

Article

Dual-Dataset Deep Learning for Improved Forest Fire Detection: A Novel Hierarchical Domain-Adaptive Learning Approach

Ismail El-Madafri ^{1,*} , Marta Peña ²  and Noelia Olmedo-Torre ¹ ¹ Department of Graphic and Design Engineering, Universitat Politècnica de Catalunya, C. Eduard Maristany 16, 08019 Barcelona, Spain; n.olmedo@upc.edu² Department of Mathematics, Universitat Politècnica de Catalunya, Av. Diagonal 647, 08028 Barcelona, Spain; marta.pena@upc.edu

* Correspondence: ismail.el.madafri@upc.edu

Abstract: This study introduces a novel hierarchical domain-adaptive learning framework designed to enhance wildfire detection capabilities, addressing the limitations inherent in traditional convolutional neural networks across varied forest environments. The framework innovatively employs a dual-dataset approach, integrating both non-forest and forest-specific datasets to train a model adept at handling diverse wildfire scenarios. The methodology leverages a novel framework that combines shared layers for broad feature extraction with specialized layers for forest-specific details, demonstrating versatility across base models. Initially demonstrated with EfficientNetB0, this adaptable approach could be applicable with various advanced architectures, enhancing wildfire detection. The research's comparative analysis, benchmarking against conventional methodologies, showcases the proposed approach's enhanced performance. It particularly excels in accuracy, precision, F1-score, specificity, MCC, and AUC-ROC. This research significantly reduces false positives in wildfire detection through a novel blend of multi-task learning, dual-dataset training, and hierarchical domain adaptation. Our approach advances deep learning in data-limited, complex environments, offering a critical tool for ecological conservation and community protection against wildfires.



Citation: El-Madafri, I.; Peña, M.; Olmedo-Torre, N. Dual-Dataset Deep Learning for Improved Forest Fire Detection: A Novel Hierarchical Domain-Adaptive Learning Approach. *Mathematics* **2024**, *12*, 534. <https://doi.org/10.3390/math12040534>

Academic Editors:

Viacheslav Voronin, Yigang Cen and Evgenii Semenishchev

Received: 28 November 2023

Revised: 4 February 2024

Accepted: 7 February 2024

Published: 8 February 2024



Copyright: © 2024 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/>).

Keywords: forest fire detection; remote-sensing forest monitoring; deep learning; domain adaptation; hierarchical learning; convolutional neural networks; transfer learning; dual-dataset training; multi-task learning; false positives

MSC: 68T05; 68T07; 68T45

1. Introduction

Forests cover nearly one-third of Earth's land, playing a critical role in maintaining ecological balance as significant carbon sinks and biodiversity repositories [1]. They are crucial in climate regulation and supporting community livelihoods [2]. Nevertheless, these ecosystems are increasingly threatened by more frequent and severe forest fires, undermining their ecological services [3]. Addressing this issue is pivotal to achieving the United Nations' Sustainable Development Goals, specifically SDG 13, which targets climate change action [4], and SDG 15, focusing on sustainable terrestrial ecosystems management [5]. Prompt and effective fire detection is essential for protecting these vital natural resources, aligning with these global objectives [6].

The importance of rapid-fire detection in forest management is paramount. Timely identification enables a quick response, minimizing fire damage and resource loss. Conversely, delays or inaccuracies in detection can result in extensive habitat destruction, ecosystem disruption, land destruction, and loss of unique species [7], adversely affecting human communities, their livelihoods, and long-term health due to smoke and pollutants.

In response to these challenges, there has been a notable transition from traditional methods to more advanced technologies in forest fire detection [6,8]. While historical

reliance was on lookout towers and aerial surveillance, the integration of deep learning (DL) [9], a complex subset of machine learning (ML), is now prevalent. DL excels in analyzing large datasets and detecting subtle patterns, capabilities often beyond the reach of traditional methods, especially in computer vision applications. This advancement is significantly enhancing wildfire detection capabilities [10].

Sensor-based systems and advanced computer vision techniques, supplementing traditional methods, show promise but face issues like false positives and limited coverage [11]. Convolutional neural networks (CNNs), fundamental in deep learning (DL), significantly enhance fire detection accuracy using imagery data, as evidenced by studies such as [12,13]. The use of multi-modal data, combining thermal and RGB imagery, further improves detection capabilities [14]. Additionally, recurrent neural networks (RNNs), which excel in processing sequential data, enable more nuanced analyses of fire progression over time. The synergy of CNNs and RNNs with multi-modal data marks a significant leap in fire detection technology [15].

However, deploying DL models in forests faces unique challenges from weather, lighting, and seasonal variability, affecting fire and smoke visibility and, thus, model performance [15]. These models' adaptability to the dynamic and unpredictable forest environment—and the risk of overfitting to controlled datasets—underscores a crucial area for research [16,17]. Challenges such as false positives from confounding natural elements like fog and sunlight, and the sporadic nature of forest fires causing data scarcity and imbalance, significantly impact CNN training [15,16]. This necessitates models that can swiftly and accurately process complex visual data in real time, adapting to new fire scenarios.

The unpredictable behavior of forest fires, including intensity shifts and variability in fire sizes and occlusions by terrain, demands training on diverse fire behaviors to ensure models recognize fires from various perspectives, overcoming uniform training's limitations [6,15,18,19].

To address these challenges, novel advancements and the integration of innovative approaches such as multi-task learning (MTL) models [20], multi-dataset training, and domain adaptation techniques have been proposed [21].

MTL models trained to concurrently execute multiple tasks offer a nuanced understanding of fire dynamics. This approach has the potential to mitigate the tendency of models to become overly specialized, a notable challenge when training is confined to controlled datasets [22]. Recent developments in this field are exemplified in the study detailed in Reference [15]. The authors, working on the wildfire dataset, have proposed an MTL framework that incorporates sub-classes of naturally occurring elements, such as fog, clouds, and sunlight. The empirical evaluation of this model demonstrated superior performance over traditional classification methods in key metrics, most notably in significantly reducing false alarm rates. This reduction is crucial, as false alarms in forest fire detection can lead to not just resource misallocation, but also critical delays in emergency response. Table 1 gives an overview of the model approach, performance results, and the study's limitation of the study.

Table 1. Overview of El-Madafri et al. (2023)’s model approach, performance results, and limitations [15].

Approach	Results	Limitations	Limitations
Employs a multi-task learning approach based on MobileNetV3 and integrating multi-class confounding elements [15].	Accuracy: Mean = 0.8766, Std = 0.0147 Precision: Mean = 0.7974, Std = 0.0206 Recall: Mean = 0.9171, Std = 0.0273 F1-score: Mean = 0.8526, Std = 0.0146 ROC-AUC: Mean = 0.8842, Std = 0.0127	Time and Computational Resources Consideration: The study does not explicitly account for the time and computational resources required for training, and the inference of the different used models, which are crucial for assessing the model’s practical applicability and efficiency in real-world scenarios. Model Performance Enhancement: The performance metrics on the test dataset of the wildfire dataset indicate a need for further improvement in model effectiveness. Regularization and Overfitting: There is a lack of consideration for regularization strategies and the potential issue of overfitting, which are critical for the model’s reliability and generalizability.	4,052,133

The results are based on five independent runs.

Domain adaptation techniques [21] are central to our approach, focusing on minimizing feature distribution differences across domains. This process, involving feature alignment, enables the model to concentrate on fundamental fire and smoke characteristics consistent across environments, rather than overfitting to dataset-specific features. Techniques include shared layers within neural networks, encouraging the extraction of relevant features across both source and target domains. This approach, serving as a form of regularization, helps reduce overfitting and ensures robustness in diverse real-world scenarios [23]. Specifically, in wildfire detection, domain adaptation becomes relevant due to significant environmental variability. For instance, Reference [23] demonstrates an effective approach in video smoke detection by aligning features of synthetic and real-world smoke images, thereby improving model accuracy in varied conditions. This study exemplifies the practical benefits of domain adaptation in enhancing DL models’ applicability.

Another significant contribution is from Reference [24], which explores maintaining CNN performance on original tasks while applying transfer learning to new, smaller datasets. Implementing transfer learning on models like VGG16, InceptionV3, and Xception, with Xception showing notable improvements, the study demonstrates the adaptability of these models to new tasks without losing original capabilities. This approach is particularly relevant for region-specific forest contexts, addressing the risk of overfitting to homogeneous data.

Despite these advancements, the unpredictable nature of forest environments continues to pose significant challenges, prompting ongoing research and development. The present study introduces a groundbreaking hierarchical domain-adaptive learning approach, leveraging the synergies of multi-task learning, transfer learning [25], and domain adaptation principles.

Our novel contribution lies in the development of a dual-task deep-learning model that utilizes dual datasets: one auxiliary set featuring fire versus no-fire images from non-forest contexts, and a primary set within forest contexts. This approach is designed to overcome the limitations of traditional single-dataset training methods, particularly under varied and challenging forest fire scenarios. By incorporating the EfficientNetB0 [26] architecture, our model benefits from shared layers for general feature extraction across datasets, alongside dedicated layers for capturing forest-specific nuances, thereby preventing early overspecialization and fostering a balanced learning environment.

The significance of our work extends beyond the technical innovation of the model's architecture. It addresses critical questions in wildfire detection: the efficacy of hierarchical domain-adaptive learning compared to traditional training methods, the role of shared layers in leveraging dual-dataset strengths, and the model's potential in significantly reducing false positives and negatives. We anticipate that our findings will not only enhance wildfire detection within forest contexts but also offer insights into model generalization that could be applied to other complex detection scenarios.

Moreover, this research's implications are broad, touching on various aspects of computer vision and DL in data-scarce and complex environments. It also contributes to wildfire management and conservation strategies by supporting more accurate and timely detection, enabling efficient firefighting resource deployment, and reducing the overall impact of wildfires.

In summarizing, our study's main contribution is the innovative use of hierarchical domain-adaptive learning to significantly improve wildfire detection, marking a pivotal step forward in the application of DL techniques to environmental conservation and disaster management.

The subsequent section details the methodology, encompassing the dual-dataset approach, model architecture, and evaluation strategies, to validate the effectiveness of the hierarchical domain-adaptive learning framework in enhancing wildfire detection.

2. Materials and Methods

2.1. Datasets Description and Preparation

This study employs a dual-dataset strategy, integrating an auxiliary dataset for non-forest scenarios alongside a primary dataset focused on forest environments. This approach is designed to reduce overfitting to the primary dataset and improve the accuracy of image classification within forest contexts.

Both datasets were meticulously balanced in terms of class representation, with the auxiliary and primary training and validation sets containing 2312 and 402 images, respectively. This balance is critical for effective training, allowing for the simultaneous processing of image batches from both datasets without cross-mixing. The performance of various model configurations is assessed using a test dataset comprising 410 images from the primary dataset, serving as a benchmark for comparative analysis.

