Classification of the water volume of dams using heterogeneous remote sensing images through a deep convolutional neural network

Mateus de Souza Miranda[1], Renato de Sousa Maximiano[1], Valdivino Alexandre de Santiago Júnior[1], Thales Sehn Körting[1], Leila Maria Garcia Fonseca[1]

[1]Instituto Nacional de Pesquisas Espaciais (INPE)
Avenida dos Astronautas, 1758, Jardim da Granja - 12227-010,
São José dos Campos, SP, Brazil.

{mateus.miranda, renato.maximiano, valdivino.santiago,

thales.korting, leila.fonseca}@inpe.br

Abstract. *Deep Convolutional Neural Networks (DCNN) have played an important role in several application domains and also in remote sensing image classification and object detection. In this article, we extend a previously proposed model, used to classify forest areas as preserved or non-preserved, in order to classify the water volume of dams in the state of São Paulo, Brazil, using remote sensing images. Our revised DCNN addresses a multi-class classification problem while our previous one was devised for binary classification. Moreover, our model relies on heterogeneous images, considering different sensors and also different spatial resolutions regarding the data sets. Results show that the overall accuracy of our model was 85.56% considering images from the Atibainha and Jaguari dams of the Cantareira water supply system to compose the testing set, demonstrating the feasibility of our approach to these types of applications. This is an indication of the good generalization capabilities of our model.*

1. Introduction

Climate change, population increase and water consumption are pointed out as the main factors of the water crisis in Southeast Brazil [INPE 2015], which prolonged drought brings significant impacts not only to the society, with the containment of water supply and increase in rates of electricity, but also to the environment, extinction of aquatic species, the disappearance of springs and rivers. The water volume and flow are monitored not only by rainfall stations, as well as by satellites, through remote sensing images, enabling the observation of activities on the Earth's surface.

Recently, Deep Learning (DL) techniques have often been implemented to monitor the water flow in rivers and dams, especially Deep Neural Networks (DNN), due to the efficiency in extracting and detecting patterns in the data, even as by the ability to classify and segment objects in images, group data sets and make predictions [Barino and dos Santos 2020]. The most popular type of DNN is the deep Convolutional Neural Network (DCNN) which is based on the human visual system [Géron 2019]. For image processing, convolutions work as filters that extract low and high-level features, such as edge and texture, making the model capable of classifying and segmenting images, were also breaking down the image, following by its reconstruction, emphasizing the object, a process called encoder-decoder [Ghassemi and Magli 2019].

The volume, variety and velocity of water-related data are growing due to increased attention to topics such as disaster response, water resource management and climate change. With the widening availability of computing resources and the popularity of DL, the data is transformed into practical knowledge, revolutionizing the water industry [Sit et al. 2020]. There are several works using the application of DL, highlighting the CNN, being used in, extraction of water bodies from remote sensing images [Chen et al. 2018] and [Namikawa et al. 1], segmentation separating water from land, snow, ice, clouds and shadows [Isikdogan et al. 2017], water reservoir recognition [Fang et al. 2019], reservoir volume simulation and prediction [Baek et al. 2020].

Therefore, in this article we evaluate the performance of a deep convolutional neural network, previously used for binary classification of preserved and non-preserved areas in the context of Cerrado on the Brazilian states of Tocantins and Goiás [Miranda et al. 2021], for classification of the volume of water in the Atibainha and Jaguari dams in the state of São Paulo, as *normal*, *low* and *critical*, using satellite images with different spatial resolution, since it was trained with 10 meters of spatial resolution and tested with 2 meters of spatial resolution, in order to assess the performance of the model.

This paper is organized as follows. Section 2 introduce related works. Section 3 presents Material and Methods. Results and discussion are presented in Section 5. Section 5 presents conclusions and future directions.

