Indonesian auction dataset. Moreover, a confusion matrix was used to evaluate further and analyse the model's performance.

These configurations were applied to all seven selected transfer learning models.

Table 2. Global Model Summary

Layer type	Output Shape
(Transfer Learning Model Architecture with ImageNet weight)	
Dense	(None, 128)
Dropout	(None, 128)
Dense	(None, 4)

Table 2 shows the global model summary used in this research, where the selected pre-trained transfer learning architecture determines the initial shape of the model. Then a dense layer and a dropout layer are added, followed by a final dense layer with four neurons designed to adapt the model to the four auction object classification tasks.

Table 3. Total parameters from each model architecture

No	Model	Total Parameters	Trainable Parameters	Non-trainable Parameters
1.	MobileNetV2	2,422,468	164,484	2,257,984
2.	NASNetMobile	4,405,528	135,812	4,269,716
3.	EfficientNetV2B0	6,083,796	164,484	5,919,312
4.	DenseNet121	7,169,220	131,716	7,037,504
5.	Xception	21,124,268	262,788	20,861,480
6.	InceptionV3	22,065,572	262,788	21,802,784
7.	ResNet50V2	23,827,588	262,788	23,564,800

Table 3 details the total parameters, including trainable and non-trainable parameters from each model. The total parameters sum all model parameters, including trainable and non-trainable parameters. The number of total parameters affects the model's complexity and memory requirements for training the model. Trainable Parameters are the parameters that will be updated during the training process. Fine-tuning these parameters allows the model to adapt to the specific task or dataset. Non-trainable Parameters are frozen and will not be updated during training. These Non-trainable parameters come from the transfer learning architecture chosen, which preserves previously learned knowledge and features.

The experimental analysis was conducted on a Windows 11 operating system using Python 3.9.19 and TensorFlow 2.10.1, with CUDA version 11.2 and cuDNN 8.1. The setup ran on an Intel Core i7-12700 processor, an NVIDIA GeForce RTX 3070 GPU with 8 GB of VRAM, and 32 GB of RAM.

3. Results and Discussions

This research compares the performance of seven transfer learning models against a baseline model from Sujak and Rofiq [5] to assess their adaptability for auction object classification, where high accuracy is critical for its implementation. The transfer learning models evaluated were MobileNetV2, NASNetMobile, EfficientNetV2B0, DenseNet121, Xception, InceptionV3, and ResNet50V2. For each model, the one with the lowest validation loss during training was selected as the final model. Each model reached optimal performance at different epochs, depending on the complexity of the architecture and its ability to adapt to the Indonesian auction dataset.

For example, MobileNetV2, which has fewer parameters, around 2.4 million, reached its best performance in just eight epochs and produced a relatively small model size of 11.1 MB. In contrast, more complex architectures like ResNet50V2 have significantly more parameters, around 23.8 million, and took longer to train, resulting in much larger model sizes of 95.5 MB but with better validation loss. This illustrates how models with more parameters generally lead to larger model sizes and require more training time to reach optimal performance. Table 4 provides a detailed summary of the training results for each model, including metrics such as training accuracy, training loss, validation accuracy, validation loss, model size, and the best epoch.

The best epoch represents the point at which the model achieves the lowest validation loss during training. Once this is achieved, training is stopped if there is no further improvement in validation loss for ten consecutive epochs by using early stopping. The varying best epochs among the models reflect differences in training times and learning capabilities. These differences underscore the importance of model selection in transfer learning, as the complexity and architecture of a model can influence its training time, generalisation ability, and performance on specific tasks like Indonesian auction object classification.

