

Forest Fire Detection: Cross-sensor Transfer Learning with DenseNet-121 from Landsat-8 to Sentinel-2 Satellite Imagery in Benguet, Philippines

John Paul Q. Tomas ^{1,*}, Lance Asher M. Elloso ¹, Estefanee A. N. Garcia ¹, Khristine F. Mecija ¹,
John Christian S. Diamante ¹, Tiong Yew Tang ², and Bonifacio T. Doma, Jr. ^{1,3}

¹ School of Information Technology, Mapúa University, Makati City, Philippines

² Department of Research Centre for Human-Machine Collaboration (HUMAC), Faculty of Engineering and Technology, Sunway University, Bandar Sunway, Malaysia

³ School of Chemical, Biological, and Materials Engineering and Sciences, Mapúa University, Manila, Philippines

Email: jptomas@mapua.edu.ph (J.P.Q.T.); lamelloso@mymail.mapua.edu.ph (L.A.M.E.);

eangarcia@mymail.mapua.edu.ph (E.A.N.G.); kfmecija@mymail.mapua.edu.ph (K.F.M.);

jcsdiamante@mapua.edu.ph (J.C.S.D); tiongyewt@sunway.edu.my (T.Y.T.); btdoma@mapua.edu.ph (B.T.D.J.)

*Corresponding author

Abstract—This study investigates the application of DenseNet-121 for forest fire detection using cross-sensor transfer learning between Landsat-8 and Sentinel-2 satellite imagery in the Philippines. The primary objectives are to evaluate the impact of a locally fine-tuned Sentinel-2 dataset on model performance and to explore the model's capacity for cross-sensor transfer learning from Landsat-8. DenseNet-121 was selected for its dense connectivity architecture, which enables efficient feature propagation, and was further enhanced using Gradient-weighted Class Activation Mapping (Grad-CAM) for spatial interpretability. The model achieved an F1-Score of 89.47% on unseen Sentinel-2 data, demonstrating strong generalization performance. Grad-CAM visualizations provided interpretable heatmaps for approximate fire localization and severity estimation, classifying alerts from low to extreme. These findings underscore the model's potential as a scalable and explainable solution for forest fire detection, supporting timely decision-making in environmental monitoring and disaster response.

Keywords—forest fire detection, cross-sensor transfer learning, DenseNet-121, convolutional neural network, remote sensing, Landsat-8, Sentinel-2, satellite imagery

I. INTRODUCTION

Forest fires are uncontrolled combustion of natural vegetation that can rapidly destroy woodland areas, leading to irreversible consequences such as habitat loss, deforestation, and environmental degradation [1, 2]. In recent years, increased land use, rising temperatures, and climate change have contributed to the rising frequency of forest fires. Remote sensing using satellite imagery has

become crucial for forest fire monitoring due to its wide spatial coverage and regular acquisition intervals. Techniques that identify fire pixels using hyperspectral images have proven useful for detecting active fires.

Several studies have applied deep learning models for fire detection, achieving promising results. For instance, Rostami *et al.* [3] used Landsat-8 data and deep multiple kernel learning, achieving a fire detection precision of 95.71%. However, implementing these systems often requires significant resources, which are not always available, especially in developing countries like the Philippines.

Despite this, the Philippines still lacks a centralized forest fire detection system. Reports from the Asian Forest Cooperation Organization (AFoCO) in 2021 highlighted challenges such as limited technological capacity, bureaucratic inefficiencies, and prevalent anthropogenic causes like slash-and-burn farming [4]. Currently, national monitoring relies heavily on weather data from the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA), rather than dedicated fire detection systems.

Globally, Ghali and Akhloufi [5] emphasize the need for deep learning systems that can detect wildfires early to reduce ecological and economic damage. Pereira *et al.* [6] also noted that while deep learning approaches using satellite imagery such as Landsat-8 show promise, the scarcity of labeled datasets limits model robustness and transferability.

To address this, the study investigates the use of cross-sensor transfer learning using DenseNet-121, a Convolutional Neural Network (CNN) pretrained on

ImageNet, to adapt fire detection features learned from Landsat-8 data to Sentinel-2 imagery. Since Sentinel-2 offers similar multispectral bands and spatial resolution (10–60 m) with a faster revisit time of 5 days, this cross-sensor approach may enhance fire detection capabilities in regions like Benguet, Philippines. The study seeks to develop a more generalizable fire detection system that can perform effectively despite limited local training data, helping to reduce unreported incidents and improve environmental monitoring capacity.

II. LITERATURE REVIEW

A. Forest Fire Detection Using Satellite Imagery

Early research by Alkhatib [7] in forest fire detection relied on traditional remote sensing techniques, employing threshold-based algorithms on thermal and multispectral data. While these approaches provided foundational insights, they often detected fires only after they had become widespread and uncontrollable, leading to devastating environmental and economic losses. It emphasized the limitations of these early systems, which primarily utilized optical sensors, digital cameras, and wireless sensor networks alongside satellites, underscoring the urgent need for more effective early detection methods.

To address these shortcomings, Pereira *et al.* [6] developed a large-scale dataset for active fire detection using Landsat-8 imagery, comprising over 150,000 image patches processed with established handcrafted algorithms from Schroeder *et al.* [8], Murphy *et al.* [9], and Kumar and Roy [10]. This dataset enabled the application of deep learning models on a global scale, significantly advancing the accuracy and scalability of fire detection. Building on this foundation, Rostami *et al.* [3] introduced a deep multiple kernel learning model, MultiScale-Net, which leveraged spectral-spatial features from Short-Wave Infrared (SWIR) and blue bands. Their approach achieved a precision of 95.71%, demonstrating the potential of CNN-based architecture for detecting active fires across diverse regions.