2. Related work

There is a great interest in the remote sensing community to rely on DL techniques to help to develop their systems. In [Ma et al. 2019], authors presented a meta-analysis and review of DL applied to remote sensing and they concluded that DL models have been used for several remote sensing image analysis land use and land cover (LULC) classification, and segmentation. Despite the success of DL, the authors mentioned that its performance in LULC classification is still inferior compared to scene classification and object detection. This remark is just to emphasize the need for more experimentation, in different contexts, to perceive the performance of DL.

With this, CNN has been used for several tasks in this area of study, which encourages their use in different applications, precisely because of the ability to process and learn about complex and large data. For example, [Khryashchev et al. 2018] CNN was applied to perform the detection of geographic objects with the help of experts, to carry out the validation of the results, where the segmentation of images has found application in urban planning, forest management, and climate modelling. In addition, [Ai et al. 2020] proposes using different remote sensing images in four spectral bands, red, green and infrared, to retrieve the water depth, taking into account the non-linear relationship between the value of radiance and the water depth value of adjacent and central pixels. Quantitative analysis and experimental results showed that the accuracy of the CNN model in retrieving shallow sea areas is improved by more than 50%, where, the RMSE accuracy can reach 0.9485.

In [Fernandes et al. 2020], the authors studied and evaluated two distinct approaches for detecting water tanks and swimming pools in satellite images, which can be useful in monitoring water-related diseases. The first method uses a support vector machine to classify into positive and negative a discretized colour histogram of a certain segment of the original image, while the second method used the Faster R-CNN structure to detect these objects, built with a training set composed by swimming pools and water tanks on the city of Belo Horizonte, Brazil. The results demonstrated that the DL method using CNN outperformed the shallow strategy, achieving an accuracy of more than 93% in the pool detection task and 73% for water tanks.

In the work of the [Pan et al. 2020], the authors performed a comparative study of water indices and image classification algorithms to map water bodies using 24 high-resolution Landsat images, using two unsupervised methods, the zero water index threshold H0 method and Otsu automatic threshold selection method, and one supervised, the K-nearest neighbours (KNN) method. The study showed that the unsupervised classification achieved results comparable to supervised, showing that in some cases we can reduce the computational cost applying an unsupervised classification.

3. Material and methods

3.1. Study area

The study area is composed of nine dams from the state of São Paulo which is one of the Brazilian states more affected by drought, and such dams are shown in Figure 1. According to [SABESP 2021], the Atibainha, Jacareí, Jaguari dams are from the Cantareira system of water; Billings and Pedro Beicht dams respectively belong to the Guarapiranga and Alto Cotia systems water. The Itupararanga and Barra Bonita dams belong to Sorocaba and Médio Tietê Hydrographic Basin, the Serraria and Serraria dams belong to Ribeira de Iguape and South Coast Hydrographic Basin, according to [SIGRH 2020].

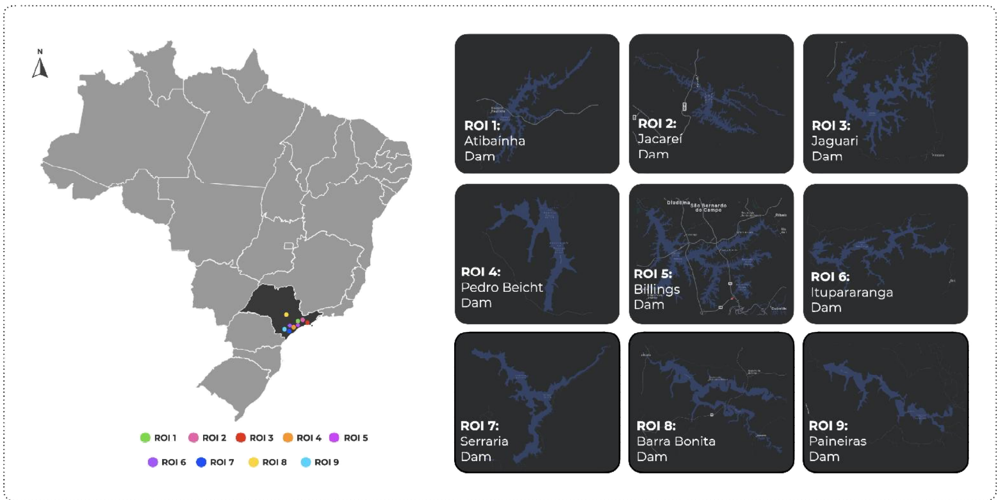

Figure 1. Study areas.

