

Received October 1, 2020, accepted October 13, 2020, date of publication October 15, 2020, date of current version October 26, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3031533

Methodology of Data Fusion Using Deep Learning for Semantic Segmentation of Land Types in the Amazon

JOEL PARENTE DE OLIVEIRA¹, MARLY GUIMARÃES FERNANDES COSTA^{1,2}, (Member, IEEE), AND CÍCERO FERREIRA FERNANDES COSTA FILHO^{1,2}, (Member, IEEE)

¹Operations and Management Center of the Amazon Protection System (CENSIPAM), Manaus 69049-630, Brazil

²Federal University of Amazonas, Manaus 69080-900, Brazil

Corresponding author: Joel Parente De Oliveira (joelparente@gmail.com)


This work was supported in part by the Samsung Electronics of Amazonia Ltd., through the terms of Federal Law no. 8.387/1991, by the agreement no. 004, assigned by the Center for Research and Development in Electronics and Information from the Federal University of Amazonas—CETELI/UFAM, in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Funding Code 001, and in part by the Academic English Solutions (<https://www.academicenglishsolutions.com>) revised this paper.

ABSTRACT This study proposes a methodology using deep learning and a multi-resolution segmentation algorithm to perform the semantic segmentation of remote sensing images. Initially the image is segmented using a CNN, and then an image with homogeneous regions is generated using a multi-resolution segmentation algorithm. Finally, a data fusion process is performed with these two images, generating the final classified image. The field of study was the Brazilian Amazon region. The proposed methodology classifies images in the following classes: forest, pasture and agriculture. The input data used were LANDSAT-8/OLI images. The reference data were extracted from the results of the TerraClass project in 2014. Two datasets were evaluated: the first with six bands and the second with three bands. Three CNN architectures were evaluated together with three optimization methods: SGDM, ADAM, and RMSProp and the dropout and L_2 regularization methods as methods for generalization improvement. The best model, CNN + optimization method + technique for generalization improvement, evaluated in the validation set, was submitted to a 5-fold cross validation methodology, and the results were compared with pre-trained networks using the learning transfer methodology; in this case the networks used for comparison were ResNet50, InceptionResnetv2, MobileNetv2 and Xception. The proposed methodology was evaluated through image segmentation of some regions of the Amazon. Finally, the proposed methodology was evaluated in regions used by other authors. The accuracy values obtained for the images evaluated were over 99%.

INDEX TERMS Deep learning, convolutional neural networks, remote sensing, image segmentation.

I. INTRODUCTION

Remote sensing consists of collecting information about an object or geographic area without direct contact. This information was compiled through various technologies such as sensors installed on-board aircraft, satellites, or other platforms [1]. The data generated from remote sensing systems are very useful for several applications, for example: urban - demographic interventions, government registrations, urban planning, support for the real estate sector; agricultural - monitoring crop conditions, crop forecasting, soil erosion; geological - minerals, oil and natural gas

The associate editor coordinating the review of this manuscript and approving it for publication was Aneel Rahim .

exploration; ecological - monitoring flooded regions, soils, oceans and continental waters; forest - lumber production, biomass estimation, monitoring of deforestation) [2].

With regard to monitoring deforestation in the Amazon through remote sensing, the National Institute for Space Research (INPE) is a world reference. Among the various projects developed by INPE for this monitoring, the Program for Monitoring Deforestation in the Legal Amazon (PRODES) and TerraClass stand out [3], [4]. PRODES provides data through annual maps of deforestation in the Amazon region. The TerraClass project uses the data generated by PRODES to perform a land use and land cover classification in the following classes: forest, pasture, agriculture, urban areas, mining and others. This information helps government

agencies develop public prevention policies to contain deforestation [5]. In 2018 and 2019, deforestation in the Amazon affected 7,535 km² and 10,129 km², respectively. Therefore, the deforested area showed an increase of 34% from 2018 to 2019. In 2019, 84.56% of all deforestation observed in the Amazon occurred in the states of Pará, Mato Grosso, Amazonas, and Rondônia [6]. Although PRODES and TerraClass are large projects and provide very reliable data, they still rely on a significant portion of work done by human intervention. In order to speed up and make the process more efficient, several studies have been done using machine learning tools [7]–[18].

In [11], [12], [15], the authors performed image segmentation using threshold classification. The results obtained by these authors had an accuracy above 90%. Some works classified images through selecting random forests, as seen in [8], [13], [16]. In [13], the authors obtained an accuracy above 98%. In [8] and [16] the accuracy was 85% and 88%, respectively.

The works of [7] and [17] employed unsupervised classification to segment remote sensing images. In [17], the authors used Bayesian Updating of Land Cover: Unsupervised (BULC-U) and obtained an accuracy of 80%. In [7], the authors used K-Nearest Neighbors (KNN) and obtained an accuracy of 75.95%.

According to [19], from the first decade of this century, the neural network was reborn under the new label known as deep learning. Deep learning makes use of computational models with hierarchical architectures composed of multiple processing layers, in order to “learn” representations of data in very different formats: audio, images, and text [20]. According to [19], the large amount of data available in recent years, together with the increase in computational power, allowed the use of deeper architectures that were not previously possible.

The works of [9], [10], [18] used deep learning to detect deforested areas in regions of the Brazilian Amazon. The results obtained by these authors were quite satisfactory. However, the authors evaluated a specific region of the Amazon. Another criticism we make of the previously published studies is that they have not made available the dataset used for benchmarking.

A major challenge in training CNNs with remote sensing images is that the classes are unbalanced. In other words, in a region captured by a satellite image, there is a great imbalance in terms of the area of various types of soils. This problem can lead, in CNN training, to optimization methods performing better in the most frequent classes. This work proposes a method to solve the difficulties in the segmentation of land use and coverage in different regions of the Amazon using CNNs. To circumvent the imbalance problem, the mosaic image technique is proposed in this work for training CNN in remote sensing applications. In this technique, small rectangular samples of agriculture, forest and pasture are extracted from satellite images. From these samples, a large image is

created, with almost the same number of small rectangles of forest, agriculture, and pasture.

Another contribution of this study is that it boosts the results obtained with CNN in soil classification by fusing its output with the output of a segmentation algorithm of homogeneous regions. The fusion results outperform the results obtained only with CNN.

Finally, we compare the results obtained with customized CNN architectures trained from scratch with the results obtained with transfer learning (TL), using shallow training of the following architectures: ResNet50 (RN50), InceptionResnetv2 (IR2), MobileNetv2 (MN2), and Xception (XC) [21]–[24].

Therefore, the main contributions of this paper are the following:

- 1) Proposing customized CNN architectures for semantic segmentation of different types of use and land cover, in different regions of the Amazon.
- 2) Proposing a training, validation and testing methodology of CNN networks, for the remote sensing area, using the mosaic image concept.
- 3) Comparing the results of the proposed CNN architectures, trained from scratch, with the results of transfer learning architectures.
- 4) Proposing a fusion technique using CNN data and data from a segmentation algorithm of homogeneous regions for improving the soil classification in remote sensing applications.
- 5) Availability of a dataset of types of land use/cover (forest, agriculture and pasture) in the Amazon region, which may be used as a base for another research in the Amazon region.

II. LITERATURE REVIEW

In [7], [16], [17], the authors carried out the classification of images in different regions of the state of Mato Grosso, Brazil. The classes used by the authors were agriculture, forest, and pasture. In [7], the authors used LANDATA-5/TM images as input dataset. The training and testing dataset were built from information from the TerraClass2008 project. The authors used the following classifiers: Decision Tree (DT), Naive Bayes (NB), KNN, Support Vector Machine (SVM), and Optimal Path Forest (OPF). The best results were obtained using the SVM classifier in the region corresponding to image 227/68 (LANDSAT-5/TM image). For this region, an accuracy of 86.04% was obtained. In [16], the authors used LANDSAT-8/OLI and PROBA-V C1 images. The classifier used by the authors was the random forest, and the accuracy obtained was 88%. In [17], the region of the state of Mato Grosso evaluated by the authors included the municipalities of Alto Boa Vista, Querência, Ribeirão Cascalheira, and Canarana. The images used were LANDSAT-5, and the data from the GlobCover project of 2009 were used as the gold standard. The authors used the unsupervised classifier called BULC-U. Using this classifier, the accuracy obtained was 80%.

In [8], the authors used random forests to classify regions of the state of Rondônia in the following classes: deforestation, degradation, and forest. The authors used LANDSAT-5/TM and LANDSAT-8/OLI images. The accuracy obtained was 85%.