Further innovation was presented by Hong *et al.* [11], who proposed FireCNN using Himawari-8 satellite data. Their model integrated multi-scale convolution and residual learning, improving detection accuracy by 35.2% compared to traditional thresholding methods, even when trained on a relatively small dataset. Similarly, Nabi and Bordanov [12] highlighted the inefficiency of segmenting entire fire regions due to their rapidly changing shapes and sizes. Instead, they argued that confirming the presence of a fire source is more critical for operational decision-making. Their study, which evaluated multiple Artificial Neural Network (ANN) architectures on Landsat-8 and Uruguay datasets, achieved 99% classification accuracy by prioritizing minimal input channels, demonstrating CNNs' adaptability in resource-constrained environments.

B. Deep Learning for Fire Detection

Deep learning has emerged as a transformative approach for remote sensing-based fire detection,

particularly using CNNs. Huang *et al.* [13] introduced DenseNet, a deeply supervised architecture designed to mitigate vanishing gradients and improve feature propagation via dense connectivity. Trained on the ImageNet dataset [14], DenseNet demonstrated superior performance over deeper ResNet models while requiring fewer computational resources. These attributes make DenseNet particularly suitable for high-dimensional satellite imagery analysis, where efficiency and accuracy are paramount.

Priya and Vani [15] utilized Inception-v3 combined with transfer learning to classify fire and non-fire satellite images, further enhancing detection through local binary pattern analysis. Their approach reduced false positives and achieved high precision and recall scores, illustrating the effectiveness of combining deep architectures with feature-based refinements.

Additionally, Minetto *et al.* [16] and Li *et al.* [17] validated DenseNet's robustness by applying it to large-scale benchmark datasets such as Functional Map of the World (FMOW) [18] and the Northwestern Polytechnical University Remote Sensing Image Scene Classification 45 (NWPU-RESISC45) dataset [20], confirming its strength as a universal feature extractor across multiple remote sensing domains.

C. Transfer Learning in Remote Sensing

While deep learning models require large amounts of labeled data, the scarcity of annotated satellite imagery poses a significant challenge. Iman *et al.* [19] discussed how Deep Transfer Learning (DTL) addresses this limitation by reusing pre-trained models, reducing training costs, and improving performance in low-data scenarios. This strategy is particularly advantageous in forest fire detection, where compiling diverse and labeled datasets can be costly and time-consuming. Priya and Vani [15] demonstrated this advantage by employing transfer learning with Inception-v3 to achieve improved fire classification accuracy using limited datasets.

D. Cross-Sensor Adaptation and Limitations

Cross-sensor adaptability remains a critical area of research, given the variability in spectral and spatial characteristics across satellite platforms. Lopez-Puigdollers *et al.* [21] explored this challenge in cloud detection, evaluating models trained on Landsat-8 and testing them on Sentinel-2. Their findings confirmed the feasibility of cross-sensor transfer when using fully convolutional architectures, provided that dataset diversity and spatial variability are adequately addressed. Extending these principles, Zhao *et al.* [22] proposed a model for detecting fire smoke across multispectral datasets from Landsat and Sentinel-2. By incorporating an Input Amplification module, their framework resolved band discrepancies between sensors without requiring core model retraining, achieving superior cross-domain performance and offering a scalable solution for multi-sensor fire detection.

III. MATERIALS AND METHODS

This study proposes a framework for forest fire detection using DenseNet-121 with Landsat-8 and Sentinel-2 imagery. Landsat-8 patches served as the base training dataset, while Sentinel-2 scenes from Benguet, Philippines were used for fine-tuning, enabling cross-sensor transfer learning.

DenseNet-121, initialized with ImageNet weights, was trained in two phases: initial training on Landsat-8 for feature extraction, followed by fine-tuning on Sentinel-2 to adapt to local conditions. Model performance was

evaluated using Accuracy, Precision, Recall, F1-Score, Intersection-over-Union (IoU), and confusion matrices.

To enhance interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) generated heatmaps highlighting regions most influential in fire prediction. These heatmaps were post-processed to estimate active fire area, expressed in hectares, and categorized into severity levels. This integrated approach combines detection, visual explainability, and impact assessment, supporting effective wildfire monitoring and decision-making as illustrated in Fig. 1.

