ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research

Vol. 10, Issue, 12, pp.42807-42811, December, 2020

https://doi.org/10.37118/ijdr.20650.12.2020

RESEARCH ARTICLE **OPEN ACCESS**

EVALUATING METHODS TO CLASSIFY SUGARCANE PLANTING USING CONVOLUTIONAL NEURAL NETWORK AND RANDOM FOREST ALGORITHMS

Flávio Henrique dos Santos Silva*[1], João Pedro dos Santos Verçosa[2] and Arthur Costa Falcão Tavares[2]

[1]Campus de Engenharias e Ciências Agrárias, Universidade Federal de Alagoas
[2]Campus de Engenharias e Ciências Agrárias, Universidade Federal de Alagoas

ARTICLE INFO

Article History:

Received 14th September, 2020
Received in revised form
06th October, 2020
Accepted 29th November, 2020
Published online 30th December, 2020

Key Words:

Deep Learning;
Automatic Classification;
Remote Sensing.

Corresponding author:
Flávio Henrique dos Santos Silva

ABSTRACT

This work aimed to develop an algorithm based on Convolutional Neural Networks and to compare it to a GIS plugin for automatic classification of sugarcane planting areas in medium spatial resolution images from Remote Sensing (RS) at municipality of Coruripe/AL, Brazil. We used Qgis software to process Land Remote Sensing Satellite (Landsat) images and to calculate sugarcane class areas in 2018, followed by Python, OpenCV, Keras and Tensorflow for developing algorithm to images training and classification, and finally, used Random Forest Classifier (RF) algorithm. Neural network algorithm classified samples obtained from 2018 image and mapped 172,65 km² of Sugarcane class and 71,71 km² of Not Sugarcane class. Já para classificação utilizando as amostras de 1986 + 2018 delimitou 155,95 km² de classe SUGARCANE e 88,41 km² como NOT SUGARCANE. Both Neural Network and Random Forest algorithms presented similar results, mainly in relation to the value of the area classified as Sugarcane and Not Sugarcane. Random Forest method performance for classification proven can validate algorithms developed in convolutional neural network for Remote Sensing data classification.

Citation: *Flávio Henrique dos Santos Silva, João Pedro dos Santos Verçosa and Arthur Costa Falcão Tavares.* 2020. "Evaluating methods to classify sugarcane planting using convolutional neural network and random forest algorithms", *International Journal of Development Research*, 10, (12), 42807-42811.

INTRODUCTION

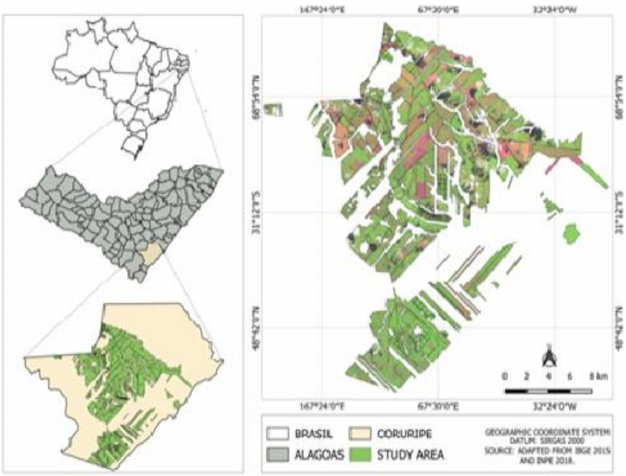
Sugarcane has important function and impact in world economy, due to factors like the adoption of precision agriculture practices, planning and grinding schedules, in addition to the use of Geographic Information Systems (GIS), algorithms and geotechnologies to productivity increase (RAHMAN, 2010; JAMIL, AHMED; SAJJAD, 2018; YAWSON; ADU; OSEI, 2018). Medium spatial resolution satellite images, as Landsat, are efficient tools for crops managing and monitoring, cause allows to obtain data in anywhere, from large to small areas (BENEDETTI, et al., 2018; CUÉ, et al., 2017; MAURO et al., 2017; SHENDRYK et al., 2019). In addition, they have been very useful for generating information on agricultural areas, for monitoring and preventing climate change (USEYA; SHENGBO, 2019).

