IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

The Impact of Image Classification Using Convolutional Neural Network on Wildfire Detection

Shaunak Thamke, Elliot Blin, Kevin Diao

Bronx High School of Science, 75 W 205th St, Bronx, NY 10468; shaunakthamke@gmail.com, elliotblin2509@gmail.com, diaokevin2@gmail.com

ABSTRACT: Wildfires pose significant threats to human populations, ecosystems, and infrastructure worldwide, with early detection remaining a critical challenge. The purpose of this study was to investigate the use of a convolutional neural network (CNN) for wildfire detection through image classification, to optimize factors such as dataset size, number of epochs, and image resolution to enhance accuracy. A CNN model was trained on a wildfire dataset of 2,700 images, categorized into "fire" and "no-fire," utilizing key libraries such as TensorFlow, Pillow, and Sci-Kit Learn for preprocessing, training, and evaluation. The model was iteratively optimized by varying the number of photos, epochs, and image resolutions, with larger datasets and increased epochs significantly improving performance. Results showed that 100x100-pixel images provided the best balance between detail and computational efficiency. These findings highlight the potential of CNNs in wildfire detection, offering a faster and more reliable solution compared to traditional methods. The approach can save lives, protect property, and reduce environmental damage. Future research should focus on enhancing the model's robustness and exploring real-time applications for early wildfire detection using drones or satellites.

KEYWORDS: Earth and Environmental Sciences, Environmental Effects on Ecosystems, Convolutional Neural Networks, Image Classification, Wildfire Detection.

Introduction

In the current state of the world, wildfires are undoubtedly some of the biggest threats to the human population. Wildfires are unpredictable and uncontrolled, and usually start in natural areas, such as forests, grasslands, or prairies. The environment suffers quite a great deal from the impact of wildfire. According to *iii.org*, in 2020 alone, over 10 million acres were burned down due to wildfires. Furthermore, these wildfires leave numerous people without a roof over their heads.

Research on wildfires has been conducted in the past. Research about wildfires has examined the effects of inhaling smoke from wildfires on individuals' health on a more general scale. Earlier research has shown that inhaling such smoke can lead to direct toxicity, oxidative stress, inflammatory reactions, immune dysregulation, genotoxicity, mutations, and other adverse effects.[1] Similarly, other research has been conducted to find the impact of specific wildfire smoke on cardiovascular health. It was found that inhaled wildfire smoke, which contains particulate matter—the primary pollutant in wildfire smoke—can lead to numerous cardiovascular issues. These issues can primarily affect the elderly, pregnant women, young individuals, and the healthy, with symptoms including systemic inflammation and vascular activation.[2] These papers illustrate the destructiveness of wildfires on people. As stated before, the environment also gets damaged by wildfires. Research was done on the impact of wildfire on surface water quality. The study was specifically conducted to examine the differences between pre- and post-wildfire water. They found that within a year, the pH of the water changed, but it returned to normal after a year. The concentration of contaminants in the water also increased significantly after the wildfire. Total Suspended Solids also increased in waters with a flow of 10 m3s–1, while higher flow decreased the total suspended solids, possibly by dilution. This research demonstrates how wildfires can impact water quality, potentially affecting the environment and human health.[3] We can dramatically lower this number by effectively distinguishing real threats of wildfires from non-threatening fires.

Past studies have found the dangers of wildfires to people, their homes, and animals. This research included developing a machine learning model using neural networks that could accurately predict where wildfires could happen. More specifically, this machine learning model can recognize wildfires before they fully develop; then, at such a small scale of a fire, we can quickly put it out. The goal of the research was to see if the CNN model could detect wildfires. The hypothesis was that if certain factors for the CNN model are changed, can the model's accuracy in predicting wildfires from a picture increase? This research could benefit anybody living in wildfire-prone areas, animals in those areas, and nature.

Methods

Libraries needed:

For the project, the Sci-Kit Learn library was utilized. It employed various machine learning tools, including utilities for preprocessing data, splitting datasets, and evaluating models. Using Scikit-learn, data was loaded and prepared to train a CNN model that could make predictions. The model was evaluated for its predictive ability (accuracy). Thus, it was en-

DOI: 10.36838/v8i3.97

sured that data preparation, model training, and evaluation were done most efficiently.