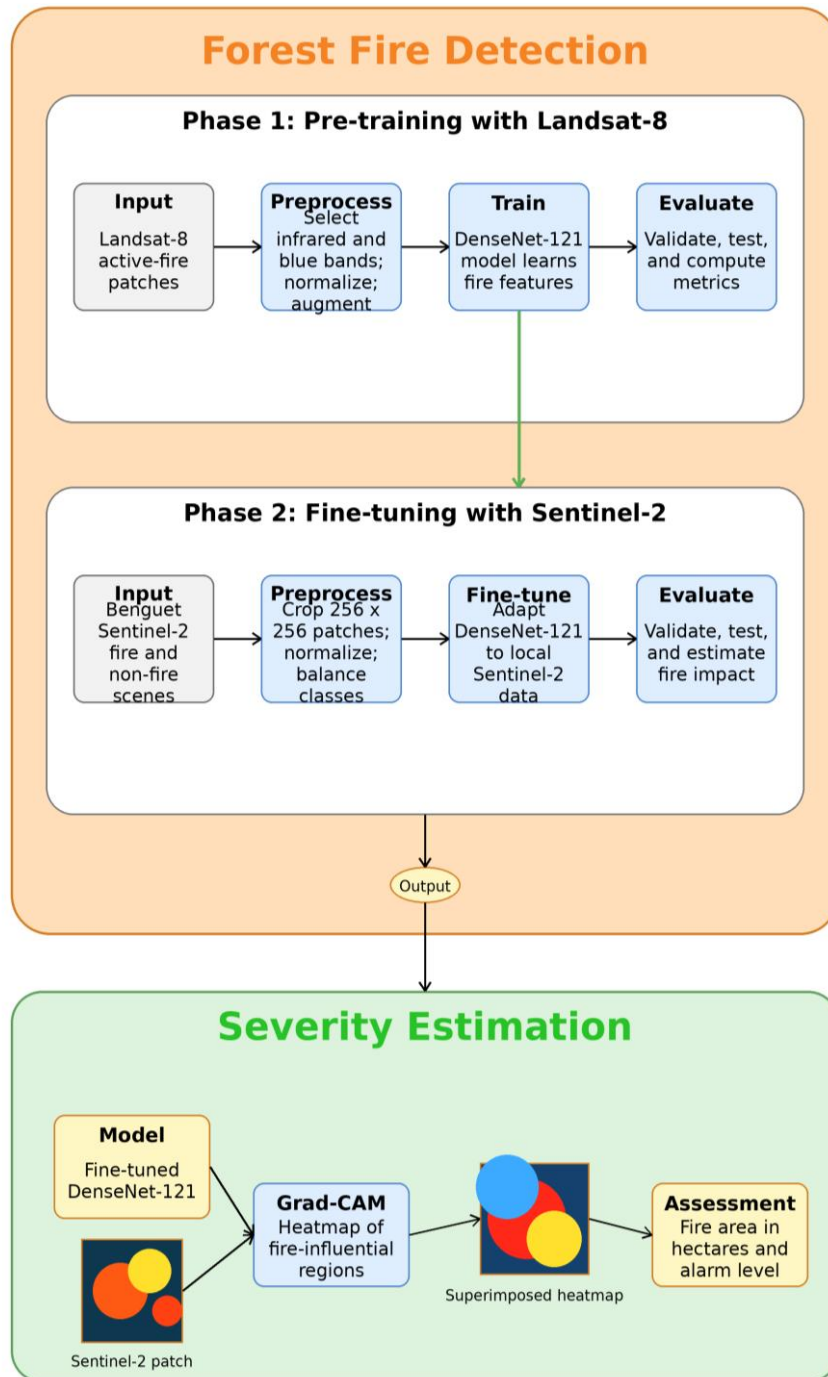


Fig. 1. Overall workflow for forest fire detection and Grad-CAM-based severity estimation.

A. Data Acquisition

Two primary datasets were employed in this study:

1) Landsat-8 active fire patches dataset

A publicly available dataset containing labeled image patches representing fire and non-fire regions. Each patch includes multiple spectral bands, particularly SWIR and blue, which are crucial for detecting thermal anomalies [6].

2) Sentinel-2 local dataset

Sentinel-2 imagery of forest fires in Kabayan, Benguet (2019–2021) was obtained from the Copernicus Open Access Hub via the Provincial Environment and Natural Resources Office (PENRO)-Benguet. Level-2A products with SWIR and blue bands were pre-processed and cropped into 256×256 non-overlapping patches for model input. These patches were analyzed using the pre-trained DenseNet-121 model for binary fire detection.

Both datasets were used to evaluate cross-sensor transfer learning capability, where the model trained on Landsat-8 was adapted for Sentinel-2 imagery.

B. Data Preprocessing

Sentinel-2 imagery was segmented into 256×256 patches, while Landsat-8 patches were used as provided. For both datasets, Short-Wave Infrared band 2 (SWIR2), Short-Wave Infrared band 1 (SWIR1), and blue bands were selected to capture fire-related spectral signatures while reducing computational cost. All pixel values were normalized (0–1), and data augmentation (horizontal, vertical, and horizontal-vertical flips) was applied to address class imbalance and improve model generalization.

C. Model Architecture and Training

DenseNet-121 was selected for its dense connectivity, which promotes feature reuse and mitigates vanishing gradients [13]. The model was initialized with ImageNet weights and adapted for binary classification (fire vs. non-fire) by adding Global Average Pooling [23], a fully connected layer (256 units, Rectified Linear Unit (ReLU), L2 regularization) [24], batch normalization, dropout (0.5) [25], and a sigmoid output layer [26].

Training involved two phases:

- (1) Pre-Training: On Landsat-8 patches using Adam (learning rate (LR) = 0.0001, batch size = 32).
- (2) Transfer Learning: Fine-tuning on Sentinel-2 patches for local adaptation.

Binary cross-entropy was used as the loss function, with early stopping [27], learning-rate reduction [28], and class weights [29, 30] applied. Model performance was evaluated using Accuracy, Precision, Recall, F1-Score, IoU, and Confusion Matrix [31].

D. Explainability and Severity Estimation

To ensure explainability, Grad-CAM [32] was applied to visualize the spatial regions that influenced the model's predictions. For each 256 × 256 input patch, Grad-CAM computed gradients of fire-class scores with respect to convolutional feature maps (e.g., conv5_block16_concat), generating a heatmap resized to input dimensions [33–35].

