Evaluation of burning detection using modified VGG19 for LULC classification changes

Mateus Freire Roberto[1], Clodoaldo Souza Faria Junior[2], João Domingos Augusto dos Santos Pereira[1], José Guilherme Magalini dos Santos Decanini[1], Moisés José dos Santos Freitas[3]

[1] Federal Institute of Education, Science and Technology of São Paulo - IFSP, Presidente Epitácio Campus, Pres. Epitácio - SP, 19470-000, Brazil - mateus.roberto@aluno.ifsp.edu.br, joao.pereira@ifsp.edu.br, guilhermedecanini@ifsp.edu.br
[2] Faculty of Science and Technology, São Paulo State University - UNESP, Presidente Prudente Campus, Pres. Prudente - SP, 19060-900, Brazil - clodoaldo.souza@unesp.br
[3] Aeronautics Institute of Technology, São José dos Campos - SP, 12228-900, Brazil - freitas@ita.br

Keywords: LULC, Remote Sensing, CNN, VGG19, Grad-CAM, Burn detection.

Abstract

This study aims to evaluate the effectiveness of a modified architecture of convolutional neural network (CNN) VGG19 for detecting fires and changes in land use and land cover classification (LULC). Remote sensing data from the Landsat 8 Operational Land Imagery (OLI) satellite was used to collect images from two distinct regions, one of which was used to obtain a dataset containing 1000 labeled images, and the other region was used to perform inference and verify the generalization of the model in an area with a high annual occurrence of fires. Analyses were conducted using a time series of normalized difference vegetation index (NDVI) and complementary cumulative distribution function (CCDF), to determine the potential for analysis in that area and define the periods of burning, pre-burning and post-burning. The VGG19 architecture was modified to maintain the input sizes of the images, resulting in a significant increase of 20.90 percentage points in the F1 score compared to the original architecture, as well as a 68.76% reduction in convergence time. In addition, the Gradient-weighted Class Activation Mapping (Grad-CAM) technique was used to improve the interpretability of the model at the moment of inference. The proposed methodology offers an approach for detecting burns by altering the LULC classification, and the modified VGG19 showed superior results.

1. Introduction

A change in land cover refers to the modification in the terrestrial environment, including alterations in hydrological resources, soil composition, and atmospheric contamination. Meanwhile, land use pertains to how human behavior affects physical changes on the Earth’s surface (Atef et al., 2023; Voelsen et al., 2023). Bridging these concepts, the classification LULC emerges as a pivotal task, aiming to categorize Earth’s surface into specific land use and land cover classes (Siddamsetty et al., 2023).

The advancement of Remote Sensing (RS) technology has simplified the image acquisition process and reduced the associated costs, resulting in an abundance of available data. Consequently, RS imagery has become a crucial data source for LULC classification, offering benefits such as wide coverage and continuous monitoring. This enables the collection of time-series data, essential for comprehending the dynamics of land use and land cover changes over time (Zhao et al., 2023).

In the recent decades, there has been an exponential increase in the satellite images and data, driven by advancements in the RS technology and the launch of numerous satellites. Notably, the Landsat mission, a collaborative effort between National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS), currently comprises nine Earth Observation (EO) satellites that provide freely accessible data. EO has made it possible to assess the impacts of the tragedy that occurred in 2019 with the collapse of the Brumadinho tailings dam (Rotta et al., 2020), estimate the depth of the euphotic zone and the depth of the Secchi disk, contributing significantly to water quality monitoring (Gomes et al., 2020), among other diverse applications, capable of predicting or monitoring phenomena without the need for physical contact with the target of study.

The assessment of variations in LULC changes enables a comprehensive evaluation of their effects on ecosystems, biodiversity, carbon equilibrium, and water resources, among other environmental aspects (Wang et al., 2023). These changes influence land-atmosphere interactions, the surface energy budget, and the hydrological cycle, which are vital for the management of agriculture, forests, and water resources (Tariq et al., 2023). In the face of adversities, wildfires emerge as a volatile and ruinous element, capable of profoundly transforming the landscape (Gajendiran et al., 2024).

In recent years, the integration of convolutional neural networks (CNNs) in RS has contributed to significant advances (Wang et al., 2022). In the work proposed by Hang et al. (2020), a new framework designed for the integration of hyperspectral imagery and light detection ranging (LiDAR) data using two coupled CNNs was presented. This framework uses one CNN to learn spectral features from hyperspectral data, and the other is used to capture elevation information from LiDAR data. Using the EuroSAT dataset from the Sentinel-2 satellite, Dewangkoro and Arymurthy (2021) take advantage of CNNs to advance LULC classification. Renowned CNN architectures such as VGG19, ResNet50 and InceptionV3 were employed for feature extraction, enhanced by the integration of a Channel Squeeze and Spatial Excitation (sSE) block.

