




ORIGINAL

Pre-trained CNNs: Evaluating Emergency Vehicle Image Classification

CNN preentrenadas: Evaluación de la clasificación de imágenes de vehículos de emergencia

Ali Omari Alaoui¹ , Omaima El Bahi² , Mohamed Rida Fethi¹ , Othmane Farhaoui¹ , Ahmad El Allaoui¹ ,
Yousef Farhaoui¹ 

¹L-STI, T-IDMS, FST Errachidia/Moulay Ismail University, Meknes, Morocco.

²L-MAIS, T-MIS, FST Errachidia/Moulay Ismail University, Meknes, Morocco.

Cite as: Omari Alaoui A, El Bahi O, Rida Fethi M, Farhaoui O, El Allaoui A, Farhaoui Y. Pre-trained CNNs: Evaluating Emergency Vehicle Image Classification. Data and Metadata. 2023;2:153. <https://doi.org/10.56294/dm2023153>


Submitted: 06-08-2023

Revised: 11-10-2023

Accepted: 29-12-2023

Published: 30-12-2023

Editor: Prof. Dr. Javier González Argote 

Guest Editor: Yousef Farhaoui 

Note: Paper presented at the International Conference on Artificial Intelligence and Smart Environments (ICAISE'2023).

ABSTRACT

In this paper, we aim to provide a comprehensive analysis of image classification, specifically in the context of emergency vehicle classification. We have conducted an in-depth investigation, exploring the effectiveness of six pre-trained Convolutional Neural Network (CNN) models. These models, namely VGG19, VGG16, MobileNetV3Large, MobileNetV3Small, MobileNetV2, and MobileNetV1, have been thoroughly examined and evaluated within the domain of emergency vehicle classification. The research methodology utilized in this study is carefully designed with a systematic approach. It includes the thorough preparation of datasets, deliberate modifications to the model architecture, careful selection of layer operations, and fine-tuning of the model compilation. To gain a comprehensive understanding of the performance, we conducted a detailed series of experiments. We analyzed nuanced performance metrics such as accuracy, loss, and training time, considering important factors in the evaluation process. The results obtained from this study provide a comprehensive understanding of the advantages and disadvantages of each model. Moreover, they emphasize the crucial significance of carefully choosing a suitable pre-trained Convolutional Neural Network (CNN) model for image classification tasks. Essentially, this article provides a comprehensive overview of image classification, highlighting the crucial significance of pre-trained CNN models in achieving precise outcomes, especially in the demanding field of emergency vehicle classification.

Keywords: Pre-trained CNN Models; Image Classification; Computer Vision; Emergency Vehicles.

RESUMEN

En este artículo pretendemos ofrecer un análisis exhaustivo de la clasificación de imágenes, concretamente en el contexto de la clasificación de vehículos de emergencia. Hemos llevado a cabo una investigación en profundidad, explorando la eficacia de seis modelos preentrenados de redes neuronales convolucionales (CNN). Estos modelos, a saber, VGG19, VGG16, MobileNetV3Large, MobileNetV3Small, MobileNetV2 y MobileNetV1, se han examinado y evaluado exhaustivamente en el ámbito de la clasificación de vehículos de emergencia. La metodología de investigación utilizada en este estudio está cuidadosamente diseñada con un enfoque sistemático. Incluye la preparación minuciosa de conjuntos de datos, modificaciones deliberadas de la arquitectura del modelo, una cuidadosa selección de las operaciones de capa y el ajuste fino de la compilación del modelo. Para conocer a fondo el rendimiento, realizamos una serie detallada de experimentos. Analizamos métricas de rendimiento matizadas, como la precisión, la pérdida y el tiempo de entrenamiento, teniendo en cuenta factores importantes en el proceso de evaluación. Los resultados

obtenidos en este estudio proporcionan una comprensión exhaustiva de las ventajas e inconvenientes de cada modelo. Además, ponen de relieve la importancia crucial de elegir cuidadosamente un modelo de red neuronal convolucional (CNN) preentrenado adecuado para las tareas de clasificación de imágenes. Esencialmente, este artículo proporciona una visión global de la clasificación de imágenes, destacando la importancia crucial de los modelos CNN preentrenados para lograr resultados precisos, especialmente en el exigente campo de la clasificación de vehículos de emergencia.

Palabras clave: Modelos CNN Preentrenados; Clasificación de Imágenes; Visión por Computador; Vehículos de Emergencia.

INTRODUCTION

In the rapidly changing field of computer vision,^(1,2) image classification^(3,4,5) plays a crucial role with extensive practical applications. It is used in various domains such as object recognition,^(6,7) autonomous driving, medical imaging, and surveillance systems, and has become an integral part of real-world systems.⁽⁸⁾ Convolutional Neural Networks (CNNs),⁽⁹⁾ a type of deep learning model, have gained significant popularity as the go-to method for image classification tasks. These models have the remarkable ability to automatically learn hierarchical representations from raw pixel data. They excel at detecting complex patterns and extracting features that are crucial for achieving high accuracy in classification tasks.^(10,11)

The introduction of pre-trained CNN architectures, such as those honed on extensive image datasets like ImageNet,⁽¹²⁾ has significantly accelerated progress in image classification.

These pre-trained models have already learned valuable features from a large dataset, allowing them to quickly identify patterns and objects in new images.⁽⁹⁾ Additionally, the concept of transfer learning has further enhanced their efficacy. Through fine-tuning, pre-trained models can be adapted to specific datasets, enabling them to perform even better on targeted tasks. This approach not only reduces training time but also produces impressive performance outcomes by leveraging the knowledge gained from the initial training.

This study significantly broadens its scope by incorporating the data of emergency vehicles into the analysis, acknowledging the crucial role they play in real-world situations.^(13,14,15) The classification of emergency vehicles, due to its central importance, holds significance in various practical applications, including traffic management, public safety, and urban planning.

The main objective of this article is to thoroughly compare and analyze the performance nuances of different CNN models trained on distinct datasets. Specifically, we will focus on the VGG19,⁽¹⁶⁾ VGG16,⁽¹⁷⁾ MobileNetV3Large, MobileNetV3Small,⁽¹⁸⁾ MobileNetV2,⁽¹⁹⁾ and MobileNetV1⁽¹²⁾ models in the extensive field of image classification. Our aim is to provide a detailed examination of their capabilities and understand how they perform in this domain. Our primary focus is on detecting emergency vehicles, as we recognize the vital role they play. We are conducting a thorough analysis to provide in-depth understanding of the strengths and limitations of each vehicle classification model. Our aim is to offer valuable insights specifically in the context of classifying emergency vehicles, considering the dynamic nature of this task.

This research aims to serve as a valuable resource to help researchers and practitioners make informed decisions when choosing a pre-trained Convolutional Neural Network (CNN) model for image classification tasks, specifically related to the detection of emergency vehicles.⁽¹³⁾ The focus is on selecting a model that is well-suited for the complex and dynamic environments in which these vehicles operate.

Related Works

In the dynamic realm of image recognition, the impact of deep convolutional neural networks (CNNs) is indisputably profound, as depicted in figure 1. This figure serves as a testament to the intricate architecture that underlies the unparalleled capabilities of CNNs, especially when addressing large-scale visual recognition challenges. It acts as a visual guide, unraveling the intricacies that reshape the landscape of image recognition.