1) Computation of activated blocks

Define a binary activation mask first:

$$A_{i,j} = \begin{cases} 1, & H_{i,j} \geq T \\ 0, & H_{i,j} < T \end{cases} \quad (1)$$

Then count activated locations:

$$N_{\text{active}} = \sum_i \sum_j A_{i,j} = \sum_i \sum_j 1(H_{i,j} \geq T) \quad (2)$$

where $A_{i,j}$ is the binary activation indicator at pixel (i, j) , $H_{i,j}$ is the heatmap value at pixel (i, j) , T is the user-defined threshold (default = 0.5), and N_{active} is the total number of activated heatmap pixels.

2) Conversion to fire area

Recommended, if using the resized 256 × 256 Grad-CAM heatmap:

$$N_{\text{fire pixels}} = \sum_{i=1}^{256} \sum_{j=1}^{256} 1(H_{i,j} \geq T) A_{\text{pixel}} = r^2 \quad (3)$$

where $N_{\text{fire pixels}}$ denote the estimated activated fire-pixel count, $A_{\text{pixel}} = r^2$ is the area represented by each resampled pixel, and $r = 10$ only if all bands/patches were resampled to a 10 m grid.

Alternatively, if counting the final feature-map blocks before resizing, $N_{\text{active blocks}}$ is the number of activated feature-map blocks, while h and w are the feature-map height and width:

$$N_{\text{fire pixels}} = N_{\text{active blocks}} \times \left(\frac{256}{h} \right) \left(\frac{256}{w} \right) \quad (4)$$

For an 8 × 8 Grad-CAM feature map, this becomes $N_{\text{active blocks}} \times 32^2$.

$$\text{Fire Area (ha)} = \frac{N_{\text{fire pixels}} \times r^2}{10000} \quad (5)$$

where r is the pixel size in meters. If 10 m resampling is used, then:

$$\text{Fire Area (ha)} = \frac{N_{\text{fire pixels}} \times 100}{10000} = \frac{N_{\text{fire pixels}}}{100} \quad (6)$$

IV. RESULTS AND DISCUSSION

A. Dataset

This study utilized two primary datasets: (1) a publicly available Landsat-8 Active Fire Detection dataset [6] containing 18,548 labeled patches after augmentation, and (2) a locally collected Sentinel-2 dataset from Benguet, Philippines, comprising 356 augmented patches.

For model development, each dataset was divided into 70% for training and validation and 30% for testing. The combined datasets capture a wide range of fire intensities, atmospheric conditions, times of acquisition, and smoke patterns.

B. Landsat-8 Baseline Performance

The DenseNet-121 model, pre-trained on ImageNet, was fine-tuned on the Landsat-8 dataset to establish baseline performance for cross-sensor transfer learning. Training converged at Epoch 57 with early stopping, demonstrating effective learning dynamics as shown in Fig. 2.

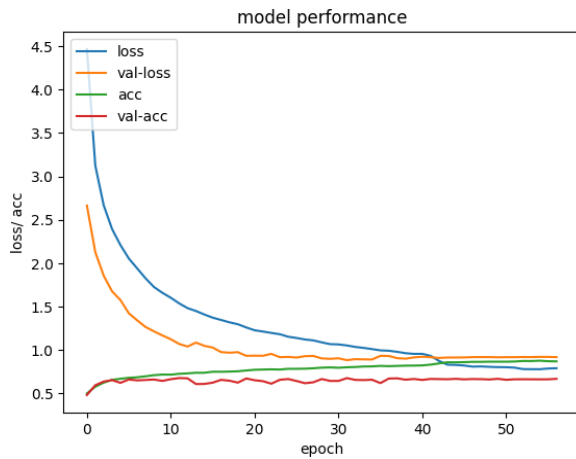


Fig. 2. Model's training performance on the base Landsat-8 dataset.

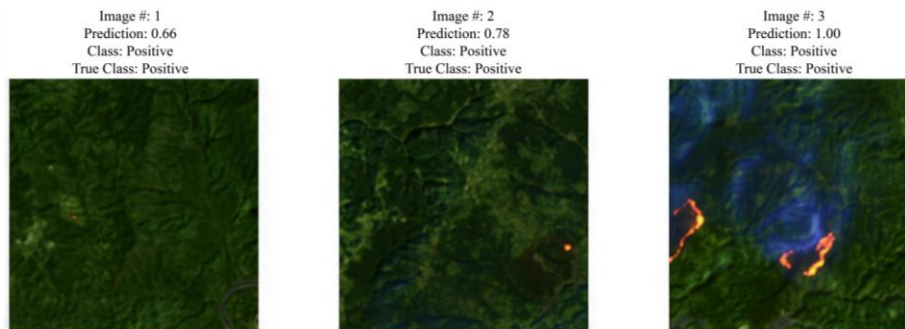


Fig. 3. Output of images above 50% classification threshold.

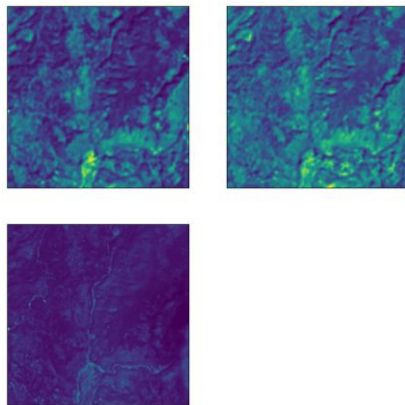


Fig. 4. Visualization of feature maps in 'input_3' layer of the false positive image (60%) of the test set.

The model achieved 66.99% accuracy on the test dataset with a precision of 0.6218, recall of 0.5624, F1-Score of 0.5906, and IoU of 0.4190. The confusion matrix shows 541 true positives, 981 true negatives, 329 false positives, and 421 false negatives.

