energies

MDPI

Article

Finetuned Deep Learning Models for Fuel Classification: A Transfer Learning-Based Approach

Hemachandiran Shanmugam [1,*] and Aghila Gnanasekaran [1,2]

[1] Department of CSE, National Institute of Technology Puducherry, Karaikal 609609, India
[2] National Institute of Technology Tiruchirappalli, Tamil Nadu 620015, India
* Correspondence: spahhema@gmail.com

Abstract: Accurately classifying petrol and diesel fuel using an image processing method is crucial for fuel-related industries such as petrol pumps, refineries, and fuel storage facilities. However, distinguishing between these fuels using traditional methods can be challenging due to their similar visual characteristics. This study aims to enhance the accuracy and robustness of existing fuel classification by utilizing the transfer learning-based finetuned pre-trained deep learning models and ensemble approaches. Specifically, we upgrade pre-trained deep models like ResNet152V2, InceptionResNetV2, and EfficientNetB7 by incorporating additional layers. Through transfer learning, these models are adapted to the specific task of classifying petrol and diesel fuels. To evaluate their performance, the upgraded deep model and an ensemble of these models are tested on a synthetic dataset. The results indicate that the ensemble of upgraded ResNet152V2, InceptionResNetV2, and EfficientNetB7 achieves recall, precision, f-score, and accuracy scores of 99.54%, 99.69%, 99.62%, and 99.67%, respectively. Moreover, a comparative analysis reveals that the upgraded models outperform state-of-the-art baseline models.

Keywords: deep transfer models; ensemble approaches; image classification; pre-trained models

check for updates

Academic Editor: Fernando Sánchez Lasheras

Received: 19 January 2025
Revised: 8 February 2025
Accepted: 21 February 2025
Published: 27 February 2025
Corrected: 20 January 2026

Citation: Shanmugam, H.; Gnanasekaran, A. Finetuned Deep Learning Models for Fuel Classification: A Transfer Learning-Based Approach. *Energies* **2025**, *18*, 1176. https://doi.org/10.3390/en18051176

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, the significance of energy sources and their environmental impact has become increasingly important to society. Accurately classifying fuel types, such as petrol and diesel, is crucial for industries and regulatory bodies. Traditional methods relied on human expertise or rule-based models for prediction. However, the advancements in artificial intelligence and machine learning have provided highly accurate and automated classification models. Traditional machine learning methods for fuel classification rely on feature engineering. It requires manual extraction and selection of relevant features such as spectral properties, density, and texture [1]. This extraction is not always optimal. Because it depends on domain expertise and may fail to capture subtle variations in complex datasets [2]. ML models face scalability issues, with performance often degrading when handling large datasets. Especially dataset containing diverse conditions, such as varying lighting, different fuel containers, and environmental factors. Another major limitation is their sensitivity to noise. Fuel properties can be affected by impurities, temperature fluctuations, and measurement inconsistencies, leading to potential misclassification. Traditional ML models struggle with generalization. It tend to perform well on the specific datasets they were trained on but fail to adapt effectively to new, unseen samples due to dataset biases. These challenges highlight the need for more robust and automated approaches.

Deep learning, a subset of machine learning, has revolutionized numerous domains by enabling computers to learn complex patterns and representations from vast amounts of data. By leveraging deep neural networks, these models can automatically extract meaningful features, both low-level and high-level, from input data [3]. This characteristic makes them particularly well-suited for computer vision applications, including image classification. When it comes to classifying petrol and diesel fuels, deep learning models excel in analyzing various characteristics to differentiate between the two types. Unlike traditional approaches that rely on explicit rules or manually-crafted features [4,5] (such as Scale-Invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Histogram of Oriented Gradients (HOG), Gray-Level Co-occurrence Matrix (GLCM), etc.), deep models learn directly from the data. This data-driven approach enables more robust and accurate classification, as the models can uncover intricate patterns and relationships present in the input data.

In the study conducted by Adige et al. [6], the classification of multi-class images was explored using both support vector machines and deep residual networks. The research investigated two distinct approaches: one involving hand-crafted features (specifically, Bag of Visual Words) combined with a machine learning model (Support Vector Machine), and the other utilizing deep learning with ResNet. The results obtained from this study demonstrated the superiority of the deep learning model, specifically when employing training algorithms such as Adaptive Moment (Adam), Root-Mean-Square propagation (RMSProp), and Stochastic Gradient Descent (SGD). By leveraging deep learning techniques, the researchers achieved improved classification performance compared to the hand-crafted feature approach [7].

Asif et al. [8] introduced an enhanced deep learning model by incorporating multiple architectures, namely Xception, DenseNet201, DenseNet121, ResNet152V2, and Inception-ResNetV2 [9,10]. In their study, they extended these architectures by adding an additional deep dense block in the final layer. Through a comprehensive comparison analysis, they evaluated the performance of each architecture and identified Xception as the most suitable method for the classification task.

Iman et al. [11] conducted a comprehensive review of deep learning, covering concepts, definitions, and the taxonomy of deep transfer models and their learning. Their work involved experimental analyses of various scenarios, evaluating the advantages and disadvantages of different pre-trained models on large-scale datasets. In a related study, Punuri et al. [12] proposed an innovative transfer learning technique based on EfficientNet and Extreme Gradient Boosting (XGBoost) [13–15]. They focused specifically on face emotion recognition and fine-tuned the final dense layer of the EfficientNet architecture using an XGBoost classifier. This approach resulted in improved performance in the face emotion recognition task.

Karaddi et al. [16] utilized eight pre-trained convolutional neural network (CNN) models, including Alexnet, Darknet-19, Darknet-53, Densenet-201, Googlenet, Inception-ResnetV2, MobilenetV2, and Resnet-18, for the classification of various lung diseases. Through their analysis, they achieved optimal results using the Densenet-201 model. In a separate study, Sengupta et al. [17] introduced a novel approach for generating deep architecture Spiking Neural Networks (SNNs) based on VGG16 [18] and Residual networks. This research demonstrated the effectiveness of the proposed approach in extracting complex features and patterns.

State-of-the-art CNN models, including EfficientNetB0, ResNet50 [19], and Xception [20], have been effectively utilized by Joshi et al. [21] in their breast cancer detection research. Similarly, Subathra et al. [22] employed ResNet50, InceptionV3, and Xception pre-trained models to identify defects in breast thermographic images accurately. These

chosen architectures represent some of the most powerful and widely recognized models in deep learning, highlighting their suitability for tackling complex medical imaging tasks.

There are significant advantages to employing deep learning models for fuel classification. Firstly, these models are capable of effectively processing extensive and diverse datasets, accommodating the diverse range of fuel patterns. Secondly, deep learning models have the ability to continuously learn from new data, allowing them to adapt and improve their performance over time [3]. Additionally, these models have the capacity to identify intricate patterns that may elude human observers, thereby potentially enhancing the reliability and precision of fuel classifications.

However, it is important to note that the development of deep learning models for fuel classification demands meticulous deliberation and expertise. Several factors come into play, including data quality, the implementation of pre-processing techniques like data scaling, data normalization, and train-test split, the careful selection of model architecture, the utilization of data augmentation techniques [23], and the adoption of optimization strategies. These elements collectively contribute to achieving precise and trustworthy results. It is worth mentioning that the interpretability of deep models can pose a challenge since they often function as black-box models, making it difficult to discern the specific features or patterns they rely on for classification.