The layered structure of CNNs, particularly the convolutional layers, proves highly effective in extracting essential features from images. CNNs use a mathematical process known as convolution. This process is a specific type of linear operation. Essentially, CNNs belong to a category of neural networks that

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-2024-363-2024 | © Author(s) 2024. CC BY 4.0 License.

incorporate convolution instead of the typical matrix multiplication in at least one layer of the network. CNNs are designed to automatically learn spatial hierarchies of features from input images in an adaptive manner. This is facilitated through the utilization of various building blocks, including convolutional layers, clustering layers, and fully connected layers (Zhao et al., 2024). Convolutional layers act as feature extractors, traversing the input image and conducting convolution operations to identify patterns such as edges, textures and shapes. Additionally, CNNs demonstrate robustness in handling the considerable variability inherent in satellite imagery, arising from fluctuations in lighting, weather conditions, and seasonal variations. Their ability to generalize effectively from training data to novel, unseen images renders them dependable large-scale and automated LULC mapping tasks (Tsenov et al., 2023; Goodfellow et al., 2016).

Thus, the aim of this study is to propose the identification of fires by detecting changes in the LULC classification using the modified VGG19 architecture. In addition, a quantitative analysis will be carried out, comparing machine learning metrics, between the original VGG19 model and this adapted version. The paper is organized as follows: section 2 presents the methodology followed, including study area, dataset and pre-processing, preliminary analysis and architecture modeling. Section 3 presents the results and discussions obtained through the methodology used. Section 4 presents the conclusions reached from the results as well as proposals for future studies.

2. Methodology

2.1 Study area and criteria

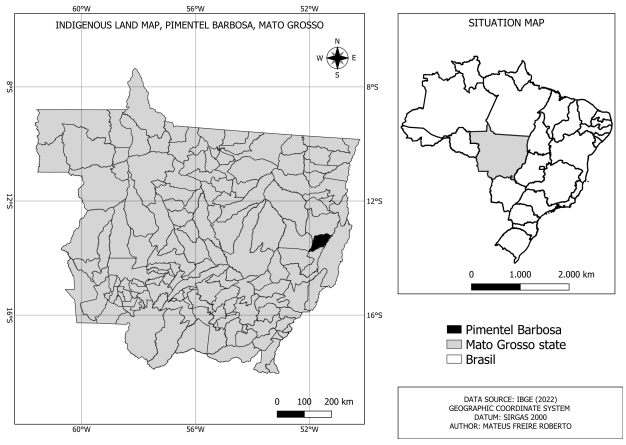


Figure 1. Pimentel Barbosa IL, Geographical location.

The study area under analysis encompasses a Brazilian Indigenous Land (IL) situated between the municipalities of Canarana and Ribeirão Cascalheira, with a population of 2,369 native peoples (Socioambiental, 2022). This region was selected due to its strong annual burning activity, situated in the southern portion recognized for having the highest annual burning activity globally, commonly referred as “Arc of Deforestation in Brazil” (Giglio et al., 2006).

This study used digital remote sensing data from the OLI sensor available on the Landsat 8 satellite, obtained through the Earth Explorer free acquisition service. The acquired scenes were collected from September 2015 to July 2017. Figure 1 shows the geographical location of the study area. In the acquisition of the orbital images, a criterion of maximum interval of three months was adhered to.

The minimum interval between scenes was set at 16 days, based on the revisit time of the Landsat 8 OLI satellite. This satellite has a total of 11 spectral bands, of which 9 belong to the OLI sensor and 2 are obtained by the Thermal Infrared Sensor (TIRS), which will not be the subject of this study. Another point to note is its spatial resolution of 30 meters, with the exception of band 8 (panchromatic) which has a spatial resolution of 15 meters.

2.2 Dataset and pre-processing

The on-demand acquisition service provided by Earth Explorer of the USGS was used. To build the training and validation data set, this service provided a scene of the São Francisco River region, located between the states of Pernambuco and Bahia, captured by the LANDSAT 8 OLI satellite. To perform the model’s inference and check its potential for generalization, scenes were also obtained from the Pimentel Barbosa IL region, from September 2015 to July 2017, which were not correlated temporally or spatially. Both scenes obtained from the Pimentel Barbosa IL and São Francisco River regions were subjected to preliminary pre-processing in order to reduce the complete scene into a smaller one, since the scene obtained covers a much larger area than that required for this study. In addition, to obtain the NDVI in the inference region, the offset method was used to obtain the radiation levels.