While the balanced precision-recall indicates reasonable fire detection capability, the moderate IoU (0.4190) and significant false negatives (421) highlight opportunities for improvement through cross-sensor transfer learning. These results establish the baseline for evaluating Sentinel-2 transfer performance.

C. Cross-Sensor Transfer Learning to Sentinel-2

The pre-trained Landsat-8 model was adapted to Sentinel-2 through progressive fine-tuning. Initially, the last 10 layers were made trainable (converging at epoch 71), followed by the last 20 layers (converging at epoch 32, validation loss: 0.2759).

Cross-sensor transfer learning achieved remarkable performance gains on Sentinel-2: 97.06% accuracy, 80% precision, 100% recall, 88% F1-Score, and 0.8 IoU. This represents a substantial improvement over the Landsat-8 baseline (66.99% accuracy, 0.4190 IoU).

Visualization of learned features in Fig. 3 confirms the model correctly identifies fire pixels in high-confidence predictions. Error analysis on Fig. 4 reveals the model occasionally misclassifies water bodies as fire regions, indicating spectral similarity challenges.

D. Independent Validation

To assess real-world applicability, an unseen Sentinel-2 scene was evaluated independently. The fine-tuned model maintained robust performance with 94.44% accuracy, 89.47% precision, 89.47% recall, 89.47% F1-Score, and 0.8095 IoU.

E. Comparative Analysis and Transfer Learning Impact

The DenseNet-121 model demonstrated significant performance enhancement through cross-sensor transfer learning. The Landsat-8 baseline established initial feature learning (66.99% accuracy), while Sentinel-2 fine-tuning improved cross-sensor performance to 94.44% accuracy on unseen data.

Table I summarizes the quantitative improvements before and after Sentinel-2 fine-tuning. Accuracy

increased from 83.34% to 94.44% (+11.10%). Most notably, precision increased from 61.29% to 89.47% (+28.18%), indicating superior false-positive reduction, critical for operational fire detection systems. IoU improved by 19.66% (61.29% to 80.95%), demonstrating better spatial overlap between predicted and actual fire regions.

While recall declined slightly (−10.53%), this conservative shift resulted in dramatically reduced false

positives, enhancing system reliability. The F1-Score improvement (+13.47%) confirms overall classification enhancement.

The substantial metric improvements validate the effectiveness of cross-sensor domain adaptation. The model successfully leveraged Landsat-8 learned features for Sentinel-2 imagery, demonstrating DenseNet-121's capability for multi-sensor fire detection applications.

TABLE I. PERFORMANCE COMPARISON BEFORE AND AFTER SENTINEL-2 FINE-TUNING (IoU)

Metric	Before Fine-Tuning (%)	After Fine-Tuning (%)	Difference (%)	Impact of Fine-Tuning
Accuracy	83.34	94.44	+11.10	Improved
Precision	61.29	89.47	+28.18	Improved
Recall	100	89.47	−10.53	Declined
IoU	61.29	80.95	+19.66	Improved
F1-Score	76	89.47	+13.47	Improved

F. Grad-CAM-Based Fire Severity Classification and Spatial Analysis

The integration of Grad-CAM with DenseNet-121 provided interpretable visual explanations for fire detection decisions. Grad-CAM generates heatmaps highlighting regions that most significantly contribute to classification, with red-intensity values indicating model confidence for fire presence as shown in Fig. 5 [32].

Fire-affected pixels were identified using threshold $T = 0.50$, corresponding to the DenseNet-121 decision boundary. Areas were classified as Low (<10 ha), Moderate (10–100 ha), High (100–1000 ha), and Extreme (>1000 ha) severity levels.

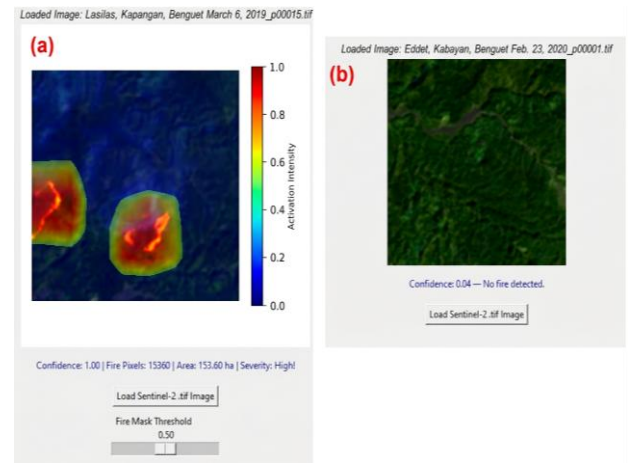
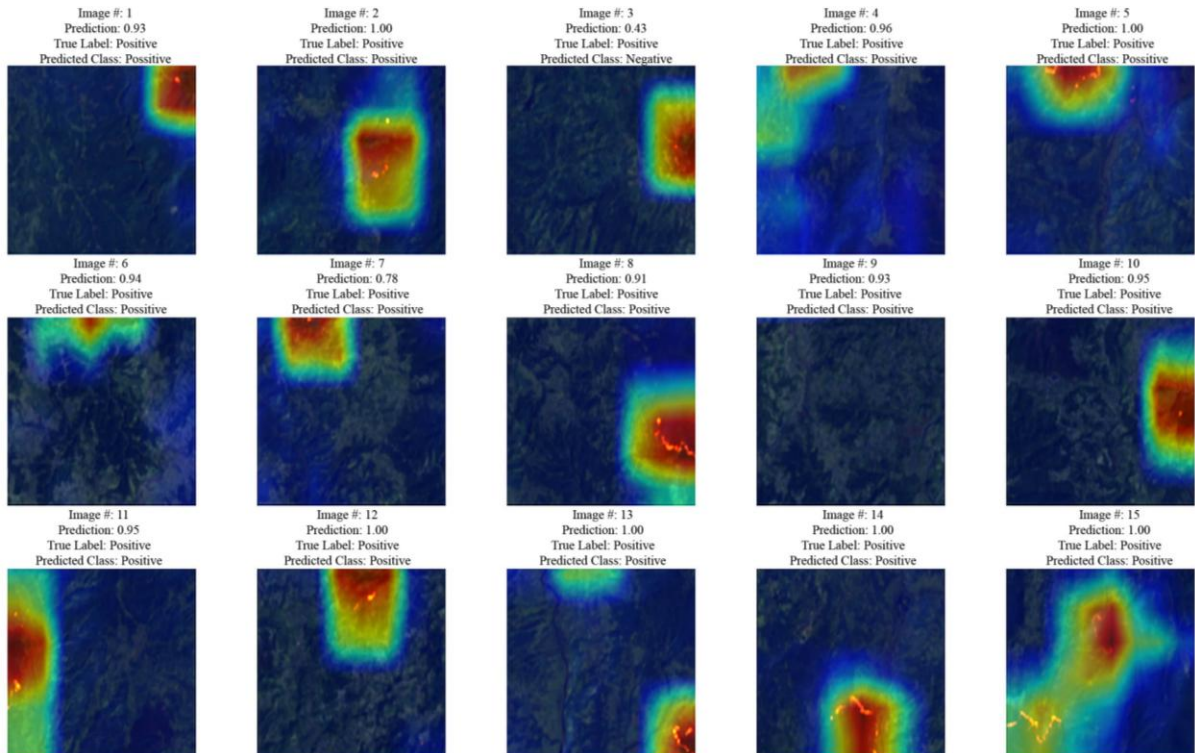