In [11], [12], [15], the authors employed a threshold segmentation for remote sensing image classification. In [11], the authors evaluated a region of the state of Mato Grosso with the objective of classifying deforestation and environmental degradation in the region. The region assessed corresponds to an area of scene 226/68 of LANDSAT. LANDSAT-5/TM images from 2005 to 2011 and LANDSAT-8/OLI images from 2013 to 2017 were used. The accuracy obtained by the authors was 94.93%. In [12], the authors used LANDSAT-8/OLI and Sentinel-2 images to classify selective cut in a region of southern Amazonas State, near the village of Santo Antonio do Matupi. The accuracy obtained by the authors was 95.7% for the LANDSAT-8/OLI images and 96.7% for the Sentinel-2 images. In [15], the authors carried out the classification of images in the region of the Carajás mines, located in the south of the state of Pará. The classes defined by the authors were cerrado, forest, water and mining area. The authors used iknos, geoeye and world view images, and the gold standard was generated by the authors themselves through visual interpretation. The accuracy obtained by the authors was 91%.

In [18], the authors used CNN to perform the classification of deforested areas in a region of southern Pará state. The authors used Landsat-8/OLI images and PRODES data as the gold standard. Using CNN U-NET, the authors obtained an accuracy of about 95%.

In [9], the authors mapped deforestation between images one year apart from the other, between 2017 and 2018 and between 2018 and 2019, using a deep learning methodology. The images used were scenes 227/63, 227/65, and 230/65 from LANDSAT-8/OLI, which correspond to regions of the state of Pará and Amazonas, Brazil. The deforestation was extracted from the PRODES project. The authors used three CNN architectures available in the literature: SharpMask, U-Net and ResUnet, and two classical machine learning algorithms: random forest (RF) and multi-layer perceptron (MLP). The best results obtained with the CNN ResUnet had an accuracy and an F1-Score of 99.93% and 94.65%, respectively.

In [10], the authors evaluated deep learning strategies for automatic deforestation detection. The experiments used two Landsat-8/OLI images, acquired at different dates. The first study area corresponds to a small region of the Amazon biome, located, more specifically, in the state of Pará, Brazil. The state of Pará comprises 26% of the Brazilian Amazon, and most of it is covered by dense tropical forest. The reference data used in this experiment refer to the deforestation that occurred between August 2016 and July 2017. The second study area is a small region of the Brazilian Cerrado biome, located in the state of Maranhão, Brazil. For this case, the deforestation took place in the years 2017

and 2018. The authors obtained an accuracy and F1-Score of 95% and 63%, respectively, in the Amazon, and 97% and 78%, respectively, in the Cerrado. The last results, obtained in the Cerrado reached higher percentages than those of the Amazon. According to the authors, this is due to the pattern of deforestation in the Cerrado biome that removes all vegetation, in contrast with the Amazon, where vegetation is not completely removed.

The papers presented in this review showed that machine learning techniques have great potential to classify remote sensing images and to perform automatic segmentation in the most varied types of vegetation cover in the Amazon region. In [9], [10], [18], the authors used deep learning techniques to classify deforestation regions in the Amazon. The results achieved were quite satisfactory. However, in each work, a specific region of the Amazon was used. Another criticism we make of the previously published works is that they have not made available the dataset used for the benchmark. This limits the diversification of solutions to the real world. No papers were found that presented an adequate model in deep learning for land use/cover classification for the Amazon region, encompassing the many varieties of existing environments.

In this work we make available the dataset used, which covers different areas of the Amazon. These areas cover the Brazilian states of Amazonas, Mato Grosso, Pará, and Rondônia.

III. MATERIALS AND METHODS

This work aims proposing a method to solve the difficulties in the segmentation of land use and coverage of different regions of the Amazon region. Figure 1 shows a flowchart of the methodology used for segmentation of land use and coverage. As shown, there are two branches: The left one corresponds to CNN processing and the right one to segmentation with homogeneous regions. In the first block of the left branch, the band's selection is performed. After, in the semantic segmentation block, the CNN performs the identification of the different soil coverage. In the right branch, we perform the segmentation of homogeneous regions. At the end, uniting the two branches, we have the fusion block. In the following sections, each block of this flowchart will be explained in detail.

A. LANDSAT IMAGE

In this work, LANDSAT-8/OLI images from the Amazon region, from 2014, were used. These images are available for free at [25]. As shown in Table 1, the images used cover the Brazilian states of Amazonas, Mato Grosso, Pará and Rondônia.

For generating the gold-standard images for the classifier, the results of the TerraClass project, from 2014, were used. The data were acquired free of charge on the INPE website [4]. The data generated in the TerraClass project delimit the regions of the Brazilian Amazon in the following classes: forest, agriculture, pasture, unobserved area, urban area, mining, others, non-forest and hydrography.

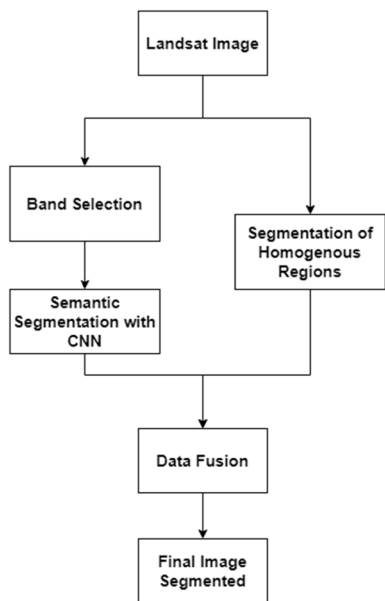


FIGURE 1. Flowchart of methodology used.

TABLE 1. List of images used.

O/P*	Date	O/P*	Date	O/P*	Date
001/65	08/12/2014	229/66	08/09/2014	231/66	08/23/2014
001/66	08/12/2014	230/64	08/16/2014	232/64	08/30/2014
001/67	08/12/2014	230/65	08/16/2014	232/65	08/30/2014
002/65	08/03/2014	230/66	07/15/2014	233/65	08/21/2014
002/66	08/19/2014	231/64	09/08/2014	233/66	08/21/2014
229/65	08/09/2014	231/65	08/23/2014	233/67	08/21/2014
223/68	08/15/2014	226/67	08/04/2014	227/71	08/11/2014
224/67	08/06/2014	226/68	08/04/2014	228/67	08/18/2014
224/68	07/21/2014	226/69	08/04/2014	228/68	08/18/2014
224/69	08/06/2014	226/70	08/04/2014	228/69	07/01/2014
224/70	08/22/2014	226/71	08/20/2014	228/70	08/18/2014
225/67	08/13/2014	227/67	07/26/2014	228/71	08/18/2014
225/69	08/13/2014	227/68	08/11/2014	229/68	08/09/2014
225/70	08/29/2014	227/69	08/11/2014	229/70	08/09/2014
225/71	08/13/2014	227/70	08/11/2014	229/71	08/09/2014
222/62	10/27/2014	224/63	07/05/2014	225/62	10/27/2014
223/62	09/16/2014	224/64	07/05/2014	225/65	10/27/2014
223/66	07/30/2014	224/65	08/22/2014	226/62	08/20/2014
223/67	08/15/2014	224/66	08/06/2014	227/65	08/11/2014
230/68	08/16/2014	231/67	08/23/2014	232/66	08/14/2014
230/69	07/31/2014	231/68	08/23/2014	232/67	08/14/2014

*Orbit/Point

For this work, information on forest, pasture and agriculture areas were used corresponding to a problem with three classes. The classes of unobserved area, urban area, mining and others were not considered because they represent a very small percentage area, and it would not be possible to extract a significant amount of samples to be used in the process of semantic classification with CNNs. The nonforest and hydrography areas were not used because they are part of an INPE exclusion mask. Thus, these regions are not considered in the systematic mapping.

B. BAND SELECTION

According to [26] the bands Blue (B2), Green (B3), Red (B4), Near Infrared (B5), Shortwave Infrared 1 (B6), and Shortwave Infrared 2 (B7) are the most suitable for vegetation analysis. According to [27], B4, B5, and B6 is the best



FIGURE 2. Image patches in color composition of bands B6-B5-B4 corresponding to: (a) forest (b) pasture (c) agriculture.

TABLE 2. Splitting image patches.