The Base Model, which refers to the previous research by Sujak and Rofiq [5], with a size of 692.4 MB, achieved a validation accuracy of 96.50% at its best epoch at 39, demonstrating a solid performance with a moderate model size. However, the majority of the transfer learning models outperformed the baseline model in both accuracy and model size, with several achieving perfect training and validation accuracy. Notably, NASNetMobile, DenseNet121, Xception, InceptionV3, and ResNet50V2 reached 100% in training and validation accuracy, indicating their architectures are suitable for classifying auction objects. Larger architectures like ResNet50V2, InceptionV3, and Xception, although having the same 262,788 trainable parameters, contain significantly more total and non-trainable parameters compared to other models, as shown in Table 3. This indicates greater complexity, which leads to larger model sizes, as seen in Table 4.

Table 4. Training Results

No	Model	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Size (MB)	Training Time
1.	Base Model	96.13%	0.12030	96.50%	0.10470	692.4	unknown
2.	MobileNetV2	99.75%	0.01550	100%	0.01331	11.1	3 min 54 s
3.	NASNetMobile	100%	0.00255	100%	0.00029	19.5	15 min 10 s
4.	EfficientNetV2B0	20.62%	1.38642	25.00%	1.38629	25.5	8 min 2 s
5.	DenseNet121	100%	0.00870	100%	0.00211	29.7	7 min 12 s
6.	Xception	100%	0.00037	100%	0.00018	84.8	14 min 41 s
7.	InceptionV3	100%	0.00119	100%	0.00011	88.8	15 min 3 s
8.	ResNet50V2	100%	0.00055	100%	0.00008	95.5	8 min 26 s

Despite their size, these models achieved the overall lowest validation loss and 100% accuracy in both training and validation. Notably, from the transfer learning model with a larger architecture, ResNet50V2 achieved its optimal performance with a faster training time, requiring only 25 epochs to produce the best model on the Indonesian auction dataset with a validation loss of 0.00008. However, a more complex architecture does not always guarantee better results. NASNetMobile, with the second smallest total parameter count, achieved comparable performance, including a low validation loss, 0.00029 and 100% accuracy in training and validation, demonstrating that simpler architectures can be equally effective.

MobileNetV2 also demonstrated excellent performance by achieving 100% validation accuracy with a significantly smaller model size of just 11.1 MB and within only eight epochs, making it one of the most efficient in terms of size and training speed. Although its training accuracy was slightly lower at 99.75%, it still outperformed the baseline model. In contrast, EfficientNetV2B0 underperformed, with training and validation accuracies of only 20.62% and 25%, respectively. This poor performance suggests that EfficientNetV2B0's architecture, while known for its efficiency in some tasks, is not well-suited for the auction object classification problem. The EfficientNetV2B0's inability to perform in the Indonesian Auction dataset underscores the importance of comparing multiple transfer learning architectures to identify the most suitable model for a specific task.

The base model lacks training time information, as it was not provided in the previous research, making a direct comparison of training time from the base model with transfer learning models impossible. However, the transfer learning architecture used in this research exhibited varying training times. MobileNetV2 was the fastest, achieving its best result in just 3 minutes and 54 seconds, making it the most efficient in terms of speed and accuracy. In contrast, NASNetMobile, despite having fewer parameters than ResNet50V2, took significantly longer at 15 minutes and 10 seconds. Interestingly, ResNet50V2, with the largest total parameters in this research, required 8 minutes and 26 seconds to achieve its best result. This suggests that training time depends not only on parameter count but also on the model’s architecture and efficiency.

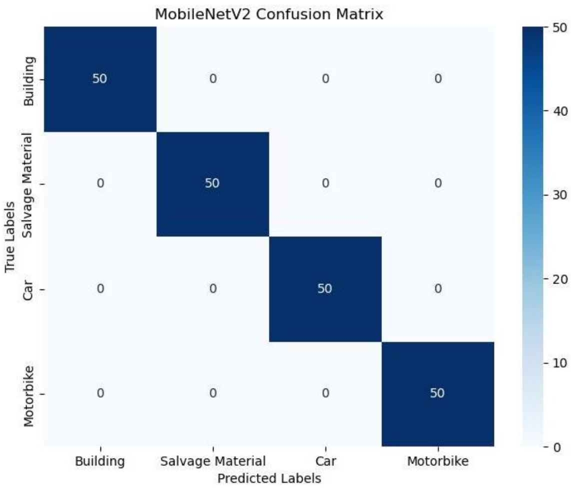

Figure 2. Confusion Matrix from MobileNetV2

Figure 2 presents the classification performance of the MobileNetV2 model, which achieved 100% accuracy across all categories, with every category correctly predicted without any misclassifications.