3.2. Data collection

The data collection consists of satellite images (rasters) obtained from the image catalogue of the National Institute for Space Research, comprising the study regions. Therefore, two data sets were created, one for training and other for testing.

The training image set consists of 120 images, each one with 8.562 $\times$ 12.736 pixels, recorded by the CBERS-4's PAN10M sensor, 10 meters of spatial resolution. We considered the near infrared (NIR), red (R) and green (G) bands. It is deserving noting that the nine dams were used as a study area for this dataset, looking at them during the period from 2015 to 2021, considering the dates with the lowest occurrence of cloud cover in the images.

For testing, the dataset consists of 3 images, each one with 56.842 $\times$ 58.344 pixels, from the CBERS-4A's WPM camera, whose multi-spectral and panchromatic lenses have, respectively, 8 and 2 meters of spatial resolution, considering the NIR, R, G and Panchromatic (PAN2M) bands. It is worth emphasizing that for this dataset the Juguari and Atibainha dams were used, both from the Cantareira system, observed during the dates September 3, 2020, April 8, and August 10, 2021.

3.2.1. Data Preprocessing

The training set images were composed of NIR, R and G bands in order to highlight the edges of the dams, based on the work of [Namikawa et al. 2019]. In Figure 2, colours pink, red and green represent the NIR, R, and G bands, respectively. Then, the regions of interest were cut, in the proportion of 224 $\times$ 224 pixels, generating a total of 770 images.

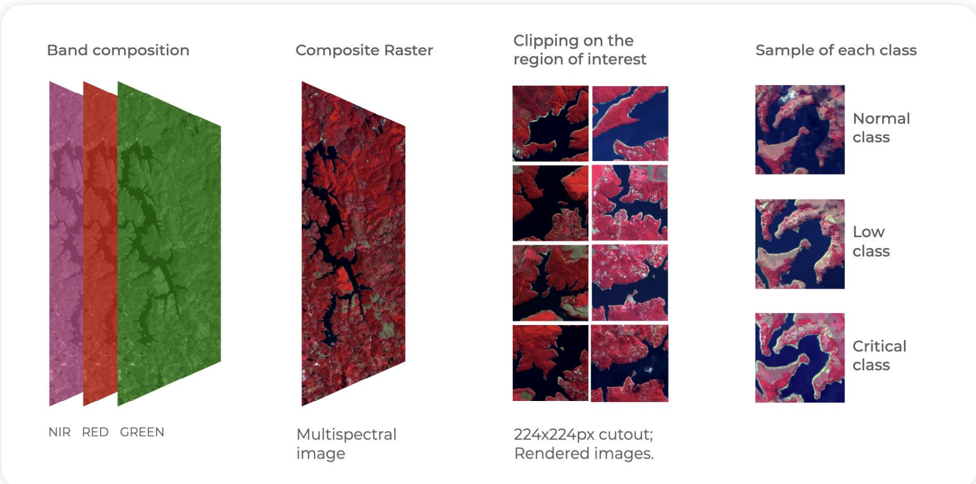

Figure 2. Training dataset: Composition of NIR, Red and Green bands, clipping interest areas, and organization of images by classes.

Based on hydrological data provided by [SABESP 2021], the images of the training and testing datasets were classified into *normal* when the volume of the dam and greater than 60% of the total capacity; *low*, when the volume is between 40% to 60% of full capacity; and *critical*, when the volume is less than 40% of the total capacity. Thus, 353, 239 and 178 images were obtained for classes *normal*, *low* and *critical*, respectively, as illustrated in Figures 2 and 3.