Class	Training	Validation	Test
Agriculture	2,000	1,000	1,000
Forest	75,000	75,000	75,000
Pasture	3,000	1,500	1,500

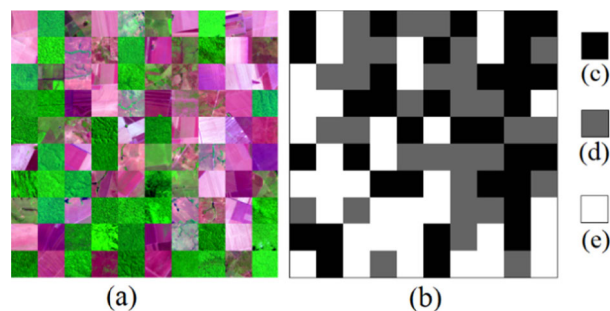


FIGURE 3. Example of mosaic image. In (a) is shown a mosaic image in color composition of bands B6-B5-B4, composed of agricultural, forest and pasture patches, in (b) the gold standard image, and in (c), (d) and (e) we have the gold standard for each image patch, corresponding to agriculture, pasture and forest area, respectively.

combination of three bands for remote sensing applications whose purpose is to perform soil classification. Thus, two versions of image datasets were assembled: the first with the six bands (B2, B3, B4, B5, B6, and B7) and the second with three bands (B4, B5, B6).

Using the reference data from the TerraClass project for the year 2014, 40 × 40 pixel-sized patches were generated for each class. Figure 2 presents three examples of these patches, corresponding to forest, pasture, and agriculture. Using the LANDSAT-8/OLI images, 4,000 agricultural patches, 225,000 forest patches, and 6,000 pasture patches were generated. These patches were divided into three sets: training, validation and testing, as shown in Table 2.

Using the patches previously described, large images, defined as mosaic images, are generated. Each mosaic image has dimensions of 400 × 400 pixels and is formed by 100 patches. Each unit of the mosaic image is randomly generated selecting a patch of agriculture, forest, or pasture. Each mosaic image has a gold standard image. For building the gold standard image, the pixels corresponding to forest regions were marked with the value 255, the pixels corresponding to pastures region were marked with the value 100, and the pixels corresponding to agriculture regions were marked with the value 1. Figure 3 shows an example of a mosaic image and the corresponding gold standard. It can be observed that, out of the 100 patches, 34 are from agriculture, 35 from forest, and 31 from pasture regions.

Two versions of image datasets were generated with a total of 4,000 mosaic images each. For each version, 70% of the images (2800) were destined for the training set, 15% (600),

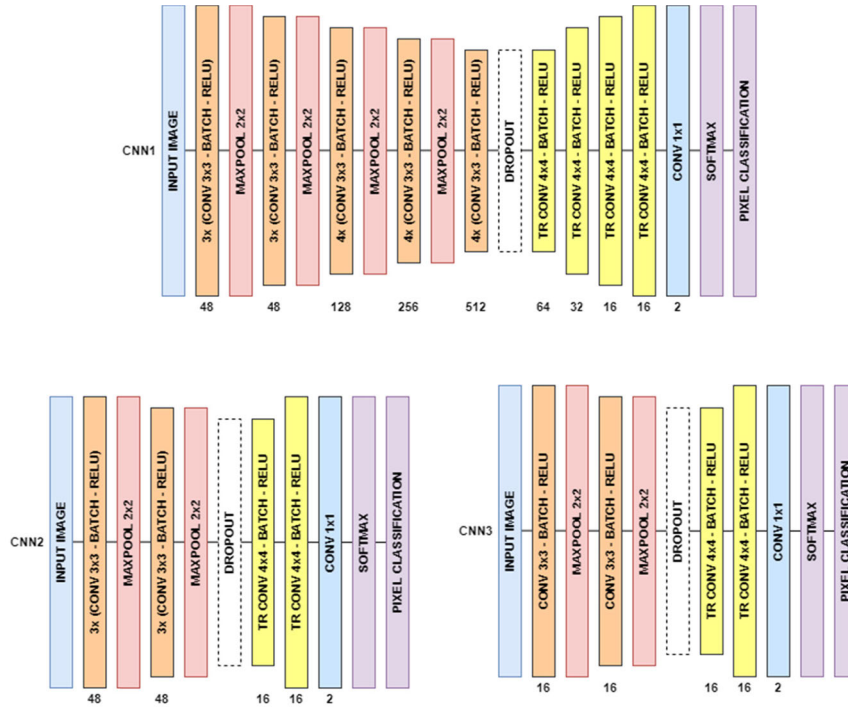


FIGURE 4. CNN architectures used in this work.

TABLE 3. Mosaic image dataset generated with the quantity for training, validation, and testing sets.

Class	Bands	Total	Training	Validation	Test
1	B2, B3, B4, B5, B6, B7	4,000	2,800	600	600
2	B4, B5, B6	4,000	2,800	600	600

for the validation set and 15% (600), for the test set. The first version consists of 6-band images: B2, B3, B4, B5, B6 and B7. The second version is composed of three-band images: B4, B5, and B6. Table 3 presents the image datasets generated used in this work.

C. CNN SEMANTIC SEGMENTATION

Three CNN architectures (Figure 4) used in this work were based on the architecture proposed in the work of [28]. In this work, the authors performed the lumen segmentation in intravascular optical coherence tomography (IVOCT) images.

As shown in Figure 4, CNN1 has four sub-sampling steps (maxpooling). Before and after the first subsampling step, there are three sequences of 3 × 3 convolutive layers, batch normalization layer, and ReLU. After the second, third and fourth subsampling steps, there are four sequences of 3 × 3 convolutive layers. In addition, CNN1 is composed of four oversampling steps. CNN2 has two subsampling steps (maxpooling) and two oversampling steps. Before each subsampling, there are three sequences of 3 × 3 convolutive layers, a batch normalization layer, and ReLU. Finally, CNN3, like CNN2, also has two subsampling steps (maxpooling) and two oversampling steps. However, before each subsampling there is only one sequence of 3 × 3 convolutive layer, batch normalization layer, and ReLU.

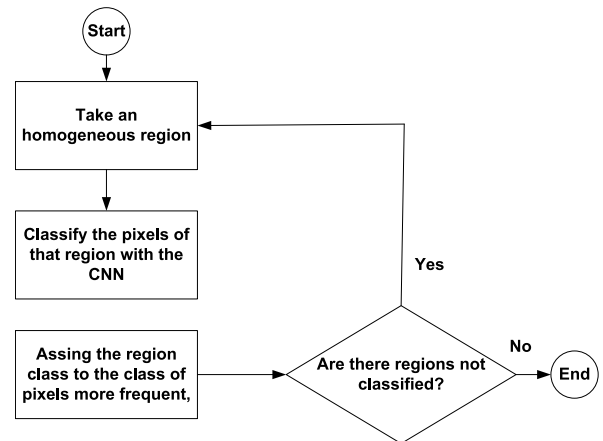


FIGURE 5. Flowchart of the fusion process.

The optimization methods evaluated were SGDM, ADAM and RMSProp. For each of these methods the following methods were used to improve generalization: dropout, L₂ regularization and dropout with L₂ regularization. Mosaic images were used as the input dataset. Two groups of simulations were done. Both groups are comprised of 36 simulations (3 architectures × 3 optimization methods × 4 improving generalization methods). In the first group, six-band mosaic images were used as CNN inputs, while in the second group, three-band mosaic images were used as CNN inputs.

For each dataset version, the model (architecture + optimization method + improving generalization method) with best performance, evaluated in terms of accuracy in the validation set was selected.

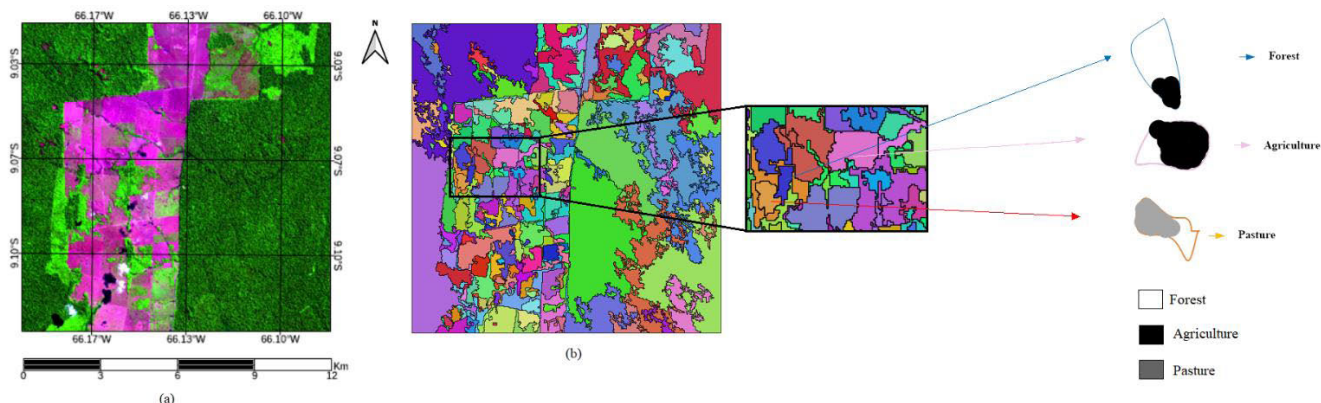


FIGURE 6. Illustration of fusion process: (a) original image; (b) result image of homogeneous region segmentation algorithm. Inside the small rectangle, three small regions are selected and magnified. The upper region has more gray pixels so it is classified as pasture. The central region has more black pixels so it is classified as agriculture, and because there are more white pixels in the lower region, it is classified as forest.

