

Deep Chromatic Feature Engineering For High-Accuracy Wildfire Prediction Using Enhanced Convolutional Neural Network Architectures

D. Crystal Jaba Kani¹ and S. Saudia²

^{1,2}Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli-12, Tamil Nadu, India.
Email: ¹crystaljebakani@yahoo.com and ²saudiasubash@msuniv.ac.in

Abstract

Wildfires constitute a major global environmental hazard, causing extensive ecological damage, economic losses, infrastructure disruption, and threats to public safety. Although deep learning has substantially improved automated wildfire image classification, conventional spatial-domain models may be affected by illumination changes, atmospheric interference, shadows, and background clutter. This chapter presents a chromatic feature-engineering framework for wildfire classification based on the YUV color space and enhanced Convolutional Neural Network (CNN) architectures. Images from the Fire and No Fire Images (FFNI) subset of the DeepFire benchmark are spatially standardized and transformed from RGB into YUV representations. The luminance (Y) and chrominance (U and V) components are analyzed as complementary representations of image structure and color information. A baseline LeNet-5 model is compared with a structurally modified LeNet-5 designed to improve the extraction of discriminative features from chromatic representations. The reported experiments show that the modified LeNet-5 achieves 96.0% classification accuracy on the YUV features, compared with 44.0% for the baseline LeNet-5. These findings indicate that an appropriate combination of chromatic representation and architecture design can substantially improve wildfire image classification. The chapter also reviews related deep learning approaches, describes the dataset and preprocessing pipeline, formulates the YUV transformation mathematically, and discusses the experimental implications of the proposed approach.

Keywords: Chromatic Feature Engineering; Convolutional Neural Networks; Computer Vision; Hyper-parameter Optimization; LeNet-5; Wildfire prediction; YUV color space.

1. INTRODUCTION

Wildfires pose a multidimensional threat to ecological systems, human infrastructure, public health, and regional economies. Their occurrence and propagation are influenced by interacting environmental factors, including fuel availability, temperature, wind conditions, vegetation moisture, and broader climatic variability [1]. The dynamic and nonlinear nature of wildfire behavior makes reliable early detection and prediction a challenging problem. Consequently, machine learning (ML) and deep learning (DL) methods have increasingly been investigated for automated wildfire detection and classification. Convolutional Neural Networks (CNNs) are particularly suitable for image-based wildfire analysis because they can learn hierarchical spatial representations directly from image data [2]. Nevertheless, the effectiveness of a CNN depends not only on its architecture but also on the representation of the input data. Wildfire images frequently contain shadows, smoke, clouds, vegetation, terrain, and illumination changes. These factors can make fire-related visual patterns difficult to distinguish from non-fire background information when conventional RGB representations are used directly [3]. This paper therefore investigates two complementary strategies. First, chromatic-space optimization is used to transform RGB images into a representation that separates luminance information from chrominance information. Second, the CNN architecture is structurally modified to improve the extraction of discriminative patterns from the resulting feature representations. The principal focus is the YUV color model and its integration with LeNet-5. Previous studies have examined both handcrafted and deep feature representations for wildfire analysis. Sathishkumar et al. (2023), for example, compared handcrafted features with CNN-based approaches on the BoWFire dataset using the YUV color space and reported an accuracy of 87.10% [4]. Vipin et al. (2012) investigated RGB and YCbCr representations and reported benefits from separating luminance information under variable outdoor illumination. These studies motivate further investigation of chromatic representations for robust wildfire classification. Research has also moved toward lightweight object-detection architectures for real-time applications [5]. Geng and Su et al. (2024) developed LUFFD-YOLO with GhostNetV2 for YUV-based processing, while Yang and Cao (2025) proposed YOLOFM using a FocalNext backbone and QAHA Rep FPN for multi-scale feature integration across RGB and YUV inputs. Such approaches demonstrate the value of alternative color representations, while also highlighting the continuing difficulty of detecting small fire regions in complex forest environments. Other studies have addressed

computational efficiency and contextual information [6]. Premal and Vinsley (2014) investigated a lightweight SqueezeNet-inspired CNN using the YCbCr color space [7]. Liu and Cheng et al. (2019) reported that retaining whole-image context could outperform classification based only on segmented fire regions, emphasizing the importance of spatial structure and boundary information [8]. Against this background, the present work evaluates YUV-based image representations with baseline and modified LeNet-5 architectures. The objectives are to (i) examine the classification capability of YUV chromatic features, (ii) assess the effect of structural modifications to LeNet-5, and (iii) compare the resulting models using accuracy and AUC-ROC measures. The reported results provide evidence that chromatic representation and CNN architecture should be considered jointly rather than as independent design choices.