Pillow, a library specializing in image-processing tasks, was also used. With Pillow, important image manipulation features, such as cropping, rotating, and applying filters to images, were possible. This was specifically helpful for the project as it involved many photos.

Specifically, Pillow was also used in the project to resample the picture. Resampling changes the image's size and, therefore, quality. This was helpful for the project, as the size of an image could correlate with the correctness of the model predictions.

TensorFlow was another library used that offered all the tools to define, train, and deploy models. It was specifically helpful in making the CNN model. TensorFlow has been helpful in training neural networks in the past. Training and testing of the neural network have mostly been done using this library.

These libraries were imported at the start of the Python code. Without these libraries, essential steps could not be taken, and the creation of the CNN model was not possible.

Dataset used:

"The Wildfire dataset" was used for the project because of its diversity, scale, and quality. The dataset spanned 2,700 aerial and ground-based images collected from various online sites and databases such as Flickr and Unsplash, capturing a wide range of forest types and geographical accuracy. The dataset is organized into clear 'fire' and 'no-fire' categories to simplify processing and model training. Additionally, the high-resolution images ensured that the detailed data was input into the model.

Figure 1: A picture of a real wildfire in the Wildfire Dataset named "fire" for modeling. Note. From The Wildfire Dataset, by Ismail El Madafri,[4] updated a year ago, Kaggle. https://www.kaggle.com/datasets/elmadafri/the-wildfire-dataset.

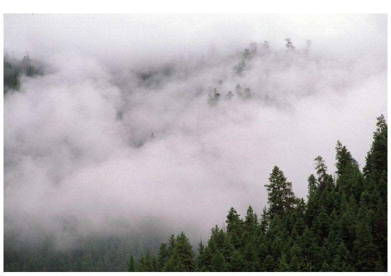

Figure 2: A picture of a forest without a wildfire from the Wildfire Dataset named "no-fire" for modeling. Note. From The Wildfire Dataset, by Ismail El Madafri[4], updated a year ago, Kaggle. https://www.kaggle.com/datasets/elmadafri/the-wildfire-dataset.

Importing Images:

Images from the wildfire dataset were stored in two files. The file containing the pictures of real wildfires was named "fire," and the one without was named "no-fire," as shown in (Figure 1) and (Figure 2), respectively; just as examples. These files were input into the code by specifying their paths and assigning a numerical label. The pictures from the "fire" folder were assigned the label 0, while the pictures from the "no-fire" folder were assigned the label 1. Labeling the two files allowed the model to map the input pictures to the correct output. This helped the model understand the difference between the two categories, allowing the model to predict whether a new image contained a wildfire. Photos were rescaled to see if they affected the model's accuracy.

Splitting the data into training and testing sets:

After importing the images, a very important step was to split the data into training and testing sets, as illustrated in Figure 3. Splitting data into training and testing sets is incredibly important for evaluating the model's performance. The training set is used to teach the model to identify patterns that enable classification. The testing set acts as the unseen data to see the model's ability to classify the new images correctly. This can also help identify overfitting, where a model performs well on the training data but not on the new data. The dataset was split 80:20 for training and testing, respectively; ideally, a validation set would be included, but limited data prevented this.

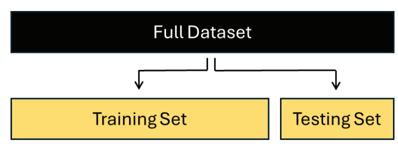

Figure 3: A Model's Division of Data into Training and Testing Data by Sydorenko.[5] This figure shows the standard machine learning practice of splitting a full dataset into training and testing sets. The training set is used to teach the model, while the testing set evaluates its performance on unseen data.