After obtaining the scene of the São Francisco River region, Quantum Geographic Information System (QGIS) software was used to create a shapefile containing 1000 points of interest, which were labeled with five different types of LULC. Each point of interest was assigned to a corresponding class, allowing geospatial information to be associated with the respective types of features through the shapefile. The shapefile was used to extract and crop 36x36 pixel images of the regions of interest (ROI) belonging to the classes: ”Agriculture”, ”Water”, ”Unvegetated area”, ”Forest” and ”Soil”. This process was carried out using a Python script on the Google Colab platform. This approach made it possible to create a dataset with 1000 labeled images. First of all, it should be noted that spectral confusion in the context of small-scale images can be a significant challenge in remote sensing and image classification. This is particularly true for ROIs, where the spectral signatures of different materials or land covers are very close together. In such images, a single pixel can contain mixed signals from several sources, leading to ambiguities in classification and analysis. This phenomenon is exacerbated when the pixel is placed in a wider spatial context, where it has to be differentiated from a diverse set of surrounding elements (Liu et al., 2020).

2.3 Preliminary analysis

To ensure the analysis potential of this region, statistical evaluations were conducted, including the Complementary Cumulative Distribution Function (CCDF). The NDVI was extracted from each scene to assess vegetation behavior. Subsequently, a time series of the NDVI index was generated, by index. Recognizing the analytical opportunities presented by this region, modeling of the VGG19 architecture, as shown in Figure 2.

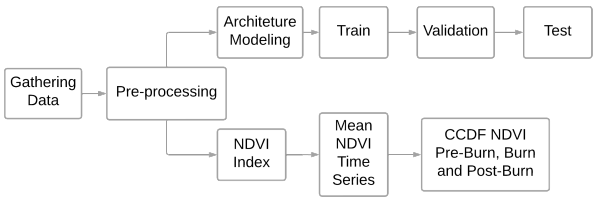


Figure 2. Workflow for LULC classification.

2.4 Architecture modeling and evaluation methods

Based on the database, two architectures were trained: VGG19 and its adapted version. Both architectures were trained using the same hyperparameters, which included a Batch Size of 32 and Momentum of 0.9 (Nesterov, 1983). However, empirical tests were conducted on both architectures to determine the optimal learning rate. Thus, it was adjusted individually for each architecture, being set at 0.0009 and 0.0045 for the VGG19 and the adapted VGG19, respectively. It is important to note that both configurations adopted the Stochastic Gradient Descent (SGD) optimizer during the training process (Robbins, 1951; Kiefer, Wolfowitz 1952). To analyze the performance of both models and compare them, cross-validation was applied with k = 5 folds. In the end, a total confusion matrix of all the folds was obtained, which was then used to extract the machine learning metrics: Accuracy, Precision, Recall and F1-score, represented by equations (1), (2), (3), (4) respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F_1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{4}$$

where TP refers to the true positive rate, TN to the true negative rate, FP to the false positive rate and FN to the false negative rate.

In addition, in this study, both models were evaluated using a visual explanation called Gradient-weighted Class Activation Mapping (Grad-CAM) proposed by Selvaraju et al. (2020). This approach introduces a technique that enhances the transparency and interpretability of decisions made by a wide range of CNN based models. The authors propose this technique by leveraging gradients from target concepts flowing into the final convolutional layer. These gradients create a coarse localization map, highlighting significant regions in an image that contribute to predicting a specific concept. Notably, Grad-CAM is applicable to various CNN model families, including those with fully-connected layers, structured outputs, and multimodal inputs. The paper demonstrates its effectiveness in image classification, captioning, and visual question answering tasks, shedding light on model failure modes, robustness to adversarial images, and improved localization. Additionally, Grad-CAM helps users establish trust in model predictions, even for non-attention-based models, by providing interpretable visualizations

3. Results and discussion

AfterNDVI Time Series and CCDF analysisobtaining In the first analysis, the average NDVI Index was observed from September 2015 to July 2017, as can be seen in Figure 3. An upward trend in NDVI values can be seen from September 2015 to March 2016, indicating a recovery in vegetation during this period. However, from that date onwards, the NDVI index showed a significant increase in the magnitude of the decay rate in July 2016 until it reached its minimum inflection point in September 2016, which was considered the burning period in this study. After this fall, the index recovered again until May 2017, when it fell again.