1.1 Deep Learning for Wildfire Prediction

The recent literature demonstrates a progression from conventional statistical classifiers toward deep neural architectures for wildfire ignition modelling, risk mapping, image classification, and real-time surveillance. Zhang et al. (2019) compared a CNN with Random Forest, Kernel Logistic Regression, Support Vector Machine, and Multilayer Perceptron models for forest-fire ignition prediction. The reported CNN accuracy was 87.92%, compared with 84.36% for Random Forest, 81.23% for Kernel Logistic Regression, 80.04% for Support Vector Machine, and 78.47% for Multilayer Perceptron [9]. Bahadori et al. (2023) used multispectral remote-sensing imagery together with historical spatiotemporal information for regional wildfire susceptibility mapping and reported an AUC of 83% [10]. Jindal and Kunwar (2020) integrated a Long Short-Term/Recurrent Convolutional Network approach with a Markov Decision Process to model temporal and meteorological information in the Rocky Mountain region, reporting an overall accuracy of 82%. Unsupervised anomaly-detection strategies have also been investigated [11]. Choi et al. (2019) combined a deep autoencoder with Isolation Forest and One-Class SVM classifiers using vegetation and meteorological variables. The reported baseline accuracy of 71% illustrates the difficulty of detecting wildfire-related anomalies in noisy environmental conditions without sufficiently discriminative representations [12].

1.2 Real-Time Wildfire Surveillance

Saponara and Elhanashi et al. (2021) investigated lightweight YOLOv2-based real-time anti-fire surveillance and reported an initial confidence score of 58% with an inference window of approximately 1–2 seconds. More recent surveillance systems combine CCTV streams with weather information and employ backbones such as ResNet, VoVNet, FBNetV3, Robust Granular Neural Networks, and Bayesian Neural Networks [13]. These developments emphasize the importance of balancing recognition accuracy with computational cost.

1.3 Chromatic Representations for Fire Detection

Color-Space transformation is a complementary strategy for reducing the dependence of fire detection on raw RGB values. Venugopal (2012) used RGB-to-YCbCr transformation to isolate high-energy fire pixels and reported a pixel-classification accuracy of 92.5%. The broader literature suggests that separating luminance from chrominance can be advantageous when images are captured under changing outdoor lighting [14]. Despite the progress of object-detection and deep-learning systems, there remains a need for compact architectures that can effectively process chromatic representations while preserving useful spatial context. The present study addresses this gap by combining YUV feature engineering with a modified LeNet-5 architecture and by directly comparing the modified model with its baseline counterpart.

2. METHODOLOGY AND DATA PRE-PROCESSING

The methodology describes the different stages of the proposed work: pre-processing including resizing, feature engineering, and label encoding, data splitting, LeNet-5 architectural variants, prediction, and model evaluation architecture shown Figure 1.