The CNN:

The CNN model used in this research is an artificial intelligence model specializing in processing data with a grid-like topology, such as images. The model is frequently used for tasks like image recognition, object detection, and natural language processing. To achieve this, the CNN model employs a range of convolutional layers, pooling layers, and fully connected layers as shown in Figure 4. The primary function of the convolutional layers is to apply filters to the input of the image to extract its features. In the code, 32 features were detected by a 3x3 filter slid across the image. The output of this process is the production of 32 feature maps, which are slightly smaller than the input images due to the use of 3x3 filters. This is what the first convolutional layer did. A pooling layer was then used to reduce the spatial dimensions of the feature maps. The spatial dimensions were the height and width of a feature map. The output of the pooling layer was that the width and height of the feature map were reduced by half while maintaining the depth of 32. A second convolution

DOI: 10.36838/v8i3.97

layer was used. For the second convolution layer, 64 different features were detected instead of 32. The size of the filters remained 3x3. The output produced 64 feature maps based on the reduced-size input from the previous pooling layer. These 64 different features were many high-level features than those seen in the first pooling. Another pooling layer was used. The spatial dimensions were further reduced with the use of this pooling layer. The final feature maps were therefore created. The feature maps were three-dimensional (3D). The maps needed to be flattened to make the 3D feature maps into a 1D vector. The CNN contained 128 neurons in each layer. After all of this, the output layer for classification was produced. The model was then compiled. The weights associated with each input were optimized during the training process. The accuracy was also tracked during the training. After the model was compiled, it was ready for training. While training the model, epochs and batch size were identified as the two factors that could affect the accuracy of the model predictions. The epoch was the number of times the model iterated over the entire training dataset for a certain number of times. The batch size was the number of samples into which the dataset was divided. The model updated weights after processing each batch. The weights were then used later when the model was tested on new images for correct classification.

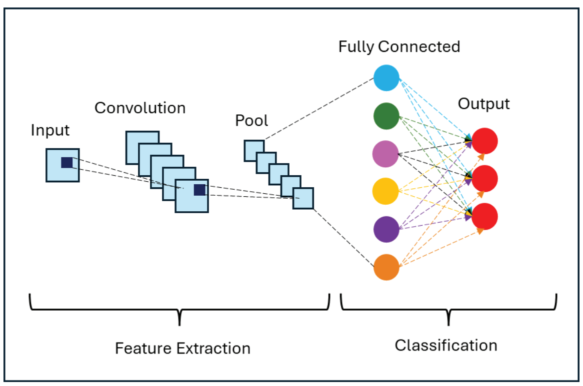

Figure 4: Schematic of CNN architecture (Sthece – ResearchGate), that processes input data through feature extraction layers (convolution and pooling operations) before passing the extracted features to fully connected layers for final classification. The network transforms raw input through hierarchical feature learning, ultimately producing multiple output classifications represented by the red nodes on the right.

Testing the CNN Model:

After being trained, the CNN model was tested on a new data set that it had never seen before. The accurate labels were also stored so that the model could check its accuracy. The model processed the input through its already trained layer to make predictions. The predicted output from the model was compared with the accurate labels. The percentage of classifications was then used to calculate the model's accuracy in detecting correctly. To further evaluate the model's performance, we adjusted various factors in the code during training and testing, including the number of epochs, the number of photos, and the size of the photos.

Results and Discussion

The study's findings demonstrated the potential of CNNs in wildfire detection, as evidenced by varying levels of accuracy based on changing factors, including dataset size, epochs, and image resolution. By optimizing these variables, the CNN model achieved high accuracy in predicting wildfires, as shown in Table 1. The graph in Figure 5 illustrates the relationship between the size of the training dataset (x-axis) and the accuracy (y-axis) of a CNN in identifying wildfires. As the dataset size increased, the model's accuracy improved, demonstrating the importance of larger datasets in enhancing performance in wildfire detection.

Table 1: The average accuracy of each epoch depends on how many photos were in the training and testing data.

	Epoch = 5	Epoch = 10	Epoch = 15	Epoch = 20	Epoch = 25
10 photos	50.00%	50.00%	50.00%	50.00%	50.00%
50 photos	65.00%	60.00%	65.00%	65.00%	70.00%
100 photos	65.00%	67.50%	67.50%	62.50%	67.50%
500 photos	70.50%	71.50%	72.50%	72.00%	73.50%
700 photos	73.50%	72.43%	73.57%	72.36%	74.62%

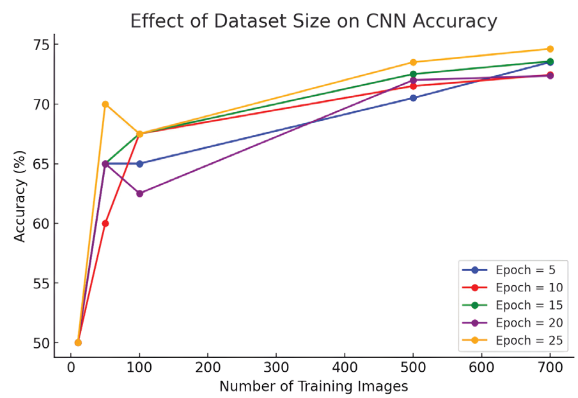

Figure 5: Effect of Dataset Size on CNN Accuracy in Wildfire Identification. Increasing the dataset led to higher model accuracy, highlighting the role of data scale in improving wildfire detection.