Fig. 5. Grad-CAM-based fire detection examples: (a) satellite image with detected fire and Grad-CAM fire mask; (b) satellite image with no fire detected.



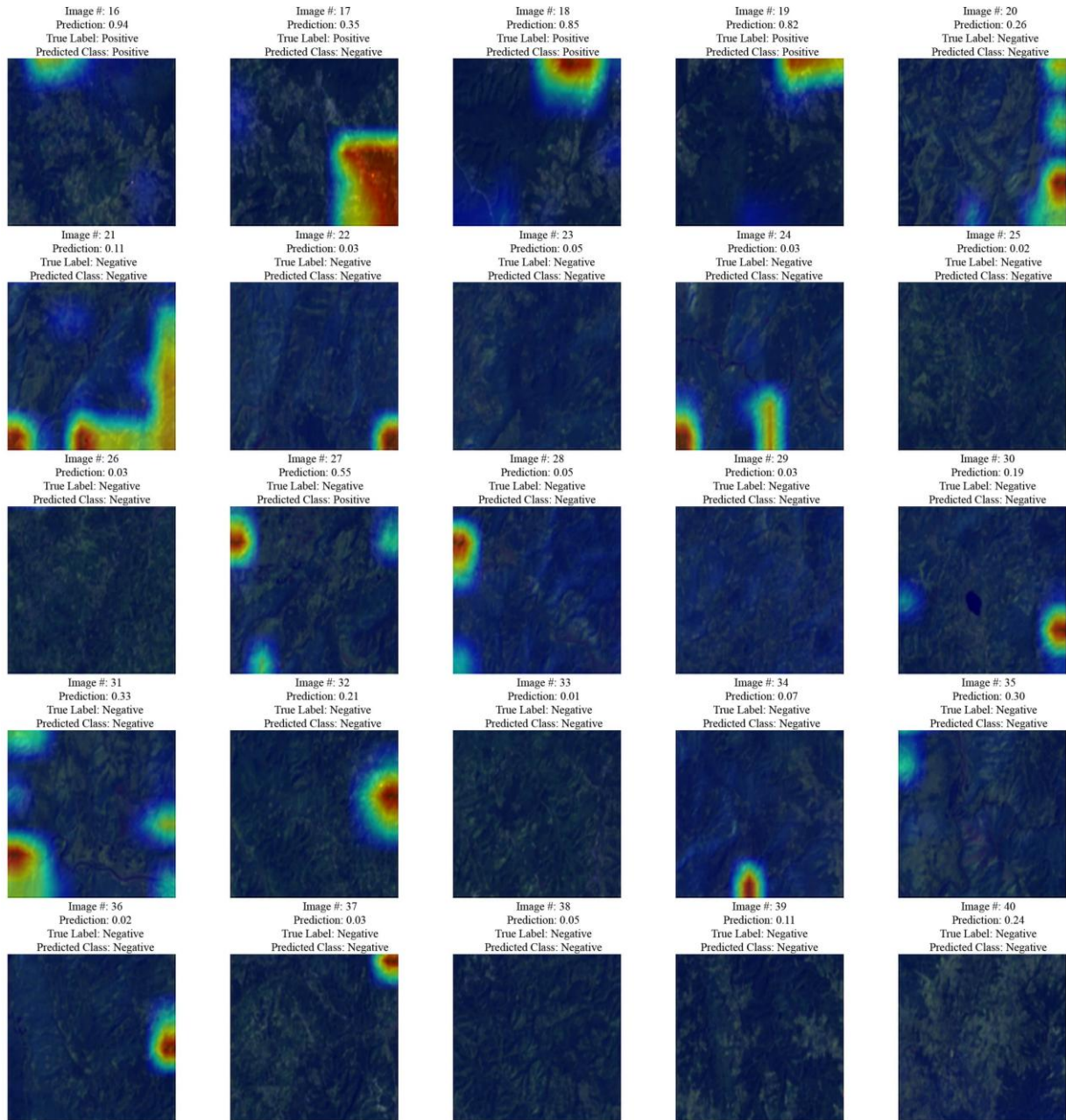


Fig. 6. Model outputs for 40 Sentinel-2 patches with superimposed Grad-CAM heatmaps.