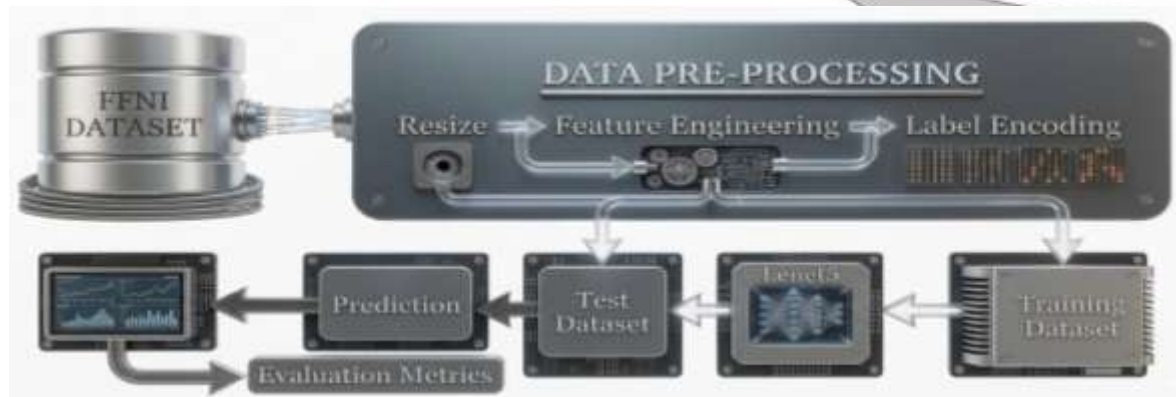


Figure 1: Proposed Techniques

2.1 DeepFire and FFNI Dataset

The proposed framework uses the DeepFire [15] benchmark and a balanced Forest Fire and No Fire Images (FFNI) research subset. The source material contains 1,900 images, including 950 fire images and 950 no-fire images. The selected FFNI subset contains 1,518 images, with 759 fire images and 759 no-fire images. The images represent forest-fire and mountain-fire scenes across varied terrain and background conditions shown in Table 1.

Table 1. Summary of the DeepFire repository and the selected FFNI research subset.

Parameter	DeepFire repository profile	FFNI research subset
Total visual samples	1,900	1,518
Fire images	950	759
No-fire images	950	759
Resolution	Variable source-image dimensions	1200×700, 750×450, and 400×300 pixels
Environmental scenarios	Forest fires, mountain fires, varied terrain	Balanced woodland and mountain-background configurations

2.2 Image Standardization

The preprocessing technique source images have heterogeneous spatial dimensions. To provide a consistent tensor shape for batch processing, all FFNI images are resized to 256×256 pixels. This standardization prevents dimension mismatches during batch formation and neural-network training. Where applicable, the preprocessing pipeline also includes image-format conversion and other deterministic operations required by the experimental implementation.

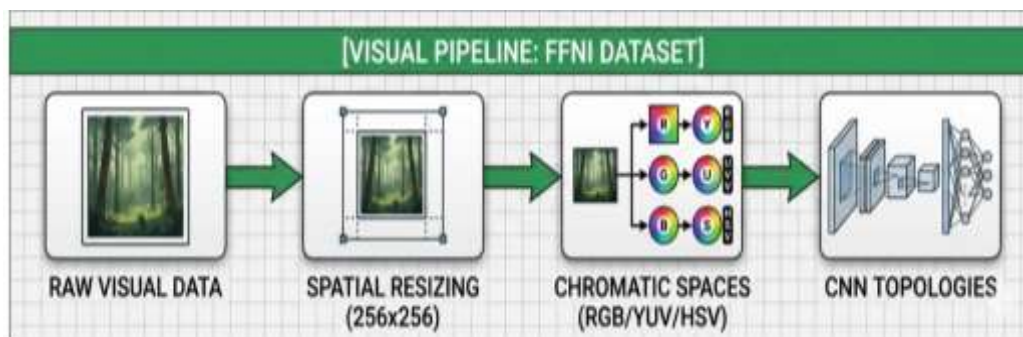


Figure 2: Data Preprocessing

2.3 RGB Standardization and Feature Domains

Images read through computer-vision libraries may be represented internally in BGR order. For a uniform experimental baseline, the channels are reordered into RGB before chromatic transformation. The resulting images are then considered across feature domains chromatic model. The chromatic domain represents color relationships converted into numerical values within this tabulated dataset.

2.4 Tabular Preprocessing

The tabulated dataset is also integrated into the broader experimental framework. Records containing missing, null, or erroneous zero values are filtered according to the preprocessing procedure. The target variable, area, is converted into a binary representation corresponding to burned and unburned states for downstream classification. This tabular workflow is conceptually distinct from the FFNI image classification pipeline and should therefore be reported separately in a final publication if both experiments are retained.