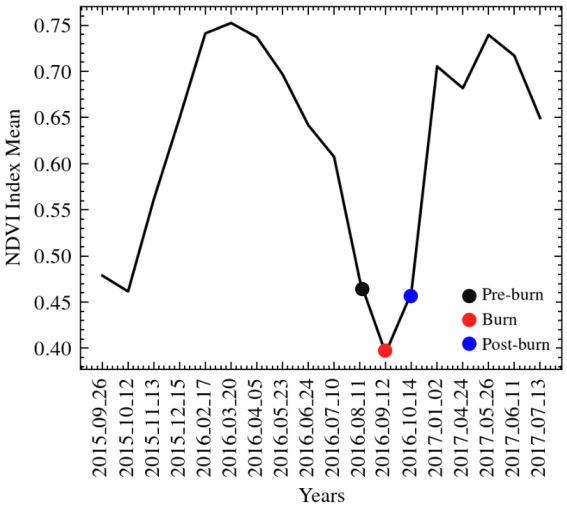


Figure 3. Average NDVI Index Time Series.

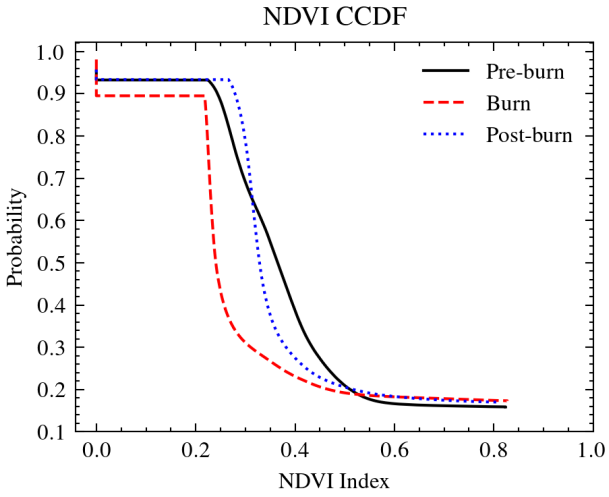


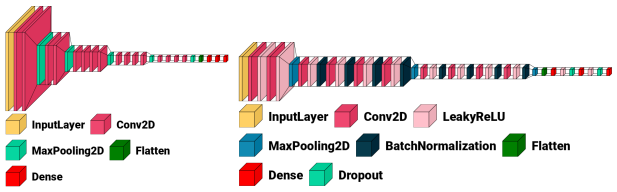
Figure 4. CCDF in Pre-burn, Burn and Post-burn.

After determining the burn period, the CCDF curves were obtained for the pre-burn, burn and post-burn periods, as shown in Figure 4. According to the appropriate NDVI index ranges presented by Akbar et al. (2019), it was observed that the pre-burn curve showed that the probability of occurrence of NDVI indices corresponding to sparse vegetation was 79% and 53% for dense vegetation. During the burning period, the probability of NDVI values classified as sparse vegetation is 36% and only 27% for dense vegetation. The post-burn curve shows 93% sparse vegetation and 36.5% dense vegetation. There-

fore, during the burning period, the probability of NDVI values classified as sparse vegetation and dense vegetation decreases dramatically compared to the burning and pre-burning periods, confirming that the fires at this time significantly affected the density of vegetation in the region. The high probability of sparse vegetation in the post-burn curve shows that the region’s vegetation is in the recovery phase, given the high probability of sparse vegetation in this region, making it clear that the time interval was not enough to “recover” the region’s vegetation.

3.1 Modified VGG19

Initially, it was observed that in the VGG19 model, the size of the feature maps decayed rapidly, which resulted in a significant loss of relevant information and led the training to underfitting. To overcome this challenge, it became necessary to resize the images during training in VGG19. However, resizing the images also introduces more spectral confusion, which can have a negative impact on the model’s performance. Given this complexity, the adaptation of the architecture focused on maintaining the original size of the input images (36x36 pixels). In order to achieve this goal, it was decided to remove the initial and final convolution blocks, reducing the number of blocks from five to three, as illustrated in Figure 5. Additionally, in order to avoid overfitting, dropout layers, L2 regularization were incorporated into the adapted VGG19. Moreover, it was observed that both architectures adjusted slowly at certain times. To mitigate this issue, the inclusion of Batch Normalization layers had a significant impact.



(a) VGG19. (b) Modified VGG19.

Figure 5. VGG19 and Adapted VGG19 architecture.

3.2 Training performance comparison

Cross-validation with k = 5 was carried out on both architectures to evaluate performance. Performance metrics such as precision, recovery, accuracy and f1 score were also obtained. The loss curve for both architectures can be seen in Figure 6. Both curves represent the average training loss over the 5 folds. It was observed that the modified VGG19, despite starting with a higher loss, reaches a loss close to zero approximately after epoch 150. In comparison, the VGG19 reaches these same values approximately after epoch 220. This shows that the modified VGG19 was able to reduce loss more quickly than its original counterpart. In addition, both architectures were trained on 250 epochs for comparison. The number of epochs was adjusted to be unbiased, carefully evaluating the evolution of the models through validation sets.