Unraveling the intricate architecture of Convolutional Neural Networks (CNNs), the backbone of transformative capabilities in image recognition.

Pioneering this trajectory, Simonyan and Zisserman laid the foundation with the introduction of Very Deep Convolutional Networks (VGG).⁽¹⁶⁾ Figure 2 provides a closer look at the milestone that is VGG16, emphasizing the increased network depth brought forth by Szegedy et al.⁽¹⁷⁾ This figure showcases how such depth contributes to heightened accuracy in the intricate tasks of image recognition.

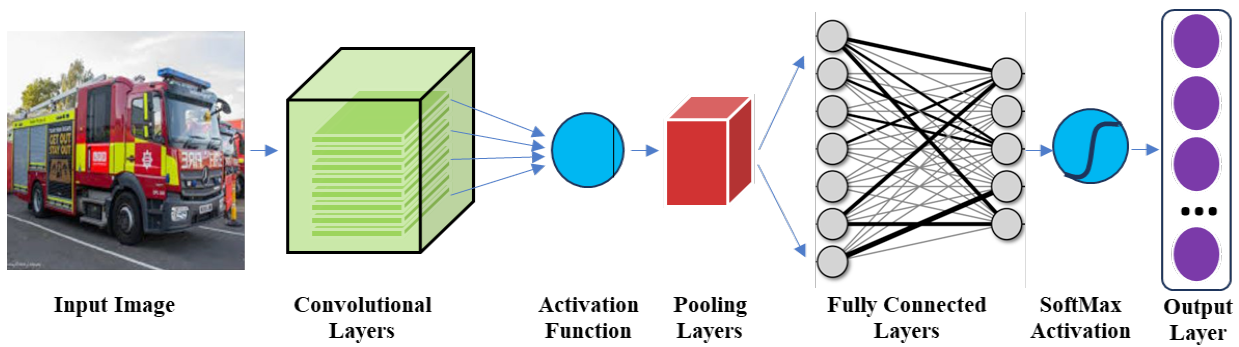


Figure 1. Convolutional Neural Networks Architecture

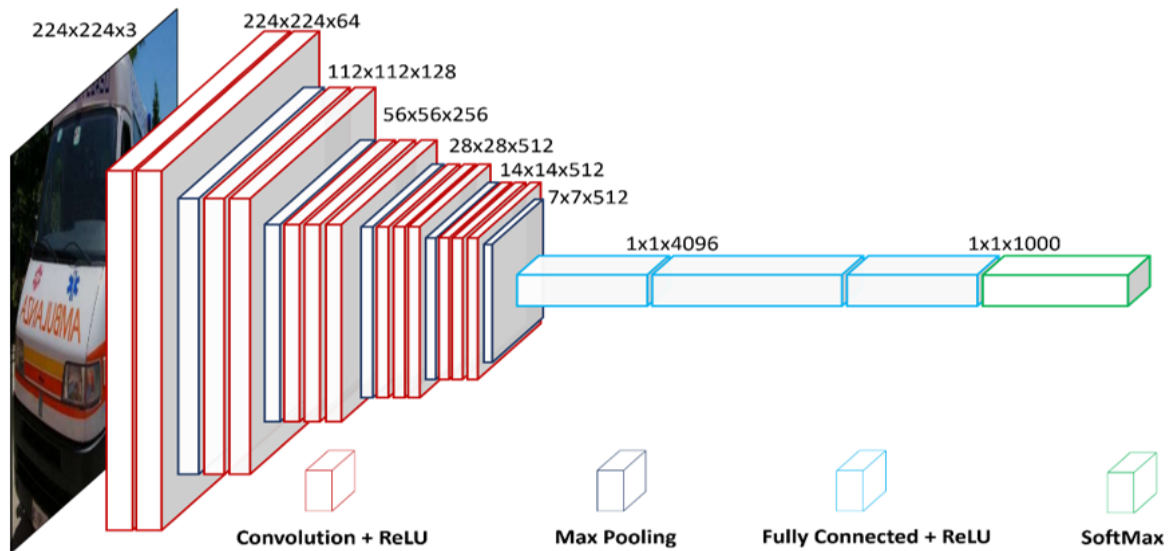


Figure 2. VGG16 Architecture

Continuing the exploration of VGG architecture, figure 3 shifts focus to VGG19. Building upon the concepts introduced in VGG16, this figure offers a detailed examination of the architecture's nuances. The increased network depth in VGG19 is highlighted, illustrating its pivotal role in achieving superior accuracy. The VGG architecture incorporates Rectified Linear Unit (ReLU)^(20,21) activation functions and SoftMax⁽²²⁾ for efficient feature extraction and final class probability computation.

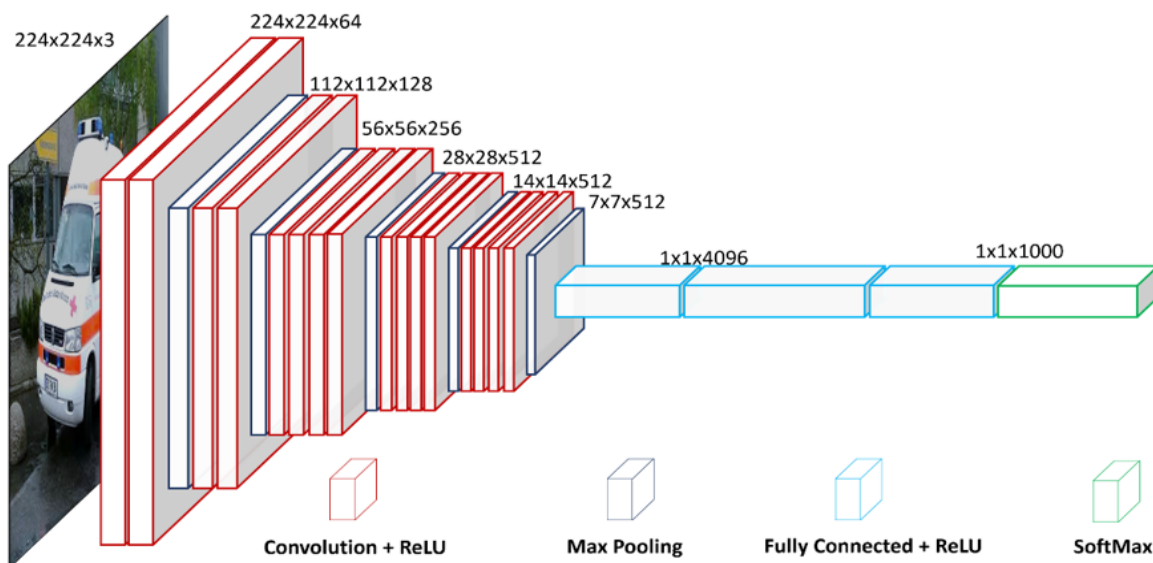


Figure 3. VGG19 Architecture

Figure 3 focuses on VGG19, showcasing increased network depth and its pivotal role in achieving superior accuracy.

