RUDN Journal of Ecology and Life Safety
Вестник РУДН. Серия: Экология и безопасность жизнедеятельности
ISSN 2313-2310 (Print); ISSN 2408-8919 (Online)
**2026 Vol. 34 No. 3 500–514**
http://journals.rudn.ru/ecology

DOI: 10.22363/2313-2310-2026-34-3-500-514
EDN: EVYAWA
UDC 504.004:004.9

Research article / Научная статья

# Experimental multi-criteria evaluation of convolutional neural network architectures for early forest fire detection using UAV data

**Serafim S. Zagorodnii[1] ✉, Irina K. Belova[2], Svetlana N. Nikulina[1]**

[1]*RUDN University, Moscow, Russian Federation*
[2]*Kaluga Branch of Bauman Moscow State Technical University, Kaluga, Russian Federation*
✉zagorodniy.serafim@mail.ru

**Abstract.** The objective of this study is to conduct a comparative analysis of four modern convolutional neural network architectures — ResNet50, MobileNetV2, MobileNetV3 Large, and EfficientNet-B0 — for wildfire detection on multimodal RGBT images captured by UAV-mounted cameras, with emphasis on practical applicability under resource-constrained conditions. The object of study consists of deep learning algorithms trained on the FireMan UAV RGBT dataset, comprising 12,000 annotated RGB and thermal images. The methodology is based on transfer learning with ImageNet-pretrained weights, the use of Focal Loss to address class imbalance, and 5-fold cross-validation. Objective evaluation employed non-parametric statistical tests (Kruskal — Wallis, Mann — Whitney) and a multi-criteria ranking using TOPSIS with bootstrap estimation, considering four key metrics: accuracy (F1-score), sensitivity, inference speed (FPS), and computational complexity (MACs). Experiments were conducted on the NVIDIA Jetson Xavier NX platform. Results showed that all architectures achieve high accuracy (F1-score > 0.92), but MobileNetV2 delivers the optimal balance: F1 = 0.934, 18.2 FPS, and minimal computational load (1.3 G MACs), making it the most suitable for autonomous UAV-based monitoring systems. ResNet50 and EfficientNet-B0 offer comparable accuracy but require 2–3 times more computational resources. MobileNetV3 Large achieved the highest sensitivity but lagged in speed. The scientific novelty lies in a statistically rigorous evaluation of CNN architectures not only by accuracy but also by real-world computational constraints, enabling a transition from laboratory experiments to practical deployment on UAVs. The findings provide a robust basis for selecting optimal architectures in resource-limited environmental monitoring systems.

**Keywords**: multimodal classification, transfer learning, MobileNetV2, ResNet50, EfficientNet B0, MobileNetV3 Large, focal loss function

**Authors' contribution**. *S.S. Zagorodnii* — conceptualization, methodology development; implementation of computational code and auxiliary algorithms; conducting experiments;

writing the original draft of the manuscript. *I.K. Belova* — conceptualization, co-authoring and editing the manuscript; visualization; supervision of the research. *S.N. Nikulina* — project administration. All authors have read and approved the final version of the manuscript.

**Conflicts of interest.** The authors declare no conflicts of interest.

**Statement on the use of artificial intelligence technologies.** Generative artificial intelligence technologies (ChatGPT) were used in the creation of this article. The usage period is July — August 2025. The stage of domain analysis, the study of analogues of the system under development. The authors declare that they fully understand and accept full responsibility for the content of this manuscript. The use of artificial intelligence tools was of a supporting nature and did not replace the authors' expert judgment, critical thinking, or scientific integrity.

**Data availability.** The data that support the findings of this study are available from the corresponding author upon request.

**Article history:** received 14.01.2026; revised 17.03.2026; accepted 30.03.2026.

**For citation:** Zagorodnii SS, Belova IK, Nikulina SN. Experimental multi-criteria evaluation of convolutional neural network architectures for early forest fire detection using UAV data. *RUDN Journal of Ecology and Life Safety*. 2026;34(3):500–514. http://doi.org/10.22363/2313-2310-2026-34-3-500-514 EDN: EVYAWA

# Экспериментальная многокритериальная оценка архитектур сверточных нейронных сетей для задачи раннего обнаружения лесных пожаров по данным БПЛА

**С.С. Загородний[1] ✉, И.К. Белова[2], С.Н. Никулина[1]**

[1]*Российский университет дружбы народов, Москва, Российская Федерация*
[2]*Калужский филиал МГТУ им. Н.Э. Баумана, Калуга, Российская Федерация*
✉zagorodniy.serafim@mail.ru