2.1.1. Auxiliary Dataset

The auxiliary dataset, integral to the study's methodology, is composed of images from two public repositories, chosen for their relevance to non-forest fire and smoke scenarios. The first, 'Dataset for Fire and Smoke Detection' by FireGasSmoke from Roboflow Universe [27], offers a diverse range of urban fire and smoke images, including burning buildings and vehicles (see Figure 1). The second, the FiSmo collection from mtcazzolato's GitHub [28], provides additional variety with fire and smoke images in varied non-forest settings. Despite not being forest-fire-specific, the dataset, recommended by the review article 'Deep Learning Approaches for Wildland Fires Remote Sensing: Classification, Detection, and Segmentation' [11], is commonly used in the field.

The selection of these sources was driven by the difficulty in sourcing publicly available, comprehensive datasets of non-forest fire and smoke images, crucial for our model's training. They provide a wide spectrum of fire and smoke scenarios, including large-scale urban fires and burning structures, as well as no-fire images in similar settings. This combination of

sources enriches the training data, making the auxiliary dataset a crucial component in the context of the study's approach (see Section 2.3).

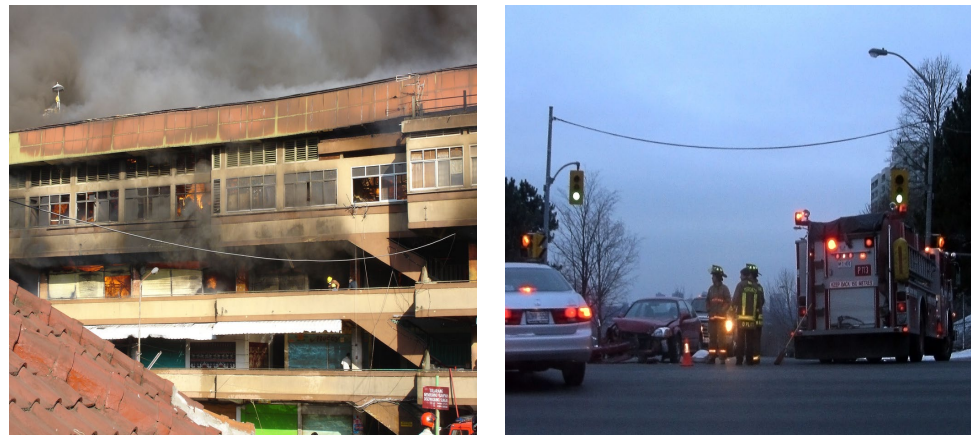


Figure 1. It displays images from the auxiliary dataset, featuring two scenarios: one with fire and one without, both set in a non-forest context.

2.1.2. Primary Dataset

The study leverages the comprehensive wildfire dataset [15] as its primary dataset, consisting of 2700 RGB images from aerial and ground-based perspectives, gathered from diverse online platforms including government agencies, Flickr, and Unsplash. This dataset encompasses a wide array of forest fire images, capturing diverse environmental conditions, forest types, and geographical regions, and includes a range of environmental and situational variability like different topographies, weather conditions, fire characteristics, and camera perspectives (see Figure 2). The images in this dataset vary significantly in resolution, with an average of 4057×3155 pixels, a range spanning from 153×206 to $19,699 \times 8974$ pixels, offering detailed visuals essential for effective deep-learning analysis in forest fire detection. To counteract the initial class imbalance (1047 fire instances vs. 1653 no-fire instances) [29], augmentation techniques [30] were employed. Employing augmentation techniques to balance the training set is pivotal in enabling more generalizable conclusions from the experiments, as it ensures a fair representation of both fire and no-fire instances, mitigating any skew in the learning process [31]. Additionally, the dataset integrates key confounding elements such as atmospheric conditions, vegetation changes, and lighting variations, enhancing its realism and challenge for more accurate forest fire detection.



Figure 2. It displays images from the primary dataset, featuring two scenarios: one with fire and one without, both set in a forest context.

2.1.3. Consistent Data Preparation

All images across both datasets were uniformly preprocessed, including normalization and resizing to 224×224 pixels. This standardization is crucial for ensuring consistent input data quality, which significantly improves the efficiency of the DL models.

2.2. Base Model Selection: EfficientNetB0

In addressing the complex challenges of forest fire detection, this study employs a customized version of the EfficientNetB0 model. The choice of EfficientNetB0 over other pre-trained networks is motivated by several key factors that align closely with the specific requirements of our application.

EfficientNetB0 is renowned for its balance of computational efficiency and accuracy [32], a crucial trait for real-world deployment in forest fire scenarios. Furthermore, unlike many state-of-the-art architectures that feature complex interconnections between different blocks, EfficientNetB0's architecture is uniquely streamlined. This simplicity is vital for the study, where the model is divided into shared layers and task-specific layers. The absence of intricate block interconnections in EfficientNetB0 reduces the complexity of managing these layers, thereby enhancing the practicality and scalability of the system [26].

Another standout feature of EfficientNetB0 is its scalable architecture, which uniformly scales network depth, width, and resolution [26]. This scaling approach is particularly advantageous as it optimizes performance without significantly increasing the computational load. The model's depthwise separable convolutions and squeeze-and-excitation blocks further enhance its capability to efficiently capture relevant image patterns, crucial for varied forest fire detection scenarios.

Furthermore, the design of EfficientNetB0 facilitates effective feature extraction, aligning with our study's objective of developing a practical and scalable forest fire detection system based on computer vision. Its lightweight design is of paramount importance for real-world deployment, where computational resources and response times are critical factors [26].

In comparison to other pre-trained models, EfficientNetB0 offers a unique combination of lightweight architecture and effective feature extraction capabilities. This makes it particularly suited for the novel training approach proposed in the study, which emphasizes efficiency and practicality in a real-world setting.

2.3. Architecture Customization of EfficientNetB0 for Dual-Task Learning

This section details the adaptation of the EfficientNetB0 framework into a dual-branch architecture, designed to support a dual-task learning strategy. The primary branch, central to the study, is trained on the wildfire dataset comprising both fire and no-fire forest images. In parallel, an auxiliary branch processes a dataset encompassing non-forest fire and no-fire scenarios. A key innovation in this design is the incorporation of shared layers, utilized by both models. These shared layers are engineered to extract general features applicable to both datasets, enhancing this part of the model's versatility and efficacy in varied fire detection scenarios. Each branch independently computes its loss function, allowing for specialized learning while maintaining a cohesive overall structure. The shared layers, adept at extracting general features beneficial for both datasets, are updated simultaneously using the combined losses from each model. This strategy of loss integration for the shared layers is devised to mitigate the risk of overfitting the primary model to its specific dataset. By integrating diverse learning inputs from the auxiliary dataset, the model is better equipped to generalize, enhancing its performance in a variety of forest fire detection scenarios. This architectural customization underpins our approach to tackling the complexities of forest fire detection, blending domain-specific learning with general feature extraction. Figure 3 illustrates the training steps of the main dual-task learning process and the information flow through both shared and task-specific layers.

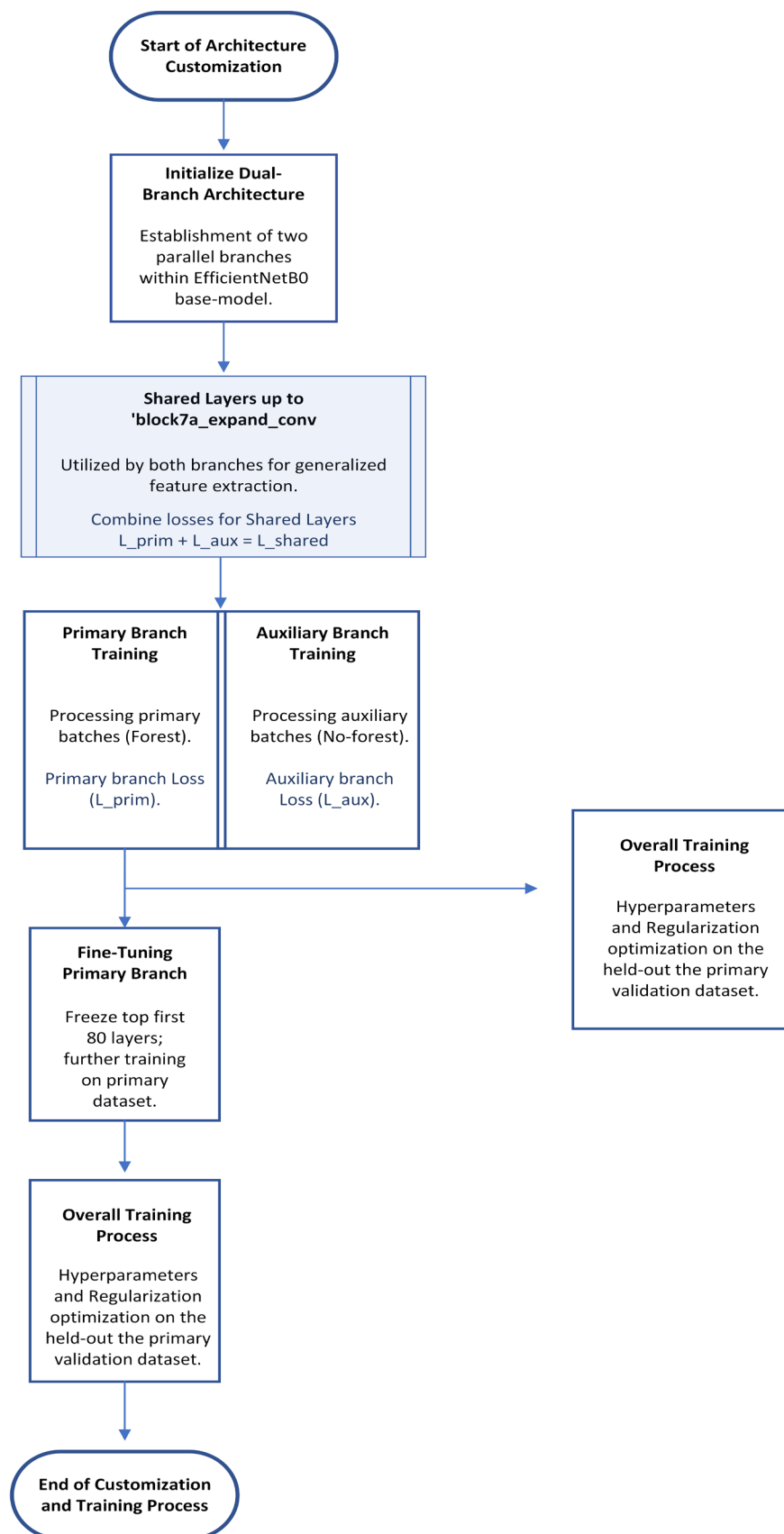


Figure 3. It displays an overview of the main dual-task learning steps' training and the flow of information through shared and task-specific layers.