1.1. Objective

The primary objective of this study is to develop and evaluate state-of-the-art deep learning models for the accurate classification of petrol and diesel images by leveraging various deep learning architectures, including Xception, VGG, ResNet, Inception, MobileNet, DenseNet, and EfficientNet, as well as proposed upgraded models and ensemble approaches [22,24–27]. This study aims to enhance classification accuracy and improve model efficiency. A key focus is on extracting meaningful features from fuel data using advanced neural networks and transfer learning approach to ensure robust and reliable classification under diverse real-world conditions. This study seeks to address challenges related to data quality, feature extraction, and model generalization, ultimately contributing to advancements in automated fuel analysis.

1.2. Impact of This Study

The aim of this research is to design an optimized deep learning framework. It outperforms traditional machine learning methods in fuel classification. The optimized deep learning framework is leading to more precise and scalable solutions. This study strives to achieve higher performance by leveraging the strengths of different models and by integrating multiple architectures through ensemble learning. It aims to explore the impact of activation functions, optimizers, and data pre-processing techniques on classification accuracy. The ultimate goal is to develop a high-performing model that can be effectively used in fuel quality assessment, sustainability efforts, and intelligent fuel monitoring systems. Added, contributing to energy efficiency and resource optimization. This study could make a significant impact in the following areas:

Automation of the Classification Process: The utilization of a machine learning model for image classification has enabled the automation of the process, leading to swift and accurate results. This reduces the reliance on manual labor and streamlines the overall workflow.

Improved Efficiency: The automation of the classification process enhances efficiency and productivity for businesses. By eliminating manual tasks, resources can be allocated more effectively, resulting in improved operational performance.

Cost Savings: The automated classification process brings cost savings in terms of reduced labor and time requirements. By optimizing resource allocation and minimizing human intervention, businesses can achieve financial benefits.

Enhanced Data Analysis:Image classification provides valuable insights into patterns and trends within the data. This enables businesses to make informed decisions and enhance their operations by leveraging the information gained from classifying images.

Enhanced Customer Service: Through image classification, businesses can improve their customer service by providing accurate and relevant information to their customers. This ensures a higher level of customer satisfaction and facilitates better communication.

Increased Safety: Image classification plays a vital role in identifying potential safety hazards associated with vehicles. By promptly identifying and addressing these hazards, businesses can ensure a safer operating environment.

2. Proposed Methodology

This section presents the detailed workflow employed in this study, outlining the proposed methodology. Figure 1 provides an overview of the approach utilized for the classification of petrol and diesel images, employing a deep transfer learning technique. Section 2.1 provides comprehensive information regarding the datasets used in this study. The preprocessing of database images, data augmentation techniques, and dataset splitting are discussed in Section 2.2. These steps address the challenge of limited datasets and aim to enhance the classification performance. In Section 2.3, we delve into the incorporation of proposed upgraded deep layers into the pre-trained model using transfer learning, emphasizing their role in feature extraction and classification.

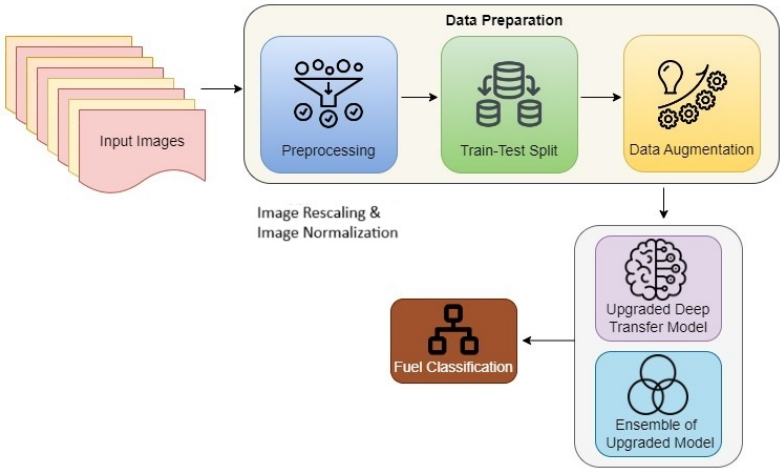


Figure 1. Comprehensive flowchart of the proposed methodology for classifying petrol and diesel images based on deep transfer learning. Image preprocessing includes image rescaling and normalization.

2.1. Dataset Description

When curating a self-collected dataset of petrol and diesel images, it is of utmost importance to ensure the data's diversity, representativeness, and accurate labelling. A high-resolution camera and smartphone were used to capture the sample images of petrol and diesel. It was ensured that the dataset encompasses a broad range of variations, including different brands, fuel containers, lighting conditions, and angles. Figure 2 shows the sample images from different fuel types, containers and lighting conditions [28]. Each image was carefully labelled as either "petrol" or "diesel". In order to maintain consistency, all samples were uniformly resized to 32 $\times$ 32 pixels, facilitating standardized inputs for the model.

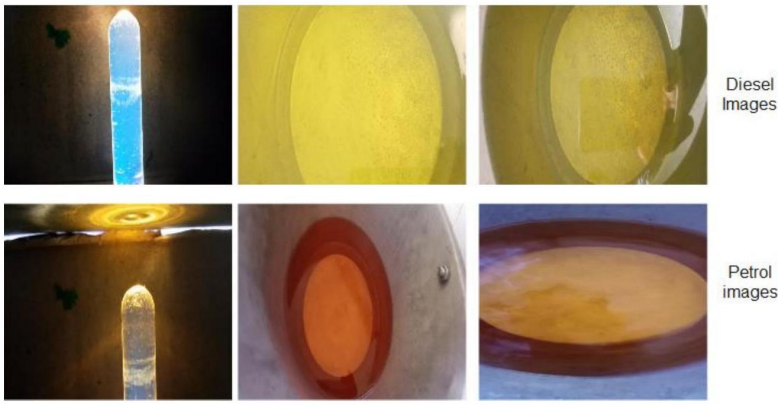


Figure 2. Sample images from the synthetic dataset, grouped by fuel type: The top row displays diesel samples, while the bottom row shows petrol samples. Each sample was collected using a different container and captured under varying lighting angles. Figure reused from our previous work [28] with permission under CC BY license.

2.2. Data Preparation

Pre-processing of images is a critical step in data preparation for fuel classification using deep learning models. In this work, dataset is carefully cleaned and preprocessed to ensure high-quality data for accurate fuel classification. First, blurry, low-resolution, and noisy images are removed which may negatively impact model performance. Then, examined for duplicate images and mislabeled samples. These are manually removed to maintain data integrity. Normalization techniques applied to standardize pixel values. Data augmentation methods such as random rotations, flips, brightness adjustments, and translations were used to introduce variability, helping the model generalize better to different real-world conditions.

Handling outliers was another important step in the fuel classification process. The excessive glare, extreme shadows, or unclear fuel samples are either corrected or removed to prevent misclassification. When it came to feature extraction, instead of relying on handcrafted features like texture or color intensity, we leveraged deep learning models to automatically learn important visual patterns. These models identified key characteristics such as subtle color variations, texture details, and shape differences between petrol and diesel images.