The CNN performances were compared with pre-trained network performances, using transfer learning, with shallow training. The pre-trained networks used in this work were ResNet50, InceptionResnetv2, MobileNetv2, and Xception combined with SGDM, ADAM, and RMSProp optimization methods, six-band mosaic images as input. A total of twelve experiments were performed.

A Dell Precision 5820 Workstation with Intel Xeon W-2102 2.9GHz, 64GB RAM, Windows 10, Matlab 2019a and NVIDIA Quadro GV100 32GB was used in the experiments. The following parameters were used for the CNN training: initial learning rate = 0.001, learning rate dropout factor = 0.5, number of epochs = 200, batch size = 2, dropout layer parameter = 0.3, L₂ smoothing factor = 0.001.

To evaluate the results, the Global Accuracy (GA), Average Accuracy (AA), Jaccard Similarity Coefficient (J), Weighted Jaccard Similarity Coefficient (WJ) and F1 Score (F1) were calculated [29].

In this work, CNN was trained using mosaic images consisting of 40 × 40-pixel patches that tend to generate images cut out with patches of the same size. To neutralize this trend and improve classification/segmentation, a fusion process using the CNN output and a multi-resolution homogeneous region segmentation algorithm was used. This algorithm and the fusion process are described in the sequences.

D. SEGMENTATION OF HOMOGENOUS REGIONS

The segmentation of homogeneous regions consists of creating an image with regions of the same homogeneity based on the value of the pixel of the image. In this work, the algorithm called a multiresolution segmentation proposed by [30] was used, which performs a general segmentation based on homogeneity definitions combined with local and global optimization techniques. A scale parameter is used to control the average size of the image object. According to the authors, multi-resolution segmentation is a technique for merging regions starting with one-pixel objects. In several subsequent steps, smaller image objects are merged into larger objects.

Throughout this pairwise grouping process, the underlying optimization procedure minimizes the heterogeneity of the resulting image objects. At each step, this pair of adjacent image objects is merged, resulting in the smallest growth of the defined heterogeneity. If the smallest growth exceeds the limit defined by the scale parameter, the process stops. In this way, we can say that multiresolution segmentation is a local optimization procedure. In this work, the use of this algorithm was performed through the e-Cognition Software. The scale factor used was 200.

E. DATA FUSION

The data fusion consists of using the image resulting from the multi-resolution algorithm and the image resulting from the semantic classification made by the CNN model. The data fusion used in this work is illustrated by means of the flowchart presented in Figure 5. For each region generated by the homogeneous region segmentation algorithm, we evaluated the pixel classification by the CNN. The most frequent class of pixels in a region is considered the class of that region. Figure 6 illustrates this process: Figure 6(a) shows an original image; Figure 6(b) shows the results of the homogeneous region algorithm applied in Figure 6(a). In Figure 6(b), inside a black rectangle, three small neighbor regions, which are magnified in the right side, are selected. These selected regions are filled with a hypothetical pixel classification by the CNN. The upper region has more white pixels, so it is classified as forest. The central region has more black pixels, so it is classified as agriculture, and the lower region has more gray pixels, so it is classified as pasture.

IV. RESULTS AND DISCUSSION

A. CNN SIMULATION

Tables 4 and 7 present the performance metrics of the three CNN architectures proposed in this work using the six-band and three-band mosaic image dataset, respectively. For the six-band dataset, the model with the best performance was the one in which CNN1 was used, together with the RMSProp optimization method and without using methods for

TABLE 4. Performance of CNN1 with six-band dataset.

Experiment	GA (%)	AA (%)	J (%)	WJ (%)	F1 (%)	Train Time (h)
CNN1/SGDM	99.83	99.83	99.65	99.65	99.45	94.14
CNN1/SGDM/Dropout	99.87	99.87	99.75	99.75	99.64	58.99
CNN1/SGDM/L ₂	99.93	99.93	99.85	99.85	99.82	99.50
CNN1/SGDM/Dropout/L ₂	99.94	99.94	99.88	99.88	99.85	96.09
CNN1/ADAM	99.87	99.87	99.75	99.75	99.70	142.30
CNN1/ADAM/Dropout	99.95	99.95	99.89	99.89	99.91	130.46
CNN1/ADAM/L ₂	99.84	99.84	99.68	99.68	99.76	139.95
CNN1/ADAM/Dropout/L ₂	99.84	99.84	99.68	99.68	99.76	138.28
CNN1/RMSProp	99.98	99.98	99.96	99.96	99.95	123.41
CNN1/RMSProp/Dropout	99.96	99.96	99.93	99.93	99.91	133.00
CNN1/RMSProp/L ₂	99.89	99.89	99.78	99.78	99.84	134.26
CNN1/RMSProp/Dropout/L ₂	99.91	99.91	99.81	99.81	99.85	134.34
CNN2/SGDM	97.52	97.53	95.21	95.20	91.30	52.89
CNN2/SGDM/Dropout	97.48	97.49	95.13	95.12	92.08	50.01
CNN2/SGDM/L ₂	98.33	98.33	96.73	96.72	93.29	40.22
CNN2/SGDM/Dropout/L ₂	98.20	98.20	96.49	96.48	93.60	74.84
CNN2/ADAM	98.92	98.92	97.86	97.86	94.64	95.02
CNN2/ADAM/Dropout	98.79	98.80	97.63	97.62	94.63	52.68
CNN2/ADAM/L ₂	99.36	99.36	98.73	98.72	96.47	94.95
CNN2/ADAM/Dropout/L ₂	99.11	99.11	98.24	98.24	95.76	91.50
CNN2/RMSProp	98.90	98.90	97.83	97.83	94.50	92.61
CNN2/RMSProp/Dropout	98.75	98.75	97.54	97.54	94.50	51.62
CNN2/RMSProp/L ₂	99.23	99.23	98.47	98.47	95.92	105.30
CNN2/RMSProp/Dropout/L ₂	99.02	99.02	98.06	98.06	95.39	81.47
CNN3/SGDM	95.38	95.40	91.33	91.31	86.01	45.21
CNN3/SGDM/Dropout	94.40	94.41	89.63	89.60	86.44	48.49
CNN3/SGDM/L ₂	95.29	95.31	91.17	91.15	86.00	32.75
CNN3/SGDM/Dropout/L ₂	93.88	93.88	88.73	88.70	82.67	42.50
CNN3/ADAM	96.32	96.34	93.02	93.00	86.73	67.73
CNN3/ADAM/Dropout	95.70	95.71	91.88	91.86	87.98	49.92
CNN3/ADAM/L ₂	96.30	96.31	92.97	92.95	87.27	57.11
CNN3/ADAM/Dropout/L ₂	95.73	95.75	91.96	91.94	88.13	68.47
CNN3/RMSProp	96.20	96.21	92.80	92.78	86.59	67.39
CNN3/RMSProp/Dropout	95.30	95.31	91.19	91.17	87.19	49.50
CNN3/RMSProp/L ₂	96.17	96.18	92.74	92.72	86.51	57.40
CNN3/RMSProp/Dropout/L ₂	73.32	73.24	58.55	58.54	50.24	68.50

improving generalization. This model obtained a GA of 99.98%. For the three-band dataset, the model with the best performance was the one in which CNN1 is used together with the RMSProp optimization method and dropout method to improve generalization. This model obtained a GA of 99.92%. Tables 5 and 8 present the confusion matrix for the best CNN model for the six-band and three-band dataset, respectively. From Tables 5 and 8, we observe that these models presented a greater classification error in regions where the gold-standard identified areas as forest but were classified as pasture. These CNN models were evaluated in a 5-fold cross-validation methodology. The dataset used for these experiments was the training and test dataset. Tables 6 and 9 present the performance obtained for this experiment using the six-band and three-band dataset, respectively.