**Аннотация.** Цель исследования — сравнительный анализ четырех современных архитектур сверточных нейронных сетей: ResNet50, MobileNetV2, MobileNetV3 Large и EfficientNet-B0 — для детекции лесных пожаров на мультимодальных RGBT-изображениях, снятых с бортовых камер БПЛА, с акцентом на практическую применимость в условиях ограниченных ресурсов. Объектом исследования выступают алгоритмы глубокого обучения, обученные на датасете FireMan UAV RGBT, включающем 12 000 аннотированных изображений в RGB и тепловом диапазонах. Методология основана на трансферном обучении с предобученными весами ImageNet, использовании фокальной функции потерь для компенсации дисбаланса классов и 5-фолд кросс-валидации. Для объективной оценки применены непараметрические статистические тесты (Краскела — Уоллиса, Манна — Уитни) и многокритериальная ранжировка по методу TOPSIS с бутстрэп-оценкой, учитывающей четыре ключевых параметра: точность (F1-score), чувствительность, скорость инференса (FPS) и вычислительную сложность (MACs). Эксперименты проведены на платформе NVIDIA Jetson Xavier NX. Результаты показали, что все архитектуры достигают высокой точности (F1-score > 0,92), однако MobileNetV2 демонстрирует оптимальный баланс: F1 = 0,934, скорость 18,2 FPS при минимальной вычислительной нагрузке (1.3 G MACs), что делает ее наиболее подходящей для автономных систем мониторинга. ResNet50 и EfficientNet-B0 обеспечивают сопоставимую точность, но требуют в 2–3 раза больше ресурсов. MobileNetV3 Large показала наилучшую чувствительность, но уступает по скорости. Научная новизна работы заключается в комплексной, статистически обоснованной оценке CNN-архитектур не только по точности, но и по реальным ограничениям вычислительной эффективности, что позволяет перейти от лабораторных экспериментов к практическому внедрению на БПЛА. Полученные данные формируют основу для выбора оптимальной архитектуры в системах экологического мониторинга с ограниченными ресурсами.

**Ключевые слова:** мультимодальная классификация, трансферное обучение, MobileNetV2, ResNet50, EfficientNet B0, MobileNetV3 Large, фокальная функция потерь

**Вклад авторов.** *Загородний С.С.* — разработка методологии исследования; реализация компьютерного кода и вспомогательных алгоритмов; проведение экспериментов; написание первоначального текста рукописи. *Белова И.К.* — концептуализация, создание рукописи и ее редактирование; визуализация; руководство исследованием. *Никулина С.Н.* — администрирование проекта. Все авторы ознакомлены с окончательной версией статьи и одобрили ее.

**Заявление о конфликте интересов.** Авторы заявляют об отсутствии конфликта интересов.

**Заявление об использовании технологий искусственного интеллекта.** При создании настоящей статьи использовался ChatGPT. Период использования — июль-август 2025. Этап анализа предметной области, изучение аналогов разрабатываемой системы. Авторы заявляют, что полностью осознают и принимают на себя всю полноту ответственности за содержание данной рукописи. Использование инструментов искусственного интеллекта носило вспомогательный характер и не заменяло экспертного суждения, критического мышления и научной добросовестности авторов.

**Заявление о доступности данных.** Данные, использованные в этом исследовании, могут быть предоставлены авторами по запросу.

**История статьи:** поступила в редакцию 14.01.2026; доработана после рецензирования 17.03.2026; принята к публикации 30.03.2026.

**Для цитирования:** *Загородний С.С., Белова И.К., Никулина С.Н.* Экспериментальная многокритериальная оценка архитектур сверточных нейронных сетей для задачи раннего обнаружения лесных пожаров по данным БПЛА // Вестник Российского университета дружбы народов. Серия: Экология и безопасность жизнедеятельности. 2026. Т. 34. № 3. С. 500–514. http://doi.org/10.22363/2313-2310-2026-34-3-500-514 EDN: EVYAWA

## Introduction

Forest fires are among the most destructive natural phenomena, causing significant damage to ecosystems, infrastructure, and economies. In recent decades, climate change has increased the frequency and area of fires in many regions of the world, from Europe and North America to countries in the Asia-Pacific region [1–4]. The effectiveness of measures for locating and extinguishing fires directly depends on the timeliness of detecting fodder, even a delay of a few minutes can lead to irreversible consequences.

Traditional monitoring methods for ground-based sensors and satellite systems have several limitations: high cost, low spatial resolution, limited coverage area and data delay [5; 6]. Unmanned aerial vehicles (UAVs), which can rapidly obtain high-quality real-time RGBT images, are increasingly important in these conditions. However, the effectiveness of such systems is determined by the accuracy of the image classification: false triggers overload communication channels and missed hotbeds reduce the reliability of the entire system [7–10].

Particular attention is being paid to the introduction of compact and energy-efficient models of neural networks (CNNs) capable of operating on board UAVs. This allows for a preliminary filtering of images "on the fly", transmitting the most informative data to the ground server, which is crucial when working in remote areas with a secure connection [11].