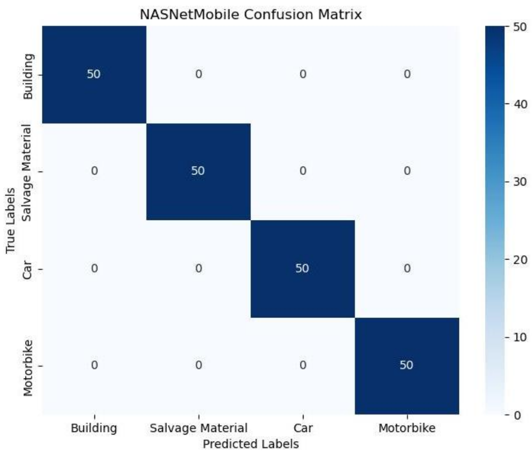

Figure 3. Confusion Matrix from NASNetMobile

Figure 3 presents the classification performance of the NASNetMobile model, which achieved 100% accuracy across all categories, with every category correctly predicted without any misclassifications.

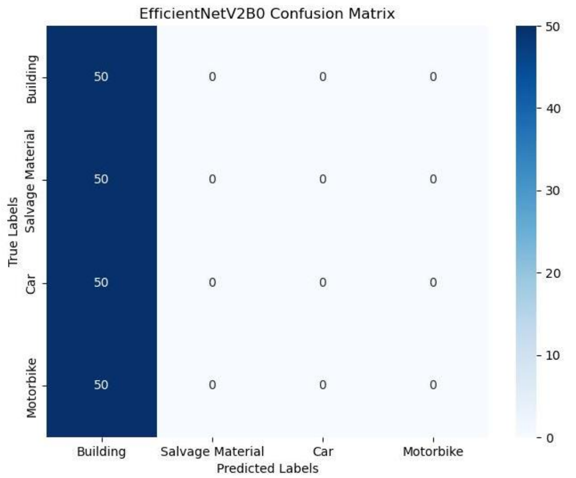

Figure 4. Confusion Matrix from EfficientNetV2B0

Figure 4 shows that EfficientNetV2B0 model failed to differentiate between the categories and did not generalize well to the task, leading to a 100% misclassification rate for all categories except building.

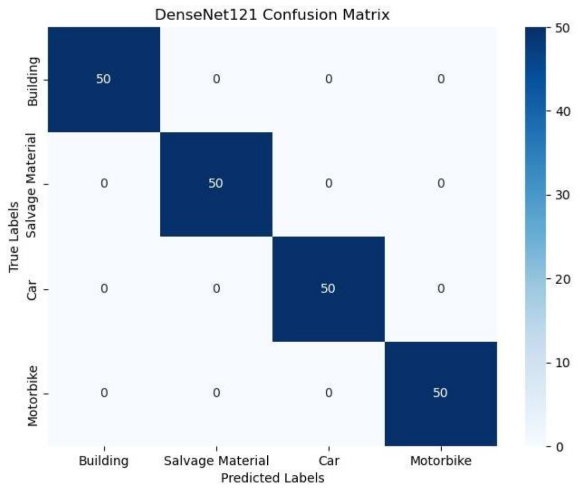

Figure 5. Confusion Matrix from DenseNet121

Figure 5 presents the classification performance of the DenseNet121 model, which achieved 100% accuracy across all categories, with every category correctly predicted without any misclassifications.

Figure 6 presents the classification performance of the MobileNetV2 model, which achieved 100% accuracy across all categories, with every category was correctly predicted without any misclassifications.