The MobileNet series, born from the innovation of Howard et al.⁽²⁰⁾ represents an evolutionary stride in efficient CNN architectures tailored explicitly for mobile vision applications. Figure 4 visually represents the architecture of MobileNetV1, emphasizing its contributions to efficient mobile vision. The innovative features introduced in MobileNetV1 set the stage for subsequent advancements in mobile-friendly CNN architectures.

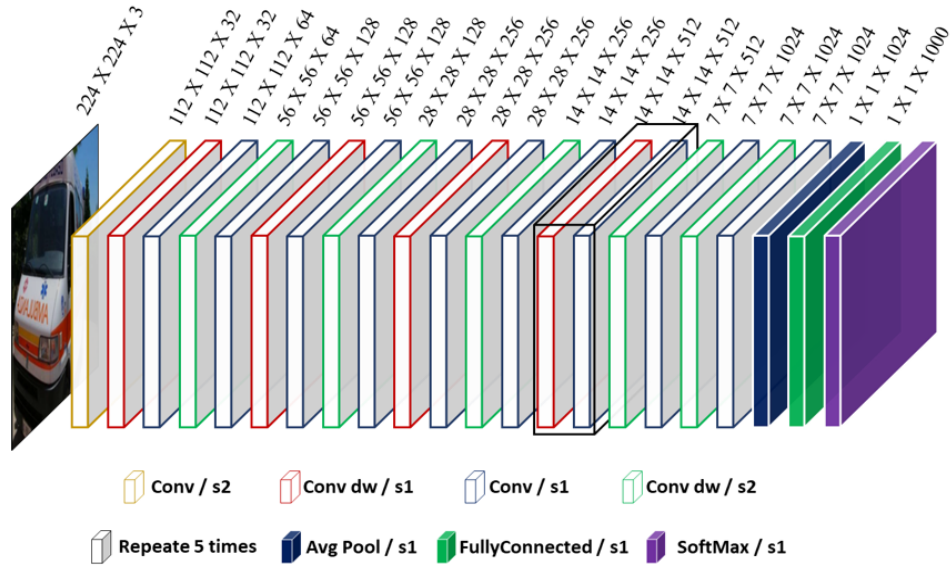


Figure 4. MobileNet Architecture

This figure presents a visual representation of an evolutionary stride in efficient CNN architectures tailored explicitly for mobile vision applications.

Figure 5 delves into the advancements brought about by MobileNetV2, introduced by Sandler et al. This figure provides insights into the architecture's features, including inverted residuals and linear bottlenecks, designed to enhance feature extraction efficiency. MobileNetV2 stands as a notable refinement in the landscape of efficient CNNs.

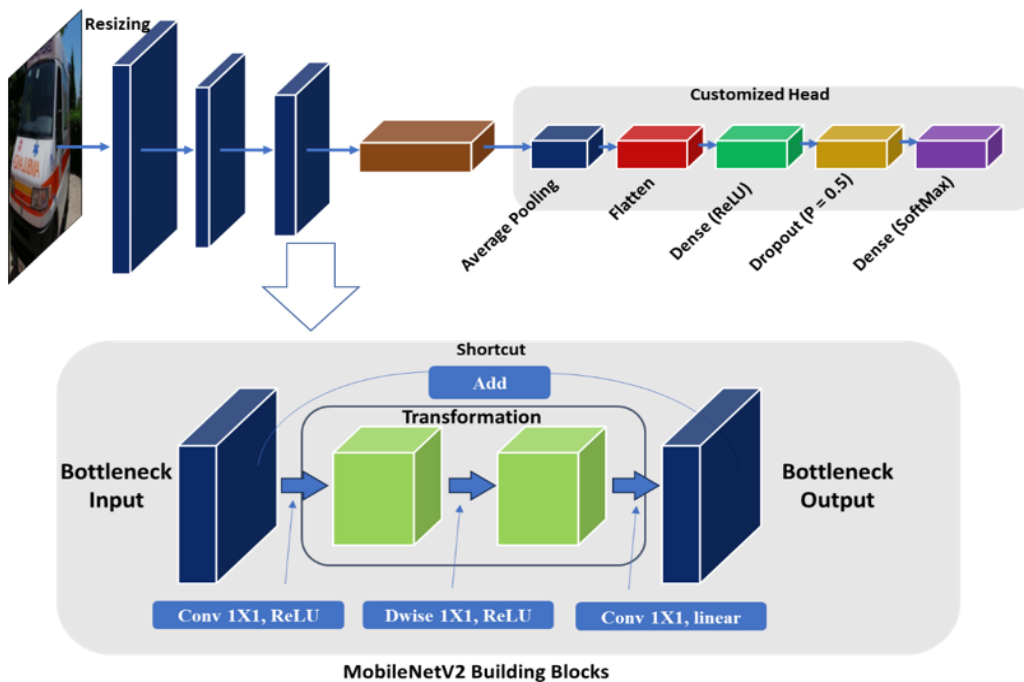


Figure 5. MobileNetV2 Architecture

Figure 6 introduces MobileNetV3, showcasing further advancements in efficient CNN architectures. This figure delves into the intricacies of MobileNetV3,⁽¹⁸⁾ underlining its contributions to the evolving field of CNNs. MobileNetV3 exemplifies ongoing efforts to optimize CNN architectures for efficiency and performance.

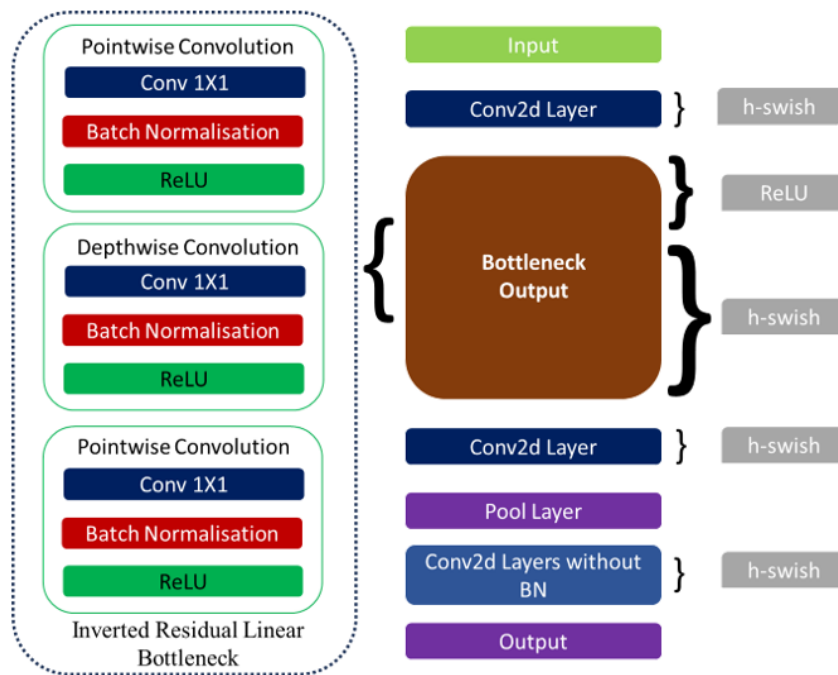


Figure 6. MobileNetV3 Architecture

In figure 7, a comprehensive analysis contrasting MobileNetV2 with MobileNetV3 reveals notable distinctions and advancements in their respective architectures. This visual examination offers valuable insights into the evolutionary changes of efficient Convolutional Neural Networks (CNNs), providing a deeper understanding of their strengths and capabilities. Specifically, the expansion layer in MobileNetV3 is reimagined, as illustrated in figure 7, building upon the original design from MobileNetV2. This comparison visually highlights the modifications in the last stage, showcasing the evolution in design and functionality.