Fig. 6 demonstrates consistent Grad-CAM activation patterns across the sampled Sentinel-2 patches shown, validating the model's spatial attention mechanism and providing quantitative fire impact assessments for operational deployment.

V. CONCLUSION

This study successfully demonstrates DenseNet-121's effectiveness for cross-sensor forest fire detection using Landsat-8 and Sentinel-2 satellite imagery in Benguet, Philippines. The research achieved both primary objectives: local fine-tuning improved model performance to 89.47% F1-Score on unseen data, while cross-sensor transfer learning proved highly effective across different satellite platforms.

The integration of cross-sensor transfer learning with Grad-CAM creates a novel framework bridging binary classification and operational fire assessment. This approach transforms neural network outputs into interpretable spatial visualizations and quantitative severity estimates from low to extreme levels.

The developed system provides high detection accuracy with spatial interpretability, addressing critical gaps in real-time fire monitoring. The severity classification framework enables prioritized emergency response and resource allocation decisions.

This work establishes a scalable methodology for remote sensing environmental monitoring in data-scarce regions. The cross-sensor approach opens possibilities for leveraging multi-platform satellite data, extending to other environmental applications. The combination of high

performance with explainable artificial intelligence creates a foundation for next-generation disaster response systems.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

John Paul Q. Tomas supervised the study, contributed to conceptualization and project administration, and reviewed and edited the manuscript. Lance Asher M. Elloso carried out feasibility analysis, methodology design, data curation, DenseNet-121 CNN implementation, experimentation, and result interpretation. Estefanee A. N. Garcia assisted with data preparation, manuscript drafting, proofreading, cross-referencing, and final revisions. Khristine F. Mecija contributed to writing—review and editing, proofreading, cross-referencing, and manuscript revision. John Christian S. Diamante contributed to technical validation, review and editing, and research supervision. Tiong Yew Tang contributed to methodology review, technical validation, and manuscript review and editing. Bonifacio T. Doma, Jr. contributed to supervision, project administration, validation, and final manuscript review. All authors read and approved the final version.