Figure 7 presents the classification performance of the InceptionV3 model, which achieved 100% accuracy across all categories, with every category correctly predicted without any misclassifications.

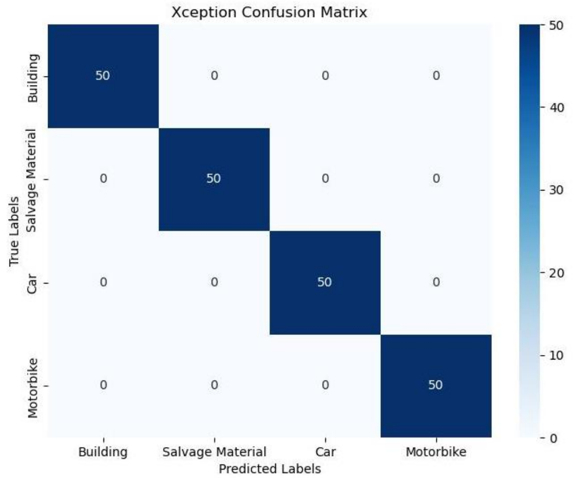

Figure 6. Confusion Matrix from Xception

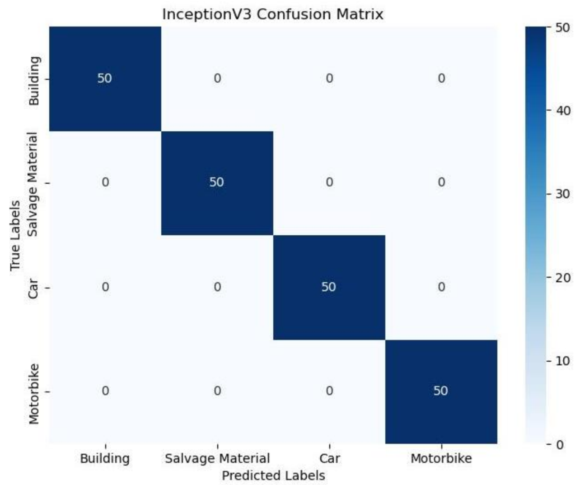

Figure 7. Confusion Matrix from InceptionV3

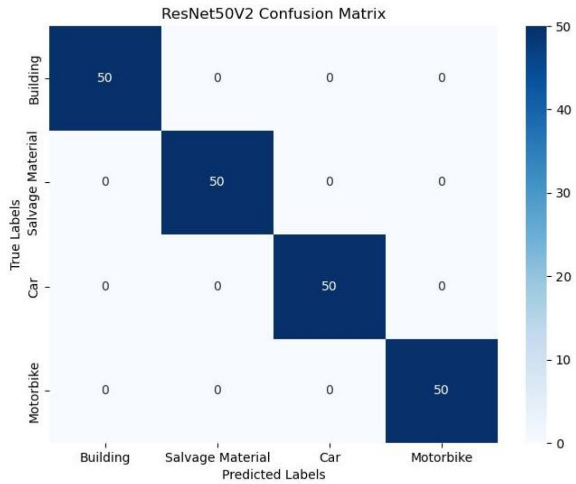

Figure 8. Confusion Matrix from ResNet50V2

Figure 8 presents the classification performance of the ResNet50V2 model, which achieved 100% accuracy across all categories, with every category correctly predicted without any misclassifications.

Table 5. Results from Previous Research [5]

	Classification Result			
	Buildings	Salvage Materials	Cars	Motorbikes
Buildings	50	0	0	0
Salvage Materials	3	44	1	2
Cars	0	0	50	0
Motorbikes	0	1	0	49

Figure 2 to 8 presents the confusion matrix for each transfer learning model. As most transfer learning models achieved 100% accuracy, the confusion matrices show perfect classification across all categories. The only exception is EfficientNetV2B0 (Figure 4), which failed to classify objects outside of the buildings category. In comparison to the previous research results in Table 5, the base model frequently misclassified salvage materials. However, this issue was effectively solved by applying transfer learning methods, which improved the classification of these challenging objects.