Random Forest Classifier algorithm has been applied to supervised classification of Landsat satellite images, which can be defined as a set of classifiers based on a decision tree, where each classifier assigns a weight to a class relative to the input vector. (BIRDI & KALE, 2018; NATTESHAN & KUMAR, 2019; BHOSLE & MUSANDE, 2019; YUAN et al., 2020; KUSSUL et al., 2017) Besides, this algorithm has shown high precision and robustness, as theoretically, a set of classifiers performs the classification more efficiently compared to an individual classifier (GALIANO et al., 2012; CASTRO et al., 2017). Comparing with other classification techniques, Kamilaris and Prenafeta-Boldú, (2018) obtained high precision results by using neural network in agricultural management.

It was necessary an organized database hailing from remote sensing images in a hierarchical manner, and respected the steps sequence: selection of raw data, training, validation and classification (YALOVEHA, HLAVCHEVA; PODOROZHNIAK, 2019; GIRSHICK et al., 2016; LAKSHMI & NARESH, 2018; SHOBANA; SUGUNA & YAMUNATHANGAM, 2018). The neural network training requires a great data organization cause its applications include the construction of complex empirical relationships to recover variables from agricultural crops and model emulators. Despite being a technique for accurate data classification, the neural network is limited by its database, which must have great spatial and temporal variability (WEISSA; JACOBB; DUVEILLERC, 2020; ZHANG, Q. et al., 2018; SUN et al., 2019; BHOSLE & MUSANDE, 2020). Thus, the work aimed to develop and use algorithms from Convolutional Neural Network and GIS to classify and quantify sugarcane plantation areas using Landsat satellite images of Coruripe/AL municipality.

STUDY AREA AND MATERIALS

Coruripe is a municipality located on the north coast of Alagoas state, 91 km away from Maceió, state's capital, between the geographical coordinates 10° 08′ 01″ S and 36° 10′ 34″ W. Its altitude varies from 50 to 60m in a territorial area of 897,780 km², has a population of 52,130 inhabitants and demographic density of 56.77 inhabitants per km² (IBGE, 2019). The study area map, was obtained by cutting out the RGB composition with bands 6-5-4 Landsat 8, from the sugarcane planting shapefile (Figure 1).



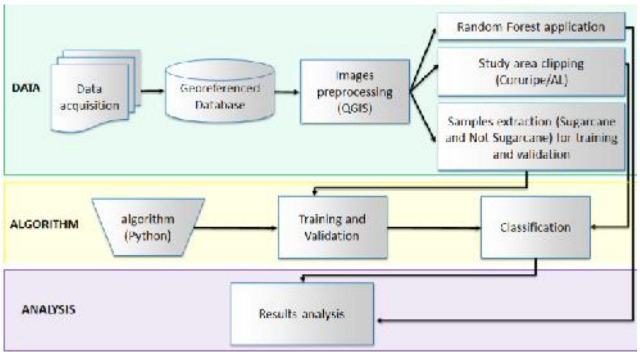
Source: Authors, 2019.

Figure 1. Study area map

The work database was composed of two Landsat satellite images, a Landsat 5 image (sensor Thematic Mapper - TM) of 1986, and a Landsat 8 (sensor Operational Land Imager - OLI) of 2018. Both images were taken in May, with a maximum clouds coverage of 10%, referring to orbit 214, point 67, and were freely downloaded from the National Institute for Space Research (INPE) site. The database also presented shapefile information (.SHP) about municipality of Coruripe / AL limit downloaded from Brazilian Institute of Geography and Statistics (IBGE) site and the sugarcane fields limits obtained from Usina Coruripe. Qgis software was used to carry out images digital pre-processing including geometric correction and clipping from the vector files into municipality of Coruripe/AL limit, and then, into sugarcane fields limit of 2018. These cutouts were essential for separate the results obtained by the algorithms for identifying and mapping the Sugarcane and Not Sugarcane areas. Python software and the libraries OpenCV, for computer vision, Keras, for machine learning, and Tensorflow, for the development of neural networks, were used to create an algorithm for image training and classification.