**The aim of this study** was a comparative analysis four modern CNN architectures in terms of accuracy, stability, speed and computational complexity. The evaluation was carried out on FireMan UAV RGBT dataset, a unique source of multimodal data specifically created for aerial monitoring tasks [12]. Its use is due to both practical relevance and scientific novelty. This set is rarely found in publications.

The scientific novelty of the work is to apply a complex multi-criteria assessment based on the TOPSIS method, bootstrap analysis and sensitive weight analysis. Unlike most studies, where models are compared on a metric basis (e.g., accuracy), the balance between quality, stability and performance has been taken into account to identify an architecture that is optimal for implementation in real-world on-board systems.

## Materials and methods

Dataset, which was used in the study, contains synchronized pairs of images in visible (RGB) and infrared (IR) ranges obtained from the UAV during controlled passes in 2022–2023. The RGB and IR flows are synchronized by time and pixel-to-pixel by field of view, which ensures an exact match of scenes in both channels. All images are in JPG format with resolution 948×754 pixels and 24 bits deep color. Before training, infrared images were subjected to intensity normalization and scaled to the network input size with relative temperature gradients.

Images are annotated manually and semi-automatically, each pair of images (RGB+IR) is provided with a binary label indicating the presence or absence of fire in the frame. The data cover different scenarios of forest and man-made fires, as well as fire-free backgrounds, making it a suitable set for image classification tasks.

To ensure the reproducibility and objectivity of the experiments, the sample was divided into two non-overlapping subsamples:

– Training + validation — 81% of all images;

– Testing — 19% of all images.

Within the test sample, the image pair is divided into classes Fire and No_Fire at a ratio of 80.7 to 19.3% respectively. In the training and validation subsample, there is a similar but more pronounced shift — 95 to 5% in favor of the Fire class. This imbalance is due to the structure of the source dataset, where images are grouped into 14 synchronized RGB-IR video sequences. The different duration of videos and uneven development of fires lead to a different number of frames belonging to each class [14].

The identified class imbalances create a number of limitations in teaching classification models. The predominance of images with fires can lead to a shift in the model towards the dominant class, reducing the accuracy of predictions for images without fire. Moreover, the difference in class ratios between the training and test samples makes it difficult to assess objectively the quality of models and their generalizability.

To minimize the impact of class imbalances, the training sample was balanced by increasing the number of images of an underrepresented class using augmentative methods, with the aim of balancing not being a full alignment of classes. As a result of the processing, the ratio of images in the instructional and validation samples was proportioned to be close to the distribution in the test sample and amounted to 77% of Fire class frames and 23% of No_Fire class frames.

Additionally, in the training process, weighting coefficients of classes, which are inversely proportional to their representation in the training sample, were used and included in the loss function. This combination of data balancing and the weighted loss function ensured that the contribution of classes to model optimization was equalized, and reduced the model's shift toward dominant classes.

Overoptimization control was performed by splitting the data into non-intersecting folds, preserving the integrity of video sequences so that images from the same video sequence did not overlap between the instructional, validation, and test samples. Additionally, a comparison of the quality metrics calculated on validation and test samples was carried out, which made it possible to assess the model's stability and the correctness of its generative capacity.

Four pre-trained neural network architectures implemented in the TensorFlow Keras Applications library are selected for experiments: ResNet50, MobileNetV2, MobileNetV3 Large and EfficientNet B0. The ResNet50 architecture has been selected as a classical deep model with a large number of parameters, allowing it to be used as a reference for quality control and stability in classification. The remaining architectures are in a class of lightweight models optimized to work under conditions of limited computing resources, which corresponds to the target task of implementation on board unmanned aerial vehicles [13].

In all cases, a transfer training approach was used using the pre-trained ImageNet set of weights. This choice is due to the limited sample size and the need to accelerate learning convergence, which reduces the risk of overlearning and ensures a more stable generalization of models when further training on specialized data.

Designed architecture of a multimodal superimposed network with two parallel flows for RGB and IR images. Each flow uses its own pretrained backbone, such as ResNet50, with global feature averaging and normalization. In the late fusion stage, traits are concatenated and transmitted through a two-layer fully connected classifier: first layer with 256 neurons, ReLU, BatchNorm and Dropout 0.5; second 64 neurons, ReLU and Dropout 0.3; outgoing layer with a sigmoid for fire probability. This design provides a balance between the number of parameters and the informational nature of the features. A conceptual diagram is shown in Figure 1.

In the study, the effectiveness of four superimposed neural network architectures was analyzed to solve the problem of binary classification of images obtained from unmanned aircraft in RGB and IR ranges. Each architecture is adapted for a multimodal production, in which images from two spectral channels are processed independently and then their features are combined to increase resilience against external conditions of the shoot. Such a strategy ensures the integrated use of visual and thermal information, which is particularly important for early fire detection in low-light and smoke conditions [15].