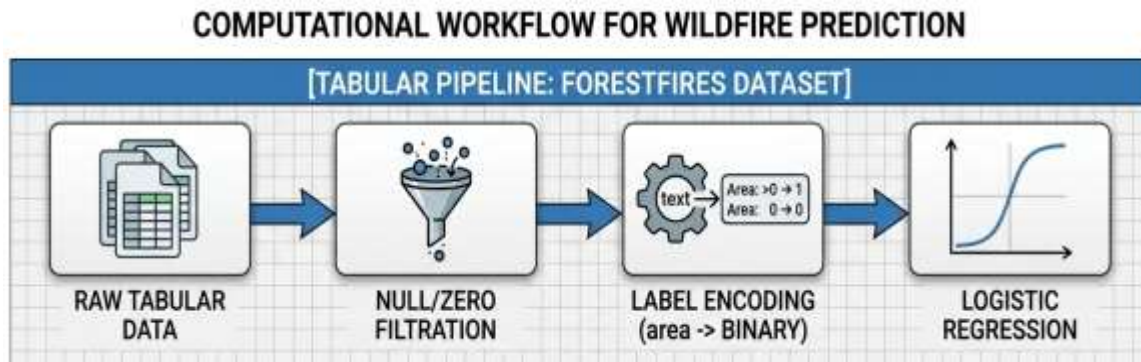


Figure 3: Data Preprocessing

3. YUV CHROMATIC FEATURE ENGINEERING

3.1 Motivation for YUV Representation

RGB represents color through three additive channels, but the channels jointly encode brightness and color information. In outdoor wildfire imagery, brightness can vary considerably because of sunlight, cloud cover, shadows, smoke, and atmospheric attenuation. A chromatic representation that separates luminance from chrominance can therefore provide a useful alternative input space for CNN-based analysis. The YUV model separates a luma component (Y) from two chrominance components (U and V). In the context of this study, Y represents image intensity and structural information, whereas U and V represent color-opponent information associated with blue and red components, respectively [16]. This separation allows the network to examine structural and color information through partially decoupled channels.

3.2 Y, U, and V Components

The three components are interpreted as follows, Y channel (luma): represents brightness and contributes strongly to the representation of structural boundaries and spatial contrast. U channel (blue chrominance): represents the color difference associated with the blue component relative to luma. V channel (red chrominance): represents the color difference associated with the red component relative to luma. For wildfire analysis, the chrominance channels may contain useful information related to the color appearance of flames and surrounding vegetation. However, the usefulness of any individual channel is ultimately an empirical question and must be established from the experimental results rather than assumed.

3.3 Mathematical Transformation

Depending on the representation used by the implementation, YUV values may be represented in normalized continuous form or as quantized 8-bit values. The normalized formulation commonly maps luma to $[0, 1]$ and chrominance to a centered interval, whereas a digital implementation may use an offset representation with integer channel values [12]. The exact range should be reported consistently with the software library and color-conversion

standard used in the experiment. For the BT.601-style digital transformation used in the supplied manuscript, the RGB-to-YUV mapping is given by:

$$Y = 0.257R + 0.504G + 0.098B + 16 \quad (1)$$

$$U = -0.148R - 0.291G + 0.439B + 128 \quad (2)$$

$$V = 0.439R - 0.368G - 0.071B + 128 \quad (3)$$

Equations (1)–(3) should be interpreted together with the exact RGB scaling and conversion convention used in the implementation. In particular, the signs of the U and V coefficients must match the selected standard and software implementation.