METHODOLOGY

All the methodological steps, contemplating the data processing, algorithm development and results analyses, are presented on the flowchart (Figure 2). Image samples were obtained from Jequiá da Praia/AL city, neighboring the study area, using slicing process set at a proportion of 5x5 pixels for each sample, which is equivalent to a territorial area of 22.5 km². The samples were manually selected and separated in Sugarcane and Not Sugarcane class, with a proportion of 70% of the total samples for training and 30% to validate the classification algorithm. Then, three training sessions of the neural network algorithm were carried out, the first one using only samples from 1986, the second using only samples from 2018 and the third mixing samples from 1986 and 2018. From these sessions, were also generated three models used to classify areas as Sugarcane and Not Sugarcane and map them based on the images of 1986 and 2018. The classification by Random Forest Classifier algorithm used Dzetsaka a QGIS plugin for image classification. Thus, was created a point-type shapefile containing in its attributes table “ID” and “class” columns of the sample. Then, the image of 2018 was classified considering Sugarcane and Not Sugarcane classes. Were collected 50 samples of each class to compose mask layer.



Source: Authors, 2019.

Figure 2. Flowchart of the methodology

RESULTS

After completing Neural Network and Random Forest algorithms training, some parameters were calculated and served to assess the quality of the classifiers, among which, Accuracy that represents precision and accuracy, that is, how closer to 1 or 100% more accurate was the training. In addition, Kappa index was also calculated and then used as a measure of agreement between real and predicted classes, adopted to estimate the accuracy. Table 4 presents the accuracy and Kappa index for the two classifiers, the closer to 1 the better the index. Other important information to validate the classifiers is the Confusion Matrix, which represents the ability of the algorithm to differentiate the classes Sugarcane and Not Sugarcane. Tables 1 and 2 show the confusion matrix of Neural Network training, and table 3 of Random Forest training.

Table 1. Confusion Matrix of Neural Network training (2018)

CLASS		PREDICTED	
		SUGARCANE	NOT SUGARCANE
REAL	SUGARCANE	4025	104
	NOT SUGARCANE	91	251

Source: Authors, 2020

Table 2. Confusion Matrix of Neural Network training (1986 + 2018)

CLASS		PREDICTED	
		SUGARCANE	NOT SUGARCANE
REAL	SUGARCANE	2621	4
	NOT SUGARCANE	26	873

Source: Authors, 2020

Table 3. Confusion Matrix of Random Forest (2018)

CLASS		PREDICTED	
		SUGARCANE	NOT SUGARCANE
REAL	SUGARCANE	53	0
	NOT SUGARCANE	2	52

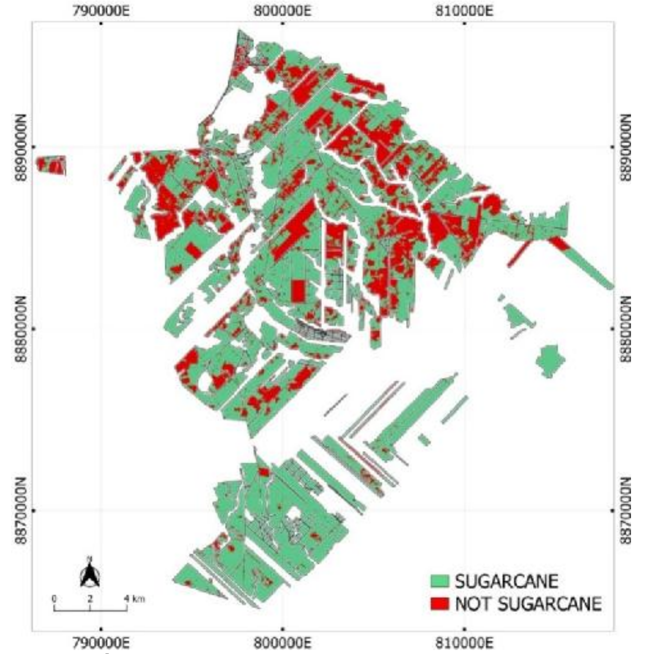
Source: Authors, 2020

Table 4. Accuracy of Neural Network training (2018; 2018+1986) and Random Forest training (2018)

Accuracy		
2018	1986 + 2018	RF
95,64%	99,15%	99,8
KAPPA Index		
0,9998	0,9996	0,9818

Source: Authors, 2020.

Analyzing the tables, it is noted that the main diagonal of the tables presents samples correctly classified during both algorithms training. Then, trained Neural Network algorithm was used to classify the images of study area (Coruripe/AL) as SUGARCANE, green color features, and NOT SUGARCNE, red color features. Figure 3 shows theresulting image after classification by using the model algorithm trained with samples extracted from Landsat image obtained in 2018, while figure 4 shows the resulting image after classification by using the model algorithm trained with samples extracted from Landsat images obtained in 1986 and 2018. Lastly, figure 5 shows resulting image after classification Landsat images obtained in 2018, by using the Random Forest algorithm.