Figure 7 shows that at a certain point, the accuracy of VGG19 seems to stagnate, remaining practically constant for a period. This may indicate that the model has reached a temporary plateau, perhaps due to a local minimum in the loss function. However, after this stagnation phase, the curve rises again.On the other hand, the modified VGG19 shows faster and more efficient progress. The accuracy curve of this modified version quickly approaches 100%, indicating that the model achieved

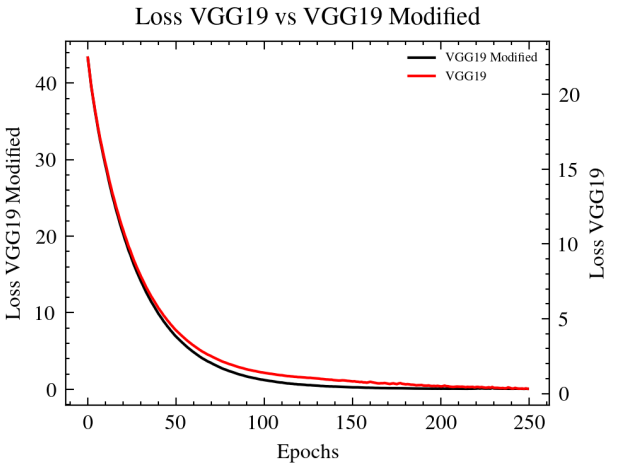


Figure 6. Average loss for 5 folds.

high accuracy well before the standard VGG19. This suggests that the modifications implemented to the network architecture have been successful in optimizing the learning process, allowing the model to learn more complex patterns more efficiently or to be better adjusted to the specifics of the data set in question.

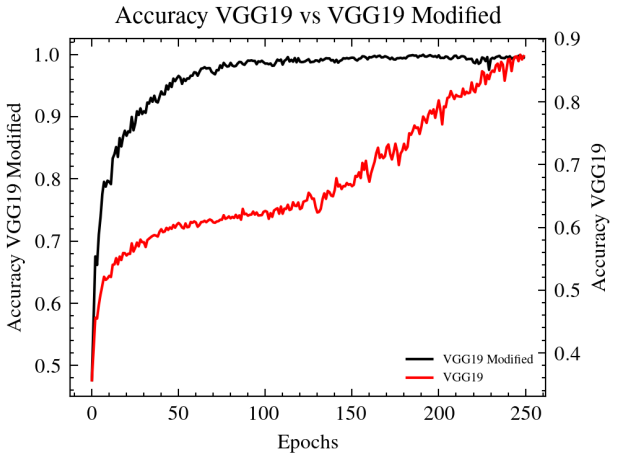


Figure 7. Average accuracy for 5 folds.

The average recall curve for training both architectures is shown in Figure 8. It can be seen that VGG19 initially starts off completely stagnant, remaining constant until epoch 70. This suggests that the model had difficulty recovering the true positives at the start of training. On the other hand, the recall curve for the modified VGG19 shows a different behavior. It starts rising from the beginning of training, which indicates that the model is already capturing true positive examples from the first few epochs. In addition, the absence of stagnation suggests that the modifications made to the network architecture have allowed the model to learn more quickly to identify true positives.

3.3 Validation performance comparison

The metrics were collected from the predictions at the end of each fold for data not seen during training, thus obtaining a total confusion matrix with all the predictions made. The confusion matrix in Figure 9 shows the prediction performance of VGG19 without adaptations, while Figure 10 shows the confusion matrix for modified VGG19

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-2024-363-2024 | © Author(s) 2024. CC BY 4.0 License.

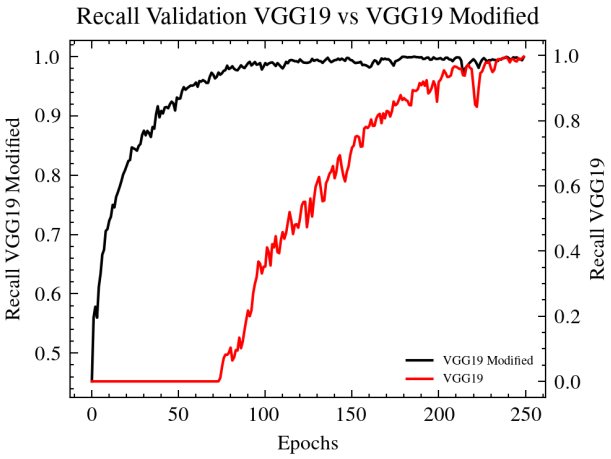


Figure 8. Average training recall for 5 folds.