The following techniques has been applied to enhance image quality and effectively split the dataset:

- Image Rescaling: Resizing images to a consistent resolution is beneficial for ensuring dataset uniformity. This step reduces computational complexity and speeds up training. Common rescaling techniques include bilinear or bicubic interpolation.
- Image Normalization: Normalizing pixel values of images to a common scale improves model convergence. Techniques such as mean subtraction or min-max scaling can be used. Normalization reduces the impact of lighting conditions and enhances the model's ability to extract meaningful features.
- Data Augmentation: Apply data augmentation techniques to increase dataset diversity and improve model generalization [29]. Techniques like random rotations, flips, translations, and brightness adjustments can be employed. Data augmentation helps the model learn robust features and reduces overfitting.
- Train-Test Split: Divide the dataset into training and testing sets. The training set is used to train the deep learning model, while the testing set is used for performance evaluation. Ensure a balanced representation of both petrol and diesel samples

in the split to avoid bias. A common split ratio is 80:20 or 70:30 for training and testing, respectively.

By applying these image pre-processing and splitting techniques, you can enhance dataset quality and diversity, improve the model's ability to learn relevant features and obtain reliable evaluations of its performance.

2.3. Upgraded Deep Transfer Models

Deep Convolutional Neural Network (DCNN) architectures, including Xception, VGG-16, VGG-19, ResNet50, ResNet50V2, ResNet101, ResNet152, ResNet152V2, InceptionV3, InceptionResNetV2, MobileNet, MobileNetV2, DenseNet121, DenseNet169, DenseNet201, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, EfficientNetB7, etc., are powerful tools for extracting features and patterns in image-based computer vision applications such as image classification. In this section, ResNet152V2, InceptionResNetV2, and EfficientNetB7 models have been upgraded by incorporating additional layers into the pre-trained models. This approach leverages transfer learning to effectively extract meaningful patterns from the fuel images and classify them accurately.

2.3.1. Upgraded ResNet152V2

ResNet152V2 is an enhanced version of the original ResNet (Residual Network) architecture, where the number "152" denotes the total number of layers, and "V2" signifies it as the second version of the ResNet152 model. ResNet152V2 is a deep convolutional neural network (DCNN) architecture widely recognized for its powerful feature extraction capabilities, particularly in image-related tasks.

This architecture addresses the issue of vanishing gradients encountered in deep neural networks, enabling the network to learn residual mappings. By incorporating residual connections, the model can effectively learn both shallow and deep features, facilitating the training of deeper networks. ResNet152V2 is typically pre-trained on large-scale image datasets like ImageNet, which consists of millions of labelled images spanning various categories. During pre-training, the model learns to extract low-level details (such as edges and textures) as well as high-level semantic information (such as object shapes and structures). It progressively extracts hierarchical features at different spatial resolutions within each layer, enabling the model to capture complex patterns and representations.

Due to its ability to learn rich image representations, ResNet152V2 is commonly employed in transfer learning scenarios. After pre-training on ImageNet, the model's weights can be transferred to a different task or domain. In this study, a transfer learning approach is utilized, applying the pre-trained model on ImageNet to the fuel classification task. By freezing the pre-trained layers and adding new layers on top, the model can be fine-tuned for specific classification, object detection, or segmentation tasks using smaller datasets. Therefore, in this study, an upgraded ResNet152V2 is fine-tuned by incorporating additional blocks, as depicted in Figure 3.

The upgraded ResNet152V2 blocks consist of the ResNet152V2 pre-trained model, which is augmented with an average pooling layer, batch normalization layers, and dense layers. The pre-trained ResNet152V2 model serves as a transfer learning approach, while the average pooling layer performs dimensionality reduction, and batch normalization is utilized to enhance the performance and stability of the deep transfer model.

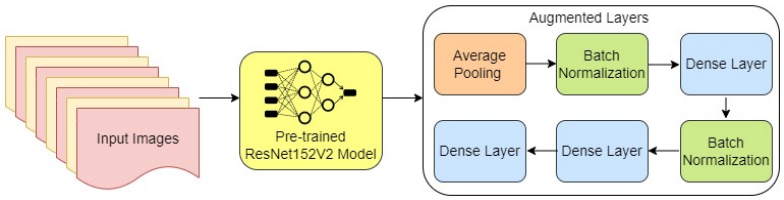


Figure 3. Schematic representation of Upgraded ResNet152V2.

2.3.2. Upgraded InceptionResNetV2

InceptionResNetV2 is a deep convolutional neural network (DCNN) architecture renowned for its power and high accuracy in various computer vision applications. This pre-trained model combines the concepts of both the Inception and ResNet models, enabling improved gradient flow and facilitating the training of deeper networks. It incorporates multiple inception modules and residual blocks, allowing it to effectively capture both local and global image features. Similar to ResNet, InceptionResNetV2 is typically pre-trained on large-scale image datasets like ImageNet, which comprises millions of labeled images from diverse categories. During pre-training, the model learns to extract high-level features and representations, enabling it to recognize a wide range of objects, textures, and patterns.

Given its exceptional performance and versatility, InceptionResNetV2 is commonly employed in transfer learning approaches. After being pre-trained on ImageNet, the model's weights can be transferred to different tasks or domains. In this study, transfer learning is applied specifically to fuel classification. The model is fine-tuned for this task and referred to as the Upgraded InceptionResNetV2, as depicted in Figure 4.

The Upgraded InceptionResNetV2 blocks consist of the InceptionResNetV2 pre-trained model, augmented with additional layers. These include an average pooling layer, batch normalization layers, a dropout layer, and dense layers. The pooling layer utilizes average pooling with a stride size of 2×2 and zero padding. Two batch normalization layers and three dense layers with ReLU activation functions (except for the final dense layer) are incorporated. A dropout rate of 0.2 is applied. The final dense layer employs the sigmoid activation function and binary cross-entropy as the loss function.

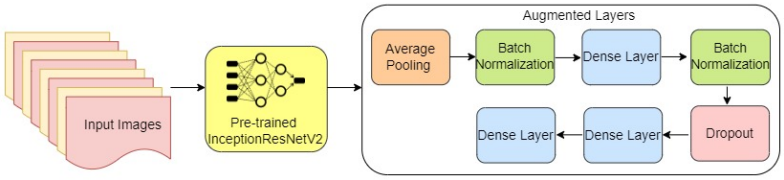


Figure 4. Schematic representation of Upgraded InceptionResNetV2.

2.3.3. Upgraded EfficientNetB7

EfficientNetB7, belonging to the EfficientNet family of DCNN architectures, stands out due to its unique compound scaling method. This approach optimizes the depth, width, and resolution of the network, striking a balance between model size and performance. By scaling these dimensions systematically, EfficientNetB7 achieves a significantly larger architecture compared to smaller variants like EfficientNetB0. This architectural expansion enables the model to capture more intricate patterns and representations, serving as a powerful feature extractor for both low-level and high-level features. As with other pre-trained models, EfficientNetB7 is typically trained on large-scale image datasets such as ImageNet, which provide millions of labelled images from diverse categories.

EfficientNetB7, like most pre-trained models, is readily available in popular deep learning frameworks such as TensorFlow and PyTorch. These frameworks provide access to pre-trained EfficientNetB7 models that can be directly utilized or fine-tuned for specific tasks. The pre-

trained weights for EfficientNetB7 are often available for download, allowing for convenient utilization of the learned representations without the need for extensive training.