Source: Authors, 2020

Figure 4. Classification, cutting out of 2018 image, using the model generated with data from 1986 + 2018

Source: Authors, 2020

Figure 4. Classification made in cutting out of 2018 image, using the Random Forest algorithm

A visual analysis allows to observe a great difference in quantity and location of the features classified as SUGARCANE and NOT SUGARCANE comparing the images resulting from the classifications with samples only from 2018 and with samples from 1986 + 2018. Now, with respect to the comparison between the images results using only samples from 2018, the only difference was a greater smoothing at features boundaries of the resulting image and classified using the algorithm developed through CNN. Therefore, the results showed an irrelevant variation in the areas values (km²) obtained with samples from 2018 only, based on the CNN (SUGARCANE: 172.65 km² and NOT SUGARCANE: 71,71 km²) and Random Forest

(SUGARCANE: 172.66 km² and NOT SUGARCANE: 71,70 km²) algorithms (Table 5).

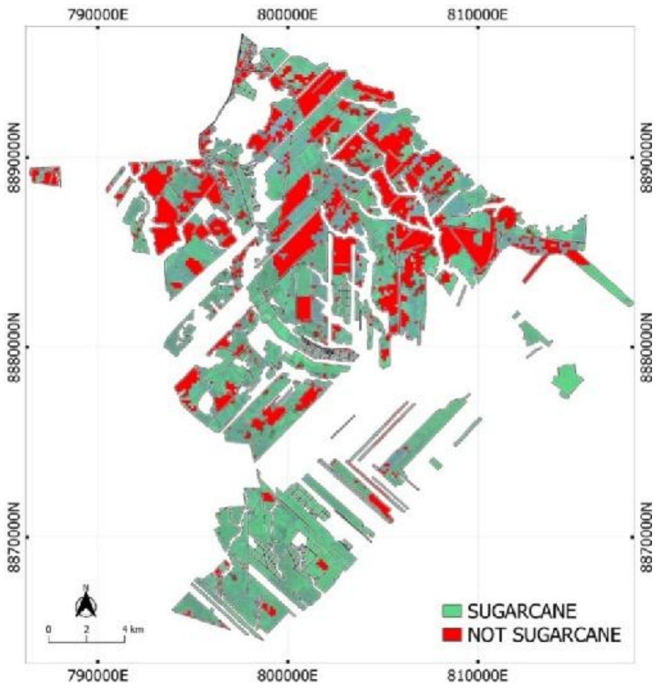
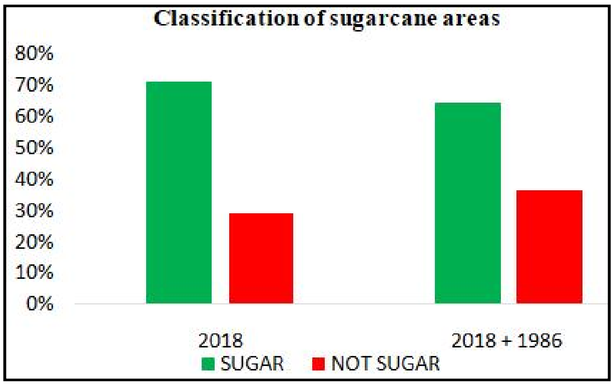


Figure 5. Classification made in cutting out of 2018 image, using the Random Forest algorithm

Table 5. Classified areas (km²) as Sugarcane and Not Sugarcane with Neural Network and Random Forest algorithms for the plantation areas in 2018

Class	Mapped area (km²)		
	2018 (NN)	1986 + 2018 (NN)	2018 (RF)
Sugarcane	172,65	155,95	172,66
Not Sugarcane	71,71	88,41	71,70
Total	244,36	244,36	244,36

NN: Neural Network algorithm; RF: Random Forest algorithm
Source: Authors, 2020



Source: Authors, 2019.