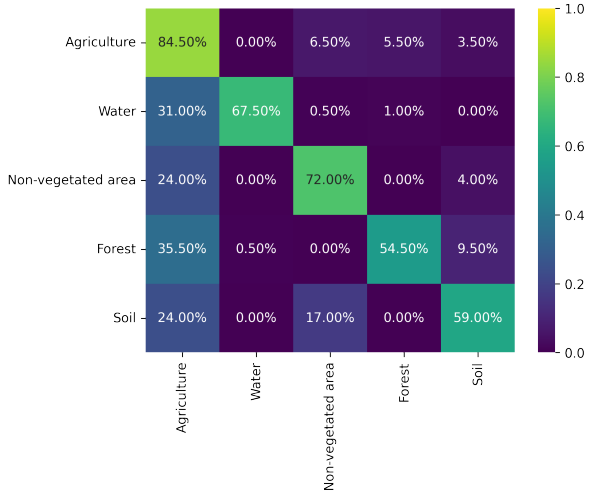


Figure 9. Confusion Matrix for VGG19.

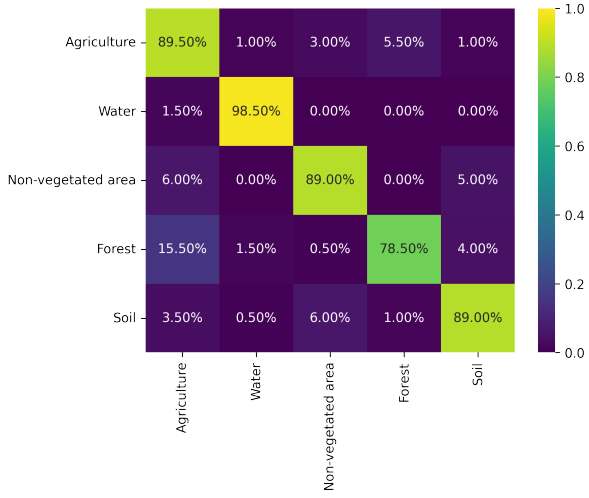


Figure 10. Confusion Matrix for VGG19 Modified.

With the confusion matrix it is possible to evaluate the metrics and compare performance in terms of accuracy, precision, recall and f1-score, as shown in Table 1. It is clear that the adapted VGG19 significantly outperformed the original VGG19 in all the metrics evaluated, indicating that the modifications made to the architecture contributed positively to the model’s performance in this specific data set.

Models	Accuracy	Precision	Recall	F1-Score
VGG19	87.00	76.74	67.50	69.02
VGG19 Modified	95.96	89.36	88.90	88.92

Table 1. Metrics comparison in percentages.

It should be noted that the adapted VGG19 achieved a significant increase of 20.90 percentage points in the f1-score. In addition, there was a reduction in training time. While the original VGG19 required 653 seconds, the adapted architecture converged in 204 seconds. In other words, the adapted version had a convergence rate three times faster than the original, representing a reduction of 68.76%.

3.4 Test performance comparison and burning detection

To test and compare the models, 3 samples were obtained from the pre-burning, burning and post-burning periods at IL Pimentel Barbosa. In addition, the Grad-CAM technique was used to better interpret the model. As can be seen in Figure 11, the modified VGG19 was able to detect the change in LULC during the burning period. Through this approach, it can be said that there was a burn, since drastic changes in LULC within forested areas in this context are highly correlated with burning. Furthermore, using Grad-CAM it is possible to verify that the activation regions that led to the classification are coherent. Normal VGG19, on the other hand, showed difficulties in classification, as was also evidenced by Grad-CAM.

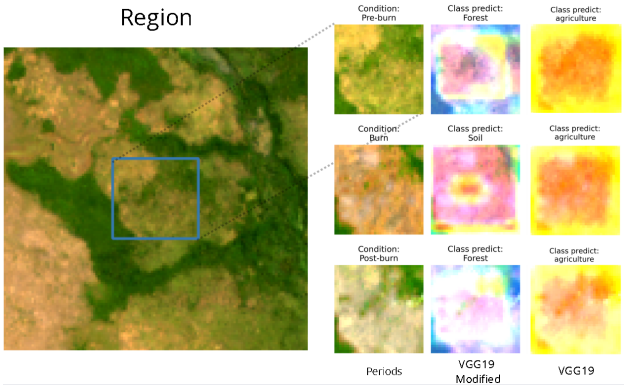


Figure 11. Comparison of the models for the first inference sample.

As can be seen in Figure 12, the modified VGG19 was able to detect the variation in LULC during the firing period. This once again demonstrates the superiority of the adapted VGG19 in terms of the complexities analyzed.