Table 4 clearly shows that most transfer learning models are highly effective in being implemented for auction object classification, outperforming the custom baseline CNN model in both accuracy and computational efficiency. NASNetMobile, DenseNet121, Xception, InceptionV3, and ResNet50V2 not only achieved 100% in training and validation accuracy but also demonstrated strong generalisation capabilities, as reflected in their low validation losses and high accuracy. Additionally, MobileNetV2 stands out for its smaller size (11.1 MB) and quick training, making it ideal for deployment in resource-constrained environments, such as mobile applications, where a smaller model is needed for seamless integration.

4. Conclusions

This research evaluated the performance of seven transfer learning models: MobileNetV2, NASNetMobile, EfficientNetV2B0, DenseNet121, Xception, InceptionV3, and ResNet50V2, on the Indonesian Auction Object Dataset. The results demonstrated that most transfer learning models significantly outperformed the baseline model, with MobileNetV2, NASNetMobile, DenseNet121, Xception, InceptionV3, and ResNet50V2 all achieving 100% in validation accuracy. Among these, MobileNetV2 and NASNetMobile stand out for their accuracy, low validation loss, and smaller model size. Moreover, MobileNetV2 is the best choice, offering a higher accuracy and the lowest model size, making it suitable for deployment in mobile applications. Moreover, EfficientNetV2B0's poor performance confirms that, although it may perform well in other domains, it is not suited for the Indonesian Auction dataset, highlighting the need to select architectures that can perform well in specific tasks carefully.

This research confirms that transfer learning improves auction object classification accuracy, but not all architectures are suitable, as EfficientNetV2B0 proved ineffective for the Indonesian auction dataset. Future research should explore the potential of other emerging transfer learning architectures, apply further fine-tuning strategies, and investigate the impact on larger, more diverse datasets to assess generalisability and robustness across different auction objects.

References

[1] Kementerian Keuangan, “Peraturan Menteri Keuangan Nomor 122 Tahun 2023 tentang Petunjuk Pelaksanaan Lelang,” 2023. Accessed: Sep. 02, 2024. [Online]. Available: https://jdih.kemenkeu.go.id/in/dokumen/peraturan/fe1bd404-e081-4ccd-5dd1-08dbf6b83bff

[2] Direktorat Jenderal Kekayaan Negara, *Laporan Kinerja Direktorat Lelang Tahun 2023*. Direktorat Jenderal Kekayaan Negara, 2024. Accessed: Sep. 02, 2024. [Online]. Available: https://lelang.go.id/assets/lap-capkin-91dcbd20.pdf

[3] A. Simanjuntak, S. B. Siahaan, D. Megaria Elisabeth, Y. Welly, and D. D. Hutagalung, “The Effect of Auction Fee, Administration Fees of State Receivables Management and Management of State-Owned Assets Against Actualization of State Non-Tax Revenue on Kantor Pelayanan Kekayaan Negara dan Lelang Medan Period 2019-2021,” *Indonesian Journal of Banking and Financial Technology*, vol. 1, no. 4, pp. 315–328, Oct. 2023, doi: 10.55927/fintech.v1i4.6581.

[4] A. S. Sulaeman and L. P. C. D. Mayasari, “PENERIMAAN BEA LELANG: EMPIRICAL FACTOR ANALYSIS PADA PROSES LELANG ASET NEGARA DI INDONESIA,” *EKUITAS (Jurnal Ekonomi dan Keuangan)*, vol. 7, no. 1, pp. 21–41, Mar. 2023, doi: 10.24034/j25485024.y2023.v7.i1.5085.

[5] G. M. M. Sujak and H. N. Rofiq, “Pemanfaatan Artificial Intelligence untuk Optimalisasi PNBP: Studi Kasus Bea Lelang pada Lelang Indonesia,” *Indonesian Treasury Review Jurnal Perbendaharaan Keuangan Negara dan Kebijakan Publik*, vol. 9, no. 2, pp. 112–124, Jun. 2024, doi: 10.33105/itrev.v9i2.873.