REFERENCES

- [1] C. Aponte, W. J. De Groot, and B. M. Wotton, "Forest fires and climate change: Causes, consequences and management options," *Int. J. Wildland Fire*, vol. 25, no. 8, i–ii, 2016. doi: 10.1071/WFv25n8_FO
- [2] R. J. Keenan, "Climate change impacts and adaptation in forest management: A review," *Ann. Forest Sci.*, vol. 72, no. 2, pp. 145–167, Mar. 2015. doi: 10.1007/s13595-014-0446-5
- [3] A. Rostami, R. Shah-Hosseini, S. Asgari, M. Aghdami-Nia, and S. Homayouni, "Active fire detection from Landsat-8 imagery using deep multiple kernel learning," *Remote Sens.*, vol. 14, no. 4, 992, Feb. 2022. doi: 10.3390/rs14040992
- [4] Asian Forest Cooperation Organization (AFoCO). (2022) Forest Fire Management Information System (FFMIS). [Online]. Available: <https://afocosec.org/wp-content/uploads/2022/05/KN2022-006-Forest-Fire-Management-Information-System-FFMIS-20220531-2.pdf>
- [5] R. Ghali and M. A. Akhloufi, "Deep learning approaches for wildland fires using satellite remote sensing data: Detection, mapping, and prediction," *Fire*, vol. 6, no. 5, 192, May 2023. doi: 10.3390/fire6050192
- [6] G. H. De Almeida Pereira, A. M. Fusioka, B. T. Nassu, and R. Minetto, "Active fire detection in Landsat-8 imagery: A large-scale dataset and a deep-learning study," *ISPRS J. Photogramm. Remote Sens.*, vol. 178, pp. 171–186, Aug. 2021. doi: 10.1016/j.isprsjprs.2021.06.002
- [7] A. A. A. Alkhatib, "A review on forest fire detection techniques," *Int. J. Distrib. Sensor Netw.*, vol. 10, no. 3, 597368, Mar. 2014. doi: 10.1155/2014/597368
- [8] W. Schroeder, P. Oliva, L. Giglio, B. Quayle, E. Lorenz, and F. Morelli, "Active fire detection using Landsat-8/OLI data," *Remote Sens. Environ.*, vol. 185, pp. 210–220, Nov. 2016. doi: 10.1016/j.rse.2015.08.032
- [9] S. W. Murphy, C. R. De Souza Filho, R. Wright, G. Sabatino, and R. Correa Pabon, "HOTMAP: Global hot target detection at moderate spatial resolution," *Remote Sens. Environ.*, vol. 177, pp. 78–88, May 2016. doi: 10.1016/j.rse.2016.02.027
- [10] S. S. Kumar and D. P. Roy, "Global operational land imager Landsat-8 reflectance-based active fire detection algorithm," *Int. J. Digit. Earth*, vol. 11, no. 2, pp. 154–178, Feb. 2018. doi: 10.1080/17538947.2017.1391341
- [11] Z. Hong *et al.*, "Active fire detection using a novel convolutional neural network based on Himawari-8 satellite images," *Front. Environ. Sci.*, vol. 10, 794028, Mar. 2022. doi: 10.3389/fenvs.2022.794028
- [12] M. Nabi and I. Bordanov, "Research of forest fire detection on satellite images," in *Proc. Int. Conf. Ind. Eng., Appl. Manuf. (ICIEAM)*, Sochi, Russia, May 2023, pp. 803–808. doi: 10.1109/ICIEAM57311.2023.10139138
- [13] G. Huang, Z. Liu, G. Pleiss, L. V. D. Maaten, and K. Q. Weinberger, "Convolutional networks with dense connectivity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 12, pp. 8704–8716, Dec. 2022. doi: 10.1109/TPAMI.2019.2918284
- [14] J. Deng, W. Dong, R. Socher *et al.*, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Miami, FL, USA, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848
- [15] R. S. Priya and K. Vani, "Deep learning based forest fire classification and detection in satellite images," in *Proc. 11th Int. Conf. Adv. Comput. (ICoAC)*, Chennai, India, Dec. 2019, pp. 61–65. doi: 10.1109/ICoAC48765.2019.246817
- [16] R. Minetto, M. Pamplona Segundo, and S. Sarkar, "Hydra: An ensemble of convolutional neural networks for geospatial land classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6530–6541, Sep. 2019. doi: 10.1109/TGRS.2019.2906883
- [17] F. Li, R. Feng, W. Han, and L. Wang, "High-resolution remote sensing image scene classification via key filter bank based on convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 8077–8092, Nov. 2020. doi: 10.1109/TGRS.2020.2987060
- [18] G. Christie, N. Fendley, J. Wilson, and R. Mukherjee, "Functional map of the world," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 6172–6180. doi: 10.1109/CVPR.2018.00646
- [19] M. Iman, H. R. Arabnia, and K. Rasheed, "A review of deep transfer learning and recent advancements," *Technologies*, vol. 11, no. 2, 40, Mar. 2023. doi: 10.3390/technologies11020040
- [20] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state of the art," in *Proc. IEEE*, vol. 105, no. 10, pp. 1865–1883, Oct. 2017. doi: 10.1109/JPROC.2017.2675998
- [21] D. López-Puigdollers, G. Mateo-García, and L. Gómez-Chova, "Benchmarking deep learning models for cloud detection in Landsat-8 and Sentinel-2 images," *Remote Sens.*, vol. 13, no. 5, 992, Mar. 2021. doi: 10.3390/rs13050992
- [22] L. Zhao, J. Liu, S. Peters, J. Li, S. Oliver, and N. Mueller, "Investigating the impact of using IR bands on early fire smoke detection from Landsat imagery with a lightweight CNN model," *Remote Sens.*, vol. 14, no. 13, 3047, Jun. 2022. doi: 10.3390/rs14133047
- [23] GeeksforGeeks. (2026). Introduction to pooling layer. [Online]. Available: <https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>
- [24] GeeksforGeeks. (2026). What is fully connected layer in deep learning? [Online]. Available: <https://www.geeksforgeeks.org/what-is-fully-connected-layer-in-deep-learning/>
- [25] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 56, pp. 1929–1958, 2014.
- [26] S. R. Dubey, S. K. Singh, and B. B. Chaudhuri, "Activation functions in deep learning: A comprehensive survey and benchmark," *Neurocomputing*, vol. 503, pp. 92–108, 2022. doi: 10.1016/j.neucom.2022.06.111
- [27] J. Brownlee. (2026). Use early stopping to halt the training of neural networks at the right time. *Machine Learning Mastery*. [Online]. Available: <https://machinelearningmastery.com/how-to-stop-training-deep-neural-networks-at-the-right-time-using-early-stopping/>
- [28] R. Heckel and F. F. Yilmaz, "Early stopping in deep networks: Double descent and how to eliminate it," arXiv Preprint, arXiv:2007.10099, 2020. doi: 10.48550/arXiv.2007.10099
- [29] B. Bakirarar and A. H. Elhan, "Class weighting technique to deal with imbalanced class problem in machine learning: Methodological research," *Türkiye Klinikleri J. Biostat.*, vol. 15, no. 1, pp. 19–29, 2023. doi: 10.5336/biostatic.2022-93961

- [30] Kamaldeep. (2026). How to improve class imbalance using class weights in machine learning? *Analytics Vidhya*. [Online]. Available: <https://www.analyticsvidhya.com/blog/2020/10/improve-class-imbalance-class-weights/>
- [31] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation," in *Proc. AI 2006: Adv. Artif. Intell.*, 2006, pp. 1015–1021.
- [32] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE International Conference on Computer Vision*, 2017, pp. 618–626.
- [33] S. Yin, L. Wang, M. Shafiq, L. Teng, A. A. Laghari, and M. F. Khan, "G2Grad-CAMRL: An object detection and interpretation model based on gradient-weighted class activation mapping and reinforcement learning in remote sensing images," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 3583–3598, 2023. doi: 10.1109/JSTARS.2023.3241405.
- [34] L. Qiao *et al.*, "Grad-CAM for network models: To support aerial vision based wildfire perception," in *Proc. 50th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Chicago, IL, USA, Nov. 2024, pp. 1–6. doi: 10.1109/IECON55916.2024.10905147
- [35] T. Sultan, M. S. Chowdhury, M. Safran, M. F. Mridha, and N. Dey, "Deep learning-based multistage fire detection system and emerging direction," *Fire*, vol. 7, no. 12, 451, Nov. 2024. doi: 10.3390/fire7120451

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](#)).