In Figure 13, it can be seen that the degradation caused by the burn has consumed a large part of the vegetation, which has led VGG19 to classify the region as soil in the post-burn period as well. In general, in all three samples, the modified VGG19 was able to detect the change in LULC, which in this context is most likely to have been caused by a burn. Normal VGG19, on the other hand, was unable to generalize or even predict the classes correctly, as evidenced by Grad-CAM.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-3-2024-363-2024 | © Author(s) 2024. CC BY 4.0 License.

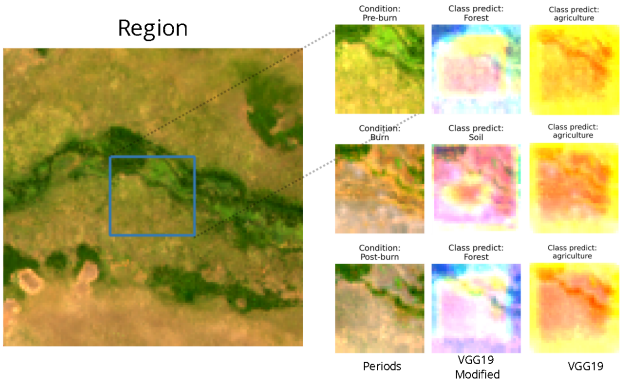


Figure 12. Comparison of the models for the second inference sample.

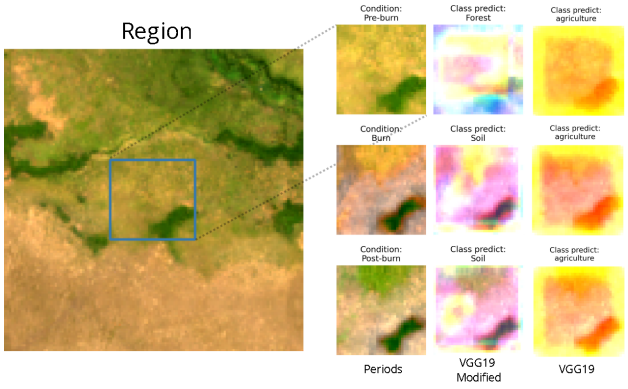


Figure 13. Comparison of the models for the third inference.

4. Conclusions and future studies

This study presents an approach to burn detection by evaluating changes in Land Use and Land Cover (LULC) classification using a modified VGG19 architecture. The study's inference area was a Brazilian IL located between the municipalities of Canarana and Ribeirão Cascalheira. The research methodology consisted of building a training and validation database, pre-processing, preliminary analysis, architecture modeling and performance comparison. To collect the data set for training and validation, Earth Explorer's on-demand acquisition service was used to obtain a satellite image of the São Francisco River region, this region being neither spatially nor temporally correlated with the inference region where the tests were conducted.

The VGG19 architecture was modified to maintain the input size of the images (36x36). The modifications included removing the initial and final convolution blocks and reducing the number of convolutional blocks from five to three. Dropout layers, L2 regularization, and the Leaky ReLU activation function were also utilized. The study then compared the performance of the modified VGG19 model with the original VGG19 in terms of accuracy, precision, recovery and f1 score. Furthermore, it should be noted that the adapted VGG19 obtained a significant increase of 20.90 percentage points in the f1 score.

The study found that the modified VGG19 outperformed the original architecture in all evaluated metrics, with a significant increase in the f1-score. Additionally, the adapted model achieved a convergence rate three times faster than its original counterpart. In the test dataset, the modified VGG19 identified changes in LULC during the pre-burning, burning, and post-burning periods. On all occasions, the model could identify the change from forest to vegetation after the burn. In contrast, the original VGG19 failed to generalize significantly for the LULC classification task. Overall, the study concludes that the proposed methodology can be used in burn classification through changes in LULC classification.

Convolutional neural networks, in particular CNNs, are occupying an essential place in LULC mapping and are proving to be highly effective in extracting essential features from images. This research also highlights the usefulness of CNNs in dealing with the considerable variability inherent in satellite images. The results of this study contribute to the field of remote sensing technology and land cover classification and offer promising potential for further applications in related research areas.

For future research, it is advisable to utilize high or medium-resolution satellite images or images obtained from drones. Additionally, consider using multispectral and thermal cameras. Explore other neural network architectures such as YOLO, Unet, and SegNet. Regarding modifications to VGG19, further enhancements can be made, such as incorporating different activation functions and including transformer blocks.

Acknowledgements

The authors of this work would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) project number 147824/2023-0. The authors would also like to thank the Federal Institute of Education, Science and Technology of São Paulo - Presidente Epitácio Campus.