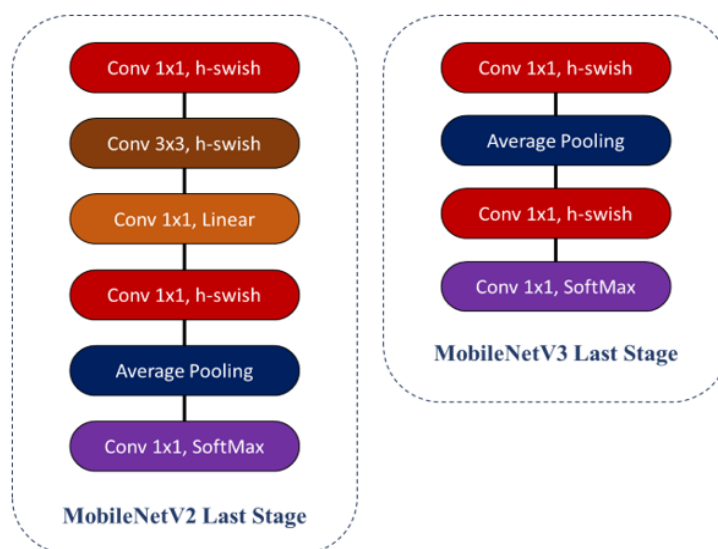


Figure 7. A comprehensive analysis contrasting MobileNetV2 with MobileNetV3 reveals notable distinctions and advancements in their respective architectures

These seminal works, coupled with techniques such as large minibatch stochastic gradient descent (SGD) training pioneered by Goyal et al.⁽⁶⁾ and the influential platform of the ImageNet Large Scale Visual Recognition Challenge,^(23,24,25,26) collectively represent a profound leap forward in the domain of image classification through the application of deep CNNs.

This comprehensive ecosystem of advancements has not only broadened our understanding of deep learning but has also significantly propelled the field toward new frontiers of exploration and achievement.

METHODOLOGY

Dataset Curation and Preprocessing

In the initial phase of our comprehensive study, our focus was on crafting a meticulously curated dataset for the evaluation of pre-trained Convolutional Neural Network (CNN) models in the domain of emergency vehicle classification.

This dataset, consisting of 2312 diverse images spanning various classes, underwent essential preprocessing steps, as illustrated in figure 8. A crucial aspect involved resizing all images to a standardized dimension of (224, 224). Leveraging the `xml.etree.ElementTree` library,⁽²⁴⁾ we parsed XML files containing image annotations, extracting vital information such as image paths, class labels, and bounding box coordinates. This foundational step ensured the dataset's readiness for subsequent analysis and model evaluation.

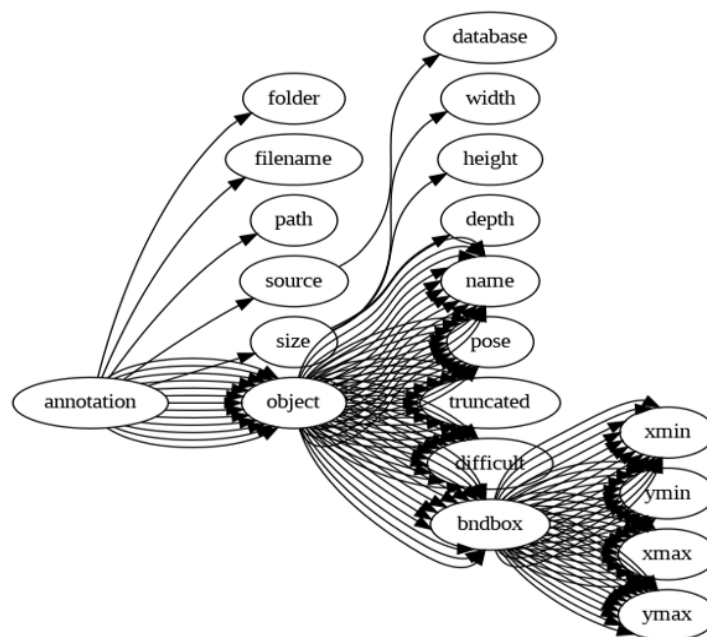


Figure 8. Xml Tree

GPU-Accelerated Training on Google Colab

To harness the computational power required for efficient training, we strategically employed Google Colab's GPUs. This not only expedited the training process but also allowed us to leverage the capabilities of these GPUs to their fullest potential. The dataset was intelligently partitioned into an 80 % training set and a 20 % validation set. This partitioning strategy ensured substantial training data while maintaining a separate subset for unbiased model performance assessment and validation, as summarized in table 1.

| Table 1. GPU Performance and Configuration Details | |
|----------------------------------------------------|------------|
| Property | Details |
| GPU Name | Tesla T4 |
| Driver Version | 525.105.17 |
| CUDA Version | 12.0 |
| Performance | P8 |
| Power Usage | 9W / 70W |

Customization of Pre-trained Models

Customizing the pre-trained models for our specific emergency vehicle classification task was a pivotal step. We loaded these models sans their top classification layers, opting to leverage the convolutional base responsible for extracting meaningful features from images. By excluding the top layers, originally designed for specific classification tasks, we tailored the models to suit our image classification problem. Additional custom

classification layers were introduced atop the base models, responsible for learning class probabilities based on the extracted features, as depicted in figure 9.

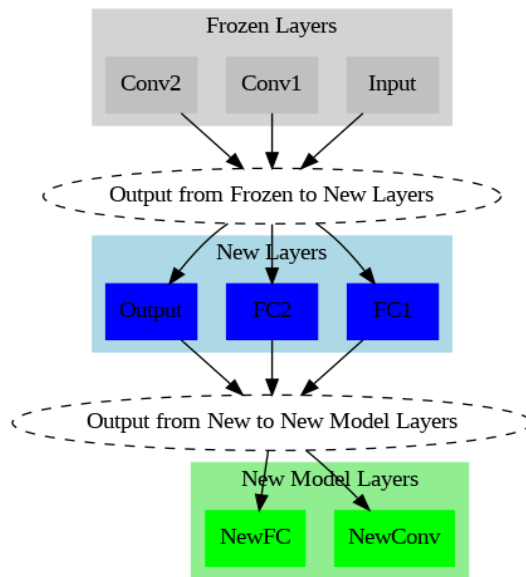


Figure 9. Customization of Pre-trained Models for Emergency Vehicle Classification

Model Training Strategy

In our pursuit of an effective training strategy, we decided to freeze specific layers in the models. This strategic decision ensured that the weights of these frozen layers remained unchanged during training, preserving the valuable pre-trained representations they had acquired. For our experiments, we selectively froze all layers except for the last four, striking a harmonious balance between capitalizing on pre-trained knowledge and facilitating adaptation to our specific classification tasks. Model compilation involved the use of the Adam optimizer with a learning rate of 0,0001, a well-established optimizer for deep learning models. The chosen loss function was sparse categorical cross-entropy, aptly suited for multi-class classification problems, as illustrated in figure 10.