Finding various factors that affected the accuracy of the CNN model was very important. The epoch during which the highest accuracy was recorded is noted, as shown in Table 2. The graph in Figure 6 illustrates the highest accuracy achieved by a model at different epochs (x-axis), with accuracy (y-axis) plotted against the number of epochs. It also demonstrates how model performance varies with the number of training epochs. If the model reached 100% multiple times, the epoch (=15), where 100% was first reached, was recorded.

Table 2: Highest accuracy of a certain number of epochs, on detecting a fire or not, using 700 photos for the training and testing data. The epoch numbers shown in brackets in the table below correspond to the epoch in Figure 6 against which the accuracy is noted.

	Epoch = 5	Epoch = 10	Epoch = 15	Epoch = 20	Epoch = 25
700 photos	90.29% (epoch=5)	97.63% (epoch=9)	100.00% (epoch=15)	100.00% (epoch=19)	100.00% (epoch 17)

DOI: 10.36838/v8i3.97

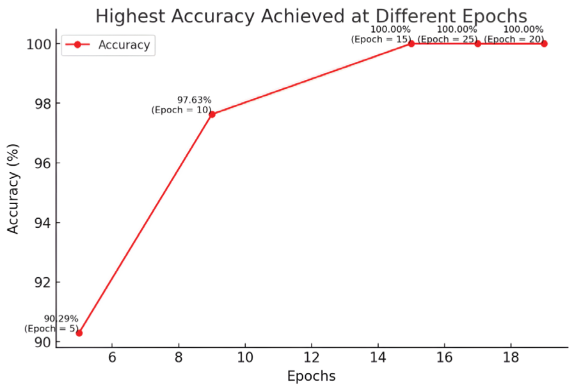

Figure 6: Plots model accuracy against training epochs, showing how performance changes across epochs; if 100% accuracy is reached multiple times (epoch = 15, 20, 25), the first epoch at which it is achieved (e.g., 15) is recorded.

Dataset and Epochs:

The relationship between dataset size and the model's accuracy is shown in Table 1. It was observed that when the model was trained on a larger number of images, such as 500 or 700 photos, the model's accuracy significantly increased. For example, the accuracy increased from an average of 50% with 10 pictures to 74.62% with 700 pictures over 25 epochs. It is suggested, based on these results, that a larger data set provides better training data for the model, enabling the model to generalize better and distinguish between images of fires and non-fires.

The number of epochs also significantly influences the model's performance, with the highest epochs being observed at several intermediate epochs. For example, using 700 photos, accuracy peaked at 74.50% at 5 epochs and slightly fluctuated as epochs increased. Although the accuracy varied depending on the number of epochs, for every row in the table, the accuracy was always higher for the greatest number of epochs used compared to when the fewest number of epochs were used. From this information, it is evident that finding the right balance is crucial, as too few epochs may result in underfitting the model, while too many epochs may lead to overfitting. It is suggested to start with a higher number of epochs, as it usually leads to good accuracy for the model.

Performance at Specific Epochs:

The model achieved the highest accuracy by using 700 photos for training and testing data across different epochs, as shown in Table 2. The model could also reach 100% accuracy as early as epoch 15 and maintain this level throughout the rest of the epoch. The CNN could effectively learn to classify wildfire images accurately, starting from the range around epoch 15. This is extraordinary, as we know the model can reach perfect accuracy, with a sufficient dataset and a suitable number of epochs.