Given the difference in the amount of data between the threes classes, we used the static data augmentation technique, which applies transformations to images such as rotate and flip. In addition, it was possible to increase and balance the number of images in each category, also, helping to equalize the process of training, avoiding that model learns more about one of the classes. We obtained 1,527 images per class, in total 4,581 training samples.

Regarding the test dataset, the images were composed using the NIR, R and G bands. Each multi-spectral image generated in the composition, with a spatial resolution of 8 meters, was used in the fusion with its respective panchromatic raster, thus generating a multi-spectral image with a spatial resolution of 2 meters. Finally, we cropped the images with the dimensions 224×224 pixels and obtained 100 images in total, where 32 are for the *critical*, 34 for the *low* and 34 for the *normal* classes, as shown in Figure 3.

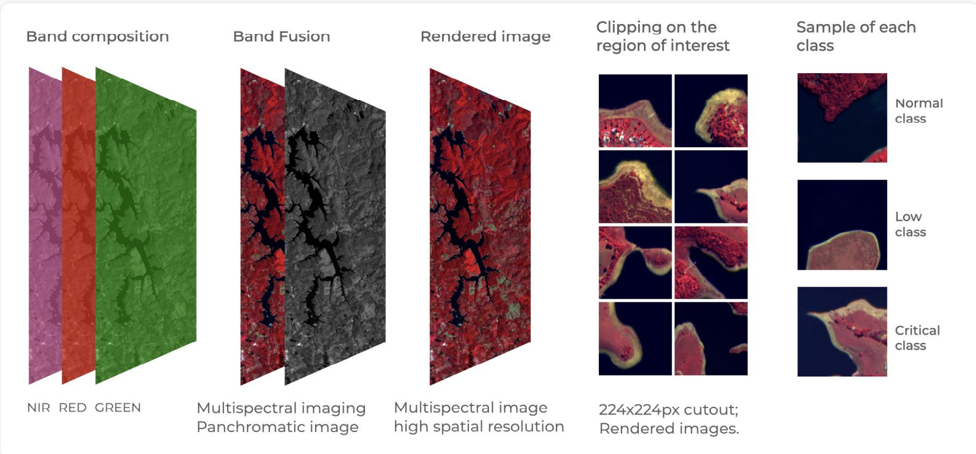

Figure 3. Testing dataset: Composition of NIR, R and G bands, fusion of the composite image with the panchromatic band, clipping of areas of interest and organization of images by classes.

3.3. The model

A DCNN model was extended from our previous work [Miranda et al. 2021] for the binary classification of vegetated regions preserved and non-preserved on the Brazilian Cerrado. However, for this work, we made adjustments in the hyper-parameters in our network architecture for the multi-class task. The model is shown in Figure 4.

The training images are inserted in the convolutional layers, with 224×224 input format, 3×3 kernel, activated by the ReLu function, after each convolution a MaxPooling2D layer, pooling (2, 2) was added, and a Dropout layer, with a probability of 25%. After the convolution layers, we have 4 fully connected layers, each with 256 neurons, activated by the ReLu function, and 4 dropout layers with a probability of 25%. The output