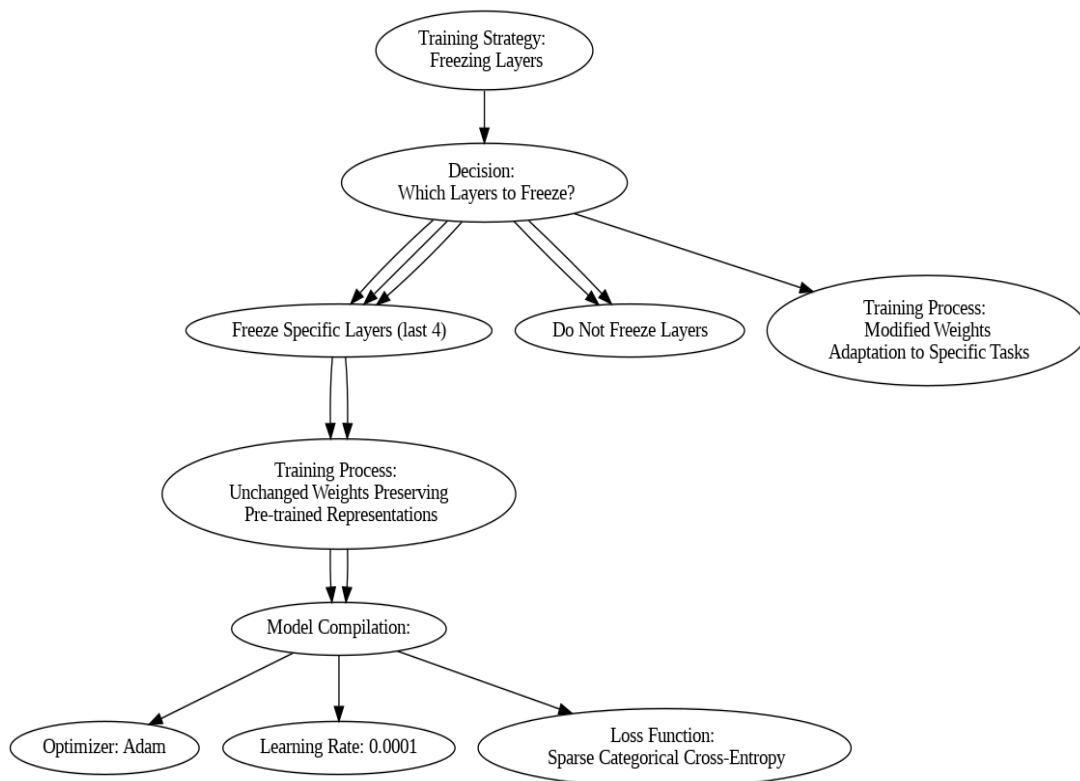


Figure 10. Model Training Strategy

RESULTS

In this pivotal section, we unveil the outcomes derived from our meticulous comparative analysis of various CNN models' object detection performance in the context of emergency vehicle classification. Our scrutiny delves deep into the intricate interplay of accuracy and speed exhibited by models such as MobileNetV1, VGG19, VGG16, MobileNetV3Large, MobileNetV3Small, and MobileNetV2, unraveling the nuanced challenges and advantages associated with the task at hand.

Model Comparison

Let us first scrutinize the results encapsulated in table 2, a comprehensive tableau capturing the essence of our study. The models, including MobileNetV1, VGG19, VGG16, MobileNetV3Large, MobileNetV3Small, and MobileNetV2, are appraised based on accuracy, loss, and training time metrics. Each entry in this table represents a snapshot of the models' prowess in handling the intricacies of emergency vehicle detection.

| Model | Accuracy | Loss | Training time |
|------------------|----------|--------|---------------|
| MobileNetV1 | 92,21 | 0,2688 | 32,47 |
| VGG19 | 87,01 | 0,4864 | 80,58 |
| VGG16 | 89,61 | 0,4472 | 69,00 |
| MobileNetV3Large | 58,44 | 0,8772 | 36,99 |
| MobileNetV3Small | 38,96 | 1,0491 | 33,59 |
| MobileNetV2 | 90,91 | 0,2766 | 34,30 |

This table provides a snapshot of pre-trained CNN models' comparative performance in emergency vehicle detection. Metrics include Accuracy (%), Loss, and Training Time (s). MobileNetV1 stands out with an impressive 92,21 % Accuracy, a low Loss of 0,2688, and a swift Training Time of 32,47 seconds. Notable variations across VGG19, VGG16, MobileNetV3Large, MobileNetV3Small, and MobileNetV2 highlight their distinct capabilities in addressing the challenges of object detection.

Visualizing the Learning Journey

Figures 12, 13, and 14 serve as visual testaments to the learning progress of the models. Each graph encapsulates the dynamic trajectory of accuracy against training time, unraveling the distinct learning curves etched by VGG19, VGG16, MobileNetV3Large, MobileNetV3Small, MobileNetV2, and MobileNetV1. These visualizations offer a nuanced perspective, illustrating how each model refines its understanding over the training epochs.

Interpreting the Performance

Diving deeper into the numerical specifics, the VGG19 model demonstrates a commendable accuracy of 87,01 % within a training time of 80,58 seconds. VGG16, not far behind, attains an accuracy of 89,61 % in 69 seconds. Meanwhile, MobileNetV3Large and MobileNetV3Small carve unique trajectories, securing accuracies of 58,44 % and 38,96 %, respectively, within their distinct training durations. Noteworthy is MobileNetV2, boasting a robust 90,91 % accuracy in 34,30 seconds. However, the crown of efficiency rests on MobileNetV1, achieving an astounding 92,21 % accuracy in a mere 32,47 seconds.

The Canvas of Insights

These results, meticulously curated and presented, offer a panoramic view of each model's performance landscape. The interplay of accuracy and training time delineates a nuanced narrative, unveiling the distinctive strengths and limitations of the models in the context of emergency vehicle classification tasks. This contribution to the scientific discourse not only enriches our understanding of these pre-trained CNN models but also serves as a beacon guiding researchers and practitioners in selecting optimal models for image classification challenges, particularly those pertaining to emergency vehicle detection.

Visual Insights

Adding a layer of visual richness to our discourse, figures 12, 13, and 14 encapsulate the essence of VGG19 and VGG16, MobileNetV3Large and MobileNetV3Small, MobileNetV2 and MobileNetV1. These graphs traverse the intricate terrain of accuracy and loss, providing a visual symphony that harmonizes with our quantitative analysis. Figure 4, the crowning jewel, intricately weaves together accuracy, loss, and training time, with MobileNetV1 emerging as the virtuoso of our ensemble, showcasing not only the highest accuracy but also a finely tuned balance between complexity and training efficiency.