As represented in Figure 5 The Upgraded EfficientNetB7 blocks consist of the pre-trained model augmented with additional layers. These include an average pooling layer, batch normalization layers, a dropout layer, and dense layers. The pooling layer utilizes the average pooling technique with a stride size of 2 × 2 and zero padding. Three batch normalization layers and two dense layers with ReLU activation functions are added, along with a dropout rate of 0.2. The final dense layer employs the sigmoid activation function and binary cross-entropy as the loss function.

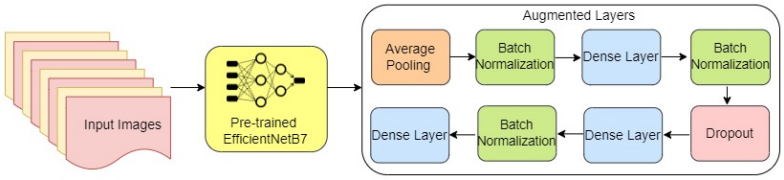


Figure 5. Schematic representation of Upgraded EfficientNetB7.

2.3.4. Ensemble of Deep Transfer Models

This study utilizes ensemble learning techniques to improve the performance and generalization capabilities of the fuel classification system. Specifically, a combination of upgraded ResNet152V2 + InceptionResNetV2, upgraded ResNet152V2 + EfficientNetB7, upgraded InceptionResNetV2 + EfficientNetB7, and upgraded ResNet152V2 + InceptionResNetV2 + EfficientNetB7 are employed to form an ensemble of models. Each model is trained on a subset of the dataset, and their predictions are aggregated using voting methods to determine the final fuel classification decision.

To evaluate the proposed approach, we have curated a comprehensive dataset of petrol and diesel images. This dataset is meticulously labeled and augmented to ensure diversity and balance. Additionally, appropriate data splitting and cross-validation strategies are employed to accurately assess the performance of the ensemble models.

Preliminary experiments indicate that the ensemble of upgraded deep models, using transfer learning, achieves higher classification accuracy compared to individual or baseline models. The ensemble models effectively capitalize on the complementary strengths of different architectures, enhancing the overall robustness of the system. The results underscore the efficacy of ensemble learning in improving the classification of petrol and diesel images, which will be discussed in detail in the following section.

3. Experimental Results and Analysis

3.1. Dataset Description

The performance analysis of the proposed ensemble learning approach, utilizing deep block models and transfer learning, is evaluated on a synthetic self-collected dataset consisting of petrol and diesel images. These images are collected from real-world environments and represent a more challenging and realistic classification scenario for petrol and diesel. The synthetic dataset comprises 1515 petrol samples and 918 diesel samples, and all of these samples are utilized in this work for experimentation. The experiments were conducted on a Windows 10 64-bit system with an Intel Core i7 processor, and the algorithms were implemented using the Python programming language.

To address the class imbalance during training, transformations include random rotations, flips, translations, and brightness adjustments to improve generalization and handle class imbalance. These augmentations were carefully controlled to preserve critical fuel characteristics. We conducted comparative experiments to address this issue on

augmented vs. non-augmented datasets and analyzed the impact on model performance. Our experiments showed that the performance differences were negligible (0.5% increase in accuracy). This experiment ensured that our model was not biased toward the majority class and the data loss introduced through the data augmentation is also negligible.

3.2. Evaluation Metrics

To assess the effectiveness of the proposed approach, it is compared with other state-of-the-art deep learning methods for the classification of petrol and diesel images. Experimental factors such as Recall, Precision, F-score, and Accuracy [30] are calculated using metrics such as False Positive (FP), True Positive (TP), True Negative (TN), and False Negative (FN). These metrics provide insights into the performance and effectiveness of the proposed approach compared to other methods.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{F-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{4}$$

3.3. Ablation Study: 30 DL Models

In this experiment, we tested and compared various state-of-the-art DL models with our proposed approach. The results of this comparison can be found in Table 1, where the proposed approach outperformed the other methods in four key experimental factors. The Upgraded Resnet152V2 method achieved recall, precision, f-score, and accuracy scores of 96.01%, 95.23%, 96.65%, and 96.10%, respectively. Similarly, the Upgraded InceptionResNetV2 method obtained scores of 95.05%, 94.36%, 95.10%, and 95.12% for recall, precision, f-score, and accuracy, respectively. Additionally, the Upgraded EfficientNetB7 method yielded scores of 95.41%, 94.98%, 95.63%, and 95.61% for recall, precision, f-score, and accuracy, respectively. The ensemble of the upgraded models (ResNet152V2 + InceptionResNetV2) achieved recall, precision, F-score, and accuracy scores of 96.28%, 97.22%, 91.10%, and 97.06%, respectively.

Table 1. The performance comparison of various transfer learning deep models and the proposed upgraded deep models for fuel classification. Bold font indicates the highest score.

Models	Recall (%)	Precision (%)	F-Score (%)	Accuracy (%)
Xception	93.1	92.84	93	93.64
VGG-16	91.57	91.94	92.02	92.68
VGG-19	86.87	87.54	87.33	90.4
ResNet50	82.14	81.76	82.66	86.02
ResNet50V2	85.02	88.58	87	89.27
ResNet101	85.03	87.16	86	89.11
ResNet152	93.56	92.45	93.12	93.55
ResNet152V2	95.13	94	95.6	95.59
InceptionV3	93.69	93.12	93.3	93.89
InceptionResNetV2	95.54	95.4	95.36	95.56

Table 1. *Cont.*

Models	Recall (%)	Precision (%)	F-Score (%)	Accuracy (%)
MobileNet	91.63	90.77	91	92.52
MobileNetV2	93.1	92.84	93	93.64
DenseNet121	94.49	95.32	95	95.43
DenseNet169	95.75	95.82	95.66	96.25
DenseNet201	96.28	97.53	97.1	97.06
EfficientNetB0	91.48	94.25	92.82	93.11
EfficientNetB1	93.9	94.33	94.11	93.72
EfficientNetB2	94.5	94.86	94.57	94.33
EfficientNetB3	94.97	94.5	94.49	94.71
EfficientNetB4	94.53	94.3	94.35	94.49
EfficientNetB5	94.79	94.54	94.61	94.75
EfficientNetB6	94.94	95.32	95.1	95.82
EfficientNetB7	95.6	96.25	95.92	95.82
Upgraded ResNet152V2 (M1)	96.01	95.23	96.65	96.1
Upgraded InceptionResNetV2 (M2)	95.05	94.36	95.1	95.12
Upgraded EfficientNetB7 (M3)	95.41	94.98	95.63	95.61
Ensemble upgraded (M1 + M2)	96.28	97.22	91.1	97.06
Ensemble upgraded (M1 + M3)	95.75	95.82	95.66	96.25
Ensemble upgraded (M2 + M3)	95.32	95.43	94.49	95
Ensemble upgraded (M1 + M2 + M3)	**99.54**	**99.69**	**99.62**	**99.67**