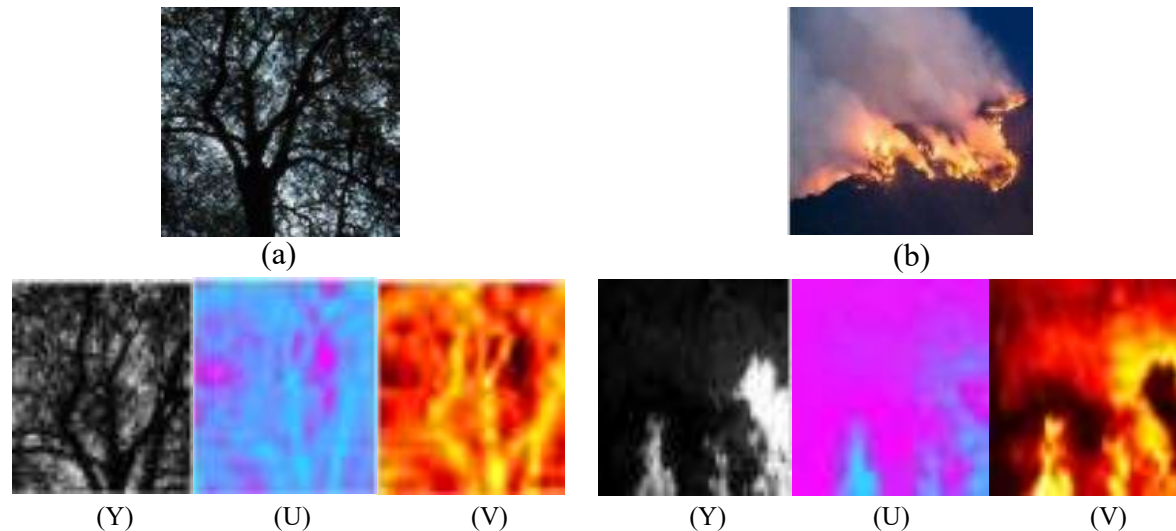


Figure 4: Representative Y, U, and V channel representations for No-Fire and Fire images.

The transformation produces three decoupled image matrices that can be processed independently or combined according to the experimental design. In the present study, the YUV representations are used as chromatic inputs to the CNN models to investigate whether the transformed feature space improves wildfire classification.

4. CNN ARCHITECTURES AND EXPERIMENTAL METHODOLOGY

4.1 Baseline LeNet-5

LeNet-5 is used as the principal baseline CNN architecture. Its original design consists of convolutional and subsampling stages followed by fully connected classification layers [17]. The architecture is compact and therefore provides a useful baseline for evaluating whether improvements arise from the input representation and targeted architectural modifications rather than from a substantially larger network.

4.2 Modified LeNet-5

The modified LeNet-5 is designed to improve the extraction of discriminative features from the transformed chromatic inputs. According to the supplied experimental description, the principal modifications concern the number of filters, receptive-field configuration, and kernel settings. The modified network therefore retains the conceptual simplicity of LeNet-5 while adapting its feature-extraction capacity to the YUV input representation.

Table 2. Conceptual comparison of the baseline and modified LeNet-5 architectures. Exact layer-wise values should be reproduced from the experimental configuration table.

Architecture aspect	Baseline LeNet-5	Modified LeNet-5
Network design	Original compact LeNet-5 topology	LeNet-5-derived topology with targeted structural changes

Filter configuration	Baseline filter counts	Adjusted/adaptive filter counts
Receptive fields	Original configuration	Localized/tuned receptive-field configuration
Kernel settings	Baseline kernel sizes	Modified/tuned kernel sizes
Purpose	Reference model	Improved representation learning for chromatic features

The supplied manuscript does not contain the complete numerical layer-by-layer specification of the modified network. For book publication, the final version should therefore include the exact number of filters, kernel dimensions, strides, activation functions, pooling operations, dense-layer sizes, optimizer, learning rate, batch size, number of epochs, and regularization settings used in the experiments. These values should be copied directly from the actual experimental configuration rather than inferred.

4.3 Comparative CNN Models

In addition to LeNet-5, the experimental description indicates that AlexNet and VGG16 baseline and modified versions were trained using YUV chromatic representations. These models provide broader architectural benchmarks against which the compact LeNet-5-based approach can be evaluated. Their detailed hyper-parameters numerical results should be presented in the corresponding experimental values in Tables 3.