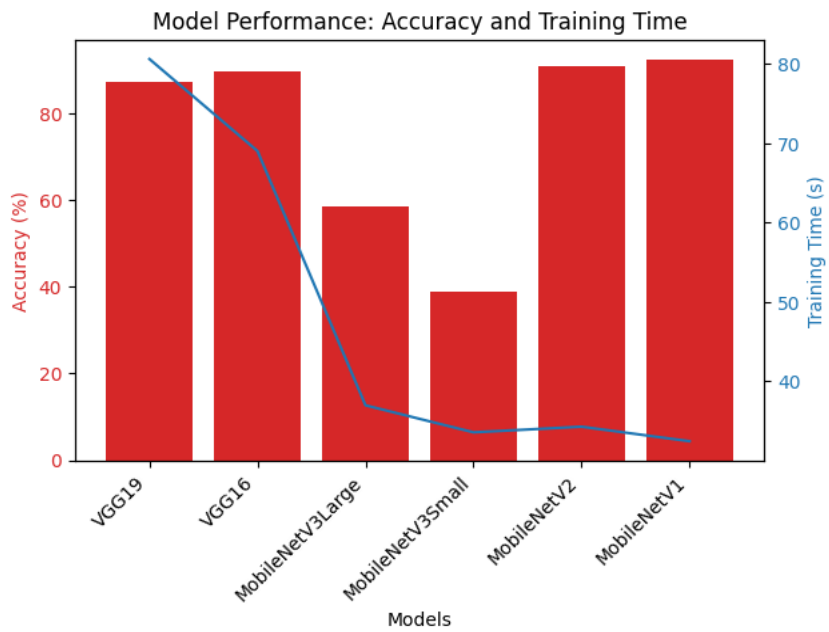


Figure 11. visually compares model performance in terms of accuracy, loss, and training time, offering concise insights into their distinctive characteristics for efficient emergency vehicle classification

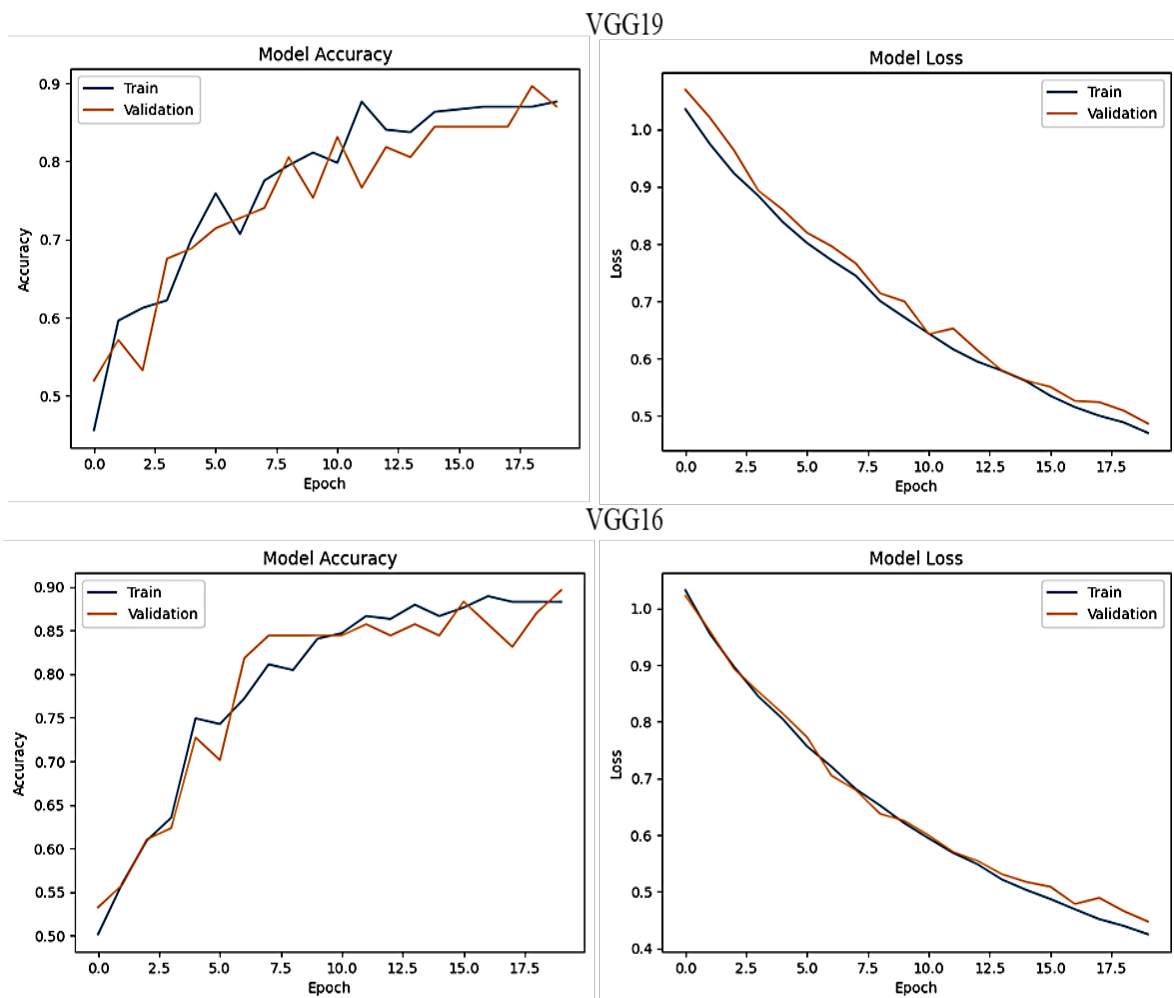


Figure 12. depicts the accuracy and loss dynamics of VGG19 and VGG16 pre-trained CNN models for emergency vehicle classification, offering a concise yet informative overview of their performance characteristics