B. SIMULATION RESULTS USING TRANSFER LEARNING

Table 10 presents the results obtained using pre-trained nets. The best performance was achieved using the Resnet50 network and the ADAM method. For this model, a GA of 94.50% was obtained. However, this performance was inferior to the CNN1 model with RMSProp optimization method. This model presented a GA of 99.98%.

TABLE 5. Confusion matrix for the best CNN model (CNN1/RMSProp) using six-band dataset.

	Agriculture	Pasture	Forest
Agriculture	1	2.19E-07	0
Pasture	9.93E-05	0.99	1.12E-04
Forest	3.14E-08	4.07E-04	0.99

TABLE 6. CNN1/RMSProp model performance using a 5-fold cross-validation methodology and six-band dataset.

Fold	GA(%)	AA(%)	J(%)	WJ(%)	F1(%)
1	99,99	99,99	99,99	99,99	99,99
2	99,99	99,99	99,97	99,97	99,97
3	99,89	99,89	99,79	99,79	99,78
4	99,99	99,99	99,99	99,99	99,99
5	99,99	99,99	99,97	99,97	99,98
Average	99,97	99,97	99,97	99,97	99,94

TABLE 7. Performance of CNN1 with three band dataset.

Experiment	GA(%)	AA(%)	J(%)	WJ(%)	F1(%)
CNN1/SGDM	99.75	99.75	99.50	99.50	99.30
CNN1/SGDM/Dropout	99.52	99.53	99.06	99.05	98.84
CNN1/SGDM/L ₂	99.00	99.00	98.02	98.02	97.95
CNN1/SGDM/Dropout/L ₂	99.73	99.73	99.45	99.45	99.46
CNN1/ADAM	99.89	99.89	99.78	99.78	99.77
CNN1/ADAM/Dropout	99.63	99.63	99.26	99.26	99.40
CNN1/ADAM/L ₂	98.51	98.52	97.08	97.08	97.56
CNN1/ADAM/Dropout/L ₂	98.70	98.70	97.44	97.43	97.92
CNN1/RMSProp	99.89	99.89	99.79	99.79	99.78
CNN1/RMSProp/Dropout	99.92	99.92	99.84	99.84	99.77
CNN1/RMSProp/L ₂	99.38	99.38	98.76	98.76	98.74
CNN1/RMSProp/Dropout/L ₂	97.38	97.40	94.93	94.92	95.93
CNN2/SGDM	95.98	96.00	92.41	92.39	86.85
CNN2/SGDM/Dropout	94.90	94.92	90.49	90.46	87.11
CNN2/SGDM/L ₂	96.37	96.38	93.09	93.08	87.98
CNN2/SGDM/Dropout/L ₂	95.95	95.96	92.33	92.31	89.08
CNN2/ADAM	97.16	97.17	94.54	94.52	89.43
CNN2/ADAM/Dropout	97.27	97.27	94.74	94.73	89.83
CNN2/ADAM/L ₂	97.10	97.10	94.42	94.41	89.29
CNN2/ADAM/Dropout/L ₂	96.82	96.84	93.91	93.89	90.23
CNN2/RMSProp	97.47	97.48	95.11	95.10	89.52
CNN2/RMSProp/Dropout	96.81	96.81	93.88	93.87	88.31
CNN2/RMSProp/L ₂	97.26	97.27	94.73	94.72	90.25
CNN2/RMSProp/Dropout/L ₂	96.53	96.55	93.39	93.37	90.27
CNN3/SGDM	93.09	93.12	87.41	87.38	81.95
CNN3/SGDM/Dropout	91.84	91.86	85.35	85.32	82.74
CNN3/SGDM/L ₂	93.16	93.18	87.52	87.49	82.07
CNN3/SGDM/Dropout/L ₂	92.25	92.28	86.01	85.97	83.94
CNN3/ADAM	94.37	94.39	89.57	89.54	81.95
CNN3/ADAM/Dropout	93.08	93.10	87.37	87.34	83.36
CNN3/ADAM/L ₂	94.07	94.08	89.03	89.00	82.92
CNN3/ADAM/Dropout/L ₂	92.94	92.97	87.15	87.12	84.47
CNN3/RMSProp	94.22	94.24	89.32	89.30	81.63
CNN3/RMSProp/Dropout	93.22	93.25	87.61	87.58	83.72
CNN3/RMSProp/L ₂	93.85	93.87	88.66	88.64	83.21
CNN3/RMSProp/Dropout/L ₂	92.43	92.45	86.32	86.29	81.60

C. CLASSIFICATION/SEGMENTATION OF IMAGES FROM AMAZON REGIONS

To evaluate the methodology proposed in this work, some images from the Amazon region were selected to perform the segmentation/classification.

These images were classified using only the model CNN2 + ADAM optimization method + L₂ regularization and using the fusion process described in sections III-D and III-E. In Figure 7, we have an example of the classification of a LANDSAT-8/OLI image using this process. As shown

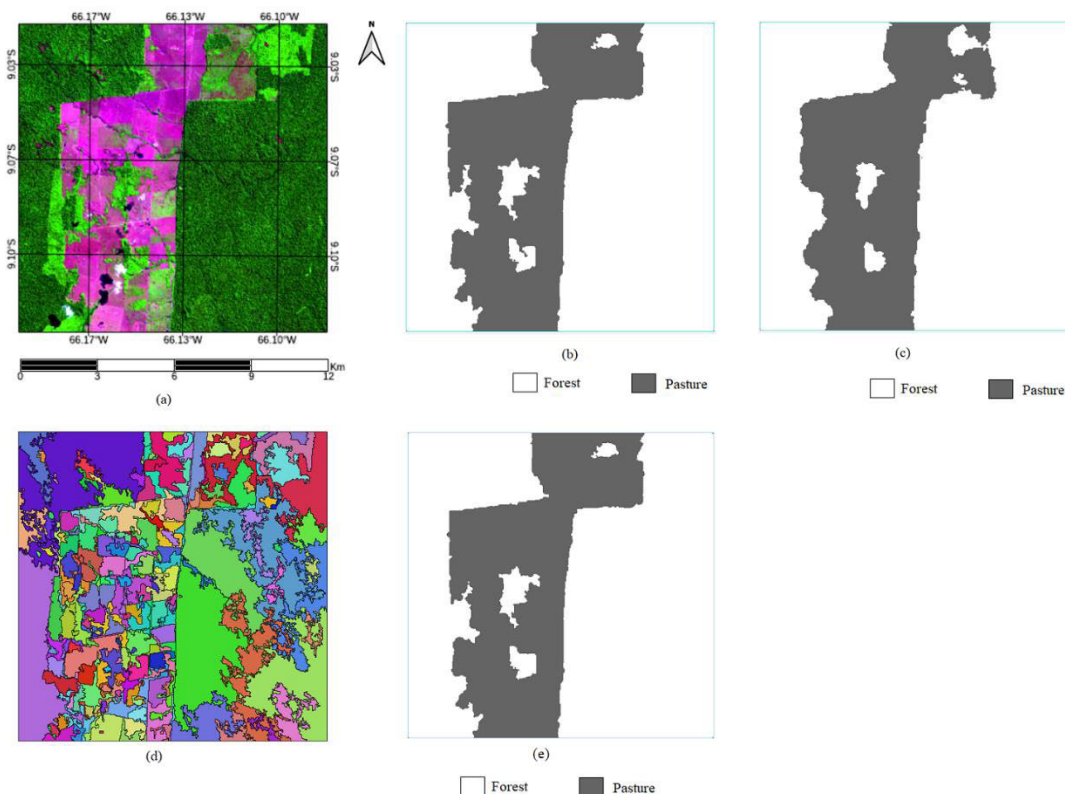


FIGURE 7. LANDSAT-8/OLI images of Amazon regions. (a) original image displayed in color composition of bands B6-B5- B4; (b) gold standard; (c) semantic classification made by CNN, with an accuracy of 96.84%. (d) output image resulting from the multi-resolution segmentation; (e) result of classified image after data fusion process, with an accuracy of 99.8%.

TABLE 8. Confusion matrix for the best CNN model (CNN1/RMSProp) using the six band dataset.

	Agriculture	Pasture	Forest
Agriculture	0.99	3.05E-04	8.82E+08
Pasture	3.03E-04	0.99	8.72E-04
Forest	2.67E-04	7.14E-04	0.99

TABLE 9. CNN1/RMSProp model performance using a 5-fold cross-validation methodology and six-band dataset.