The Impact of Image Resolution:

The model's accuracy was affected by image resolution, as illustrated in Table 3. Images resized to 100x100 pixels achieved the highest accuracy of 73.43% across three trials, outperforming smaller resolutions of 10x10 and even larger resolutions of 250x250, as evidenced in Figure 7. The graph shows the impact of image resolution (x-axis) on the accuracy (y-axis) of a CNN (CNN) across multiple trials. The graph includes data from three trials, demonstrating consistent improvement in accuracy with higher resolutions. These findings indicate that medium-resolution images provide a balance between preserving sufficient detail for feature extraction and minimizing computational demands. Extremely low-resolution images can lose critical features, while extremely high-resolution images can introduce noise or overcomplicate the model. This goes against the common belief that increasing the resolution of a photo always increases the accuracy of a machine-learning model.

Table 3: The accuracy of the model is based on the size to which the images are rescaled. The epoch was set to 10, and 700 photos were used for the training and testing set.

	10x10	50x50	100x100	250x250
Trial 1	71.43%	72.50%	70.29%	72.50%
Trial 2	73.21%	71.36%	76.07%	68.57%
Trial 3	70.71%	71.43%	73.93%	68.21%
Average	71.78%	71.76%	73.43%	69.76%

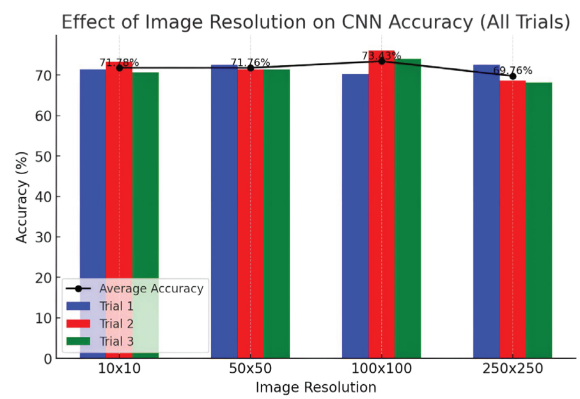

Figure 7: Effect of Image Resolution on CNN Accuracy (All Trials). The results indicate that CNN accuracy is resolution-dependent, with medium-sized images (around 100×100) generally performing best by balancing detail and computation, while very low (10*10) or very high (500*500) resolutions can degrade performance—contradicting the idea that more pixels always help.

Implications for Wildfire Detection:

It is confirmed that machine learning models, such as CNNs, can be highly effective in early wildfire detection. By changing certain factors, such as dataset size, the number of epochs, and image resolution, these models can achieve high levels of accuracy. The early classification of wildfire can save many lives, protect people's property, and reduce environmental damage.

Limitations and Potential Errors:

The project was performed on only one specific dataset. This may not account for all situations in the world, which could cause these results to vary slightly based on other datasets. We could also continue to try to optimize the CNN's structure to converge to an optimal accuracy more quickly.

Future Work:

In the future, more research can be conducted on the model to further maximize its accuracy in predicting whether something is a wildfire. We advise that future research focus on questions such as: Does the model still achieve a high percentage in predicting wildfires from lower-quality images? How well can the model perform under certain weather conditions in the place where the pictures are being taken? What are the common causes of false positives and false negatives in the predictions from the model, and how could this be minimized? Can the model be utilized for real-time detection of wildfires, using video streams?

Conclusion

The research shows the effectiveness of the CNN model in accurately detecting wildfires from images, with the optimization of factors such as dataset size, epochs, and image resolution. Unlike traditional methods that rely on delayed reporting to detect wildfires, this approach demonstrates the speed and precision of wildfire detection using a CNN model. The research addresses the research question by affirming that CNN models can serve as an effective tool for wildfire detection when properly trained and optimized. This supports the hypothesis, as changing specific factors can significantly improve the model's accuracy. For example, the experiment demonstrates that using larger datasets and increasing the number of epochs can lead to improved performance. The model can be applied to the real world, using drones and satellites, to enhance the capability of detecting wildfires at an early stage before severe damage occurs. Future advancements can lead to the increased performance of the model. To conclude, the research shows ways that machine learning models, specifically CNN models, can be an effective solution in saving lives and protecting the environment.

Acknowledgments

We would like to express our deepest gratitude to our Bronx Science Research teacher, Dr. Vladimir Shapovalov, for his invaluable support, guidance, and encouragement throughout this project. His approval of this research topic, which holds significant implications for Earth and Environmental Sciences as well as the study of Environmental Effects on Ecosystems, has been instrumental in shaping our work.