Table 3. Hyper-parameters of LeNet5

Layer	LeNet5 Original	Size	LeNet5 Modified	Size	Activation Functions
Conv1	6	5x5	64	3x3	Relu
Cov2	6	5x5	-	3x3	Relu
Pool3	-	2x2	-	2x2	Relu
Conv4	16	5x5	16	3x3	Relu
Pool5	-	2x2	-	2x2	Relu
FC6	4096		304		Relu
FC7	2		2		Softmax

4.4 Training Environment and Evaluation

The experiments were conducted in a Google Colab environment using a two-core NVIDIA GPU configuration, with Keras, TensorFlow, and OpenCV as the principal software libraries. The reported manuscript specifies a GPU frequency of 1.59 GHz. For reproducibility, the paper should additionally state the exact GPU model, Python version, random seed, train/validation/test split, optimizer, learning rate, batch size, number of epochs, and data-augmentation settings which that applicable in shown Table 4.

Table 4. Parameter values

Parameters	Values
GPU Model	Nvidia T4 GPU
Python Version	3.12
Random seed	30
Split	20
Optimizer	Adam Optimizer
Learning rate	0.00001
Batch Size	32
Number of epochs	40
Shear_ range	0.3
Zoom_ range	0.2
Horizontal flip	True
Rescale	1./255

5 RESULT AND DISSCUSSION

5.1 Evaluation Metrics

Classification accuracy is calculated as the proportion of correctly classified samples among all evaluated samples [18]. For binary classification, the Receiver Operating Characteristic (ROC) curve plots True Positive Rate (TPR) against False Positive Rate (FPR) over a range of decision thresholds. The Area Under the ROC Curve (AUC-ROC) summarizes the model's ranking performance across those thresholds. AUC-ROC must therefore not be described as a plot of TPR against True Negative Rate (TNR) [19].

5.2 Wildfire Prediction Using Chromatic Features

The experimental framework evaluates CNN models using chromatic representations. The present section focuses on the chromatic experiments based on YUV features. The reported performance is evaluated using classification accuracy and AUC-ROC, with the baseline and modified LeNet-5 models serving as the principal comparison.

5.3 Accuracy of Baseline and Modified LeNet-5

The supplied experimental results show a substantial difference between the baseline and modified LeNet-5 models when trained with YUV color features. The baseline LeNet-5 achieves an accuracy of 44%, whereas the modified LeNet-5 achieves 96%. The reported improvement indicates that the architectural adaptations substantially affect the ability of the network to learn discriminative representations from the YUV input space. The comparison of accuracy of baseline and modified LeNet-5 using YUV features shown in Table 5.

Table 5. Reported classification accuracy of baseline and modified LeNet-5 using YUV features and benchmark dataset.

Model	Reported accuracy (%)
LeNet-5 (Baseline)	44%
LeNet-5 (Modified)	96%
Sathishkumar, V. E, 2023-Xception [4]	79.23%
Geng, 2024 - YOLOv3 [6]	85%
Jin.L, 2024 – SWVR [20]	79.6%
Pang. Y, 2023 - FuF-Det [21]	78.69%
Xavier, 2023 - CNN model [22]	87.10%

5.4 AUC-ROC Analysis

The manuscript also reports ROC/AUC analysis for the baseline and modified LeNet-5 models. The modified model is described as producing the strongest ROC performance in the YUV setting. However, the supplied material uses 96.0% interchangeably for accuracy and AUC-ROC. These metrics should be reported separately in the final publication. If the measured AUC is 0.96, it should be written as AUC = 0.96 (or 96%, if the publisher permits percentage notation), while the classification accuracy should independently be reported as accuracy 96.0%.