Fold	GA(%)	AA(%)	J(%)	WJ(%)	F1(%)
1	99.86	99.86	99.72	99.72	99.65
2	99.87	99.87	99.74	99.74	99.71
3	99.76	99.76	99.53	99.52	99.59
4	99.87	99.87	99.74	99.74	99.72
5	99.94	99.94	99.89	99.89	99.87
Average	99.86	99.86	99.72	99.72	99.71

in Figure 7(c), in the results of CNN classification, there are some regions incorrectly classified as forest.

In Figure 7(e), which shows the result of the fusion process, these misclassifications are minimized. The accuracy obtained using only CNN semantic segmentation was 96.84%. After the data fusion process, the accuracy increased to 99.81%.

We can observe that the proposed methodology presented impressive performances. The best results were obtained using the CNN2 model, the ADAM optimization method, and L_2 regularization for improving generalization. The data fusion process has increased the image classification accuracy, as shown in Figure 7. Table 11 presents a comparison of

TABLE 10. Performance using transfer learning.

Fold	GA(%)	AA(%)	J(%)	WJ(%)	F1(%)
RN50/SGDM	92.26	92.28	86.03	86.00	86.31
RN50/ADAM	94.50	94.51	89.77	89.75	87.73
RN50/RMProp	94.17	94.19	89.21	89.18	88.70
IR2/SGDM	89.77	89.79	81.84	81.81	80.44
IR2/ADAM	87.62	87.66	78.14	78.11	77.75
IR2/RMSProp	88.23	88.25	79.38	79.35	79.91
MN2/SGDM	77.87	77.93	64.00	63.95	57.17
MN2/ADAM	77.41	77.44	63.80	63.76	56.27
MN2/RMSProp	77.83	77.86	64.27	64.23	58.92
XC/SGDM	92.25	92.27	85.96	85.93	82.67
XC/ADAM	91.34	91.36	84.41	84.37	83.02
XC/RMSProp	91.18	91.20	84.21	84.17	82.21

TABLE 11. Performance of image segmentation shown in Figure 7 using only CNN2/ADAM/ L_2 and the data fusion process.

Segmentation Method	GA (%)	AA (%)	J (%)	WJ (%)	F1 (%)
CNN2/ADAM/ L_2	96.84	96.42	93.09	93.90	81.08
Data Fusion Process	99.81	99.82	99.56	99.61	98.09

the image segmentation performance of Figure 7 using only CNN model and the data fusion process.

In Figure 8, we show two LANDSAT-8/OLI images of Amazon regions with their respective gold standard, and the classified image using the CNN2 + ADAM optimization method + L_2 regularization and data fusion process. Image 1 corresponds to a region of scene 228/67, and image 2 corresponds to a region of scene 226/68. The accuracies

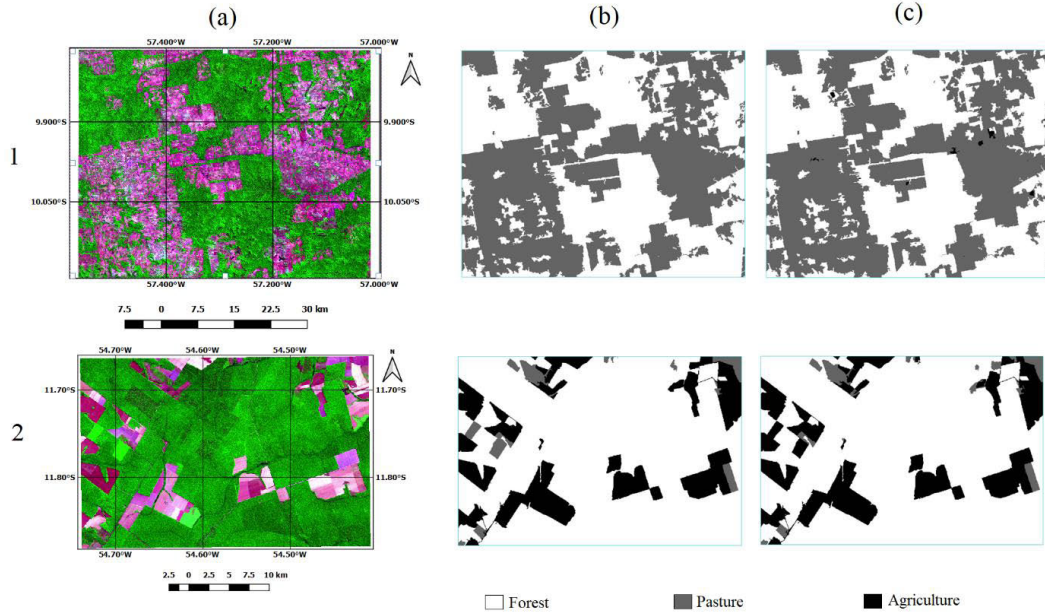


FIGURE 8. LANDSAT-8/OLI images of Amazon regions. (a) original images displayed in color composition B6-B5-B4; (b) gold-standard image; (c) images classified by CNN2/ADAM/L₂ model and fusion process. The accuracy obtained for images 1 and 2 was 99.63%, 99.08%, respectively.

TABLE 12. Datasets assembled using images evaluated by [9] and [10].

Dataset	Orbit/Point	Date	Train	Validation
1 [9]	227/63	07/08/2017	1400	600
	227/63	07/21/2017		
	227/65	08/18/2017		
	227/65	07/21/2018		
2 [10]	230/65	06/21/2017	1400	600
	230/65	06/24/2018		

TABLE 13. Performance of three best models of each CNN using the same regions used by [9] and [10].

DT*	Experiment	GA(%)	AA(%)	J(%)	WJ(%)	F1(%)
1	CNN1/RMSProp	98.48	97.98	96.26	97.03	99.68
	CNN2/ADAM/L ₂	98.44	97.94	96.89	96.93	98.68
	CNN3/RMSProp/L ₂	97.96	97.30	95.91	95.99	98.14
2	CNN1/RMSProp	98.59	98.10	96.26	97.25	99.90
	CNN2/ADAM/L ₂	99.55	99.32	98.95	99.10	99.60
	CNN3/RMSProp/L ₂	99.52	99.33	98.85	99.05	99.55

*Dataset

obtained for images 1 and 2 were 99.63% and 99.08%, respectively.

D. EVALUATION OF METHODOLOGY IN REGIONS USED BY [9], [10]

The methodology presented in this work was also evaluated in the regions evaluated by [17], [18]. Two datasets with 2000 mosaic images each were assembled using the images evaluated by [17], [18], respectively. Table 12 shows the images used for the construction of each dataset and the division of the set for training and validation. Each mosaic image is 400 × 400-pixel size and is made up of patches of 20 × 20-pixel size. Each image patch corresponds to a region of deforestation or nondeforestation. Figures 9 and 10 show examples of mosaic images and the corresponding

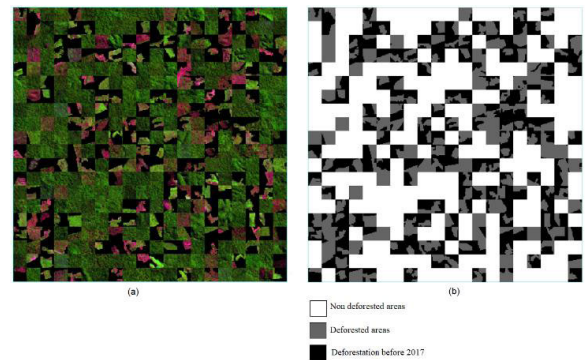


FIGURE 9. Example of a mosaic image for the region used by [9]. In (a), the mosaic image is displayed in color composition B6-B5-B4 and (b) the corresponding gold standard.