Figure 6. Correct classification percentages (2018 e 1986 + 2018)

Table 5 presents classification quantitative results obtained by the Neural Network and Random Forest algorithms, for sugarcane planting areas at Coruripe/AL, in 2018. The classification based on the model generated with training using only samples from 2018 mapped 172.65 km² as Sugarcane and 71.71 km² as Not Sugarcane class. For classification using the model trained with the 1986 and 2018 samples, it mapped 155.95 km² of Sugarcane and 88.41 km² as Not Sugarcane class. Random Forest algorithm, classified an area of 172.66 km² as Sugarcane and 71.70 km² as Not Sugarcane class. Figure 6 shows that classification algorithm made above 50% correct delimitation for Sugarcane class, being 71% of correctness using the model with samples from 2018, and 64% using the model with samples from 1986 + 2018.

The difference found in the percentage values was considered helpful since there is a 32-year great lapse (1986 and 2018) between the images, besides being obtained by different versions (5 and 8) of the Landsat satellite.

Conclusion

The Convolutional Neural Network algorithm presented good accuracy for classification of medium spatial resolution satellite images. Therefore, its application for multitemporal data analysis in other regions of Alagoas is very promising such for medium resolution satellite images as high spatial resolution aerial images from satellite, old aerial photogrammetric surveys and aerial photos obtained by Air Vehicle unmanned (UAV). The 1986 image was not classified, but the training of the algorithm with its samples presented great accuracy and can be used when the official mill map (1986), still on printed paper, is vectorized and georeferenced.

According to hits of values obtained by the algorithm when classifying and mapping Sugarcane class, it can be said that there could have been greater accuracy if there were specific information from the Mill in relation to the planting areas with freshly harvested, newly planted or in sugarcane stage soil preparation for planting the 2019/2020 harvest. These three cases have characteristics of exposed soil and, therefore, had no samples used to train the algorithm in the identification and delimitation of the Sugarcane class. The areas values (km²) may also present better accuracy, with the improvement of the algorithm in terms of edges smoothing and precision that delimit the polygons of the Sugarcane class. Comparing the results obtained using Neural Network and Random Forest, we concluded that both algorithms presented similar results, mainly in relation to the value of the area classified as Sugarcane and Not Sugarcane. In this and in several studies the Random Forest method performance for classification proven can be used to validate the classifiers developed using deep learning techniques. Thus, the developed neural network algorithm can be a good option for classification of images from Remote Sensing.

Acknowledgment

We appreciate the Federal University of Alagoas (UFAL) for the opportunity and idealization of this study, and the Pro-Rectory of Research and Graduate Studies (PROPEP/UFAL) for scholarship. And finally, the Geoprocessing sector of Sugarcane Mill Coruripe S.A. by making available georeferenced data and Topography team.

REFERENCES

BENEDETTI, P. et al. M3Fusion: A Deep Learning Architecture for Multi-{Scale/Modal/Temporal} satellite data fusion. arXiv preprint arXiv:1803.01945, 2018.

BHOSLE, K., MUSANDE, V. (2019) Evaluation of deep learning Convolutional Neural Network for Crop Classification, International Journal of Recent Technology and Engineering (IJRTE), DOI: 10.35940/ijrte.B2872.078219

BHOSLE, K., MUSANDE, V. (2020) Evaluation of CNN model by comparing with convolutional autoencoder and deep neural network for crop classification on hyperspectral imagery, Geocarto International, DOI: 10.1080/10106049.2020.1740950

BIRDI, K. P., KALE, K. Deep Convolutional Neural Network Models for Land Use and Land Cover Identification using Dataset Created From LISS-IV Satellite Images. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN: 2456 -3307, v. 3, n. 3, p.2123-2134, março-abril-2018. URL do jornal: http://ijsrcseit.com/ CSEIT18356

CASTRO, J. D. B., FEITOZA, R. Q., LAROSA, L. C., DIAZ, P. M. A., SANCHES, I. A. A Comparative Analysis of Deep Learning Techniques for Sub-Tropical Crop Types Recognition from Multitemporal Optical/SAR Image Sequences. 2017 30Th Sibgrapi Conference On Graphics, Patterns And Images (Sibgrapi), [S.L.], p. 1-8, out. 2017. IEEE. http://dx.doi.org/10.1109/sibgrapi. 2017.57.

CUÉ, L. E. et al. A comparative analysis of deep learning techniques for crop type recognition in temperate and tropical regions from multitemporal SAR image sequences. Anais do XXVII CBC, Rio de Janeiro, p. 730-734. 2017.

GALIANO, V. R. et al. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, [s. l.], ano 2012, v. 67, p. 93-104, 2012. Disponível em: https://www.sciencedirect.com/science/article/pii/S092427 1611001304#!. Acesso em: 07 maio. 2020.