Model quality assessment is a three-fold group cross validation (3 fold GroupKFold), in which all frames belonging to the same video sequence are treated as a single group and not separated between the instructional and validation samples. This approach prevents data leakage between folds and provides an objective assessment of models' ability to generalize.

Each architecture is taught in 15 independent runs, executed with different initial conditions and random data splits. This reduces the impact of random factors, such as footing or model initial weights. The result is 45 independent trained models for each architecture (3 folds of 15 runs), which provides a statistically stable sample of results.

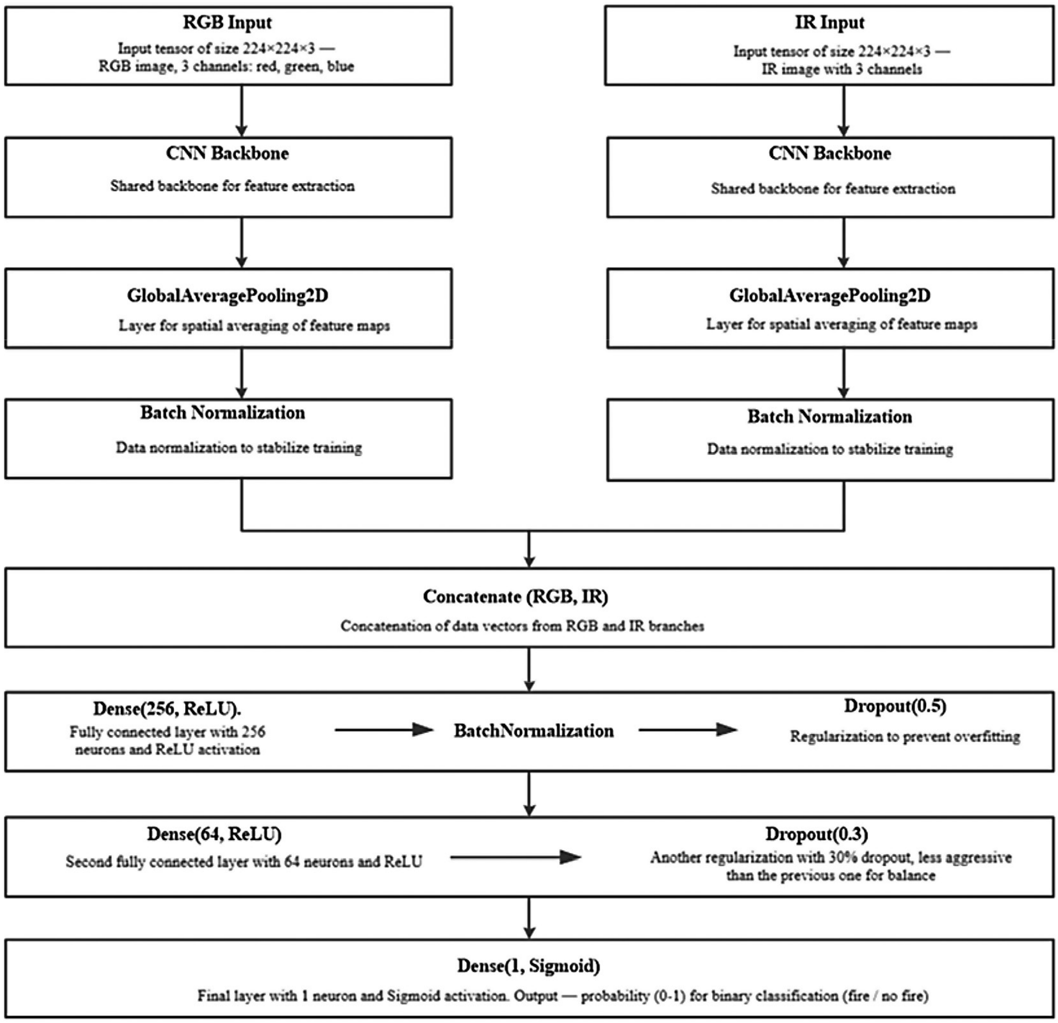


**Figure 1.** General architecture of the multimodal classification system

*Source:* compiled by the S.S. Zagorodnii.

After training and validation, each model was tested on an independent subsample; 45 representative results were obtained for each architecture, allowing the assessment of average values and variances in quality and computational efficiency metrics, and perform a strict statistical comparison between models.

Training was performed with the optimizer Adam (initial learning speed $1\times10^{-4}$) with adaptive reduction learning rate (ReduceLROnPlateau — 2 times in the absence of improvement val_loss during 3 epochs, at least $1\times10^{-6}$). Early stop was used to prevent overlearning (EarlyStopping – if there is no improvement val_loss during the 6 epochs, with the restoration of better weights). Maximum number of epochs – 40, batch size — 32, input image size — 224×224 pixels.