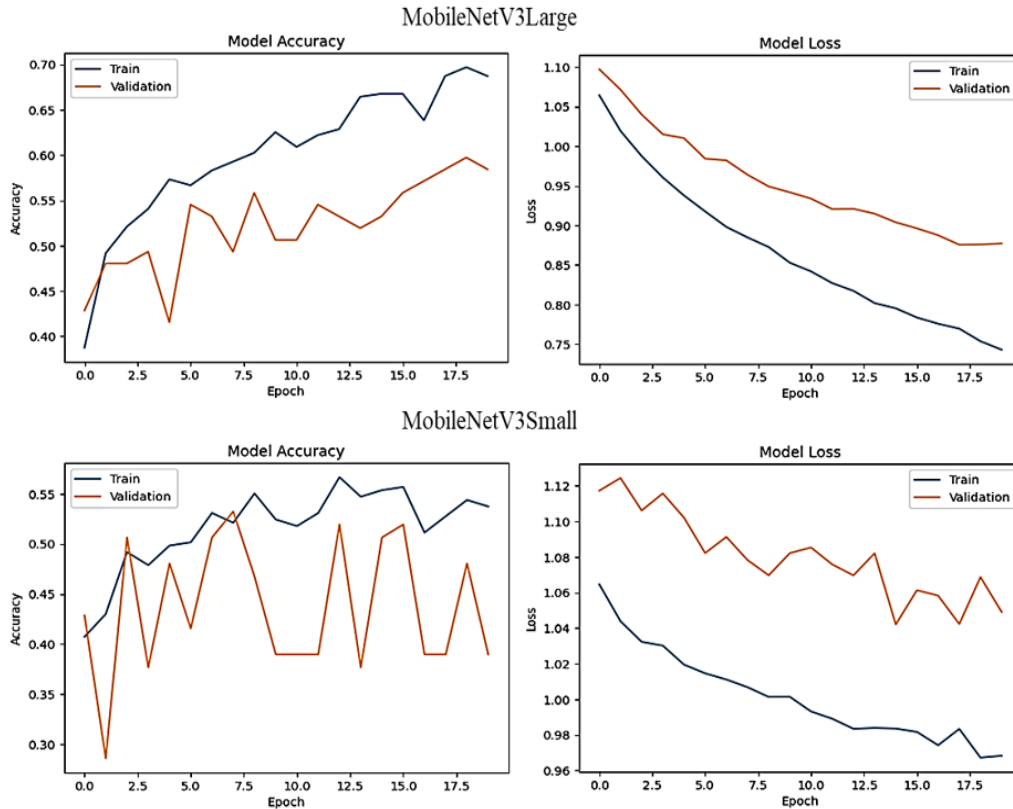


Figure 13. illustrates the nuanced interplay of accuracy and loss in MobileNetV3Large and MobileNetV3Small pre-trained CNN models, shedding light on their distinct performance attributes in the realm of emergency vehicle classification

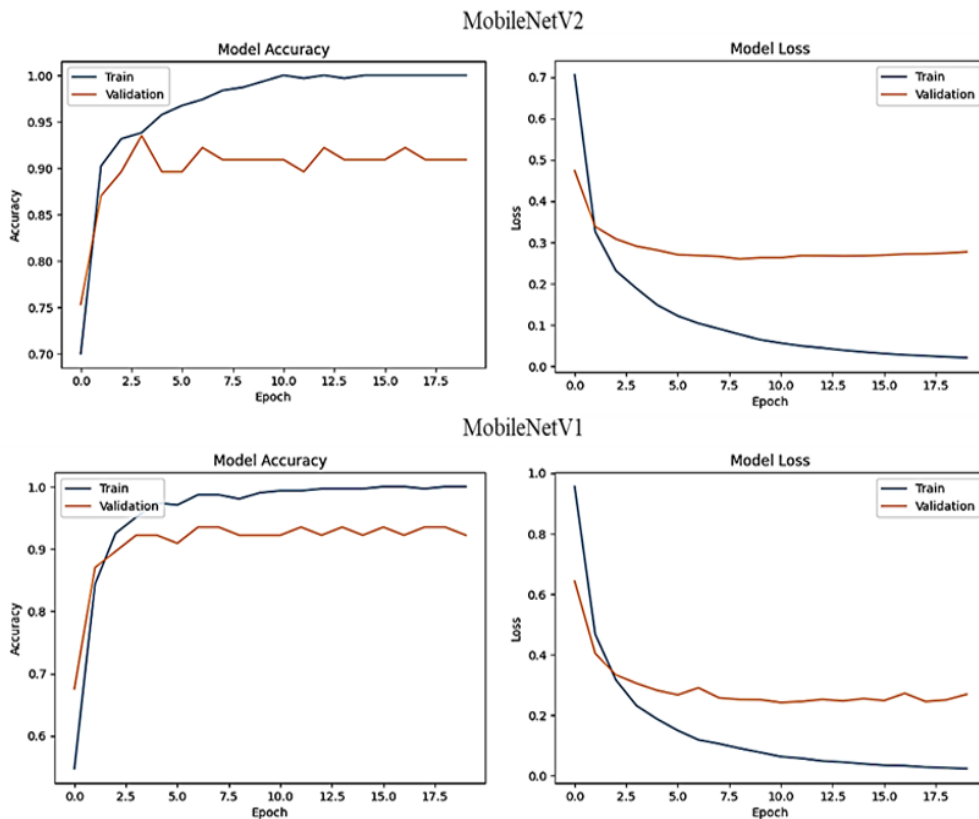


Figure 14. presents a detailed scrutiny of accuracy and loss for MobileNetV2 and MobileNetV1, unraveling their respective efficacy in emergency vehicle classification. This graphical representation enriches our understanding of the performance dynamics of these pre-trained CNN models

MobileNetV1 performed best according to table2, with interesting data displayed in Figure 11. Loss is depicted on the right axis, while Accuracy (%) and training time (s) are on the left.



Figure 15. Ambulance Classification with predicted class and confidence



Figure 16. Police Car Classification with precise class prediction and confidence showcased



Figure 17. Fire Truck Classification, highlighting predicted class and confidence levels

Among the evaluated models, our study reveals that MobileNetV1 stands out as the most accurate, boasting a sophisticated architecture that strikes a balance between precision and efficiency. Notably, it achieves this remarkable accuracy while requiring a relatively short training time, reinforcing its viability for applications demanding swift and precise emergency vehicle classification.

CONCLUSION

During our thorough investigation focused on classifying emergency vehicles using Convolutional Neural Network (CNN) models, we discovered that MobileNetV1 has emerged as the leading model, delivering the

highest level of accuracy compared to other models. Additionally, we observed that VGG16 and MobileNetV3Small exhibited strong performance, and one noteworthy aspect was their relatively shorter training times. Each model we examined has its own unique set of strengths and limitations, resulting in a diverse range of options for emergency vehicle classification. While the VGG16 and VGG19 models offer complexity, they require substantial computational resources. On the other hand, MobileNet strikes a favorable balance between accuracy and computational efficiency.

The careful selection of a suitable pre-trained Convolutional Neural Network (CNN) model plays a crucial role in achieving high accuracy and efficiency in emergency vehicle classification. Various factors need to be considered, including the intricate architectural details, computational requirements, and specific needs of the application. It is important to evaluate each model in a comprehensive manner, taking into account its unique characteristics. By having a deep understanding of these attributes, decision-makers can make informed choices, resulting in state-of-the-art classification outcomes.

This thorough investigation emphasizes the immense significance of utilizing pre-trained Convolutional Neural Network (CNN) models in order to enhance the abilities of emergency vehicle classification. It also highlights the importance of carefully selecting the most appropriate model for this purpose. By carefully weighing these factors, researchers can contribute to advancing progress in various fields by guaranteeing optimal performance in the crucial task of identifying emergency vehicles.

Perspective

In the field of emergency vehicle identification, there is a significant progression from the initial classification stage to the subsequent object detection stage. This progression is strategically important as it allows for better identification and understanding of emergency vehicles. Notably, the remarkable performance of MobileNetV1, MobileNetV2, and MobileNetV3 models in classification tasks serves as a foundation for further exploration and advancement in more complex challenges, specifically in the realm of object detection. Through these advancements, we can expect to gain a deeper understanding of the intricacies involved in identifying emergency vehicles accurately.