[6] S. Mastura Ajeng, D. K. Sugiharti, and A. Cahyadini, “Perlindungan Hukum Terhadap Pemenang Lelang Eksekusi Hak Tanggungan Atas Ketidaksesuaian Objeknya yang Dilakukan Melalui E-Auction,” *INNOVATIVE: Journal Of Social Science Research*, vol. 4, no. 4, pp. 475–487, Jul. 2024, doi: 10.31004/innovative.v4i4.12845.

[7] J. Jareño, G. Bárcena-González, J. Castro-Gutiérrez, R. Cabrera-Castro, and P. L. Galindo, “Enhancing Fish Auction with Deep Learning and Computer Vision: Automated Caliber and Species Classification,” *Fishes*, vol. 9, no. 4, p. 133, Apr. 2024, doi: 10.3390/fishes9040133.

[8] J. Bobulski and S. Szymoniak, “Deep Learning Method for Classifying Items into Categories for Dutch Auctions,” *Computer Assisted Methods in Engineering and Science*, vol. 31, no. 1, pp. 67–79, Jan. 2024, doi: 10.24423/cames.2024.979.

[9] Z. Ardalan and V. Subbian, “Transfer Learning Approaches for Neuroimaging Analysis: A Scoping Review,” *Front Artif Intell*, vol. 5, Feb. 2022, doi: 10.3389/frai.2022.780405.

[10] G. Kaur, N. Sharma, R. Chauhan, K. Joshi, and R. Gupta, “Rice Leaf Disease Detection Using InceptionV3 Transfer Learning Model,” in *2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, IEEE, Dec. 2023, pp. 1–6. doi: 10.1109/SMARTGENCON60755.2023.10442321.

[11] A. S. Paymode and V. B. Malode, “Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG,” *Artificial Intelligence in Agriculture*, vol. 6, pp. 23–33, 2022, doi: 10.1016/j.aiia.2021.12.002.

[12] S. Niu, Y. Liu, J. Wang, and H. Song, “A Decade Survey of Transfer Learning (2010–2020),” *IEEE Transactions on*

Artificial Intelligence, vol. 1, no. 2, pp. 151–166, Oct. 2020, doi: 10.1109/TAI.2021.3054609.

[13] B. Sailaja and T. VenuGopal, "Brain MRI Image Classification and Analysis Using Modified ResNet50V2 for Parkinson's Disease Detection," *SN Comput Sci*, vol. 4, no. 6, p. 854, Nov. 2023, doi: 10.1007/s42979-023-02313-y.

[14] J. Chen, D. Zhang, M. Suzauddola, Y. A. Nanehkaran, and Y. Sun, "Identification of plant disease images via a squeeze-and-excitation MobileNet model and twice transfer learning," *IET Image Process*, vol. 15, no. 5, pp. 1115–1127, Apr. 2021, doi: 10.1049/ipr2.12090.

[15] C. C. Foong, G. K. Meng, and L. L. Tze, "Convolutional Neural Network based Rotten Fruit Detection using ResNet50," in *2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC)*, IEEE, Aug. 2021, pp. 75–80. doi: 10.1109/ICSGRC53186.2021.9515280.

[16] "Keras Application Benchmark." Accessed: Sep. 02, 2024. [Online]. Available: https://keras.io/api/applications/

[17] F. J. J. Joseph, S. Nonsiri, and M. Annop, "Keras and TensorFlow: A Hands-On Experience," in *Advanced Deep Learning for Engineers and Scientists: A Practical Approach*, R. and A. S. A. and K. G. R. Prakash Kolla Bhanu and Kannan, Ed., Cham: Springer International Publishing, 2021, pp. 85–111. doi: 10.1007/978-3-030-66519-7_4.

[18] N. Nayman, A. Golbert, A. Noy, and L. Zelnik-Manor, "Diverse Imagenet Models Transfer Better," in *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, IEEE, Jan. 2024, pp. 1903–1914. doi: 10.1109/WACV57701.2024.00192.