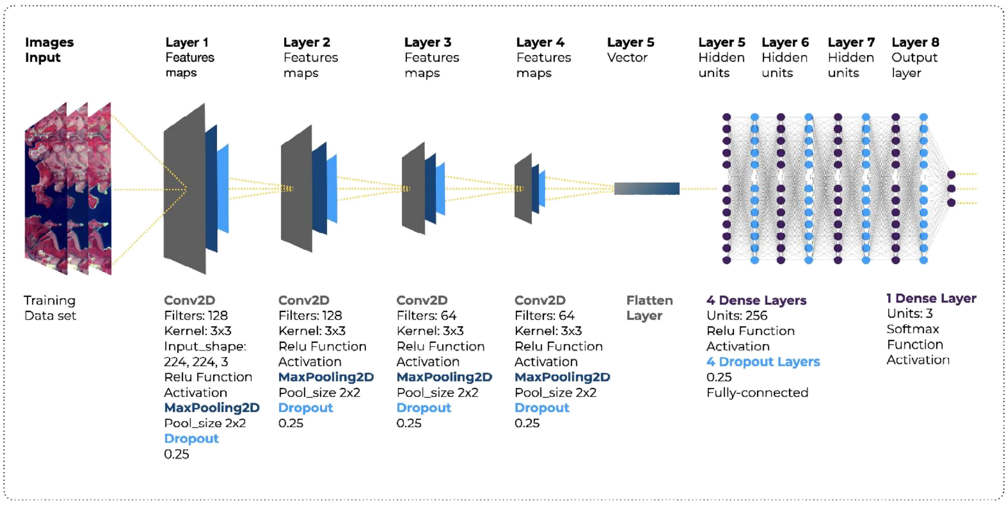

Figure 4. Deep Learning Model Used for Model Training.

layer has 3 neurons, activated by the Softmax function, returning a probabilistic distribution. The model was trained using the Adam optimizer, adjusted for learning rate=0.001, β_1=0.9, β_2=0.999, ϵ=1e-07, which consists of a descending stochastic gradient method based on adaptive moment estimation first and second-order, and the loss function was categorical-crossentropy. For training, we defined 100 epochs, saving the model at the end of the run.

There are three basic differences between this new model and our previous one. About the four convolutional layers, we had 64, 128, 64, 128 filters, respectively, now we have 128, 128, 64, 64 filters. This change was introduced because the two first layers get more pieces of information about the images input and the others get the main features, coming from the previous layers. For the hidden layers, we added more one dense and dropout layers. In the output, we have three neurons since we are dealing with three classes (multi-class problem).

3.3.1. Metric

In order to evaluate the performance of our model, we used the accuracy metric, which divides the value of correct predictions by the total number of samples:

$$accuracy = \frac{\#\text{correct predictions}}{\#\text{samples}} \tag{1}$$

We assessed the accuracy per class and also considering the entire testing set (overall accuracy).

4. Results and discussion

Figure 5 presents the accuracy by each class and the overall accuracy, where the *normal* class images had only 2.94% samples which were incorrectly labelled, and our model presented accuracies of 79.31% and 77.78% for the *low* and *critical* classes, respectively. The overall accuracy regarding the test set is 85.56%. Some classification errors can be observed in Figure 6 where, for each image, we show the true value and the incorrect prediction.

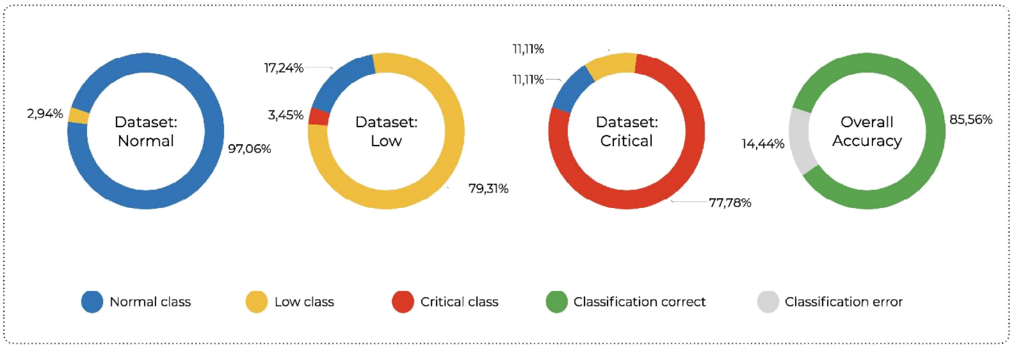

Figure 5. Overall and per class accuracy of the test images.