2.3.1. Shared Layers

In the customized architecture of the EfficientNetB0 model, the shared layers up to ‘block7a_expand_conv’ play a pivotal role. These layers were selected for their capacity to learn generalized features, effectively processing inputs from both the primary (forest) and auxiliary (non-forest) datasets. The depth of these layers is strategically chosen to enable comprehensive feature extraction across diverse environmental scenarios. This design decision, to utilize layers up to this specific last layer’s block, was guided by the interconnected nature of the EfficientNetB0 architecture. By aligning the shared layers with one of the final layers of a block constituting the base model, an optimized balance was achieved. This allows for effective integration of dual learning paths without the complexities of splitting or modifying the fundamental structure of the EfficientNetB0. Through this approach, the model is expected to exhibit enhanced adaptability and performance in detecting forest fires, leveraging the strengths of both datasets while maintaining architectural integrity.

In the customized architecture, the shared layers process data from both the primary and auxiliary models during training. In this methodology, the central focus is on the strategic optimization of weight adjustments in the shared layers, guided by the sum of losses from both the primary and auxiliary datasets (see Equation (1)). By exposing these layers to a diverse array of inputs, the primary model is less prone to becoming overly specialized to its training data’s characteristics.

This foundational approach is anticipated to enhance the model’s effectiveness in accurately detecting and classifying forest fires, including those in scenarios not directly covered by the training data. Figure 4 presents a schematic representation of how EfficientNetB0’s architecture has been tailored for dual-task learning.

$$L_{prim} + L_{aux} = L_{shared} \quad (1)$$

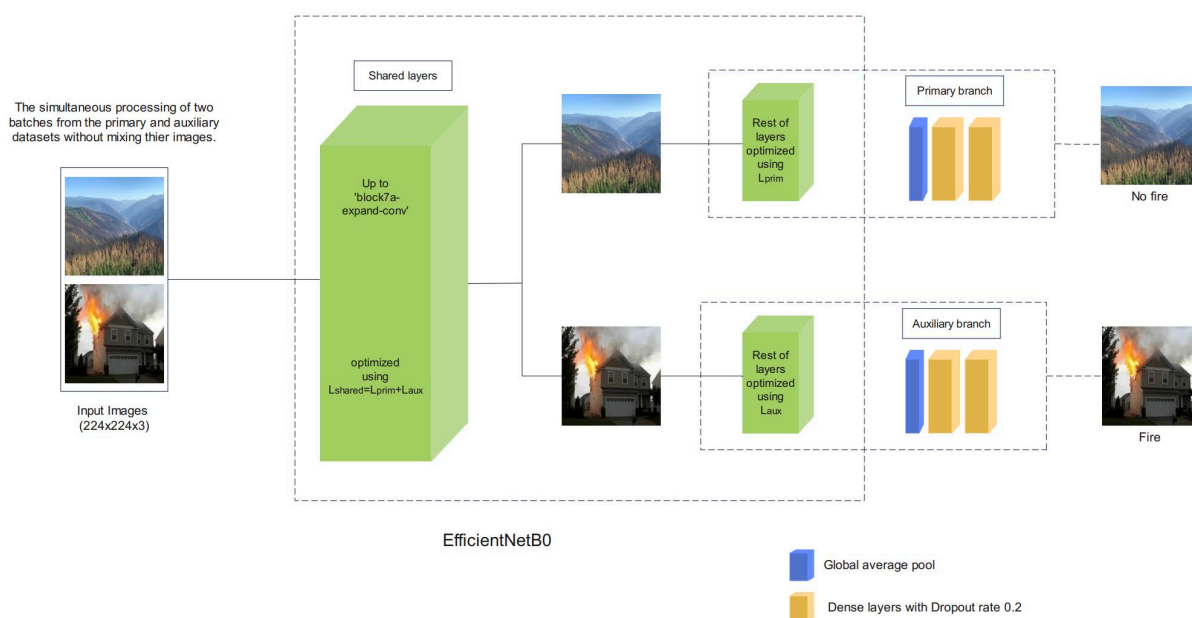


Figure 4. It displays an overview of the architecture customization of EfficientNetB0 for dual-task learning.

2.3.2. Fine-Tuning

Both branches use a sequential stack of GlobalAveragePooling2D and dense layers. Post initial training, fine-tuning is performed by unfreezing the top first 80 layers of the primary model and training it further on the primary dataset. This fine-tuning process aims to enhance the model’s specialization in forest fire detection while maintaining the learned generalized features from the top first 80 layers of the shared layers. Figure 5 showcases the refined architecture of the primary model post fine-tuning.

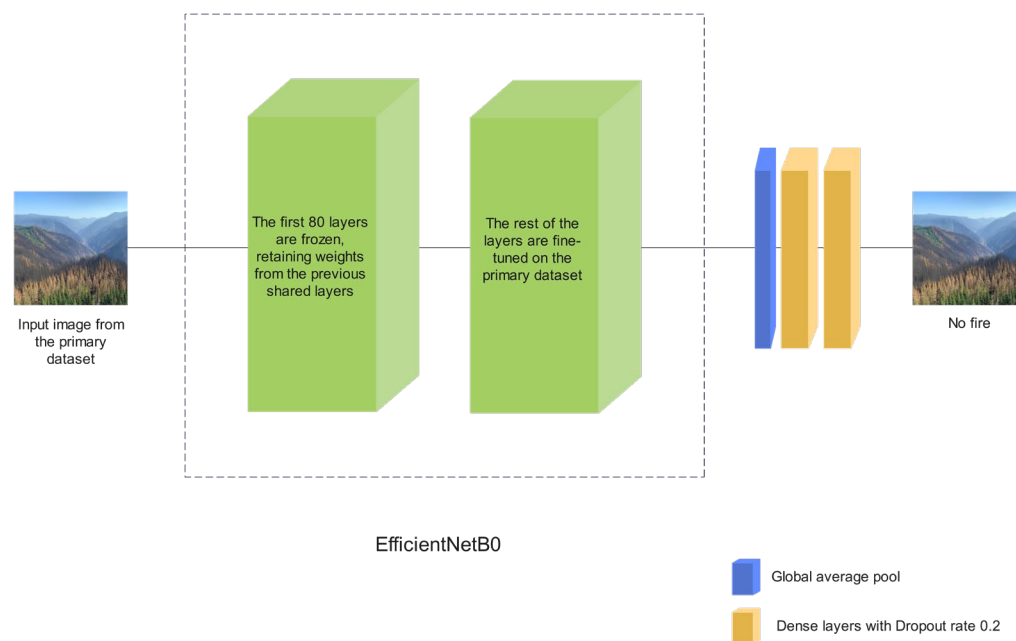


Figure 5. It displays an overview of the final primary model following its fine-tuning.

2.3.3. Non-Mixing of Datasets

In the custom training loop, both the auxiliary and primary batches are processed simultaneously, ensuring that the shared layers learn from both datasets without mixing images. This separation is crucial for maintaining the distinct characteristics of each dataset, enabling the model to learn specific features relevant to both non-forest and forest fire scenarios.

The outputs from the auxiliary and primary branches are computed separately for auxiliary and primary inputs, respectively.

2.4. Training Process

The training of the models was conducted using the Keras framework with TensorFlow and GPU support. Each training regimen was carefully tailored to maximize learning from both the auxiliary and primary datasets while preserving their distinct characteristics. After the convolutional layers of the pre-trained EfficientNetB0 base model, a global average pooling layer reduced the dimensionality of the outputs, setting the stage for the fully connected prediction layer. To enhance generalization, a dropout regularization rate of 0.2 was implemented following systematic testing of various regularization strategies on the held-out validation dataset. Indeed, once the learning rate and number of epochs were optimized for the model trained from scratch (see Section Hyperparameters Optimization), an assessment was undertaken by evaluating the impact of incremental additions of traditional regularization methods. Dropout was meticulously explored within these custom layers, testing rates from 0.1 to 0.5 in increments of 0.1, with the aim of discerning the rate that fostered the best balance between learning capacity and generalization on the held-out validation dataset. Concurrently, batch normalization was exclusively examined for these newly added layers, considering that the base model already included batch normalization in its architecture. The dropout rate of 0.2 was found to strike the optimal balance between complexity and performance, leading to its adoption across all subsequent experiments.

Model performance was evaluated using accuracy, precision, recall, F1-score, specificity, false negative rate (FNR), Matthews correlation coefficient (MCC), and ROC-AUC score. These metrics provided a comprehensive assessment of the models' effectiveness in fire detection, encompassing aspects of precision, recall balance, and class differentiation abilities.

Hyperparameters Optimization

A validation strategy using a held-out dataset assessed performance on unseen data. Various learning rates (10^{-2} , 10^{-3} , 10^{-4} , and 10^{-5}) were tested to identify the most effective one, which was then applied to the test dataset.

In fine-tuning all the EfficientNetB0-based models, three configurations were explored for the number of frozen layers: 80, 160, and 238. This division into three parts was designed to systematically evaluate the impact of fine-tuning different proportions of the model's layers, based on the architecture of EfficientNetB0 and transfer learning principles [33–35]. Furthermore, this approach considered balancing computational efficiency with performance enhancements [36]. Freezing 80 layers, approximately one-third of the total layers, aimed to balance leveraging pre-trained features with adaptability for forest fire detection. Extending the fine-tuning to 160 layers, roughly two-thirds of the model, tested the model's reliance on more extensive pre-trained features for effectiveness. Finally, freezing all 238 layers, while keeping the dense layers trainable, served as a baseline to assess the pre-trained model's capabilities without further convolutional layer adaptation, focusing on learning higher-level abstractions relevant to the task. This structured approach allowed us to evaluate the incremental benefits of fine-tuning additional layers and identify the optimal balance for our specific application.

In this study, all experiments employed the adaptive moment estimation (ADAM) optimizer as the default choice. Due to limitations in resources, we did not explore alternative optimizers, focusing instead on maximizing the efficiency and effectiveness within the constraints of our available resources.

The batch size was set to 32, optimizing computational efficiency and gradient stability. Training epochs were varied in 5-epoch increments to identify the optimal number for each configuration, adhering to the principle that fine-tuning requires fewer epochs than full training. This careful variation in epochs was crucial to achieve each model's peak performance without overfitting, ensuring robustness and generalizability. All the chosen hyperparameters for each approach, including learning rates, number of epochs, batch sizes, and the number of frozen layers, will be summarized in corresponding tables in the results section (see Section 3).