Taking into account the imbalances of classes (77% fires, 23% no), a focal loss function ($\gamma = 2.0$, $\alpha = 0.25$) was used, reinforcing the contribution of complex examples (weak fire, smoke, reflections), and adding weight of classes for balanced learning.

To enhance the model's generalizability, data learning intermingled before each epoch, which prevents memorization of example sequences and promotes more consistent weight updates.

The process of training and validation was accompanied by monitoring of the following metrics:

– Accuracy — proportion of correctly classified images;

– Precision — accuracy of the definition of the "fire" class;

– Recall — full fire detection;

– AUC — area under the ROC curve, reflecting the model's overall ability to distinguish between two classes.

Metric data were used to comprehensively assess the effectiveness of teaching.

Upon completion of the training, each of the 45 models in each architecture was tested on a pending test sample not used in the training and validation process. For analysis, all quality metrics and resource efficiency values were retained.

For each architecture, a sample of 45 results per metric was formed. In the first stage, the normality of metric distributions was tested using the Shapiro-Wilk criterion. Since the distribution of data was not normal, a non-parametric Kruskal –Wallis criterion was used to identify statistically significant differences between architects. For a detailed analysis of the differences between pairs of models, the Dunne test with Bonferroni correction was used, which allowed to control the accumulated probability of error of the first kind in multiple comparisons. Statistical processing was performed in a Python environment using libraries SciPy and scikit-posthocs.

The TOPSIS multi-criteria analysis method using the scikit-criteria library was used for complex model ranking. All metrics have been provisionally normalized by the Min-Max method, with values where minimum values are preferable (such as infensory time) inverted. The analysis excluded over-correlated or low variability. The weight of criteria was determined by a combination of statistical significance and expert assessment of their impact on practical application of models. To check the stability of the ranking, a bootstrap analysis and a sensitive weight analysis of the criteria were carried out.

This multi-step approach provides an objective assessment of architecture in both classification quality and computational characteristics, enabling the identification of architectures most suitable for practical use under conditions of limited UAV resources and mobile systems.

## Results

The results of a pilot study aimed at assessing the effectiveness of neural network architecture are presented below. The Shapiro-Wilk test for the normality of metric distributions showed that the vast majority of samples including all metrics for ResNet50, MobileNetV2 and EfficientNet B0 deviated significantly from the normal distribution. This is due to the stochastity of neural network learning, variability in the initialization of weights and computational environment. Therefore, non-parametric methods were chosen for analysis: the Kruskal—Wallis criterion for cross-group comparison and the Bonferroni-adjusted Dunna test for pair differences, which ensured stability of results with abnormalities and emissions.

The analysis revealed statistically significant differences ($p < 0.05$) between architectures for most computational performance indicators, indicating unequal optimization and suitability for resource-constrained systems. The overall quality metrics of the Accuracy, Precision, F1 score, MCC and Balanced Accuracy classification showed no significant differences ($p > 0.05$), which indicates comparable accuracy and balance of all models in the classification task.

The results of the Kruskal — Wallis criterion show significant differences in sensitivity (Recall), ranking quality (AUC, AP) and resource efficiency metrics with a similar level of overall classification accuracy between models (Figure 2).

This highlights the need for integrated model assessment, taking into account not only classification quality but also computational performance, especially when applied in real-time systems.

In the final stage of statistical analysis using the Dunna test with Bonferroni correction, it was found that the ResNet50 architecture pairs, MobileNetV2, MobileNetV3 Large and EfficientNet B0 show statistically significant differences in classification quality and computational efficiency metrics ($p < 0.05$), which confirms their heterogeneity. In particular, between MobileNetV2 and MobileNetV3 Large, as well as MobileNetV3 Large and EfficientNet B0, the differences proved to be insignificant, indicating their comparable behavior despite architectural differences, which is logical for models optimized for limited resources. In contrast, ResNet50 as a reference model with many parameters is substantially different from all the others. The TOPSIS method, which is emission-resistant and does not require data allocation assumptions, has been used to rank architectures in a comprehensive manner. Before analysis, all metrics were normalized by the Min-Max method in the range [0.1], and for indicators where lower values are preferable (such as Training Time), an inversion was performed. The correlation analysis revealed an almost complete dependence between FLOPs, Params and Model Size ($r = 0.999$), so the last two metrics are excluded. The final set for TOPSIS included: Recall, Precision, Average Precision, FLOPs, Average FPS and Training Time to provide a balance between quality and efficiency.