[19] A. Chatterjee, J. Mukherjee, and P. P. Das, "ImageNet Classification Using WordNet Hierarchy," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 4, pp. 1718–1727, Apr. 2024, doi: 10.1109/TAI.2023.3297086.

[20] R. D. Bhosale and D. M. Yadav, "Analysis of EfficientNet Family Models by Retraining for Tuberculosis Detection from Chest X-Ray Images," in *2023 Second International Conference on Informatics (ICI)*, IEEE, Nov. 2023, pp. 1–6. doi: 10.1109/ICI60088.2023.10420853.

[21] T. Ibtekar, Md. R. Haque, and A. Y. Srizon, "Comparative Study Between ResNet And EfficientNet Family For Classification of Leukemia," in *2023 26th International Conference on Computer and Information Technology (ICCIT)*, IEEE, Dec. 2023, pp. 1–6. doi: 10.1109/ICCIT60459.2023.10441438.

[22] A. D. Saputra, D. Hindarto, and H. Santoso, "Disease Classification on Rice Leaves using DenseNet121, DenseNet169, DenseNet201," *Sinkron*, vol. 8, no. 1, pp. 48–55, Jan. 2023, doi: 10.33395/sinkron.v8i1.11906.

[23] L. Yong, L. Ma, D. Sun, and L. Du, "Application of MobileNetV2 to waste classification," *PLoS One*, vol. 18, no. 3, p. e0282336, Mar. 2023, doi: 10.1371/journal.pone.0282336.

[24] I. Naskinova, "Transfer learning with NASNet-Mobile for Pneumonia X-ray classification," *Asian-European Journal of Mathematics*, vol. 16, no. 01, Jan. 2023, doi: 10.1142/S1793557122502400.

[25] M. F. Fayyad and Mustakim, "Application of AlexNet, EfficientNetV2B0, and VGG19 with Explainable AI for Cataract and Glaucoma Image Classification," in *2024 International Electronics Symposium (IES)*, IEEE, Aug. 2024, pp. 406–412. doi: 10.1109/IES63037.2024.10665856.

[26] N. Asaad Zebari, A. A. H. Alkurdi, R. B. Marqas, and M. Shamal Salih, "Enhancing Brain Tumor Classification with Data Augmentation and DenseNet121," *Academic Journal of Nawroz University*, vol. 12, no. 4, pp. 323–334, Oct. 2023, doi: 10.25007/ajnu.v12n4a1985.

[27] F. Salim, F. Saeed, S. Basurra, S. N. Qasem, and T. Al-Hadhrami, "DenseNet-201 and Xception Pre-Trained Deep Learning Models for Fruit Recognition," *Electronics (Basel)*, vol. 12, no. 14, p. 3132, Jul. 2023, doi: 10.3390/electronics12143132.

[28] T. H. Jaware, H. Kasturiwale, R. Thakur, M. D. Jakhete, M. Chavan, and M. Rane, "Deep Learning Model for Colon Cancer Classification using InceptionV3," *Journal of Electrical Systems*, vol. 20, no. 4s, pp. 132–139, Apr. 2024, doi: 10.52783/jes.1862.

[29] B. Sailaja and T. VenuGopal, "Brain MRI Image Classification and Analysis Using Modified ResNet50V2 for Parkinson's Disease Detection," *SN Comput Sci*, vol. 4, no. 6, p. 854, Nov. 2023, doi: 10.1007/s42979-023-02313-y.

[30] M. Yeung, E. Sala, C.-B. Schönlieb, and L. Rundo, "Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102026, Jan. 2022, doi: 10.1016/j.compmedimag.2021.102026.

[31] B. Baker *et al.*, "Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos," in *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., Curran Associates, Inc., 2022, pp. 24639–24654. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2022/file/9c7008aff45b5d8f0973b23e1a22ada0-Paper-Conference.pdf