2.5. Comparative Analysis of Training Approaches

In this study, we conducted a detailed comparison between the novel hierarchical domain-adaptive learning approach and several baseline training approaches. This comparison aimed to explore their relative effectiveness in forest fire detection using deep-learning methodologies. Such comparative analyses are crucial in understanding the potential advantages and limitations of our proposed method within the broader spectrum of conventional training paradigms.

2.5.1. Baseline Approaches

Training from Scratch on the Target Dataset

In the first baseline approach, the EfficientNetB0 model, initially initialized with pre-trained weights from the ImageNet database, was retrained from scratch exclusively on the wildfire dataset. This methodology allowed for a foundational benchmark against which the incremental benefits of further domain-specific fine-tuning and advanced training techniques could be evaluated. By retraining the model with the pre-existing ImageNet-based knowledge and, then, focusing solely on the target dataset, the study provided insights into how effectively the EfficientNetB0 architecture could adapt to the task of RGB forest fire detection represented by the wildfire dataset, starting with a broad knowledge base and then specializing solely on the specific dataset.

Choosing this approach as a baseline is crucial to the study's methodology. It offers a clear perspective on the model's adaptability and learning efficiency when transitioning from a generalist initialization base to a highly specialized context. This baseline is key in understanding the relative enhancements brought about by further fine-tuning and

advanced training strategies in subsequent approaches. It establishes a comparative framework to measure the improvements or changes that result from additional domain-specific training, essential for a comprehensive evaluation of different training strategies in the specialized field of forest fire detection.

Transfer Learning with Optimized Frozen Layers

This baseline approach employed transfer learning, initializing the model with pre-trained EfficientNetB0 weights from the ImageNet database [37]. A critical element of this method was the optimization of frozen layers [38], with 80 layers frozen and training conducted over 15 epochs. This optimization aimed to balance the retention of learned features from the ImageNet dataset with the necessity of adapting to the specific characteristics of forest fire images.

The choice of this baseline is justified by its relevance to the common practice in deep learning, where models are often adapted from general tasks to more specialized domains. By utilizing a well-established model pre-trained on a diverse dataset like ImageNet, the study investigates the effectiveness of transfer learning in adapting to a niche domain. This approach serves as a crucial comparison point, assessing how effectively a model, pre-tuned on a broad dataset, can be adapted through layer optimization to the specialized task of forest fire detection. Furthermore, this method allows for an exploration of the trade-off between feature retention and domain-specific adaptation, a key consideration in the application of transfer learning to specialized areas [21,33].

Sequential Transfer Learning

In the third approach, the model was initially trained for 15 epochs on the auxiliary dataset comprising non-forest fire and no-fire scenarios. This phase was followed by a fine-tuning process on the primary wildfire dataset, during which 80 layers of the model were frozen. The fine-tuning was conducted over 10 epochs, allowing the model to integrate specific characteristics of forest fire scenarios while retaining general fire and no-fire features learned from the auxiliary dataset. This method aims to facilitate a nuanced adaptation to the domain of forest fire detection and provided a basis for a detailed comparison.

By initiating the training on the auxiliary dataset and subsequently transitioning to the primary one, the study enables the comparison of this approach with the second baseline (see the second subsection of Section 2.5.1), where fine-tuning was conducted using weights pre-trained on the ImageNet database. This comparison was critical in evaluating the impact of source dataset proximity on the effectiveness of the transfer learning process, particularly in the context of forest fire detection.

The choice of this baseline was strategic, ensuring that the improvements identified in the study were not merely a result of the diversity introduced by the auxiliary dataset, but also a consequence of the specific training methodology employed. This approach enabled a thorough assessment of the hierarchical domain-adaptive learning approach, isolating the influence of the training method from the dataset characteristics, thereby strengthening the validity of the study's findings.

Mixed Dataset Training

In this last baseline approach, distinct from the other methodologies, the model was trained on a dataset that was a direct combination of images from both the auxiliary and primary datasets. Over a course of 25 epochs, the model was exposed uniformly to this integrated dataset, providing a comprehensive view of both forest and non-forest fire scenarios in a single training phase. This method stands apart from the second and third baselines and the hierarchical domain-adaptive learning approach, where the datasets are used in a sequential or hierarchical manner.

The rationale behind selecting this baseline is twofold. First, it provides a unique perspective on the model's ability to learn from a homogeneously mixed dataset, as opposed to the layered or sequential learning used in other approaches. This allows for an evaluation

of how the simultaneous exposure to both datasets affects the model's learning dynamics and its subsequent performance in forest fire detection. Second, it serves as a crucial point of comparison to discern the effectiveness of a straightforward dataset combination versus more complex training methodologies employed in the study. By training the model on this unified dataset, the study aims to isolate the impact of a diverse yet non-hierarchical dataset structure on model performance.

This mixed dataset training approach thus offers a distinct benchmark within the study. It aids in understanding whether a simple combination of datasets could be as effective as or more effective than sequential or hierarchical training methods in enhancing the model's accuracy and robustness in forest fire detection scenarios.

2.6. Statistical Analysis of Results

To ensure the reliability and generalizability of our findings, outcomes were derived from five independent runs for each model across a range of training scenarios. For each method, the reported results include the average values and standard deviation after these runs, offering a balanced view of models' performance. In emphasizing statistical robustness, bootstrapping methods were specifically applied in the critical comparison between the study's proposed approach and the most performant baseline method. This choice was guided by the need for a robust statistical framework to handle potential variances and ensure the precision of our findings. Bootstrapping allowed for the calculation of 95% confidence intervals for key performance metrics, such as accuracy, precision, recall, F1-score, and AUC-ROC. This targeted application of bootstrapping aims to enhance the thoroughness of the comparative analysis, ensuring that the conclusions drawn are not only statistically sound but also relevant and reliable.

This statistical approach enabled a detailed and balanced comparison across the various models. By doing so, it shed light on the relative strengths and weaknesses of each training methodology in the context of forest fire detection. While this analysis provided valuable insights into the effectiveness of different approaches, it primarily served as a foundation for further exploration in this critical domain. The findings contribute to the ongoing discourse in the field, offering a basis for future research and practical applications in forest fire detection and other areas where precise and reliable detection is crucial.

3. Results

Upon applying the different discussed approaches to the wildfire dataset, we observed significant potential for enhancements in the model's capacity to detect forest fires and handle false alarms. Central to these improvements is the implementation of a novel hierarchical domain-adaptive learning strategy. This approach significantly bolstered the model's robustness, markedly reduced the incidence of false alarms, and enhanced its adaptability to diverse and challenging conditions inherent in RGB forest fire detection.

This study methodically examines four foundational approaches within the EfficientNetB0 framework: training from scratch, transfer learning with pre-trained weights, sequential transfer learning, and the mixed dataset approach. Each approach has been evaluated for both its theoretical foundations and practical applicability in the realm of forest fire detection.

To facilitate a balanced and thorough comparison, the performance of these models was carefully analyzed across several metrics, notably precision and recall, averaged over five independent runs. The following subsections will delve into these metrics, offering a detailed exploration of the comparative strengths and weaknesses of each method. The aim is to shed light on their potential contributions to advancing the field of wildfire detection, as indicated by our analysis of the wildfire dataset.

3.1. Performance of Baseline Training Approaches

In this section, the findings from the metrics results are crucial in informing the selection of an effective baseline model. This approach will subsequently act as a benchmark for com-

paring the performance of the proposed hierarchical domain-adaptive learning strategy. The goal is to assess the performances of traditional approaches while simultaneously establishing a framework to understand the potential enhancements offered by the novel method.

3.1.1. Training from Scratch

Initially, we explore the training-from-scratch approach, where the EfficientNetB0 model is trained exclusively on the wildfire dataset. This method serves as a foundational benchmark, allowing us to assess the benefits and limitations of leveraging pre-trained models and domain-specific fine-tuning in subsequent approaches. The performance metrics for this approach are shown in Table 2, offering a baseline against which we can compare more advanced methodologies.

Table 2. Performance metrics means using the training-from-scratch approach.

Metrics	Training from Scratch
Accuracy	Mean = 0.9033, Std = 0.0111
Precision	Mean = 0.8372, Std = 0.0103
Recall	Mean = 0.9509, Std = 0.0200
F1-score	Mean = 0.8904, Std = 0.01312
Specificity	Mean = 0.8828, Std = 0.0078
FNR	Mean = 0.0490, Std = 0.0200
MCC	Mean = 0.8184, Std = 0.0224
ROC-AUC	Mean = 0.9168, Std = 0.01020

These results present the average values and standard deviation after five runs of performance metrics for the training-from-scratch approach.

The following Table 3 outlines the optimized hyperparameters for the approach, with configurations derived from the held-out validation dataset of the wildfire dataset.

Table 3. Hyperparameter configuration for the training-from-scratch approach.

Hyperparameter	Value
Learning rate	1×10^{-3}
Optimizer	Adam
Batch size	32
Epochs	15
Fine-tune at the layer number	1

3.1.2. Transfer Learning with Pre-Trained Weights

The second approach in the comparative analysis employed transfer learning, wherein the model was initialized with pre-trained EfficientNetB0 weights derived from the ImageNet dataset. A critical component of this method was the strategic optimization of the number of layers to be frozen. This delicate balance aimed at retaining the beneficial learned features from the ImageNet dataset while effectively adapting the model to the specific challenges posed by forest fire imagery.

The performance metrics of this approach, as illustrated in Table 4, offer insights into the efficacy of transfer learning in the context of the study.

Table 4. Performance metrics means using the transfer learning with pretrained weights approach.

Metrics	Transfer Learning with Pre-Trained Weights
Accuracy	Mean = 0.8565, Std = 0.0140
Precision	Mean = 0.7647, Std = 0.025
Recall	Mean = 0.9132, Std = 0.0262
F1-score	Mean = 0.8318, Std = 0.0123
Specificity	Mean = 0.8206, Std = 0.0290
FNR	Mean = 0.0868, Std = 0.0262
MCC	Mean = 0.7180, Std = 0.0210
ROC-AUC	Mean = 0.8669, Std = 0.0106

These results present the average values and standard deviation after five runs of performance metrics for the transfer learning with pretrained weights approach.

Table 5 outlines the optimized hyperparameters for the approach, with configurations derived from the held-out validation set of the wildfire dataset.

Table 5. Hyperparameter configuration for the transfer learning with pretrained weights approach.

Hyperparameter	Value
Learning rate	1×10^{-3}
Optimizer	Adam
Batch size	32
Epochs	15
Fine-tune at the layer number	80

3.1.3. Sequential Transfer Learning

The Sequential Transfer Learning approach represents another key facet of the comparative study. This method involves a step-wise process where the model, initially trained on a related domain, the auxiliary dataset, is further refined on the primary dataset. This approach is designed to leverage the cumulative learning from both domains, potentially enhancing the model's adaptability and accuracy in detecting forest fires.