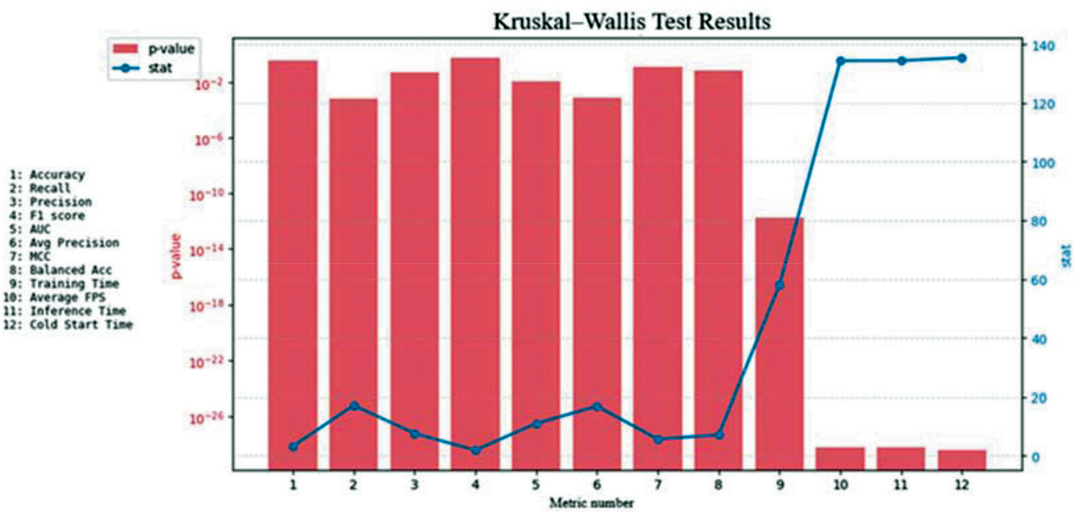


**Figure 2.** Kruskal — Wallis test results diagram

*Source:* compiled by the I.K. Belova.

To perform the TOPSIS method, it was necessary to determine the weights of the criteria reflecting their relative importance (Figure 3).

The distribution of weighting coefficients of indicators is supported by the results of statistical analysis carried out using the Kruskal — Wallis criterion and subsequent pairs comparisons. In the formation of the weight structure, the principle of equal balance between quality metrics and resource efficiency metrics (0.5:0.5) was observed, which ensured an equivalent influence of both groups of criteria on the final integrated assessment within the framework of the TOPSIS method.

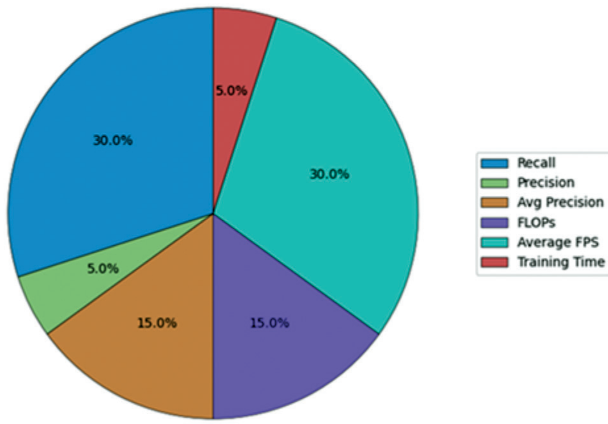


**Figure 3.** Diagram of metric weights distribution for multi-criteria ranking of model architectures using the TOPSIS method

*Source:* compiled by the I.K. Belova.

Among the quality metrics, the highest weight (0.30) is given to the Recall score, which has demonstrated statistically significant differences between architectures and is crucial for the fire detection task at hand, where minimizing event passes is a priority.

Among resource efficiency metrics, the highest weight (0.30) was given to Average FPS because this indicator demonstrated statistically significant differences between architectures and is critical for systems operating on drones, where high processing speed is important.

Thus, the proposed weighting system provides a rational relationship between qualitative and resource characteristics of models, reflecting as statistical significance differences between architects, so as their practical relevance to the task of detecting fires in real time.

After forming the weights and aggregation of metrics for each model architecture, a multi-criteria analysis was carried out using the TOPSIS method to determine the degree of approximation of each model to the ideal solution. Highest TOPSIS integral score obtained for MobileNetV2 (0.87). This demonstrates the most balanced combination of quality of detection and efficiency in the use of computing resources. MobileNetV3 Large (0.72), demonstrated a lower degree of optimality but also high-quality predictions. The

EfficientNet B0 (0.43) and ResNet50 (0.35) models are less balanced and have limited applicability in conditions requiring high processing speed and limited computing power.

To assess the reliability and sustainability of the results, a bootstrap resampling was performed based on 1,000 samples of original data. The method is used to construct confidence intervals for TOPSIS estimates totals and determine the degree of variability in estimates (Figure 4).