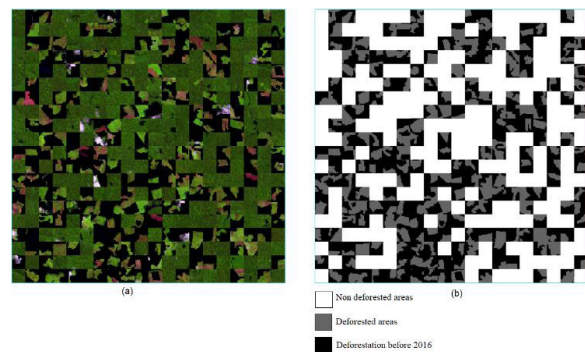


FIGURE 10. Example of a mosaic image for the region used by [10]. In (a), the mosaic image is displayed in color composition B6-B5-B4 and (b) corresponding gold standard.

gold standards for datasets 1 and 2, respectively. Table 13 represents the performance of the CNN models for each of

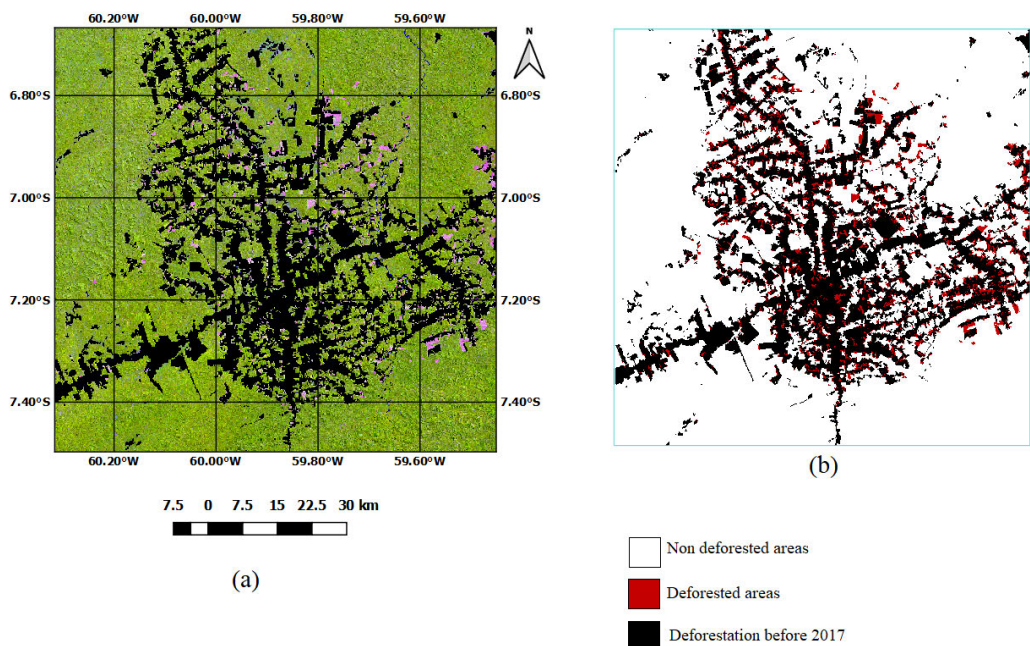


FIGURE 11. LANDSAT-8/OLI images of Amazon regions. (a) Image of a region (3200 × 3200pixels) of the LANDSAT 230/65 scene evaluated by [9] displayed in color composition B6-B5-B4. (b) shows classified image, with accuracy of 99.91%.

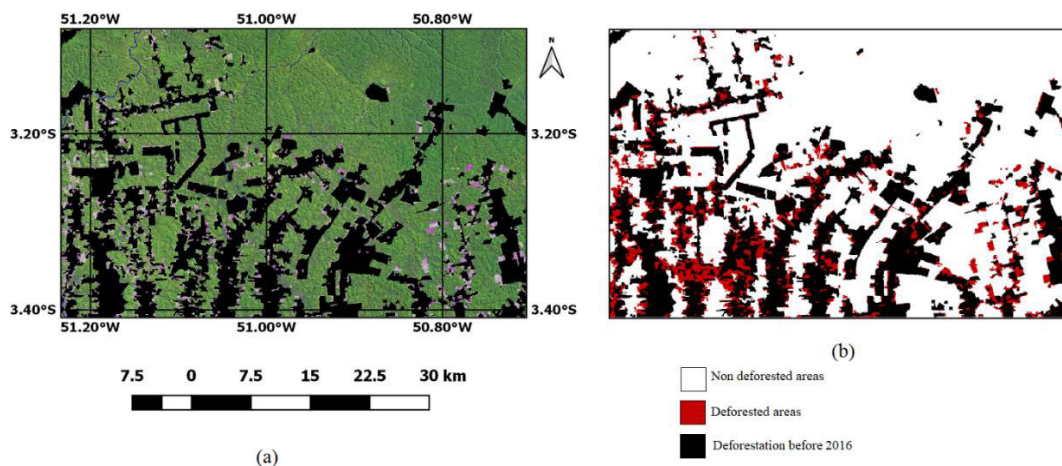


FIGURE 12. LANDSAT-8/OLI images of Amazon regions. In (a), image of a region (2000 × 1200pixels) of the LANDSAT 225/62 scene evaluated by [10] displayed in color composition B6-B5-B4. In (b), image is classified with accuracy of 99.94%.

these datasets. The best model of each CNN presented in Table 6 was used.

The CNN3/RMSProp/L₂ model and the data fusion process were used to classify images presented in Figure 11 and 12. Figure 11 corresponds to a region evaluated by [17]. For this image, the obtained accuracy was 99.91% and the value of F1-Score was 99.94%. The accuracy value was very similar to the value obtained by [17]. The F1-Score was higher than the value obtained by the authors, which was 94.65%. Figure 12 corresponds to a region evaluated by [18]. For this image the obtained accuracy was 99.94%, and the value of the F1-Score was 99.92%. The accuracy value was higher than the best results obtained by [18], which was 95%. In regard

to the value of the F1-Score, the value was higher than the one obtained by the authors, which was 63%.

V. CONCLUSION

This study proposed a methodology using deep learning to solve the difficulties in the segmentation of land use and land cover for the Amazon region for the pasture, agriculture and forest classes. The proposed methodology consisted of a data fusion process performed through semantic segmentation using CNN and a homogeneous region algorithm. LANDSAT-8/OLI optical images of the Brazilian Amazon region were used to validate the proposed methodology. The reference data used were the results obtained from the

TerraClass project of INPE in 2014. The CNN architectures used in this study were based on the architecture proposed in the work of [28]. Two datasets were evaluated: the first using six bands and the second with three bands. The results for the six-band dataset outperforms the one obtained with a three-band dataset.

In this study, CNN was trained using mosaic images consisting of 40×40 pixel patches that tend to generate images cut out with patches of the same size. To neutralize this trend and improve classification/segmentation, a fusion process using the CNN output and a multi-resolution homogeneous region segmentation algorithm was used.

The proposed method was compared to some pretrained CNN architectures using transfer learning. The comparison demonstrated the effectiveness of our proposed method against these well-known CNN architectures in the task of remote sensing applications. The proposed method proved to be more appropriate for this specific application. For the same region evaluated by [9], the results were practically the same. Regarding the results obtained by [10], our method obtained better results.

The main contributions of this work to soil classification in remote sensing applications were twofold. The first one was to propose the mosaic image technique for training CNN architectures. The second one was a fusion technique to improve the soil classification, using two types of data: the CNN output image and output image of a homogeneous region segmentation algorithm. Another contribution is making the datasets used in this work available, which may be used as a base for other research in the Amazon Region. Data will be available upon request. Given the results presented, we conclude that the methodology proposed in this work proved to be quite efficient in performing the task of segmentation/classification of remote sensing images for regions of the Amazon.

For future studies, we propose using more varied classes of land use, in addition to using other types of remote sensing images, for example, synthetic aperture radar (SAR) images.