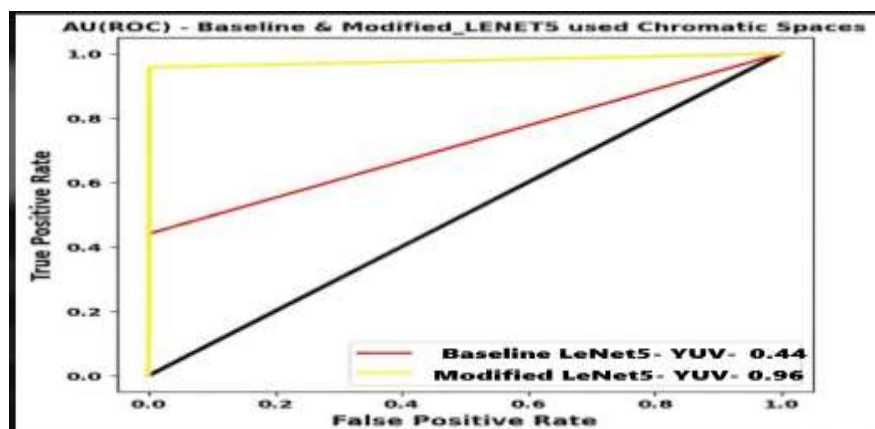


Figure 5: ROC curves for the baseline and modified LeNet-5 models using YUV chromatic features.

5.5 Interpretation of the Results

The difference between the two LeNet-5 configurations suggests that the input representation alone does not determine classification performance. The network must also have sufficient capacity and an appropriate receptive-field configuration to extract meaningful patterns from the transformed channels. The modified architecture appears to provide a more suitable feature-extraction mechanism for the YUV representation than the baseline architecture under the reported experimental conditions. Nevertheless, a high classification score should be interpreted together with the experimental protocol. The final publication should report the number of training and test samples, class balance, data-splitting strategy, confidence intervals or repeated-run statistics were available, and the complete confusion matrix. These details are important for determining whether the reported improvement generalizes beyond the evaluated subset.

6. CONCLUSION

This paper presented a chromatic feature-engineering framework for wildfire image classification using YUV representations and enhanced CNN architectures. The approach combines standardized image preprocessing, RGB-to-YUV transformation, and targeted modification of the compact LeNet-5 architecture. The FFNI subset of the DeepFire benchmark provides a balanced collection of 759 fire and 759 no-fire images, enabling a controlled comparison between the baseline and modified models. Based on the results supplied in the manuscript, the modified LeNet-5 achieved 96.0% classification accuracy using YUV features, while the baseline LeNet-5 achieved 44.0%. The modified version gave better result than baseline architecture for detection fire or none from images. The results support the central hypothesis that both feature representation and network architecture influence wildfire classification performance. Future work should the satellite surveillance real time data and deep learning architecture network use prevention of wildfire early detection, monitoring most wildfire areas. Further studies should also compare individual Y, U, and V channels, combined YUV inputs, RGB-YUV fusion, and alternative color spaces under a common experimental protocol.

7. REFERENCES

- [1] Barik A, Baidya Roy S. Climate change strongly affects future fire weather danger in Indian forests. *Communications Earth & Environment*. 2023 Dec 2;4(1):452. <https://doi.org/10.1038/s43247-023-01112-w>.
- [2] Spiller, D., Carbone, A., Amici, S., Thangavel, K., Sabatini, R., & Laneve, G. (2023). Wildfire detection using convolutional neural networks and PRISMA hyperspectral imagery: A spatial-spectral analysis. *Remote Sensing*, 15(19), 4855. <https://doi.org/10.3390/rs15194855>.
- [3] Ejaz, U., Hamza, M. A., & Kim, H. C. (2025). Channel attention for fire and smoke detection: Impact of augmentation, color spaces, and adversarial attacks. *Sensors*, 25(4), 1140. <https://doi.org/10.3390/s25041140>.
- [4] Sathishkumar, V. E., Cho, J., Subramanian, M., & Naren, O. S. (2023). Forest fire and smoke detection using deep learning-based learning without forgetting. *Fire Ecol* 19 (9). <https://doi.org/10.1186/s42408-022-00165-0>.
- [5] Vipin, V. (2012). Image processing based forest fire detection. *International Journal of Emerging Technology and Advanced Engineering*, 2(2),87-95. <https://doi.org/10.48550/arXiv.2406.04624>.
- [6] Geng, X., Su, Y., Cao, X., Li, H., & Liu, L. (2024). YOLOFM: An improved fire and smoke object detection algorithm based on YOLOv5n. *Scientific Reports*, 14(1), 4543. <https://doi.org/10.1038/s41598-024-55232-0>.
- [7] Premal, C. E., & Vinsley, S. S. (2014, March). Image processing based forest fire detection using YCbCr colour model. In 2014 international conference on circuits, power and computing technologies [ICCPCT-2014] (pp. 1229-1237). IEEE. DOI: 10.1109/ICCPCT.2014.7054883.
- [8] Liu. T, Cheng. J, Du. X, Luo. X, Zhang. L, Cheng.B and Wang. Y, 2019, "Video smoke detection method based on change-cumulative image and fusion deep network", *Sensors*, Vol.19, No.23, p.5060.
- [9] Zhang. G, Wang. M and Liu. K, 2019, "Forest fire susceptibility modelling using a convolutional neural network for Yunnan province of China", *International Journal of Disaster Risk Science*, Vol.10, No.3, pp.386-403.