Similarly, the ensemble of the upgraded models (ResNet152V2 + EfficientNetB7) attained recall, precision, f-score, and accuracy scores of 95.75%, 95.82%, 95.66%, and 96.25%, respectively. Moreover, the ensemble of the upgraded models (InceptionResNetV2 + EfficientNetB7) obtained recall, precision, f-score, and accuracy scores of 95.32%, 95.43%, 94.49%, and 95.00%, respectively. However, the ensemble approach combining the upgraded versions of ResNet152V2, InceptionResNetV2, and EfficientNetB7 achieved the highest scores for recall, precision, f-score, and accuracy, reaching 99.54%, 99.69%, 99.62%, and 99.67%, respectively. The box plot depicted in Figure 6 presents the distribution of key metrics—Accuracy, Precision, Recall, and F-Score—across 30 Deep Learning (DL) models. Notably, the optimal values for each metric are visually emphasized with a red circle, providing a clear indication of the best-performing models. The ensemble model combined the strengths of different deep learning models. It leads to improved feature extraction and classification accuracy. The key reasons for this improvement are feature extraction, generalization and robustness to variability of images. ResNet152V2 is excellent in identifying fine-grained patterns and textures in fuel images. It specializes in learning deep hierarchical features. EfficientNetB7 optimizes network depth, width, and resolution using a compound scaling technique to efficiently capture multi-scale representations. The ensemble reduces dependence on the shortcomings of any one model by combining the results from both models, utilizing a variety of feature extraction capabilities. Lighting, angles, and container reflections vary in fuel images. The ensemble model improves classification stability by combining predictions from several architectures and also reduces overfitting.

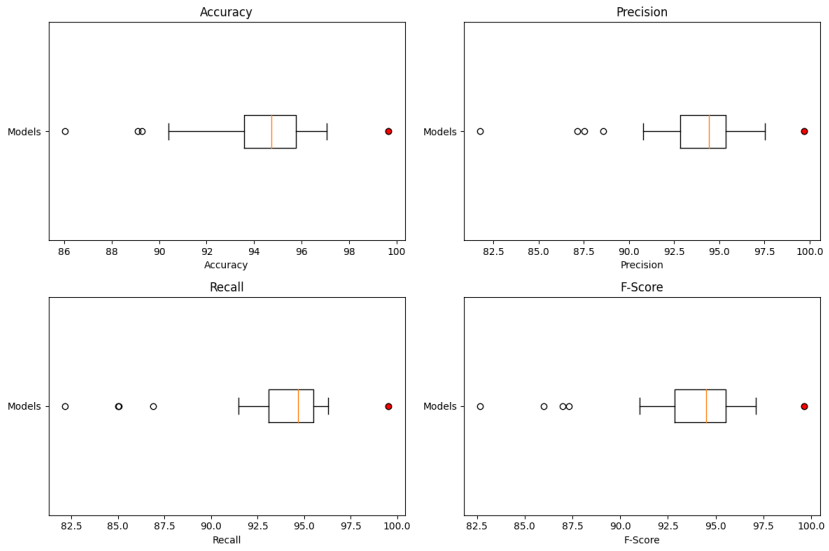


Figure 6. Performance Metrics Distribution of 30 DL Models with Highlighted Optimal Values.

3.4. Impact of Optimizer And Activation Function

In this experiment, we have employed a range of configuration parameters to train our methods. The configuration details of our deep learning models encompass various parameters, including input size, number of epochs, stride, learning rate, regularizers, batch size, loss function, dropout rate, activation function, optimizers, and pooling layers.

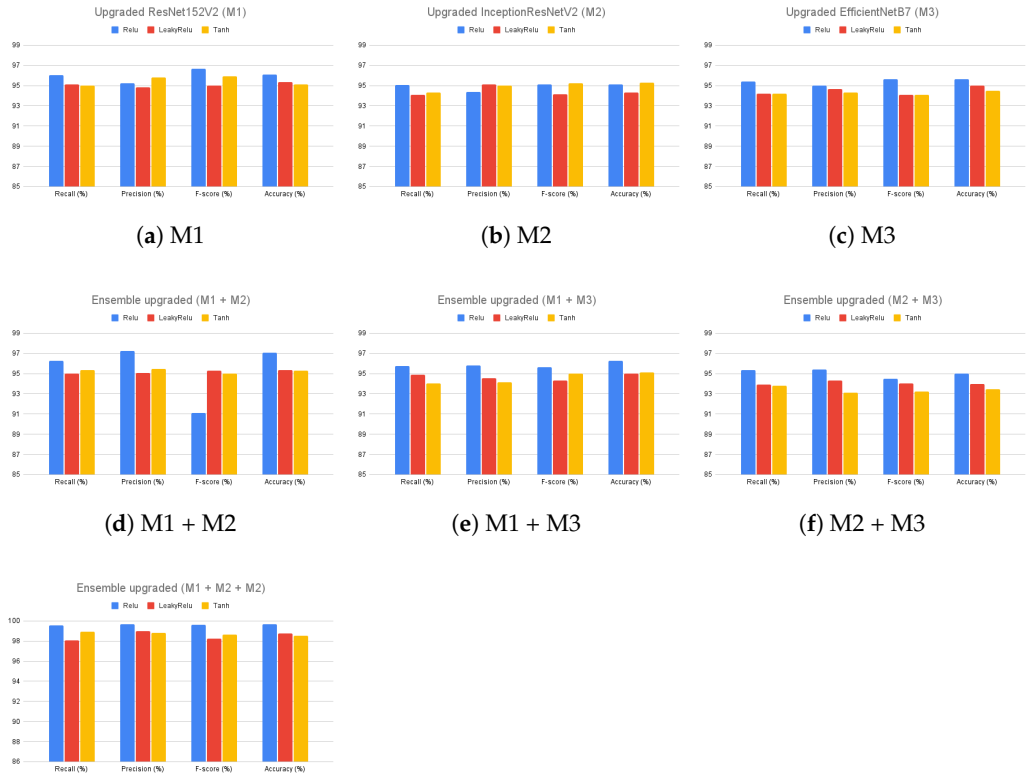
For multi-class classification tasks, the softmax activation function is typically used in the final layer, while for binary classification, the sigmoid activation function is employed. Moreover, activation functions are utilized in the intermediate layers to introduce non-linearity to the model. In our analysis, we have explored and compared the effects of ReLU, Leaky ReLU, and Tanh activation functions in the intermediate layers, as outlined in Table 2. Furthermore, we have illustrated the comparative analysis of these activation functions in the intermediate layers in Figure 7. Figure 8 illustrates the distribution of performance metrics, including Accuracy, Precision, Recall, and F-score, across all 30 DL models trained with different activation functions from Table 2. From Figure 8 and Table 2, we have compared different activation functions including Tanh, ReLU and Leaky ReLU) and their impact on model performance. ReLU was selected as the optimal activation function because it avoids vanishing gradient problem which gives the better performance on deep networks.

Table 2. The performance comparison of various Activation functions (Intermediate Layer) in the proposed upgraded deep models for fuel classification. Bold font indicates the highest score.