Analyzing Figure 5, the errors regarding the *normal* class are all related to the *low* class (i.e. the DCNN misclassified a *normal* test image as a *low* test image) while there are five times more errors related to the *normal* class compared to the *critical* class, within the accuracy of the *low* class in isolation. When looking at the errors related to the *critical* class, we see a tie between the two other classes (*normal* and *low*). Therefore, the model presents more difficult to differentiate the *low* and *critical* classes.

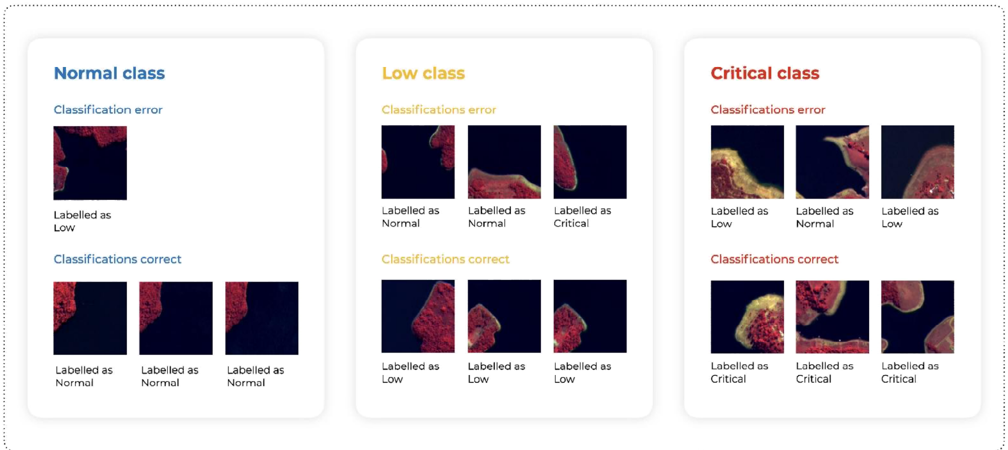

Figure 6. Classification errors by subset of test images.

Figure 6 shows some incorrectly and correctly labelled images for each class. The errors of classification are related to the difference between the training and test sets in terms of spatial resolution since the model was trained with 8 meters of spatial resolution images and tested with images with 2 meters, although the model had labelled correctly the images in their respective class. It is worth noting that a parameter used to assess the correctness of image classification was the water volume report provided by [SABESP 2021].

The revised DCNN is naturally different from the previous one. However, the general conception of our design remains the same. In the previous work, it was obtained an overall accuracy of 87% which is basically the same as our current model (85.56%). Hence, this is an indication of the good generalization capabilities of our approach, since we have a completely different context (water) compared to our previous study (vegetation). Even though we can not underestimate the pre-processing steps that, properly conducted, contribute to the success of our DL method, the network structure, hyper-parameter values seem to be robust for different contexts and satellite images.

5. Conclusion

Precisely classifying the volume of water in dams is an important task especially in locations where droughts are more frequent. In this article, we extended a previous proposed DCNN, used to classify vegetation areas, to classify the water volume of dams in the state of São Paulo, Brazil. Our revised CNN addresses a multi-class classification problem and obtained an overall accuracy of 85.56% considering the images from the Atibainha and Jaguari dams of the Cantareira water supply system to compose the testing dataset. This accuracy is close to that obtained in our previous work [Miranda et al. 2021], indicating good generalization capabilities of our model. We believe these are encouraging results, as the model achieves good performance even if in the training set we have images with less detailed spatial resolutions compared to the testing set.

Nevertheless, this research can be improved regarding the number of samples for training; use of images, such as Sentinel-3, which provide altimetry values of water bodies, as auxiliary information; image pre-processing techniques for detection or segmentation of the edges of water bodies; and more robust adjustments of the hyper-parameters of the DCNN. Also, it is possible to extract time series from the images and make forecasts of periods of the water crisis, and thus collaborate in the development of supply, consumption, and generation strategies for hydroelectric energy. Certainly, DL techniques are potential collaborators for environmental monitoring, concerning the consumption and management of water bodies. In addition, they enrich the techniques used in the area of remote sensings, such as image classification, making it more agile and diligent, especially when processing large volumes of data.