References

Atef, I., Ahmed, W., Abdel-Maguid, R. H., Baraka, M., Darwish, W., Senousi, A. M., 2023. Land use and land cover simulation based on integration of artificial neural networks with cellular automata-markov chain models applied to el-fayoum governorate. *ISPRS Ann. Photogramm. Remote Sens Spatial Inf. Sci.*, X-1/W1-2023, 771-777.

Dewangkoro, H. I., Arymurthy, A. M., 2021. Land Use and Land Cover Classification Using CNN, SVM, and Channel Squeeze and Spatial Excitation Block. *IOP Conference Series: Earth and Environmental Science*, 704(1).

Gajendiran, K., Kandasamy, S., Narayanan, M., 2023. Influences of wildfire on the forest ecosystem and climate change: A comprehensive study. *Environ. Res*, 240. Article 117537.

Giglio, L., Csiszar, I., Justice, C. O., 2006. Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. *J. Geom. Phys.*, 111.

Gomes, A. C., Alcântara, E., Rodrigues, T., Bernardo, N., 2020. Satellite estimates of euphotic zone and Secchi disk depths in a colored dissolved organic matter-dominated inland water. *Ecological Indicators*, 110. Article 105848.

Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press.

Hang, R., Li, Z., Ghamisi, P., Hong, D., Xia, G., Liu, Q., 2020. Classification of Hyperspectral and LiDAR Data Using Coupled CNNs. *IEEE Transactions on Geoscience and Remote Sensing*, 58(7), 4939-4950.

Kiefer, J., Wolfowitz, J., 1952. Stochastic Estimation of the Maximum of a Regression Function. *Annals of Mathematical Statistics*, 23(3), 462-466.

Liu, M., Yu, T., Gu, X., Sun, Z., Yang, J., Zhang, Z., Mi, X., Cao, W., Li, J., 2020. The Impact of Spatial Resolution on the Classification of Vegetation Types in Highly Fragmented Planting Areas Based on Unmanned Aerial Vehicle Hyperspectral Images. *Remote Sensing*, 12(1), 146.

Nesterov, Y., 1983. A method for solving the convex programming problem with convergence rate o(1/k^2).*Proceedings of the USSR Academy of Sciences*, 269, 543−−547.

Robbins, H. E., 1951. A Stochastic Approximation Method. *Annals of Mathematical Statistics*, 22, 400-407.

Rotta, L. H. S., Alcantara, E., Park, E., Negri, R. G., Lin, Y. N., Bernardo, N., Mendes, T. S. G., The, F. C. R. S., 2019. Brumadinho tailings dam collapse: possible cause and impacts of the worst human and environmental disaster in Brazil. *Int. J. Appl. Earth Obs. Geoinf.*, 90(2020), 10211.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2020. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *International Journal of Computer Vision*, 128(2), 336-359.

Siddamsetty, J., Stricker, M., Charfuelan, M., Nuske, M., Dengel, A., 2023. Inter-region transfer learning for land use land cover classification, ISPRS Ann. Photogramm. Remote Sens. *Spatial Inf. Sci.*, X-1/W1-2023, 881-888.

Socioambiental, I. I., 2022. Terras indígenas no brasil.

Tariq, S., Zeydan, , Nawaz, H. et al., 2023. Impact of land use/land cover (LULC) changes on latent/sensible heat flux and precipitation over Türkiye. *Theor Appl Climatol*, 153, 1237-1256.

Tsenov, R. M., Henry, C. J., Storie, J. L., Storie, C. D., Murray, B., Sokolov, M., 2022. Exploration of convolutional neural network architectures for large region map automation.

Voelsen, M., Lauble, S., Rottensteiner, F., Heipke, C., 2023. Transformer Models For Multi-Temporal Land Cover Classification Using Remote Sensing Images, ISPRS Ann. Photogramm. Remote Sens. *Spatial Inf. Sci.*, X-1/W1-2023, 981-990.

Wang, J., Bretz, M., Dewan, M. A. A., Delavar, M. A., 2022. Machine learning in modeling land-use and land cover-change (LULCC): Current status, challenges and prospects. *Science of the Total Environment*, 822. Article 153559.

Wang, Y., Sun, Y., Cao, X., Wang, Y., Zhang, W., Cheng, X., 2023. A review of regional and Global scale Land Use/Land Cover (LULC) mapping products generated from satellite remote sensing. *ISPRS J. Photogramm. Remote Sens.*, 2023(206), 311-334.

Zhao, S., Tu, K., Ye, S., Tang, H., Hu, Y., Xie, C., 2023. Land Use and Land Cover Classification Meets Deep Learning: A Review. *Sensors*, 23, 8966.

Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., Parmar, M., 2024. A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57, 99.