Table 6 details the performance metrics for this approach. By examining these metrics, we gain valuable insights into the effectiveness of sequential transfer learning in the context of forest fire detection, particularly in terms of its ability to harness previously learned features for improved detection accuracy.

Table 6. Performance metrics means using the sequential transfer learning approach.

Metrics	Sequential Transfer Learning
Accuracy	Mean = 0.8939, Std = 0.0076
Precision	Mean = 0.8383, Std = 0.0107
Recall	Mean = 0.9195, Std = 0.0188
F1-score	Mean = 0.8769, Std = 0.0128
Specificity	Mean = 0.8876, Std = 0.0077
FNR	Mean = 0.0804, Std = 0.0188
MCC	Mean = 0.7955, Std = 0.0216
ROC-AUC	Mean = 0.9035, Std = 0.0114

These results present the average values and standard deviation after five runs of performance metrics for the sequential transfer learning approach.

Table 7 outlines the optimized hyperparameters for the approach, with configurations derived from the validation set of the wildfire dataset.

Table 7. Hyperparameter configuration for the sequential transfer learning approach.

Hyperparameter	Value
Learning rate (auxiliary dataset)	1×10^{-2}
Learning rate (primary dataset)	1×10^{-4}
Optimizer (both)	Adam
Batch size	32
Epochs (auxiliary dataset)	10
Fine-tune at the layer number (auxiliary dataset)	1
Epochs (primary dataset)	10
Fine-tune at the layer number (primary dataset)	80

3.1.4. Mixed Dataset Training

The study involves training the model on the mixed dataset, comprising images from both the specific wildfire dataset and the source dataset.

Table 8 presented below showcases the results yielded by this approach.

Table 8. Performance metrics means using the mixed dataset training approach.

Metrics	Mixed Dataset Training
Accuracy	Mean = 0.7380, Std = 0.018
Precision	Mean = 0.5060, Std = 0.027
Recall	Mean = 0.5060, Std = 0.021
F1-score	Mean = 0.4900, Std = 0.024
ROC-AUC	Mean = 0.5080, Std = 0.029

These results present the average values and standard deviation after five runs of performance metrics for the mixed dataset training approach.

Table 9 outlines the optimized hyperparameters for the approach, with configurations derived from the validation dataset of the wildfire dataset.

Table 9. Hyperparameter configuration for the mixed dataset training approach.

Hyperparameter	Value
Learning rate	1×10^{-3}
Optimizer	Adam
Batch size	32
Epochs	25
Fine-tune at the layer number	1

3.2. Performance of the Hierarchical Domain-Adaptive Learning Approach: An In-Depth Analysis

In the preceding evaluations, a range of baseline training approaches for forest fire detection using deep learning was methodically assessed. Each method, distinct in its design and objectives, has established a foundation upon which the hierarchical domain-adaptive learning strategy is built. This strategy, conceived to augment the model's robustness and precision, employs hierarchical learning concepts that are finely attuned to the domain-specific nuances of forest fires. The subsequent analysis delves into the performance of this advanced approach, contrasting it with conventional methodologies. The metrics

presented herein (see Table 10), the culmination of comprehensive testing and statistical analysis, highlight the strategy's superior capability in the accurate detection of forest fires, suggesting its promising applicability in real-world settings.

Table 10. Performance metrics means using the hierarchical domain-adaptive learning approach.

Metrics	Hierarchical Domain-Adaptive Learning
Accuracy	Mean = 0.9536, Std = 0.0048
Precision	Mean = 0.96112, Std = 0.00771
Recall	Mean = 0.9633, Std = 0.029
F1-score	Mean = 0.9621, Std = 0.0031
Specificity	Mean = 0.9383, Std = 0.0129
FNR	Mean = 0.0366, Std = 0.0029
MCC	Mean = 0.9024, Std = 0.0087
ROC-AUC	Mean = 0.9917, Std = 0.00080

These results present the average values and standard deviation after five runs of performance metrics for the hierarchical domain-adaptive learning approach.

Table 11 outlines the optimized hyperparameters for the approach, with configurations derived from the validation set of the wildfire dataset.

Table 11. Hyperparameter configuration for the hierarchical domain-adaptive learning approach.

Hyperparameter	Value
Learning rate (initial training)	1×10^{-2}
Learning rate (fine-tuning)	1×10^{-3}
Optimizer (both)	Adam
Batch size	32
Epochs (initial training)	10
Fine-tune at the layer number (initial training)	1
Epochs (fine-tuning)	15
Fine-tune at the layer number (fine-tuning)	80

3.3. Comparative Analysis of Baseline Methodologies: A Critical Evaluation of Performance Metrics

This section provides an analysis of baseline methodologies used in this study, focusing on a critical evaluation of their performance metrics. The goal is to explore the relative merits and challenges of each approach, shedding light on their effectiveness in forest fire detection. This comparative assessment is crucial for understanding how different training strategies perform in a specialized context.

In the comparative analysis between the training-from-scratch approach and the sequential transfer learning methodology, statistical examination reveals a landscape where differences across most metrics do not reach statistical significance (see Table 11). This observation extends to accuracy, precision, F1-score, specificity, MCC, and ROC-AUC metrics, where the confidence intervals suggest that the disparities between the two approaches fall within the bounds of chance variation.

Notably, however, recall and the false negative rate (FNR) emerge as outliers in this pattern. While recall exhibits a statistically significant advantage for the training-from-scratch method, indicating it correctly identifies more true positives, the FNR—which measures the proportion of positives the model fails to identify—also shows a statistically significant difference. The confidence interval for FNR also does not overlap the zero

difference mark, inversely reflecting the findings for recall and further corroborating the comparative advantage of the training-from-scratch method. This result indicates a lower rate of missed positive cases (false negatives) for the training-from-scratch approach compared to sequential transfer learning. This finding is particularly salient given the critical role of recall in forest fire detection systems, where the cost of a false negative—failing to identify an actual fire—can be disproportionately high. The observed decrease in recall when comparing the transfer learning approach with 80 frozen layers to training from scratch points to an increased incidence of false negatives in detecting forest fires. These layers, originally trained on varied non-forest fire scenarios, retain a broad but non-specific set of features. While this generalist knowledge may be effective in broad applications, the freezing of these layers limits their adaptability to the unique characteristics of forest fires during subsequent training. This lack of specificity, combined with a potential negative bias towards ‘no fire’ scenarios, impairs the model’s sensitivity to forest-fire-specific cues. Consequently, the model becomes less adept at correctly identifying forest fires, leading to a higher rate of missed detections. Thus, while the intention is to preserve general knowledge, the frozen layers inadvertently diminish the model’s capacity to detect the nuanced and distinct patterns inherent in forest fires, resulting in an elevated risk of false negatives.

Further reinforcing the case for the training-from-scratch approach, mean values calculated over five experimental runs consistently show a higher performance relative to sequential transfer learning. This pattern of superiority is not only limited to recall but extends across all evaluated metrics, further emphasizing the robustness of the training-from-scratch strategy.

Lastly, the mixed dataset outcomes clearly imply that the mixed approach was unable to sufficiently adapt to the specific characteristics required for the intended task, probably due to conflicting features or patterns within the combined datasets.

The analytical results suggest that the training-from-scratch approach generally aligns with the other baselines in terms of key metrics, with the exception of recall where it demonstrates a notable advantage. Given this context, the training-from-scratch method has been adopted as the reference standard for further comparative analysis (see Table 12).

Table 12. Confidence intervals for the difference in performance between the training-from-scratch approach and the sequential transfer learning approach.

Metrics	Confidence Intervals
Accuracy	(−0.0052, 0.0173)
Precision	(−0.0132, 0.01088)
Recall	(0.0062, 0.05402)
F1-score	(−0.0022, 0.0297)
Specificity	(−0.0143, 0.0047)
FNR	(−0.0540, −0.0087)
MCC	(−0.0060, 0.0493)
ROC-AUC	(−0.001, 0.0278)

The comparative analysis between the ‘Transfer Learning with Optimized Frozen Layers’ and the ‘Sequential Transfer Learning’ approaches reveals significant insights. The notable performance enhancement of the sequential transfer learning approach across most metrics, except for recall and FNR, underscores the efficacy of fine-tuning from a more closely related dataset—the auxiliary one—compared to a more generalized dataset like ImageNet (refer to Table 13). This distinction in source datasets appears to be a pivotal factor in the model’s overall performance.

The auxiliary dataset’s closer alignment with the context of forest fires provides a more relevant feature set for the model to learn from, facilitating a more precise adaptation to the specific task of forest fire detection. Unlike the broader spectrum of images in ImageNet,

the auxiliary dataset's focus on non-forest fire and no-fire scenarios offers a more direct relevance to the target domain, enabling the model to fine-tune its detection capabilities in a way that is more attuned to the nuances of forest fire imagery.

The lack of improvement in recall for the sequential transfer learning approach, despite its fine-tuning on a more domain-specific auxiliary dataset, suggests that the method's initial broad training may still impose limitations on its ability to fully adapt to the unique aspects of forest fire detection. This could be due to the retained generalist knowledge from the initial training phase, which, despite subsequent domain-specific fine-tuning, might not provide the level of specificity required to optimize recall or equivalently to decrease the FNR.

Table 13. Confidence intervals for the difference in performance between the transfer learning approach and the sequential transfer learning approach.

Metrics	Confidence Intervals
Accuracy	(0.0237, 0.0514)
Precision	(0.0515, 0.0992)
Recall	(−0.0238, 0.0339)
F1-score	(0.0305, 0.0603)
Specificity	(−0.0956, −0.0438)
FNR	(−0.0201, 0.0377)
MCC	(−0.1039, −0.0528)
ROC-AUC	(0.0215, 0.0491)

3.4. Comparative Analysis of the Baseline with the Study's Proposed Approach

In assessing the efficacy of the present work's proposed approach, the training-from-scratch approach was employed as a baseline due to its superior performance among the four initial methodologies considered. This foundational method served as a comparative standard against which the proposed hierarchical domain-adaptive learning approach was rigorously evaluated. The analysis indicates that the proposed approach demonstrates a clear tendency toward enhanced performance metrics, although the interpretation of these results is approached with careful consideration of statistical confidence and the inherent complexity of model evaluation (see Tables 14 and 15).

Table 14. Performance metrics comparison of the baseline and the study's proposed approach.