Moving forward, the possibility of transitioning from the task of classifying entire images to the more challenging task of detecting individual objects within them is incredibly promising. This advancement holds immense potential in various fields. To achieve this, we can leverage the strengths of pre-trained CNN models like MobileNetV1, which is known for its accuracy, and VGG16, which is revered for its efficiency. By building upon these models, we can establish a strong foundation for our object detection endeavors, making our efforts more effective and precise.

- **Enhancing Precision:** Object detection allows for a more granular understanding of the visual scene, enabling precise identification of each distinct object, including emergency vehicles.
- **Complexity in Context:** The complexity of VGG16 and the efficiency of MobileNetV1, observed in classification tasks, can be further tested, and optimized in the context of object detection, where challenges like multiple object instances and varying scales come to the forefront.
- **Diversifying Methodologies:** Exploring object detection methods beyond CNN models, such as region-based or anchor-based approaches, adds diversity to the toolkit. This diversification enhances adaptability to different scenarios and contributes to the overall robustness of the detection system.

Embracing the evolution from classification to object detection not only extends the capabilities of the existing models but also opens avenues for addressing more intricate challenges in the identification of emergency vehicles within diverse visual contexts.

REFERENCES

1. Zegegy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In CVPR.
2. Taye, M.M.: Theoretical understanding of convolutional neural network: concepts, architectures, applications, future directions. *Computation (Basel)*. 11, 52 (2023). <https://doi.org/10.3390/computation11030052>.
3. Fan, J., Lee, J., Lee, Y.: A transfer learning architecture based on a support vector machine for histopathology image classification. *Applied Sciences*. 11, 6380 (2021). <https://doi.org/10.3390/app11146380>.
4. Sharma, N., Jain, V., Mishra, A.: An Analysis Of Convolutional Neural Networks For Image Classification. *Procedia Computer Science*. 132, 377-384 (2018). <https://doi.org/10.1016/j.procs.2018.05.198>
5. Goyal, P., Dollár, P., Girshick, R., Noordhuis, P., Wesolowski, L., Kyrola, A., ... & Zheng, A.: Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour. (2017).

6. Yang, H., Zhang, Y., Chao, Y., Ding, W.: Ultra-lightweight CNN design based on neural architecture search and knowledge distillation: A novel method to build the automatic recognition model of space target ISAR images. *Defence Technology*. 18, 1073-1095 (2022). <https://doi.org/10.1016/j.dt.2021.04.014>.
7. Fan, J., Lee, J., Lee, Y.: A transfer learning architecture based on a support vector machine for histopathology image classification. *Applied Sciences*. 11, 6380 (2021). <https://doi.org/10.3390/app11146380>.
8. Neelam Jaikishore, C., Podaturpet Arunkumar, G., Jagannathan Srinath, A., Vamsi, H., Srinivasan, K., Ramesh, R.K., Jayaraman, K., Ramachandran, P.: "Implementation of Deep Learning Algorithm on a Custom Dataset for Advanced Driver Assistance Systems Applications" (2022).
9. Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J.: A Survey of Convolutional Neural Networks: Analysis, applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems*. 33, 6999-7019 (2022). DOI.
10. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. In *Nature*.
11. Gonzalez-Argote J. A Bibliometric Analysis of the Studies in Modeling and Simulation: Insights from Scopus. *Gamification and Augmented Reality 2023*;1:5-5. <https://doi.org/10.56294/gr20235>.
12. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *NIPS*.
13. Omari Alaoui A, El Bahi O, Oumoulyte M, El Youssefi A, Farhaoui Y, El Allaoui A (2023) Optimizing Emergency Vehicle Detection for Safer and Smoother Passages. In: *Proceedings of the 6th International Conference on Networking, Intelligent Systems & Security*. Association for Computing Machinery, Larache, Morocco. <https://doi.org/10.1145/3607720.3607728>.
14. S. Kaushik, A. Raman and K. V. S. Rajeswara Rao: "Leveraging Computer Vision for Emergency Vehicle Detection-Implementation and Analysis" (2020).
15. S. Roy and M. S. Rahman: "Emergency Vehicle Detection on Heavy Traffic Road from CCTV Footage Using Deep Convolutional Neural Network", *International Conference on Electrical, Computer, and Communication Engineering (ECCE)*, Cox'sBazar, Bangladesh, (2019).
16. Hasan S, Rabbi G, Islam R, Imam Bijoy H, Hakim A (2022) Bangla Font Recognition using Transfer Learning Method. In: *2022 International Conference on Inventive Computation Technologies (ICICT)*. pp 57-62. DOI: 10.1109/ICICT54344.2022.9850765
17. Simonyan K, Zisserman A (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. Doi : 10.48550/arXiv.1409.1556
18. Romero-Carazas R. Prompt lawyer: a challenge in the face of the integration of artificial intelligence and law. *Gamification and Augmented Reality 2023*;1:7-7. <https://doi.org/10.56294/gr20237>.
19. Qian S, Ning C, Hu Y (2021) MobileNetV3 for Image Classification. In: *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*. pp 490-497. doi :10.1109/ICBAIE52039.2021.9389905
20. Sandler, M., Howard, A.W., Zhu, M., Zhmoginov, A., Chen, L.-C.: MobileNetV2: Inverted Residuals and Linear Bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 4510-4520. (2018).
21. Howard, A.G. Zhu, M., Chen, B., Kalenichenko, D., ... & Adam, H.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. (2017).
22. Bai, Y.: RELU-Function and Derived Function review. *SHS Web of Conferences*. 144, 02006 (2022). <https://doi.org/10.1051/shsconf/202214402006.S>

23. Pearce, T.: Understanding softmax confidence and uncertainty, <https://arxiv.org/abs/2106.04972>.

24. Russakovsky, O., Deng, J., Su, H., ... & Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211-252. (2015).

25. xml.etree.ElementTree – The ElementTree XML API – Python 2.7.18 documentation, <https://docs.python.org/2/library/xml.etree.elementtree.html>.

26. Gonzalez-Argote D, Gonzalez-Argote J, Machuca-Contreras F. Blockchain in the health sector: a systematic literature review of success cases. *Gamification and Augmented Reality* 2023;1:6-6. <https://doi.org/10.56294/gr20236>.

27. Auza-Santiváñez JC, Díaz JAC, Cruz OAV, Robles-Nina SM, Escalante CS, Huanca BA. mHealth in health systems: barriers to implementation. *Health Leadership and Quality of Life* 2022;1:7-7. <https://doi.org/10.56294/hl20227>.

FINANCING

No financing.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Ali Omari Alaoui, Omaima El Bahi, Mohamed Rida Fethi, Othmane Farhaoui, Ahmad El Allaouiand, Yousef Farhaoui.

Research: Ali Omari Alaoui, Omaima El Bahi, Mohamed Rida Fethi, Othmane Farhaoui, Ahmad El Allaouiand, Yousef Farhaoui.

Drafting - original draft: Ali Omari Alaoui, Omaima El Bahi, Mohamed Rida Fethi, Othmane Farhaoui, Ahmad El Allaouiand, Yousef Farhaoui.

Writing - proofreading and editing: Ali Omari Alaoui, Omaima El Bahi, Mohamed Rida Fethi, Othmane Farhaoui, Ahmad El Allaouiand, Yousef Farhaoui.