Proposed Model	Activation Function	Recall (%)	Precision (%)	F-score (%)	Accuracy (%)
Upgraded ResNet152V2 (M1)	Relu	**96.01**	**95.23**	**96.65**	**96.10**
	LeakyRelu	95.12	94.85	95.00	95.32
	Tanh	95.01	95.81	95.93	95.12
Upgraded InceptionResNetV2 (M2)	Relu	**95.05**	**94.36**	**95.10**	**95.12**
	LeakyRelu	94.11	95.10	94.12	94.33
	Tanh	94.30	95.02	95.25	95.30
Upgraded EfficientNetB7 (M3)	Relu	**95.41**	**94.98**	**95.63**	**95.61**
	LeakyRelu	94.20	94.64	94.10	95.00
	Tanh	94.18	94.32	94.10	94.46

Table 2. *Cont.*

Proposed Model	Activation Function	Recall (%)	Precision (%)	F-score (%)	Accuracy (%)
Ensemble upgraded (M1 + M2)	Relu	**96.28**	**97.22**	**91.10**	**97.06**
	LeakyRelu	95.01	95.06	95.28	95.33
	Tanh	95.32	95.48	95.02	95.28
Ensemble upgraded (M1 + M3)	Relu	**95.75**	**95.82**	**95.66**	**96.25**
	LeakyRelu	94.87	94.54	94.33	95.02
	Tanh	94.03	94.16	95.00	95.11
Ensemble upgraded (M2 + M3)	Relu	**95.32**	**95.43**	**94.49**	**95.00**
	LeakyRelu	93.90	94.33	94.01	93.97
	Tanh	93.82	93.11	93.24	93.44
Ensemble upgraded (M1 + M2 + M3)	Relu	**99.54**	**99.69**	**99.62**	**99.67**
	LeakyRelu	98.05	98.96	98.26	98.76
	Tanh	98.90	98.84	98.66	98.50



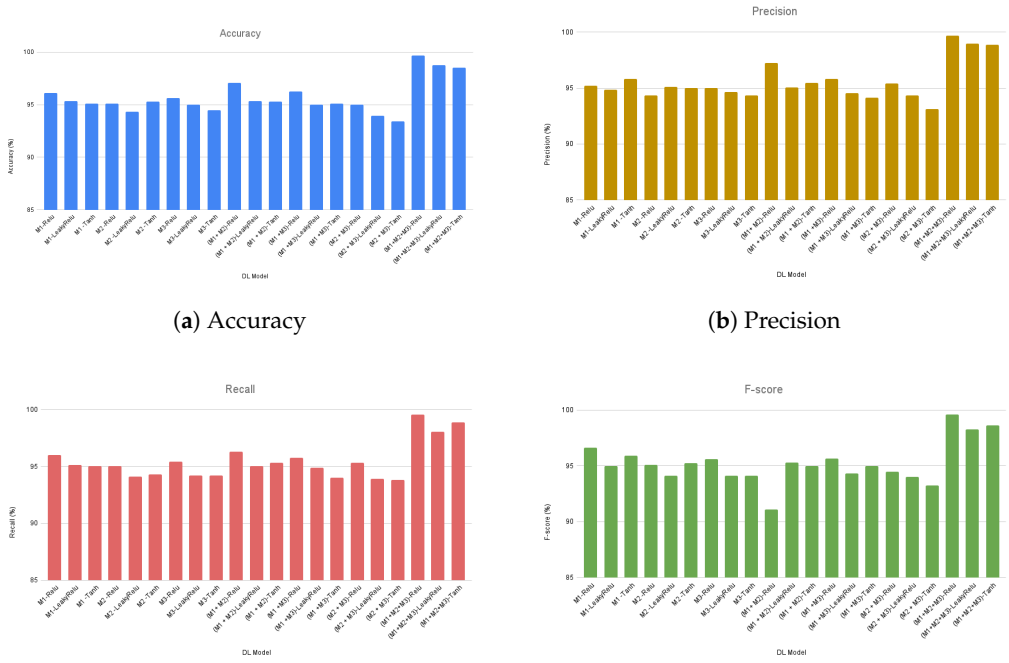
(**a**) M1 (**b**) M2 (**c**) M3

(**d**) M1 + M2 (**e**) M1 + M3 (**f**) M2 + M3

(**g**) M1 + M2 + M3

Figure 7. Comparison analysis of activation functions across upgraded and ensembled DL Models.

We have conducted a meticulous analysis of three optimizers, namely Adam, RMSProp, and SGD, and the results are presented in Table 3 and Figure 9. Figure 10 illustrates the distribution of performance metrics, including Accuracy, Precision, Recall, and F-score, across all 30 DL models trained with different optimizers from Table 3. The results from Table 3 and Figure 10 present the comparison of different optimizers, including SGD, RMSprop and Adam. Adam was best performing for our task because it combines the strengths of momentum and adaptive learning rate and more robustness to noise.



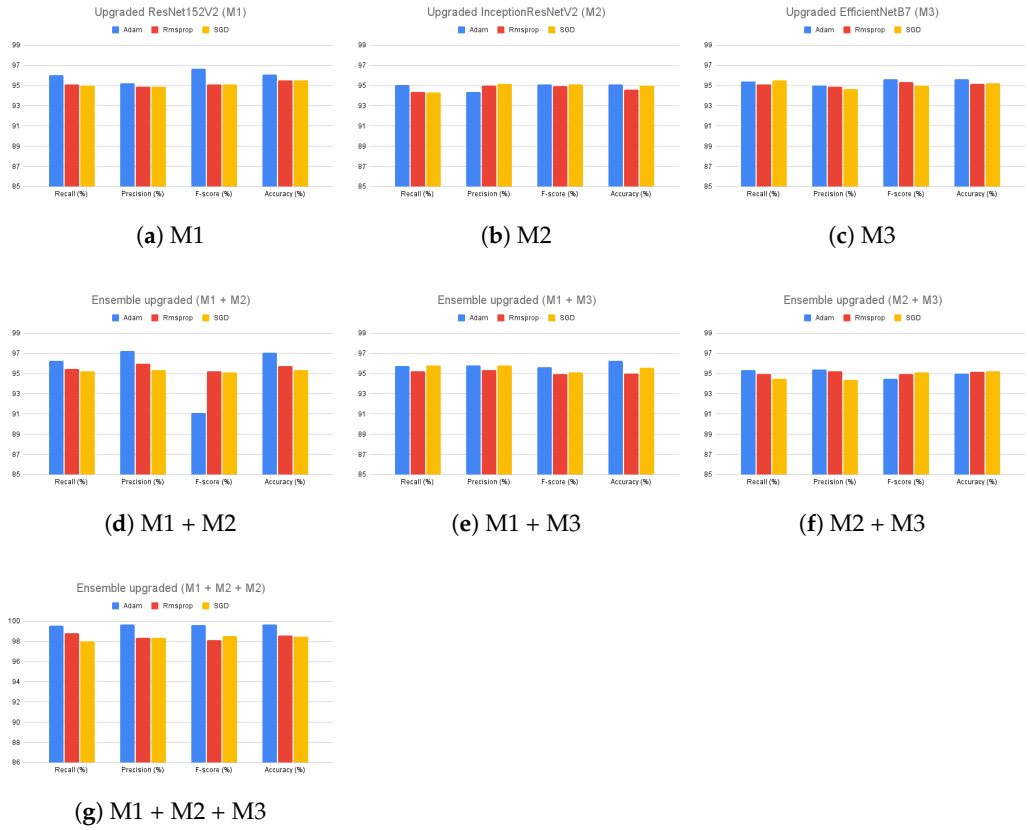
(**a**) Accuracy (**b**) Precision

(**c**) Recall (**d**) F-Score

Figure 8. Performance comparison of 30 DL models with varied activation function: Ablation study results.