- [10] Bahadori, N., Razavi-Termeh, S. V., Sadeghi-Niaraki, A., Al-Kindi, K. M., Abuhmed, T., Nazeri, B., & Choi, S. M. (2023). Wildfire susceptibility mapping using deep learning algorithms in two satellite imagery dataset. *Forests*, 14(7), 1325. <https://doi.org/10.3390/f14071325>.
- [11] Jindal, R., Kunwar, A. K., Kaur, A., & Jakhar, B. S. (2020, July). Predicting the dynamics of forest fire spread from satellite imaging using deep learning. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 344-350). IEEE. DOI:10.1109/ICESC48915.2020.9155734.
- [12] Choi, H, Kim, M, Lee, G. and Kim, W, 2019, "Unsupervised learning approach for network intrusion detection system using auto encoders". *The Journal of Supercomputing*, Vol. 75, No. 9, pp.5597-5621.
- [13] Saponara, S., Elhanashi, A., & Gagliardi, A. (2021). Real-time video fire/smoke detection based on CNN in antifire surveillance systems. *Journal of Real-Time Image Processing*, 18(3), 889-900. <https://doi.org/10.1007/s11554-020-01044-0>.
- [14] Vipin, V. (2012). Image processing based forest fire detection. *International Journal of Emerging Technology and Advanced Engineering*, 2(2),87-95. <https://doi.org/10.48550/arXiv.2406.04624>.
- [15] Khan, A., Hassan, B., Khan, S., Ahmed, R., & Abuassba, A. (2022). DeepFire: A novel dataset and deep transfer learning benchmark for forest fire detection. *Mobile Information Systems*, 2022(1), 5358359. <https://doi.org/10.1155/2022/5358359>.
- [16] Tremeau, A., Tominaga, S., & Plataniotis, K. (2008). Color in image and video processing: most recent trends and future research directions. *EURASIP Journal on Image and Video Processing*, 2008(1), 581371. doi:10.1155/2008/581371.
- [17] Poynton, C. (2012). *Digital video and HD: Algorithms and Interfaces*. Elsevier. ISBN: 978-0-12-391926-7.
- [18] Bhaiswar, R., Chawla, D., & Joshi, H. (2025). Segmenting Electronic Word-of-Mouth Consumers: A Lifestyle and Behaviour-based Study. *Global Business Review*, 09721509251374405. <https://doi.org/10.1177/09721509251374405>.
- [19] Kumar. R., & Indrayan. A. (2011). Receiver operating characteristic (ROC) curve for medical researchers. *Indian pediatrics*, 48(4), 277-287. <https://doi.org/10.1007/s13312-011-0055-4>.
- [20] Jin, L., Yu, Y., Zhou, J., Bai, D., Lin, H., & Zhou, H. (2024). SWVR: A lightweight deep learning algorithm for forest fire detection and recognition. *Forests*, 15(1), 204. <https://doi.org/10.3390/f15010204>.
- [21] Pang, Y., Wu, Y., & Yuan, Y. (2023). FuF-Det: An early forest fire detection method under fog. *Remote Sensing*, 15(23), 5435. <https://doi.org/10.3390/rs15235435>.
- [22] Xavier. K.L.B.L and Nanayakkara, V.K, 2022, "Development of an early fire detection technique using a passive infrared sensor and deep neural networks". *Fire Technology*, Vol.58, No.6, pp.3529-3552.