Metric	Baseline	Hierarchical Domain-Adaptive Learning
Primary accuracy	Mean = 0.9033, Std = 0.0111	Mean = 0.9536 , Std = 0.0048
Precision	Mean = 0.8372, Std = 0.0103	Mean = 0.96112 , Std = 0.00771
Recall	Mean = 0.9509, Std = 0.0200	Mean = 0.9633 , Std = 0.029
F1-score	Mean = 0.8904, Std = 0.01312	Mean = 0.9621 , Std = 0.0031
Specificity	Mean = 0.8828, Std = 0.0078	Mean = 0.9383 , Std = 0.0129
FNR	Mean = 0.0490, Std = 0.0200	Mean = 0.0366 , Std = 0.0029
MCC	Mean = 0.8184, Std = 0.0224	Mean = 0.9024 , Std = 0.0087
ROC-AUC	Mean = 0.9168, Std = 0.01020	Mean = 0.9917 , Std = 0.00080

These results are based on five runs for each model, with bootstrapping used to calculate 95% confidence intervals for the differences in performance metrics between the two methods.

Table 15. Confidence intervals for the difference in performance between the baseline and the study's proposed approach.

Metrics	Confidence Intervals
Accuracy	(−0.062, −0.043)
Precision	(−0.136, −0.112)
Recall	(−0.0295, 0.0514)
F1-score	(−0.083, −0.059)
Specificity	(−0.0679, −0.04302)
FNR	(−0.0055, 0.0306)
MCC	(−0.1048, −0.0621)
ROC-AUC	(−0.085, −0.068)

For accuracy, the hierarchical domain-adaptive learning approach exhibited a higher mean accuracy (0.9536) relative to the baseline (0.9033). The bootstrapped 95% confidence interval for the difference in accuracy (−0.062 to −0.043) does not encompass the null, suggesting a statistically meaningful elevation in performance.

Similarly, precision, which quantifies the correctness of positive predictions, was notably superior in the hierarchical domain-adaptive learning approach, with a mean precision of 0.96112 over the baseline's 0.8372. The associated confidence interval for this metric, extending from −0.136 to −0.112, further attests to the statistical significance of this enhancement. Such an increase in precision is crucial, as it reduces the rate of false alarms—a common challenge in wildfire detection systems.

Regarding recall, the confidence interval (−0.0295 to 0.0514) straddles zero, indicating that the evidence for improvement in sensitivity is suggestive rather than conclusive. While the mean recall of the proposed method (0.9633) is higher than that of the baseline (0.9509), the overlapping confidence interval warrants a cautious interpretation, emphasizing the nuanced nature of model comparison. The FNR metric, inversely related to recall, equivalently indicated a trend towards improvement, yet with inconclusive statistical significance.

The F1-score and ROC-AUC metrics corroborate these findings. The hierarchical approach's F1-score mean of 0.9621 vastly exceeds the baseline's 0.8904, with the confidence interval ranging from −0.083 to −0.059, reinforcing the method's statistical robustness in terms of precision and recall's harmonic mean. The hierarchical domain-adaptive learning approach improved not only the model's ability to correctly identify wildfires but also its accuracy in confirming non-fire events, as evidenced by its higher mean specificity (0.9383) compared to the baseline (0.8828). The bootstrapped 95% confidence interval for specificity (−0.0679 to −0.04302) lies entirely below zero, indicating that the improvement in correctly identifying true negatives is statistically significant. This enhancement in specificity is particularly beneficial in the context of wildfire detection, where the accurate rejection of non-fire events is just as crucial as the correct identification of actual fires to minimize unnecessary deployments of emergency services.

Furthermore, the MCC for the hierarchical domain-adaptive learning approach demonstrates a substantial leap in performance, with a mean of 0.9024 surpassing the baseline's 0.8184. The confidence interval (−0.1048, −0.0621) confirms that the increase in MCC is statistically significant, supporting the model's enhanced predictive quality. MCC, a balanced measure even with unbalanced data, reflects a comprehensive performance improvement across both classes. Its significant increase underscores the hierarchical approach's advanced ability to balance true positives and negatives, crucial for the reliability of wildfire detection systems in varied operational environments.

The ROC-AUC, a probabilistic measure of the model's discriminatory capacity, demonstrates a remarkable confidence interval from −0.085 to −0.068, indicating a near-certain

statistical significance in the advanced method's ability to distinguish between the presence and absence of wildfires.

Overall, the advanced hierarchical domain-adaptive learning approach appears to offer improvements in several key performance metrics when contrasted with the baseline training-from-scratch approach. The approach's ability to discern subtle indicators of wildfires from varied datasets suggests a model that is not only more accurate but also more versatile and reliable in different environmental conditions. This reliability is vital for early detection systems where the cost of false negatives can be devastating. The statistical significance of these findings, particularly in accuracy, precision, MCC, and ROC-AUC, underscores the potential for the hierarchical domain-adaptive learning approach to set new benchmarks for performance in this critical domain.

On the other hand, and acknowledging the importance of detecting fires at night, we have analyzed the performance of the proposed model on the 17 night images contained within the wildfire test dataset. The error analysis on the misclassified ones reveals that the proposed approach outperforms the training-from-scratch model, reducing false predictions from three to just one on these night images. This underscores the potential of the hierarchical domain-adaptive learning model under nocturnal conditions.

Finally, we have conducted an error analysis based on feature visualization using gradient-weighted class activation mapping (Grad-CAM) [39], following methodologies outlined in References [15,40]. This qualitative analysis highlights the improved ability of the proposed approach in accurately identifying distinguishing image features. Notably, the model's enhanced precision is in part attributed to its capability to detect subtle yet critical fire characteristics, such as the distinct glow of flames amidst smoke or unique patterns of spreading fires, even against complex backgrounds. However, to further enhance the interpretability of the novel hierarchical domain-adaptive learning approach, additional feature analysis is needed. Such future analysis will unravel the decision-making layers of the model, thereby ensuring transparency and trustworthiness in its application for forest fire detection. This future endeavor is crucial not only for validating the model's precision but also for deepening our understanding of its operational mechanics.

Time Complexity and Parameter Efficiency Insights

In this subsection, we delve into a parameter and time complexity analysis of the proposed hierarchical domain-adaptive learning approach as compared to the traditional training-from-scratch method. Recognizing the critical importance of both model performance and computational efficiency for practical deployment, we present a focused evaluation on the number of parameters and inference time—key indicators of a model's operational viability. This brief yet substantive examination aims to provide a clear perspective on how our approach aligns with the demands of real-time application scenarios, particularly in the context of forest fire detection where timely processing is paramount. Table 16 displays a comparative analysis of time complexity and parameter efficiency between the baseline model and the approach proposed in this study.

Table 16. Time complexity and parameters efficiency comparison between the baseline and the study's proposed approach.

Efficiency Aspect	Baseline	Hierarchical Domain-Adaptive Learning
Number of parameters	4,052,133	4,052,133
Avg. inference time/image (ms)	Mean = 5.58	Mean = 5.46
Total training time	Mean = 19 min 36 s	Mean = 57 min 45 s

These results are based on five independent runs for each model.

In embarking on the analysis, it is noteworthy to first acknowledge that, despite the innovative aspects of the hierarchical domain-adaptive learning approach, the final architecture of the model aligns precisely with the EfficientNetB0 framework used in the

training-from-scratch method. This architectural consistency, encompassing both the base-model structure and the regularization techniques, ensures that any observed variances in performance can be attributed to the training methodology rather than differences in model complexity. Such an insight is crucial for a fair and accurate comparison of the two approaches, especially in understanding their operational implications in high-stakes applications like forest fire detection. In this sense, we found that the hierarchical domain-adaptive learning approach yields a model with a total of 4,052,133 parameters, of which 4,011,010 are trainable. This parameter count is identical to that of the model developed using the training-from-scratch approach, underscoring the architectural alignment between the two methodologies despite the advanced training strategy employed.

For our time complexity analysis, we utilized Google Colab with an NVIDIA A100 GPU and a high-RAM configuration providing 53 GB, alongside 8 CPU cores for efficient processing as documented by Colab Pro+ features. Our consistent software stack included Python 3.10.12 and TensorFlow 2.15.0. Non-existing interference from background processes was insured. The models were preloaded, and the GPU was warmed up before measuring inference times using Python's time module. Across five runs, the hierarchical domain-adaptive learning approach achieved a mean inference time per image of 0.00546 s, and the training-from-scratch approach had a mean of 0.00558 s. The calculated confidence interval for the difference in inference times between approaches including image preprocessing was $(-0.00036, 0.00059)$ seconds, indicating no significant difference and underscoring the efficiency of the novel approach.

In addition to the parameter count and inference time analysis, we evaluated the training time and computational resource utilization of both models. The hierarchical domain-adaptive learning approach required a significantly longer training time, averaging 57 min, and 45 s over five runs. This extended duration can be attributed to the intricate fine-tuning process inherent in the approach. In contrast, the training-from-scratch model took an average of 19 min and 36 s. While the training time for the hierarchical domain-adaptive learning approach is notably longer than the traditional method, it is crucial to emphasize that, in scenarios like forest fire detection, the emphasis is more on the efficiency of the model during deployment and real-time inference, rather than on the initial time investment for training.

The computational resource utilization analysis during inference revealed comparable GPU load changes and stable memory consumption for both the hierarchical domain-adaptive learning approach and the training-from-scratch method. Notably, the GPU load variations were minimal and the free memory remained consistent across both models. The observed parity in resource utilization highlights the practicality of our hierarchical approach, suggesting its suitability for time-sensitive applications without imposing undue computational demands. It is important to note, however, that consistent inference times may be influenced by the GPU capabilities and the scope of our test dataset of 410 images. Future research will extend this analysis to larger datasets, providing deeper insights into the training approach's impact on inference time and learned features.

3.5. Performance Benchmarking: Hierarchical Domain-Adaptive Learning on EfficientNetB0 vs. EfficientNetB7

As we advance our understanding of deep-learning applications in critical scenarios like forest fire detection, the exploration of model architectures of varying complexities becomes essential. This subsection aims to juxtapose the performance of the hierarchical domain-adaptive learning approach when applied to two distinct models within the EfficientNet family. The comparison sheds light on the balance between model complexity, reflected in the number of parameters, and operational efficiency, particularly inference time. Such an analysis is vital for discerning whether the increased accuracy potential of EfficientNetB7 translates into tangible benefits in a real-time detection context or if the lightweight B0 variant offers a more practical solution. By systematically evaluating performance metrics, inference speed, and parameter counts over multiple runs, we provide a

data-driven argument to guide the selection of an optimal model architecture for forest fire detection applications.

Table 17 outlines the optimized hyperparameters for the approach, with configurations derived from the validation set of the wildfire dataset.

Table 17. Hyperparameter configuration for the hierarchical domain-adaptive learning approach using the EfficientNetB7 as a base model.