Table 3. The performance comparison of various optimizers in the proposed upgraded deep models for fuel classification. Bold font indicates the highest score.

Proposed Model	Optimizer	Recall (%)	Precision (%)	F-score (%)	Accuracy (%)
Upgraded ResNet152V2 (M1)	Adam	**96.01**	**95.23**	**96.65**	**96.1**
	RMSProp	95.1	94.89	95.12	95.54
	SGD	95.01	94.89	95.12	95.54
Upgraded InceptionResNetV2 (M2)	Adam	**95.05**	**94.36**	**95.1**	**95.12**
	RMSProp	94.4	95.01	94.92	94.59
	SGD	94.32	95.2	95.12	95
Upgraded EfficientNetB7 (M3)	Adam	**95.41**	**94.98**	**95.63**	**95.61**
	RMSProp	95.14	94.89	95.36	95.16
	SGD	95.5	94.63	95.01	95.21
Ensemble upgraded (M1 + M2)	Adam	**96.28**	**97.22**	**91.1**	**97.06**
	RMSProp	95.45	95.96	95.26	95.76
	SGD	95.24	95.32	95.12	95.35
Ensemble upgraded (M1 + M3)	Adam	**95.75**	**95.82**	**95.66**	**96.25**
	RMSProp	95.23	95.34	94.94	95
	SGD	95.82	95.83	95.12	95.6
Ensemble upgraded (M2 + M3)	Adam	**95.32**	**95.43**	**94.49**	**95**
	RMSProp	94.94	95.23	94.95	95.15
	SGD	94.5	94.36	95.1	95.21
Ensemble upgraded (M1 + M2 + M3)	Adam	**99.54**	**99.69**	**99.62**	**99.67**
	RMSProp	98.82	98.35	98.1	98.6
	SGD	98.01	98.34	98.53	98.45



(**a**) M1 (**b**) M2 (**c**) M3

(**d**) M1 + M2 (**e**) M1 + M3 (**f**) M2 + M3

(**g**) M1 + M2 + M3

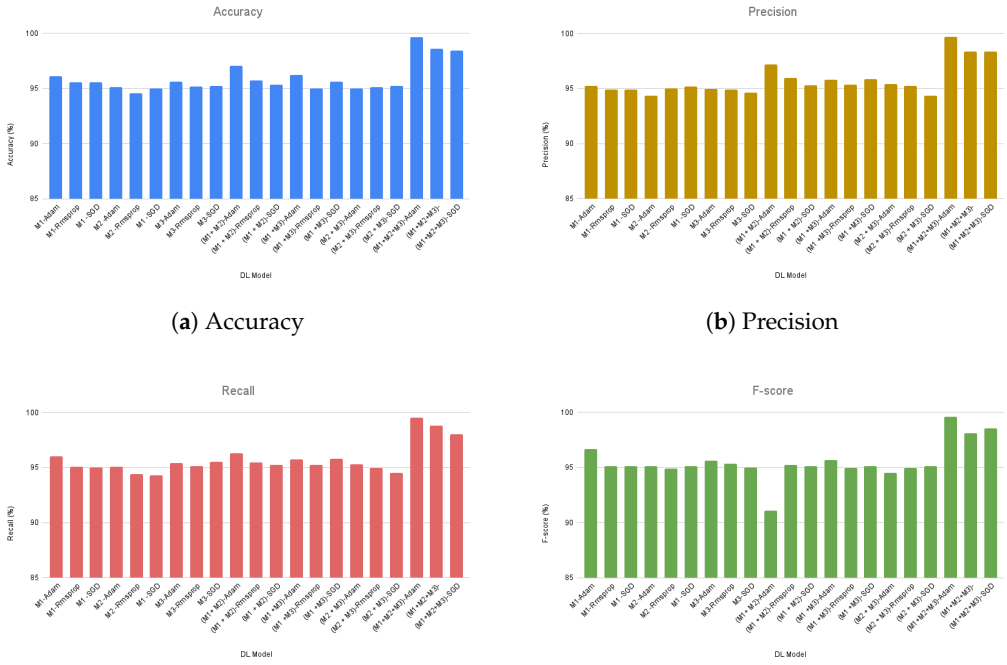
Figure 9. Comparison analysis of optimizers across upgraded and ensembled DL Models.

3.5. Results Summary

Models with larger depths and advanced architectural designs (such as ResNet152V2 and DenseNet201) outperformed shallower networks (e.g., VGG-16, ResNet50). The EfficientNet family, particularly EfficientNetB7, provided consistent improvements in precision and recall due to its balanced approach to scaling. The best-performing model, EnsembleUpgraded (M1 + M2 + M3), leveraged the strengths of multiple architectures, achieving 99.67% accuracy, far surpassing individual models. Fine-tuned versions of existing architectures (UpgradedResNet152V2, UpgradedInceptionResNetV2, and UpgradedEfficientNetB7) provided noticeable improvements, demonstrating the impact of hyperparameter tuning and transfer learning optimizations. The choice of Adam optimizer plays a crucial role. It is stable and efficient convergence in deep learning models. Unlike traditional Stochastic Gradient Descent (SGD), which requires careful tuning of learning rates, Adam dynamically adjusts learning rates for each parameter. It is well-suited for complex, high-dimensional data such as petrol and diesel. This adaptability ensures that the model can effectively learn from diverse datasets without getting stuck in local minima. ReLU activation function is used throughout the models due to its computational efficiency and ability to mitigate the vanishing gradient problem. While activation functions like sigmoid and tanh are prone to saturation. It leads to slow learning in deep networks. ReLU allows for faster and more effective training. The ensemble model with Adam optimization and ReLU activation function resultant in a highly accurate, robust, and efficient classification for fuel classification.

This study highlights that while individual deep learning models perform well, ensemble approaches significantly enhance classification performance by leveraging multiple feature extraction strategies. The combination of ResNet152V2, InceptionResNetV2, and EfficientNetB7 provided state-of-the-art results, making it a highly effective approach

for fuel classification. Future work could explore further fine-tuning, dataset expansion, and real- time deployment to validate these findings in practical applications.



(**a**) Accuracy (**b**) Precision

(**c**) Recall (**d**) F-Score

Figure 10. Performance comparison of 30 DL models with varied optimizers: Ablation study results.

4. Conclusions and Future Work

In conclusion, the proposed study yields several significant findings. Firstly, an in-depth examination of baseline pre-trained models was conducted. Secondly, the proposed upgraded versions of ResNet152V2, InceptionResNetV2, and EfficientNetB7 yield accuracy of 95.59%, 95.56% and 95.82% respectively for classification of fuel images. Thirdly, the ensemble of these three proposed approaches showcased an outstanding performance of 99.67% accuracy with 99.54% recall. The proposed approach exhibits great potential for various domains including fuel quality control and automotive industries. Future research efforts will concentrate on integrating the proposed algorithm with the hardware like Arduino / Jetson Nano based controller with built-in CMOS camera to identify fuel types. These advancements will further enhance the applicability and effectiveness of the proposed approach in real-world scenarios.