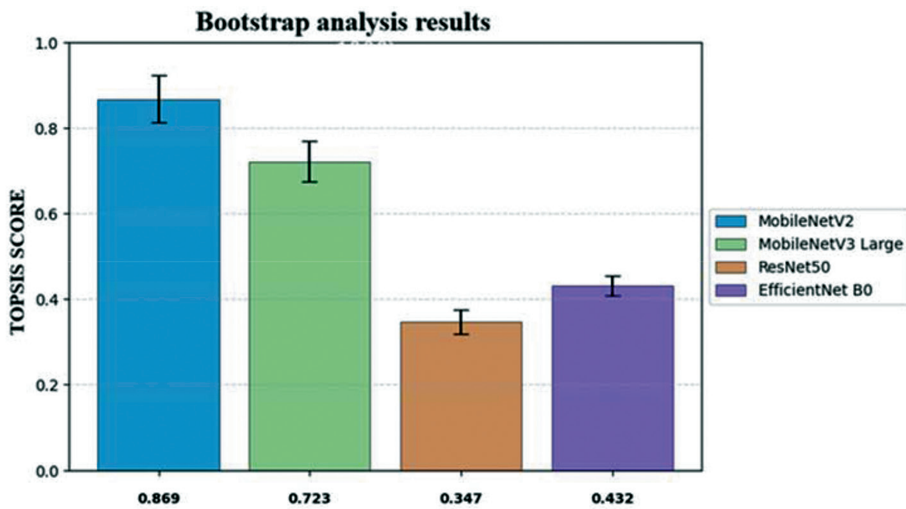


**Figure 4.** Diagram of average TOPSIS scores for model architectures
*Source:* compiled by the I.K. Belova.

The MobileNetV2 model shows the highest mean TOPSIS = 0.869 at a narrow 95% confidence interval [0.813; 0.923], indicating stability in its dominance. MobileNetV3 Large is also characterized by fairly narrow spacing and reproducibility of the result. The EfficientNet B0 and ResNet50 models show lower mean values and do not overlap with leading architectures, which confirms the statistical stability of the hierarchy.

For additional testing of the stability of the obtained conclusions, a sensitive analysis of weights (sensitivity analysis) was carried out. His goal was to assess the impact of changes in metric weights on the final ranking of architectures. Within the analysis, TOPSIS analysis was repeated several times (1,000 iterations) with weight variation within 20% of the original values, provided that quality metrics remained a priority over resource efficiency indicators (quality resource efficiency).

The results showed absolute stability of the MobileNetV2 model in 100% iterations, while MobileNetV3 Large was stable at 2nd place. Thus, the sensitive analysis confirms that TOPSIS results have a high degree of resilience even with significant changes in the priority structure between criteria. This means that MobileNetV2 dominance is not a result of the selection of a particular weighing configuration but reflects its objective advantage in terms of overall quality and resource efficiency.

The joint use of TOPSIS, bootstrap and sensitive analysis allowed not only to identify the optimal architecture but also to ensure the reliability, statistical significance and reproducibility of the results obtained, which ensures a high level of confidence in the final conclusions of the study.

## Conclusion

During the study, a comprehensive assessment of four architectures of tunneling neural networks — MobileNetV2, MobileNetV3 Large, EfficientNet B0 and ResNet50 — was carried out for the task of binary classification of images of forest fires obtained from unmanned aerial vehicles. For the multi-criteria evaluation of model performance, accuracy and resource efficiency measures, TOPSIS methods, bootstrap analysis and sensitive weight analysis of criteria were used.

The TOPSIS method allowed to identify the most balanced architecture, taking into account both qualitative and resource characteristics. Highest integral score obtained for MobileNetV2 (0.87) This architecture demonstrated an optimal combination of high fire sensitivity (Recall) and video processing speed (Average FPS), which is critical for fire monitoring systems on board the UAV. MobileNetV3 Large ranked second (0.72), while EfficientNet B0 and ResNet50 showed lower balance and limited applicability for high-speed tasks.

Thus, the joint use of TOPSIS, bootstrap and sensitive analysis has:

– to identify MobileNetV2 as the most balanced architecture for early detection of UAV fires;

– to confirm its statistical stability and reproducibility;

– to base the choice of architecture not only on accuracy but also on performance, which is critical for real systems.

The work contributes to the development of a methodology for evaluating CNN architectures for multimodal tasks by offering a comprehensive approach combining statistical validation, multi-criteria analysis and resilience to changing conditions. The results can be used in designing on-board fire monitoring systems, where a balance between quality, speed and energy efficiency is required.