References

Ai, B., Wen, Z., Wang, Z., Wang, R., Su, D., Li, C., and Yang, F. (2020). Convolutional neural network to retrieve water depth in marine shallow water area from remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:2888–2898.

Baek, S.-S., Pyo, J., and Chun, J. A. (2020). Prediction of water level and water quality using a cnn-lstm combined deep learning approach. *Water*, 12(12):3399.

Barino, F. O. and dos Santos, A. B. (2020). Rede neural convolucional 1d aplicada à previsão da vazão no rio madeira. *XXXVIII Simpoósio brasileiro de telecomunicões e processamento de sinais*.

Chen, Y., Fan, R., Yang, X., Wang, J., and Latif, A. (2018). Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning. *Water*, 10(5):585.

Fang, W., Wang, C., Chen, X., Wan, W., Li, H., Zhu, S., Fang, Y., Liu, B., and Hong, Y. (2019). Recognizing global reservoirs from landsat 8 images: A deep learning approach. *IEEE journal of selected topics in applied earth observations and remote sensing*, 12(9):3168–3177.

Fernandes, E., Wildemberg, P., and dos Santos, J. (2020). Water tanks and swimming pools detection in satellite images: Exploiting shallow and deep-based strategies. In *Anais do XVI Workshop de Visão Computacional*, pages 117–122. SBC.

Géron, A. (2019). Hands on with machine learning with sckit-learn e tensorflow: concepts, tools and techniques for building intelligent systems. *Altas Books e Consultoria Eireli*.

Ghassemi, S. and Magli, E. (2019). Convolutional neural networks for on-board cloud screening. *Remote Sensing*, 11(12):1417.

INPE (2015). Estudo internacional avalia causas da crise hidrica no sudeste do brasil. `http://www.inpe.br/urc/noticias/noticia.php?Cod_Noticia=4035`. Accessed: 2021-02-06.

Isikdogan, F., Bovik, A. C., and Passalacqua, P. (2017). Surface water mapping by deep learning. *IEEE journal of selected topics in applied earth observations and remote sensing*, 10(11):4909–4918.

Khryashchev, V., Ivanovsky, L., Pavlov, V., Ostrovskaya, A., and Rubtsov, A. (2018). Comparison of different convolutional neural network architectures for satellite image segmentation. In *2018 23rd conference of open innovations association (FRUCT)*, pages 172–179. IEEE.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., and Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152:166–177.

Miranda, M. S., Santiago Júnior, V. A., Korting, T. S. Leonardi, R., and Freitas Júnior, M. L. (2021). Deep convolutional neural network for classifying satellite images with heterogeneous spatial resolutions. In *Computational Science and Its Applications – ICCSA 2021: International Conference on Computational Science and its Applications, Cagliari, Italy*. Springer, Cham.

Namikawa, L., Castejon, E. F., and Korting, T. S. (2019). Water bodies from rapideye images. *Researchgate*.

Namikawa, L. M., KÃ¶rting, T. S., and Castejon, E. F. (1). Water body extraction from rapideye images: an automated methodology based on hue component of color transformation from rgb to hsv model. *Revista Brasileira de Cartografia*, 68(6).

Pan, F., Xi, X., and Wang, C. (2020). A comparative study of water indices and image classification algorithms for mapping inland surface water bodies using landsat imagery. *Remote Sensing*, 12(10):1611.

SABESP (2021). Dados dos sistemas produtores. *Companhia de Saneamento Básico do Estado de São Paulo*.

SIGRH (2020). Comitê de bacia hidrográfica sorocaba e médio tietê (cbh-smt). *Sistema Integrado de Gerenciamento de Recursos Hídricos do Estado de São Paulo*.

Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., and Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*, 82(12):2635–2670.