Author Contributions: H.S.—Initial Drafting of the paper—Collection of datasets and choice of their suitability—Formulation of pseudocode; A.G.—Research proposal—Construction of the workflow and model—Review—Improvisation of the proposed model. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Changsen, F.; Liang, S.; Jiaying, W.; Youbing, Z.; Fushuan, W. Short-term Load Forecasting of Distribution Transformer Supply Zones Based on Federated Model-Agnostic Meta Learning. *IEEE Trans. Power Syst.* **2025**, *40*, 31–45.
2. Mao, T.; Chenglin, H.; Jie, C.; Ling, W.; Zhengmao, L. Multi-node Load Forecasting Based on Multi-task Learning with Modal Feature Extraction. *Eng. Appl. Artif. Intell.* **2022**, *112*, 104856.
3. Aqib, M.; Allah, B.; Zulfiqar, H. Fast Learning Through Deep Multi-Net CNN Model For Violence Recognition in Video Surveillance. *Comput. J.* **2020**, *65*, 457–472.
4. Loris, N.; Stefano, G.; Sheryl, B. Handcrafted vs. non-handcrafted features for computer vision classification. *Pattern Recognit.* **2017**, *71*, 158–172.
5. Fadi, M.; Fatma, S. A hybrid method of feature extraction for signatures verification using cnn and hog a multi-classification approach. *IEEE Access* **2023**, *11*, 21873–21882.
6. Sevim, A.; Rifat, K.; Ali, D.; Ercan, K. Classification of apple images using support vector machines and deep residual networks. *Neural Comput. Appl.* **2023**, *35*, 12073–12087.
7. Veluguri, S.; Rajasomashekar, S.; Sarala, B. A Novel Representation and Prediction Initiative for Underground Water by Using Deep Learning Technique of Remote Sensing Images. *Comput. J.* **2022**, *66*, 1784–1801.
8. Sohaib, A.; Ming, Z.; Fengxiao, T.; Yusen, Z. An enhanced deep learning method for multi-class brain tumor classification using deep transfer learning. *Multimed. Tools Appl.* **2023**, *82*, 31709–31736.
9. Srikanta, D.; Prabira, K.; Santi, K. Cervical transformation zone segmentation and classification based on improved inception-resnet-v2 using colposcopy images. *Cancer Inform.* **2023**, *22*, 11769351231161477.
10. Rym, D.; Walid, B. A skin lesion classification method based on expanding the surrounding lesion-shaped border for an end-to-end inception-resnet-v2 classifier. *Signal Image Video Process.* **2023**, *17*, 3525–3533.
11. Mohammadreza, I.; Hamid, R.; Khaled, R. A review of deep transfer learning and recent advancements. *Technologies* **2023**, *11*, 40. [CrossRef]
12. Punuri, S.B.; Kuanar, S.K.; Kolhar, M.; Mishra, T.K.; Alameen, A.; Mohapatra, H.; Mishra, S.R. Efficient net-xgboost: An implementation for facial emotion recognition using transfer learning. *Mathematics* **2023**, *11*, 776. [CrossRef]
13. Jireh, Y.; Khean, T.; Steven, M.; Seuk, W.; Wai, K. State of the art: A review of sentiment analysis based on sequential transfer learning. *Artif. Intell. Rev.* **2023**, *56*, 749–780.
14. Abhisek, R.; Maheshkumar, H.; Balasubramanian, R.; Adel, H. Transfer learning enhanced vision-based human activity recognition: A decade-long analysis. *Int. J. Inf. Manag. Data Insights* **2023**, *3*, 100142.
15. Omneya, A. Tomato leaf disease classification via compact convolutional neural networks with transfer learning and feature selection. *Horticulturae* **2023**, *9*, 149. [CrossRef]
16. Sahebgoud, H.; Lakhan, D. Automated multi-class classification of lung diseases from cxr-images using pre-trained convolutional neural networks. *Expert Syst. Appl.* **2023**, *211*, 118650.
17. Abhronil, S.; Yuting, Y.; Robert, W.; Chiao, L.; Kaushik, R. Going deeper in spiking neural networks: Vgg and residual architectures. *Front. Neurosci.* **2019**, *13*, 95.
18. Guan, Q.; Wang, Y.; Ping, B.; Li, D.; Du, J.; Qin, Y.; Lu, H.; Wan, X.; Xiang, J. Deep convolutional neural network vgg-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: A pilot study. *J. Cancer* **2019**, *10*, 4876–4882. [CrossRef]
19. Shudi, W.; Manman, X.; Ying, S.; Guozhang, J.; Yaoqing, W.; Xin, L.; Zhao, G.; Fan, H.; Li, J.; Zou, G.; et al. Improved single shot detection using densenet for tiny target detection. *Concurr. Comput. Pract. Exp.* **2023**, *35*, e7491.
20. Fawwaz, A.; Tajrial, R.; Arifin, S.; Suwarno. Automated classification of multi-class human protozoan parasites using xception as transfer learning. *Int. J. Intell. Syst. Appl. Eng.* **2023**, *11*, 817–825.
21. Shubhangi, A.; Anupkumar, M.; Olof, O.; Siddhaling, U.; Deepak, D.; Arunkumar, B. Enhanced pre-trained xception model transfer learned for breast cancer detection. *Computation* **2023**, *11*, 59. [CrossRef]
22. Muthu, S.; Krishnaveni, P.K.; Joshua, V.; Malathi, S. Breast cancer detection with resnet50,inception v3, and xception architecture. *J. Pharm. Negat. Results* **2023**, *14*, 60–68.
23. Mingle, X.; Sook, Y.; Alvaro, F.; Dong, S. A comprehensive survey of image augmentation techniques for deep learning. *Pattern Recognit.* **2023**, *137*, 109347.
24. Gaurav, M. Efficient deep learning: A survey on making deep learning models smaller, faster, and better. *Acm Comput. Surv.* **2023**, *55*, 259.
25. Adnan, H.; Sareer, U.; Muhammad, F.; Sanghyun, S. An efficient and robust hand gesture recognition system of sign language employing finetuned inception-v3 and efficientnet-b0 network. *Comput. Syst. Sci. Eng.* **2023**, *46*, 3509–3525.
26. Dostdar, H.; Muhammad, I.; Israr, H.; Roobaea, A.; Saddam, H.; Syed, S. Face mask detection using deep convolutional neural network and mobilenetv2-based transfer learning. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 1536318.

27. Yannik, S.; Dieter, K.; Tobias, M. Survey on deep learning based computer vision for sonar imagery. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105157.
28. Hemachandiran, S.; Ajit, K.; Aghila, G. Advancing real-time fuel classification with novel multi-scale and multi-level MHOG and light gradient boosting machine. *Int. J. Cogn. Comput. Eng.* **2024**, *5*, 398–405.
29. Muhammad, S.; Nouman, A.; Naeem, I.; Muhammad, U.; Faisal, M.; Imran, R.; Musaddiq, U.; Baig, M.J.A.; Baig, S.; Salaria, U.A. Deep Learning in Age-invariant Face Recognition: A Comparative Study. *Comput. J.* **2020**, *65*, 940–972.
30. Morteza, R.; Mohammad, F.; Babak, A.; Navid, Y. An Ensemble-Based Credit Card Fraud Detection Algorithm Using an Efficient Voting Strategy. *Comput. J.* **2021**, *65*, 1998–2015.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.