REFERENCES

- [1] J. R. Jensen, *Remote Sensing of the Environment: An Earth Resource Perspective*, vol. 1, 2nd ed. 2014.
- [2] F. F. Sabins, *Remote Sensing: Principles and Applications*, 2008.
- [3] INPE. (2019). *Satellite Monitoring of Deforestation of the Brazilian Amazon Forest*. Accessed: Sep. 11, 2020. [Online]. Available: <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>
- [4] INPE. (2019). *National Institute of Space Research—Projects and Research—TerraClass*. Accessed: Aug. 17, 2019. [Online]. Available: http://www.inpe.br/cra/projetos%5C_pesquisas/dados%5C_terraClass.php
- [5] A. Noma, T. S. Körting, and L. M. G. Fonseca, “Uma comparação entre classificadores usando Regiões e perfis EVI para agricultura,” in *Proc. 15th Simpósio Brasileiro de Sensoriamento Remoto. Anais*, vol. 16, Apr. 2013, pp. 2250–2257. [Online]. Available: <http://urlib.net/dpi.inpe.br/marte2/2013/05.28.23.34>.
- [6] INPE. (2019). *The Consolidated Rate of Deforestation by Shallow Cut for the Nine States of Legal Amazon (AC, AM, AP, MA, MT, PA, RO, RR e TO) in 2019*. Accessed: Aug. 17, 2020. [Online]. Available: http://www.inpe.br/noticias/noticia.php?Cod%5C_Noticia=5465
- [7] D. P. Barbosa, A. Noma, T. S. Körting, and L. M. G. Fonseca, “Um estudo experimental com classificadores baseados em Regiões e Perfis EVI,” in *Proc. 17th Simpósio Brasileiro de Sensoriamento Remoto*, vol. 1, 2015, pp. 6381–6388.
- [8] E. L. Bullock, C. E. Woodcock, and P. Olofsson, “Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis,” *Remote Sens. Environ.*, vol. 238, Nov. 2020, Art. no. 110968, doi: [10.1016/j.rse.2018.11.011](https://doi.org/10.1016/j.rse.2018.11.011).
- [9] P. de Bem, O. de Carvalho Junior, R. F. Guimarães, and R. T. Gomes, “Change detection of deforestation in the Brazilian amazon using landsat data and convolutional neural networks,” *Remote Sens.*, vol. 12, no. 6, p. 901, Mar. 2020, doi: [10.3390/rs12060901](https://doi.org/10.3390/rs12060901).
- [10] M. O. Adarme, R. Q. Feitosa, P. N. Happ, C. A. De Almeida, and A. R. Gomes, “Evaluation of deep learning techniques for deforestation detection in the Brazilian amazon and cerrado biomes from remote sensing imagery,” *Remote Sens.*, vol. 12, no. 6, p. 910, Mar. 2020, doi: [10.3390/rs12060910](https://doi.org/10.3390/rs12060910).
- [11] Y. E. Shimabukuro, E. Arai, V. Duarte, A. Jorge, E. G. D. Santos, K. A. C. Gasparini, and A. C. Dutra, “Monitoring deforestation and forest degradation using multi-temporal fraction images derived from landsat sensor data in the Brazilian amazon,” *Int. J. Remote Sens.*, vol. 40, no. 14, pp. 5475–5496, Jul. 2019, doi: [10.1080/01431161.2019.1579943](https://doi.org/10.1080/01431161.2019.1579943).
- [12] T. A. Lima, R. Beuchle, A. Langner, R. C. Grecchi, V. C. Griess, and F. Achard, “Comparing Sentinel-2 MSI and landsat 8 OLI imagery for monitoring selective logging in the Brazilian Amazon,” *Remote Sens.*, vol. 11, no. 8, p. 961, Apr. 2019, doi: [10.3390/rs11080961](https://doi.org/10.3390/rs11080961).
- [13] C. Alves, S. Júnior, O. O. Bittencourt, F. Morelli, and R. Santos, “Classificação De Áreas Queimadas Por Machine Learning Usando Dados De Sensoriamento Remoto,” in *Proc. Ann. 16th Brazilian Remote Sens. Symp. SBSR*, 2019, pp. 1784–1787.
- [14] V. Leseux et al., “Cobertura do solo utilizando Sentinel-2 (sensor Msi),” in *Proc. Ann. XIX Brazilian Remote Sensing Symp. SBSR*, 2019, pp. 720–723.
- [15] F. S. Nascimento, M. Gastauer, P. W. M. Souza-Filho, W. R. Nascimento, D. C. Santos, and M. F. Costa, “Land cover changes in open-cast mining complexes based on high-resolution remote sensing data,” *Remote Sens.*, vol. 12, no. 4, pp. 1–21, 2020, doi: [10.3390/rs12040611](https://doi.org/10.3390/rs12040611).
- [16] H. L. G. Cassol, E. Arai, E. E. Sano, A. C. Dutra, T. B. Hoffmann, and Y. E. Shimabukuro, “Maximum fraction images derived from year-based project for on-board autonomy-vegetation (PROBA-V) data for the rapid assessment of land use and land cover areas in mato grosso state, Brazil,” *Land*, vol. 9, no. 5, p. 139, May 2020, doi: [10.3390/LAND9050139](https://doi.org/10.3390/LAND9050139).
- [17] J. Lee, J. A. Cardille, and M. T. Coe, “Agricultural expansion in mato grosso from 1986-2000: A Bayesian time series approach to tracking past land cover change,” *Remote Sens.*, vol. 12, no. 4, pp. 1–18, 2020, doi: [10.3390/rs12040688](https://doi.org/10.3390/rs12040688).
- [18] R. V. Maretto, L. M. G. Fonseca, N. Jacobs, T. S. Körting, H. N. Bendini, and L. L. Parente, “Spatio-temporal deep learning approach to map deforestation in Amazon rainforest,” *IEEE Geosci. Remote Sens. Lett.*, early access, Apr. 28, 2020, doi: [10.1109/lgrs.2020.2986407](https://doi.org/10.1109/lgrs.2020.2986407).
- [19] C. C. Aggarwal, *Neural Networks and Deep Learning*. New York, NY, USA: Springer, 2018.
- [20] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- [21] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 770–778, Dec. 2016, doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [22] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-ResNet and the impact of residual connections on learning,” in *Proc. 31st AAAI Conf. Artif. Intell. AAAI*, 2017, pp. 4278–4284.
- [23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520, doi: [10.1109/CVPR.2018.00474](https://doi.org/10.1109/CVPR.2018.00474).
- [24] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807, doi: [10.1109/CVPR.2017.195](https://doi.org/10.1109/CVPR.2017.195).
- [25] USGS. (2019). *Earth Explorer*. Accessed: May 19, 2019. [Online]. Available: <http://earthexplorer.usgs.gov/>
- [26] USGS. (2017). *What are the Best Spectral Bands to Use for my Study*. Accessed: May 19, 2019. [Online]. Available: <https://landsat.usgs.gov/what-are-best-spectral-bands-use-my-study>

- [27] Z. Yu, L. Di, R. Yang, J. Tang, L. Lin, C. Zhang, M. S. Rahman, H. Zhao, J. Gaigalas, E. G. Yu, and Z. Sun, "Selection of landsat 8 OLI band combinations for land use and land cover classification," in *Proc. 8th Int. Conf. Agro-Geoinform. (Agro-Geoinform.)*, Jul. 2019, pp. 1–5, doi: 10.1109/Agro-Geoinformatics.2019.8820595.
- [28] M. Miyagawa, M. G. F. Costa, M. A. Gutierrez, J. P. G. F. Costa, and C. F. F. C. Filho, "Lumen segmentation in optical coherence tomography images using convolutional neural network," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 600–603, doi: 10.1109/EMBC.2018.8512299.
- [29] MathWorks. (2017). *Evaluate Semantic Segmentation Data Set Against Ground Truth*. Accessed: Aug. 10, 2019. [Online]. Available: <https://in.mathworks.com/help/vision/ref/evaluatesemanticsegmentation.html>
- [30] M. Baatz and A. Schape, "Multiresolution Segmentation: An optimization approach for high quality multi-scale image segmentation," in *Angewandte Geographische Informations-Verarbeitung XII*, 2000, pp. 12–23.



JOEL PARENTE DE OLIVEIRA received the degree in computer engineering from the State University of Amazonas (UEA) in 2010 and the master's degree in electrical engineering from the Federal University of Amazonas (UFAM), Brazil, in 2017. He is currently a Science and Technology Analyst, working in database administration, with the Center for Management and Operation of the Amazon Protection System, Operations and Management Center of the Amazon Protection System (CENSIPAM). His research interests include pattern recognition and digital image research.



MARLY GUIMARÃES FERNANDES COSTA (Member, IEEE) received the degree in electrical engineering from the Federal University of Amazonas (UFAM) in 1983, and the master's degree in electrical engineering, in the area of telecommunications, and the Ph.D. degree in electrical engineering, automation concentration area, from the State University of Campinas (UNICAMP), in 1988, and 1996, respectively. She is currently a Professor with the Department of Electrical Engineering, UFAM. She has experience in digital image processing, pattern recognition, and biomedical engineering, focusing on the following subjects such as image automation, the development of diagnostic aid tools, and expert systems.



CÍCERO FERREIRA FERNANDES COSTA FILHO (Member, IEEE) received the degree in electrical engineering from the Federal University of Pernambuco (UFPE) in 1982, and the master's degree in electrical engineering, in the area of microelectronics, and the Ph.D. degree in automation from the State University of Campinas (UNICAMP), in 1985 and 1996, respectively. He worked with the UNICAMP and the Center for Analysis, Research, and Technological Innovation Foundation-FUCAPI. Since 1990, he has been a Professor with the Federal University of Amazonas. He is currently a full professor, teaching undergraduate (electrical engineering and computer engineering) and graduate courses (master's degree in electrical engineering-PPGEE and the Ph.D. degree in computer engineering-PPGI). His research interests include artificial intelligence, pattern recognition, digital image processing, and optimization.

• • •