Hyperparameter	Value
Learning rate (initial training)	1×10^{-2}
Learning rate (fine-tuning)	1×10^{-3}
Optimizer (both)	Adam
Batch size	32
Epochs (initial training)	15
Fine-tune at the layer number (initial training)	1
Epochs (fine-tuning)	20
Fine-tune at the layer number (fine-tuning)	80

Tables 18–20 encapsulate the empirical findings from the comparative analysis, presenting an array of performance metrics alongside inference times and parameter counts for the hierarchical domain-adaptive learning approach using both EfficientNetB0 and B7. These results provide a quantitative foundation for assessing the efficacy and efficiency trade-offs between the two models in the context of forest fire detection.

Table 18. Performance metrics comparison of the study’s proposed approach using the EfficientNetB0 and EfficientNetB7.

Metric	Hierarchical Domain-Adaptive Learning (EfficientNetB7)	Hierarchical Domain-Adaptive Learning (EfficientNetB0)
Primary accuracy	Mean = 0.9604 , Std = 0.0024	Mean = 0.9536, Std = 0.0048
Precision	Mean = 0.9659 , Std = 0.0045	Mean = 0.96112, Std = 0.00771
Recall	Mean = 0.9697 , Std = 0.0047	Mean = 0.9633, Std = 0.029
F1-score	Mean = 0.9677 , Std = 0.0019	Mean = 0.9621, Std = 0.0031
Specificity	Mean = 0.9458 , Std = 0.0075	Mean = 0.9383, Std = 0.0129
FNR	Mean = 0.0303 , Std = 0.0047	Mean = 0.0366, Std = 0.0029
MCC	Mean = 0.9166 , Std = 0.0050	Mean = 0.9024, Std = 0.0087
ROC-AUC	Mean = 0.9912, Std = 0.0026	Mean = 0.9917 , Std = 0.0008

These results are based on five runs for each model, with bootstrapping used to calculate 95%.

The analysis of the comparison between the hierarchical domain-adaptive learning approach using EfficientNetB0 and EfficientNetB7 as base models indicates that both approaches have a consistent number of parameters, which aligns with the expectation that the core architecture remains the same despite the training methodology.

However, there are noteworthy differences in the average inference time per image, with the B7-based model showing a higher average inference time. This could be due to the inherent complexity of the B7 model, which, while potentially offering higher accuracy, also requires more computational resources for inference.

Table 19. Confidence intervals for the difference in performance between the study’s proposed approach using the EfficientNetB0 and EfficientNetB7.

Metrics	Confidence Intervals
Accuracy	(−0.0112, −0.0029)
Precision	(−0.0124, 0.0027)
Recall	(−0.0103, −0.0008)
F1-score	(−0.0090, −0.0024)
Specificity	(−0.0201, 0.0050)
FNR	(0.0015, 0.0103)
MCC	(−0.0236, −0.0058)
ROC-AUC	(−0.0018, 0.0029)

Table 20. Time complexity and parameters efficiency comparison between the study’s proposed approach using the EfficientNetB0 and EfficientNetB7.

Efficiency Aspects	Hierarchical Domain-Adaptive Learning (EfficientNetB7)	Hierarchical Domain-Adaptive Learning (EfficientNetB0)
Number of parameters	64,102,809	4,052,133
Avg. inference time/image (ms)	Mean = 8.79	Mean = 5.46
Total training time	Mean = 111 min 23 s	Mean = 57 min 45 s

These results are based on five independent runs for each model.

The training time presents a significant difference, with the B7-based approach taking substantially longer. This is expected since B7 models are larger and more complex, and therefore require more time to train, even if the eventual deployment for inference does not reflect this complexity due to optimized weights.

The results shown in the table suggest that, while the B7-based approach may offer slight improvements in certain performance metrics, these come at the cost of increased computational resources during training. This trade-off is crucial in scenarios where resources and time are limited, such as in rapid response situations for forest fire detection.

The decision to use a more or less complex architecture as the base model in hierarchical domain-adaptive learning, and, indeed, in forest fire detection generally, should be guided by the specific demands of the deployment scenario, considering both performance metrics and computational efficiency.

4. Discussion

The hierarchical domain-adaptive learning approach introduced in this study represents a significant enhancement in forest fire detection using deep learning. This approach significantly improves key performance metrics, such as accuracy and precision, verified through bootstrapped 95% confidence intervals. These results are not merely incremental but substantial, highlighting the model’s improved capability in distinguishing between fire and non-fire scenarios.

Our comparative analysis underscores the importance of domain-specific adaptations in training, demonstrating how hierarchical learning tailored to forest fire characteristics yields a performance edge. This underscores the vital role of domain-adaptive methodologies in advancing this field.

Practically, the improved accuracy and precision from this approach are critical for early forest fire detection, reducing the impact of such events. The adaptability of the proposed model to diverse environmental conditions is essential for creating dependable systems, especially in scenarios where false negatives have serious consequences.

The model's distinct performance advantage arises from its novel training approach, particularly the hierarchical learning strategy and the domain-adaptive element fine-tuned for forest fire detection. Freezing specific layers during fine-tuning helps preserve general features while adapting others to forest fire specifics, balancing overfitting prevention and generalization.

This study also compares the hierarchical domain-adaptive learning framework to the sequential transfer learning approach, focusing on precision and recall in forest fire detection. The hierarchical framework's dual-branch architecture and shared layers for general feature extraction result in a more holistic learning process, potentially improving precision while maintaining recall. This contrasts with the sequential transfer learning approach, which may exhibit limitations in specificity and adjustment to forest fire characteristics due to its sequential learning process.

The hierarchical domain-adaptive learning framework's success likely stems from its balance between generalization and domain-specific adaptation—a balance not fully achieved by sequential transfer learning. This highlights the importance of the architectural design of dataset integration in developing specialized models, like those needed for forest fire detection.

However, the framework also highlights challenges in optimizing recall. The auxiliary dataset, despite aiding in generalization and precision, may not effectively enhance recall due to its lack of specificity to forest environments. This suggests a need for datasets more aligned with the primary task and training strategies that capture a broader range of forest fire scenarios.

Additionally, the analysis reveals that methods employing pre-trained, frozen layers, especially those trained on general datasets, often struggle to adapt to the unique demands of forest fire detection. Conversely, domain-specific training, as exemplified by our framework, shows marked improvements across multiple metrics. This underscores the need for carefully selected training strategies and datasets tailored to the task of optimizing the model performance effectively.

On the other hand, in our dual-task learning strategy for forest fire detection, we recognize the risk of negative transfer, especially with shared layers processing both forest and non-forest fire scenarios. While these layers aim to extract general features useful for forest fire detection, they might inadvertently learn features more specific to non-forest scenarios, potentially impacting the model's performance on its primary task.

To counter this, we limited the depth of the shared layers to the top 80, balancing the use of pre-trained features with minimizing the risk of irrelevant feature incorporation. Additionally, fine-tuning focused on the forest-specific dataset helps align the model with the primary task, reducing the chance of negative transfer. However, we acknowledge that the risk of negative transfer cannot be completely eliminated.

Future work will concentrate on further refining our model's architecture to enhance its forest fire detection capabilities while minimizing negative transfer. We plan to explore advanced techniques like the multisource maximum predictor discrepancy (MMPD) model [41], incorporating domain-adaptive methods and continual learning strategies to better tailor the model to forest fire scenarios.

Continuing the discussion, it is noteworthy to acknowledge the pioneering role of Study [15] in employing the wildfire test dataset, which establishes a benchmark for our investigation. This study employed MobileNetV3 as a base model, whereas the current work is based on EfficientNetB0. Despite the architectural differences, such a comparison remains insightful. Table 21 encapsulates a direct comparison, shedding light on the advancements achieved in detection accuracy, precision, recall, and other key metrics. This juxtaposition is critical, as it not only benchmarks our model against existing literature but also demonstrates the improvements and benefits brought forth by the approach.

Table 21. Performance metrics comparison between the model proposed by [15] and the study's proposed approach.

Metric	The Model Proposed by [15]	Hierarchical Domain-Adaptive Learning
Primary accuracy	Mean = 0.8766, Std = 0.0147	Mean = 0.9536 , Std = 0.0048
Precision	Mean = 0.7974, Std = 0.0206	Mean = 0.96112 , Std = 0.00771
Recall	Mean = 0.9171, Std = 0.0273	Mean = 0.9633 , Std = 0.029
F1-score	Mean = 0.8526, Std = 0.0146	Mean = 0.9621 , Std = 0.0031
ROC-AUC	Mean = 0.8842, Std = 0.0127	Mean = 0.9917 , Std = 0.00080

These results are based on five independent runs for each model.

Study's Limitations and Future Work

Despite these promising results, the study acknowledges the inherent limitations of any single methodology in the complex and variable context of forest fires. The performance of the model must be continually assessed against emerging datasets and evolving fire scenarios to ensure ongoing relevance and effectiveness. Additionally, while the model excels in the conditions tested, further research is needed to evaluate its performance across more region-specific geographies.

Another crucial avenue that remains unexplored is the temporal aspect inherent in the progression of wildfire events. The sequential nature of image data, portraying the evolving dynamics of fire and smoke, holds a treasure trove of information that can significantly enhance early detection and monitoring. While the current training approach has laid a solid foundation in handling spatial features effectively, it does not explicitly account for temporal dependencies. This limitation opens up a thrilling realm of possibilities for future work. In the forthcoming studies, we plan to evolve the approach to not only evaluate but also train the model on sequential data. This adaptation will entail augmenting the proposed approach to encapsulate the sequential (temporal) information inherent in wildfires, thus moving beyond the solely spatial analysis conducted in the present study. By weaving in the temporal dimension, we aim to advance towards a more robust and insightful wildfire detection model that can operate efficiently in real time, providing timely alerts that are crucial for mitigating the adverse impacts of wildfires.

This temporal adaptation and evaluation will not only address the current study's limitations but also propel the wildfire detection domain a step closer to more realistic and effective solutions. In this research, the methodology primarily focuses on the primary task of wildfire detection, emphasizing its critical importance in the model's application. While this approach aligns with our immediate objectives, it presents a limitation in fully exploring the potential synergies in dual-task learning.

On the other hand, while our study marks progress in forest fire detection, we acknowledge that night-time fire detection remains a crucial aspect requiring further investigation. Given its importance for practical implementation, more extensive research in this area is essential for enhancing the model's reliability and effectiveness under varied lighting conditions, especially during night-time scenarios.