## References


[1] Akhloufi MA, Castro NA, Couturier A. UAVs for wildland fires. *Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything*. 2018;10643:134–147. https://doi.org/10.1117/12.2304834

[2] Bouguettaya A, Zarzour H, Taberkit AM, Kechida A. A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms. *Signal Processing*. 2022;190:108309. https://doi.org/10.1016/j.sigpro.2021.108309 EDN: HTSLZA

[3] Burton C, Lampe S, Kelley DI, Hantson S, Christidis N, Gudmundsson L, Forrest M, Burke E, Chang J, Huang H, Ito A, Kou-Giesbrecht S, Lasslop G, Li W, Nieradzik L, Li F, Chen Ya, Randerson Ja, Reyer ChPO, Mengel M. Global burned area increasingly explained by climate change. *Nature Climate Change*. 2024;14(11):1186–1192. https://doi.org/10.1038/s41558-024-02140-w EDN: READTK

[4] Ghali R, Akhloufi MA, Mseddi WS. Deep learning and transformer approaches for UAV-based wildfire detection and segmentation. *Sensors*. 2022;22(5):1977. https://doi.org/10.3390/s22051977 EDN: HYRIMZ

[5] Ghali R, Akhloufi MA. Deep learning approaches for wildland fires remote sensing: classification, detection, and segmentation. *Remote Sensing*. 2023;15(7):1821. https://doi.org/10.3390/rs15071821 EDN: BRAETM

[6] Hossain FMA, Zhang YM, Tonima MA. Forest fire flame and smoke detection from UAV-captured images using fire-specific color features and multi-color space local binary pattern. *Journal of Unmanned Vehicle Systems*. 2020;8(4):285–309. https://doi.org/10.1139/juvs-2020-0009 EDN: JPMXBA

[7] Khan A, Hassan B, Khan S, Ahmed R, Abuassba A. DeepFire: A novel dataset and deep transfer learning benchmark for forest fire detection. *Mobile Information Systems*. 2022;2022(1):5358359. https://doi.org/10.1155/2022/5358359 EDN: PJBAGK

[8] Kularatne W, Sedaghat Shayegan K, Álvarez Casado C., Rajala J, Hänninen, T, Bordallo López M, Nguyen N Le. FireMan-UAV-RGBT (1.0.0-beta.2). *Zenodo*. 2024. https://doi.org/10.5281/zenodo.13732947

[9] Perevedentsev Y, Sherstyukov B, Gusarov A, Aukhadeev T., Mirsaeva N. Climate-induced fire hazard in forests in the Volga federal district of European Russia during 1992–2020. *Climate*. 2022;10(7):110. https://doi.org/10.3390/cli10070110 EDN: WIFGSO

[10] Pryadilina NK, Zinoveva IS, Skvorcov EA. Economic assessment of the consequences of forest fires (on the example of the Ural federal district). *Actual Directions of Scientific Researches of the XXI Century: Theory and Practice*. 2024;12(1):79–94. https://doi.org/10.34220/2308-8877-2024-12-1-79-94 EDN: AZXDJX

[11] Shamsoshoara A, Afghah F, Razi A. Aerial imagery pile burn detection using deep learning: the FLAME dataset. *Computer Networks*. 2021;193:108001. https://doi.org/10.1016/j.comnet.2021.108001 EDN: OOWULQ

[12] *Wonjae Lee, Seonghyun Kim, Yong-Tae Lee, Hyun-Woo Lee, Min Choi.* Deep neural networks for wild fire detection with unmanned aerial vehicle. *2017 IEEE International Conference on Consumer Electronics (ICCE).* Las Vegas, NV; 2017. p. 252–253, https://doi.org/10.1109/ICCE.2017.7889305.

[13] Nikulina SN, Belova IK, Vasiliev VV. Selection of the optimal method for training neural networks for solving environmental monitoring problems of industrial zones. In: *Current Problems of Ecology and Environmental Management: Collection of Scientific Papers of the XXVI International Scientific and Practical Conference*. In 3 vols. Moscow; 2025. p. 130–138. (In Russ.)

[14] Vasiliev VV, Belova IK. Data preparation for automatic relief recognition from unmanned aerial vehicles using neural networks. In: *Research and Innovation: Synergy of Knowledge and Practice: Collection of Articles of the II International Scientific and Practical Conference*. Moscow; 2025. p. 55–61. (In Russ.) EDN: KEJWBU

[15] Belova IK, Vidyaev KA. Using neural network technologies for assessing the condition of forest resources. *Paradigma*. 2025;(4-2):26–31. (In Russ.) EDN: YEZERT

**Bio notes:**

*Serafim S. Zagorodnii*, Postgraduate Student, RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation. ORCID: 0009-0009-7500-7307; eLIBRARY SPIN-code: 8839-5769. E-mail: zagorodniy.serafim@mail.ru

*Irina K. Belova*, Associate Professor, Candidate of Physical and Mathematical Sciences, Kaluga Branch of Bauman Moscow State Technical University, Russian Federation, 1 Universitetsky Gorodok St, Kaluga, 248002, Russian Federation. ORCID: 0009-0000-6337-5342; eLIBRARY SPIN-code: 3560-5659. E-mail: belova.ik@bmstu.ru

*Svetlana N. Nikulina*, Associate Professor, Candidate of Technical Sciences, RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation. ORCID: 0009-0001-5958-5168. E-mail: nikulina-sn@rudn.ru