We are also sincerely thankful to our research mentor, Odysseas Drosis, for his expert guidance in helping us frame our research approach and question, as well as for his assistance in testing and evaluating the accuracy of our model, using Python programming. His insights and mentorship greatly enhanced the quality of this project.

Finally, we extend our heartfelt appreciation to our parents for their constant moral support and encouragement, which motivated us to persevere throughout the course of this research.

References

1. Lei, Y., Lei, T. H., Lu, C., Zhang, X., & Wang, F. (2024). Wildfire Smoke: Health Effects, Mechanisms, and Mitigation. Environmental science & technology, 58(48), 21097–21119. https://doi.org/10.1021/acs.est.4c06653
2. Chen, H., Samet, J. M., Bromberg, P. A., & Tong, H. (2021). Cardiovascular health impacts of wildfire smoke exposure. Particle and fibre toxicology, 18(1), 2. https://doi.org/10.1186/s12989-020-00394-8
3. Raoelison, O. D., Valencia, R., Lee, A., Karim, S., Webster, J. P., Poulin, B. A., & Mohanty, S. K. (2023). Wildfire impacts on surface water quality parameters: Cause of data variability and reporting needs. Environmental pollution (Barking, Essex: 1987), 317, 120713. https://doi.org/10.1016/j.envpol.2022.120713
4. El-Madafri I, Peña M, Olmedo-Torre N. The Wildfire Dataset: Enhancing Deep Learning-Based Forest Fire Detection with a Diverse Evolving Open-Source Dataset Focused on Data Representativeness and a Novel Multi-Task Learning Approach. Forests. (2023); 14(9):1697. https://doi.org/10.3390/f14091697
5. Sydorenko, I. (2021). Machine Learning & Training Data. High quality data annotation for Machine Learning. https://labelyourdata.com/articles/machine-learning-and-training-data
6. Jonnalagadda, A. V., & Hashim, H. A. (2024). SegNet: A segmented deep learning-based convolutional neural network approach for drones wildfire detection. arXiv preprint arXiv:2405.00031. https://arxiv.org/abs/2405.00031
7. Maggioros, S., & Tsalkitzis, N. (2024). Wildfire danger prediction optimization with transfer learning. arXiv preprint arXiv:2403.12871. https://arxiv.org/abs/2403.12871
8. Martin, V., Venable, K. B., & Morgan, D. (2024). Development and application of a Sentinel-2 satellite imagery dataset for deep-learning driven forest wildfire detection. arXiv preprint arXiv:2409.16380. https://arxiv.org/abs/2409.16380
9. Phung & Rhee (2019). A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets. Applied Sciences, 9(21), 4500. https://doi.org/10.3390/app9214500
10. Sun, C. (2022). Analyzing multispectral satellite imagery of South American wildfires using deep learning. arXiv preprint arXiv:2201.09671. https://arxiv.org/abs/2201.09671
11. Zheng, Y., Zhang, G., Tan, S., Yang, Z., Wen, D., & Xiao, H. (2023). A forest fire smoke detection model combining convolutional neural network and vision transformer. Frontiers in Forests and Global Change, 6, Article 1136969. https://www.frontiersin.org/articles/10.3389/ffgc.2023.1136969/full

Authors

Shaunak Thamke - Shaunak, a rising Bronx Science junior, is a passionate STEM and AI researcher and innovator. He earned top awards at the Terra STEM Fair and STEAMinAI Championship (2025). He leads as Varsity Tennis Captain and Quantum Mechanics Club Co-President. He also volunteers as a certified schoolhouse.world tutor.

Kevin Diao - Kevin, a rising junior and avid tennis player, is a dedicated scholar with a passion for finance and STEM. Earning second place at the Terra STEM Fair (2025), he excels academically, leads as a Varsity Tennis team member, and actively contributes to his school and community.

Elliot Blin - Elliot, a rising junior at Bronx Science, is a multisport athlete, musician, and STEM enthusiast passionate about quantum mechanics. He earned second place at the 2025 Terra STEM Fair, leads as captain and is on 3 varsity teams, and serves as co-president of the Quantum Mechanics Club.

DOI: 10.36838/v8i3.97