Another limitation of the current study that could be explored in a future work is the non-benefit from experimenting with various loss-weighting strategies to balance the learning between the primary and secondary tasks on the shared layers, potentially enhancing the model's overall performance and robustness. Particularly intriguing is the prospect of adaptive loss-weighting mechanisms, where weights are dynamically adjusted during training based on specific criteria such as task performance or model confidence. Theoretical development in this area, coupled with practical, quantitative evaluations, including the development of new metrics and in-depth statistical analyses, could yield insights into the efficacy of these strategies.

Additionally, understanding the broader implications of such approaches, not only for wildfire detection but for other dual-task learning scenarios across different domains,

could substantially advance the field. This direction promises to uncover new avenues in model optimization and application, paving the way for more accurate, reliable, and versatile machine-learning systems.

Conversely, although cross-validation is a recognized method for evaluating model generalizability, given the computational intensity of training high-resolution image data and the substantial size of the datasets, we opted for a robust single hold-out validation set.

Furthermore, the parameter selection process, while based on empirical evidence and preliminary experimentation, acknowledges the potential for further optimization. Advanced hyperparameter optimization methods, such as grid search or intelligent optimization algorithms, offer a more systematic and exhaustive exploration of the parameter space. Exploring these methods in future studies, given sufficient computational resources, could potentially refine the model's performance further, enhancing accuracy and generalizability. This exploration would also provide deeper insights into the model's parameter sensitivity and stability.

5. Conclusions

In summary, the hierarchical domain-adaptive learning approach introduced in this study marks a significant step forward in the field of forest fire detection. This approach has proven its effectiveness not only through rigorous statistical validation but also in its potential for real-world applications. It exemplifies the efficacy of domain-adaptive deep-learning strategies in environmental monitoring, setting a new benchmark for innovation in this vital area.

Crucially, the approach's adeptness at processing varied datasets and adapting to specific forest fire scenarios, underscored by its enhanced performance metrics, positions it as a pivotal development for ongoing research and practical implementation. This study's findings open avenues for more precise, adaptable forest fire detection technologies, addressing a pressing need in environmental conservation and management.

Author Contributions: Conceptualization, I.E.-M., M.P. and N.O.-T.; methodology, I.E.-M., M.P. and N.O.-T.; software, I.E.-M.; validation, I.E.-M., M.P. and N.O.-T.; formal analysis, I.E.-M., M.P. and N.O.-T.; investigation, I.E.-M., M.P. and N.O.-T.; data curation, I.E.-M.; writing—original draft, I.E.-M.; writing—review and editing, M.P. and N.O.-T.; visualization, I.E.-M.; supervision, M.P. and N.O.-T.; project administration, M.P. and N.O.-T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All datasets and code necessary to corroborate the present research findings are available at the following link: <https://kaggle.com/datasets/elmadafri/the-wildfire-dataset> (accessed on 1 November 2023). Researchers are encouraged to use these resources for further studies.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Available online: <https://www.fao.org/state-of-forests/en/#:~:text=Forests%20cover%2031%20percent%20of,processes%20re%20not%20significantly%20disturbed> (accessed on 1 November 2023).
2. Available online: <https://www.worldbank.org/en/topic/forests> (accessed on 1 November 2023).
3. Seneviratne, S.I.; Zhang, X.; Adnan, M.; Badi, W.; Dereczynski, C.; Di Luca, A.; Ghosh, S.; Iskandar, I.; Kossin, J.; Lewis, S.; et al. Weather and Climate Extreme Events in a Changing Climate. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2021; pp. 1513–1766.
4. Available online: <https://www.unep.org/explore-topics/sustainable-development-goals/why-do-sustainable-development-goals-matter/goal-15#:~:text=Protect,%20restore%20and%20promote%20sustainable,degradation%20and%20halt%20biodiversity%20loss> (accessed on 1 November 2023).
5. Available online: <https://www.unep.org/explore-topics/sustainable-development-goals/why-do-sustainable-development-goals-matter/goal-13#:~:text=Climate%20change%20is%20increasing%20the,sanitation,%20education,%20energy%20and%20transport> (accessed on 1 November 2023).

6. Mohapatra, A.; Trinh, T. Early Wildfire Detection Technologies in Practice—A Review. *Sustainability* **2022**, *14*, 12270. [CrossRef]
7. Available online: <https://www.fws.gov/story/2022-10/how-does-wildfire-impact-wildlife-and-forests> (accessed on 1 November 2023).
8. Barmpoutis, P.; Papaioannou, P.; Dimitropoulos, K.; Grammalidis, N. A Review on Early Forest Fire Detection Systems Using Optical Remote Sensing. *Sensors* **2020**, *20*, 6442. [CrossRef]
9. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
10. Bouguettaya, A.; Zarzour, H.; Taberkit, A.M.; Kechida, A. A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms. *Sign. Proc.* **2022**, *190*, 108309. [CrossRef]
11. Ghali, R.; Akhloufi, M.A. Deep Learning Approaches for Wildland Fires Remote Sensing: Classification, Detection, and Segmentation. *Remote Sens.* **2023**, *15*, 1821. [CrossRef]
12. Muhammad, K.; Ahmad, J.; Baik, S.W. Early fire detection using convolutional neural networks during surveillance for effective disaster management. *Neurocomputing* **2018**, *288*, 30–42. [CrossRef]
13. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 24–49. [CrossRef]
14. Allison, R.S.; Johnston, J.M.; Craig, G.; Jennings, S. Airborne Optical and Thermal Remote Sensing for Wildfire Detection and Monitoring. *Sensors* **2016**, *16*, 1310. [CrossRef]
15. El-Madafri, I.; Peña, M.; Olmedo-Torre, N. The Wildfire Dataset: Enhancing Deep Learning-Based Forest Fire Detection with a Diverse Evolving Open-Source Dataset Focused on Data Representativeness and a Novel Multi-Task Learning Approach. *Forests* **2023**, *14*, 1697. [CrossRef]
16. Lines, E.R.; Allen, M.; Cabo, C.; Calders, K.; Debus, A.; Grieve, S.W.D.; Miltiadou, M.; Noach, A.; Owen, H.J.F.; Puliti, S. AI applications in forest monitoring need remote sensing benchmark datasets. In Proceedings of the 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 17–20 December 2022; pp. 4528–4533.
17. Diez, Y.; Kentsch, S.; Fukuda, M.; Caceres, M.L.L.; Moritake, K.; Cabezas, M. Deep Learning in Forestry Using UAV-Acquired RGB Data: A Practical Review. *Remote Sens.* **2021**, *13*, 2837. [CrossRef]
18. Sousa, M.J.; Moutinho, A.; Almeida, M. Wildfire detection using transfer learning on augmented datasets. *Expert Syst. Appl.* **2020**, *142*, 112975. [CrossRef]
19. Hao, X.; Liu, L.; Yang, R.; Yin, L.; Zhang, L.; Li, X. A Review of Data Augmentation Methods of Remote Sensing Image Target Recognition. *Remote Sens.* **2023**, *15*, 827. [CrossRef]
20. Zhang, Y.; Yang, Q. A Survey on Multi-Task Learning. *IEEE Trans. Knowl. Data Eng.* **2022**, *34*, 5586–5609. [CrossRef]
21. Yu, F.; Xiu, X.; Li, Y. A Survey on Deep Transfer Learning and Beyond. *Mathematics* **2022**, *10*, 3619. [CrossRef]
22. Lu, K.; Huang, J.; Li, J.; Zhou, J.; Chen, X.; Liu, Y. MTL-FFDET: A Multi-Task Learning-Based Model for Forest Fire Detection. *Forests* **2022**, *13*, 1448. [CrossRef]
23. Xu, G.; Zhang, Y.; Zhang, Q.; Lin, J.; Wang, J. Wang Deep domain adaptation based video smoke detection using synthetic smoke images. *Fire Saf. J.* **2017**, *93*, 53–59. [CrossRef]
24. Sathishkumar, V.E.; Cho, J.; Subramanian, M.; Naren, O.S. Forest fire and smoke detection using deep learning-based learning without forgetting. *Fire Ecol.* **2023**, *19*, 9. [CrossRef]
25. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [CrossRef]
26. Tan, M.; Le, Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 10–15 June 2019; pp. 6105–6114.
27. Available online: <https://universe.roboflow.com/firegassmoke/dataset-for-fire-and-smoke-detection> (accessed on 1 November 2023).
28. Cazzolato, M.T.; Avalhais, L.P.S.; Chino, D.Y.T.; Ramos, J.S.; Souza, J.A.; Rodrigues, J.F., Jr.; Traina, A.J.M. FiSmo: A Compilation of Datasets from Emergency Situations for Fire and Smoke Analysis. In *SBB2017—SBB2 Proceedings of Satellite Events of the 32nd Brazilian Symposium on Databases—DSW (Dataset Showcase Workshop)*; SBC: Traverse City, MI, USA, 2017; Available online: <http://sbbd.org.br/2017/wp-content/uploads/sites/3/2017/10/proceedings-satellite-events-sbbd-2017.pdf> (accessed on 1 November 2023).
29. Johnson, J.M.; Khoshgoftaar, T.M. Survey on deep learning with class imbalance. *J. Big Data* **2019**, *6*, 27. [CrossRef]
30. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]
31. Kovács, G. An empirical comparison and evaluation of minority oversampling techniques on a large number of imbalanced datasets. *Appl. Soft Comput.* **2019**, *83*, 105662. [CrossRef]
32. Elangovan, P.; Nath, M.K. En-ConvNet: A novel approach for glaucoma detection from color fundus images using ensemble of deep convolutional neural networks. *Int. J. Imaging Syst. Technol.* **2022**, *32*, 2034–2048. [CrossRef]
33. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *arXiv* **2014**, arXiv:1411.1792.
34. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 2921–2929.
35. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
36. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]

37. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Li, F.-F. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255. [CrossRef]
38. El Madafri, I.; Peña Carrera, M.; Olmedo Torre, N. Applying Artificial Intelligence Models for the Automatic Forest Fire Detection. In *A: Avenços en Recerca i Desenvolupament del Departament d'Enginyeria Gràfica i de Disseny*; OmniaScience: Barcelona, Spain, 2023; pp. 95–104. ISBN 978-84-126475-1-8. Available online: <http://hdl.handle.net/2117/386430> (accessed on 29 June 2023).
39. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626. [CrossRef]
40. Islam, A.M.; Masud, F.B.; Ahmed, M.R.; Jafar, A.I.; Ullah, J.R.; Islam, S.; Shatabda, S.; Islam, A.K.M.M. An Attention-Guided Deep-Learning-Based Network with Bayesian Optimization for Forest Fire Classification and Localization. *Forests* **2023**, *14*, 2080. [CrossRef]
41. Ma, Y.; Yang, Z.; Zhang, Z. Multisource Maximum Predictor Discrepancy for Unsupervised Domain Adaptation on Corn Yield Prediction. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4401315. [CrossRef]

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.