GIRSHICK, R., DONAHUE, J., DARRELL, T., MALIK, J. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 38(1):1-1, 142-158. 2016.

IBGE. Instituto Brasileiro de Geografia e Estatística - cidades: Coruripe - AL. Disponível em: https://cidades.ibge.gov.br/ brasil/al/coruripe/panorama Acesso em: 25 out. 2019.

JAMIL, M., AHMED, R., SAJJAD, H. Land suitability assessment for sugarcane cultivation in Bijnor district, India using geographic information system and fuzzy analytical hierarchy process. GeoJournal 83, 595–611. 2018.

KAMILARIS, A., PRENAFETA-BOLDÚ, F. X. A review of the use of convolutional neural networks in agriculture. The Journal of Agricultural Science 156, 312–322. 2018.

KUSSUL, N., LAVRENIUK, M., SKAKUN, S., SHELESTOV, A. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. Ieee Geoscience And Remote Sensing Letters, [S.L.], v. 14, n. 5, p. 778-782, maio 2017. Institute of Electrical and Electronics Engineers (IEEE). http://dx.doi.org/10.1109/ lgrs.2017.2681128.

LAKSHMI, J. V. N., NARESH, G. V. N. A Review on Developing Tech-Agriculture using Deep Learning Methods by Applying UAVs. IJSRCSEIT, V. 3, p.1858-1863. 2018.

MAURO, N., VERGARI, A., BASILE, T. M. A., VENTOLA, F. G., ESPOSITO, F. End-to-end Learning of Deep Spatio-temporal Representations for Satellite Image Time Series Classification. Dc@pkdd/ecml, [s. l], p. 1-8, 2017

NATTESHAN, N. V. S; KUMAR, N. S. Crop Detection and Classification using Remote Sensing Images. International Journal Of Innovative Technology And Exploring Engineering, [S.L.], v. 8, n. 12, p. 1165-1173, 26 dez. 2019. Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication - BEIESP. http://dx.doi.org/10.35940/ijitee.k1318.10812s19.

RAHMAN, E. M. The potential for using remote sensing to quantify stress in and predict yield of sugarcane (Saccharum spp. hybrid). Thesis (Doctor of Philosophy in Environmental Sciences) - Faculty of Science and Agriculture, Pietermaritzburg. 2010.

SHENDRYK, Y., RIST, Y., TICEHURST, C., THORBURN, P. Deep learning for multi-modal classification of cloud, shadow and land cover scenes in PlanetScope and Sentinel-2 imagery. Isprs Journal Of Photogrammetry And Remote Sensing, [S.L.], v. 157, p. 124-136, nov. 2019. Elsevier BV. http://dx.doi.org/ 10.1016/j.isprsjprs.2019.08.018.

SHOBANA, G., SUGUNA, M., YAMUNATHANGAM, D. Identification of Crop Disease by Predictive Analysis in Hadoop Environment. IJPAM, v. 118, n. 18, p. 2805-2809, 2018.

SUN, Y., LUO, J., XIA, L., WU, T., GAO, L., DONG, W., HU, X., HAI, Y. Geo-parcel-based crop classification in very-high-resolution images via hierarchical perception. International Journal Of Remote Sensing, [S.L.], v. 41, n. 4, p. 1603-1624, 8 out. 2019. Informa UK Limited. http://dx.doi.org/10.1080/01431161.2019.1673916.

USEYA, J., SHENGBO, C. Exploring the Potential of Mapping Cropping Patterns on Smallholder Scale Croplands Using Sentinel-1 SAR Data. Chinese Geographical Science, 29(4), 626–639. 2019.

YALOVEHA, V., HLAVCHEVA, D., PODOROZHNIAK, A. Usage of convolutional neural network for multispectral image processing applied to the problem of detecting fire hazardous forest areas. Advanced Information Systems, Vol. 3, No. 1, P. 116–120. 2019.

YAWSON, D. O., ADU, M. O., OSEI, K. N. Spatial assessment of sugarcane (Saccharurnspp. L.) production to feed the Komenda Sugar Factory, Ghana. Heliyon 4.11, e00903–e00903. 2018.

YUAN, Q., SHEN, H., LI, T., LI, Z., LI, S., JIANG, Y., XU, H., TAN, W., YANG, Q., WANG, J. Deep learning in environmental remote sensing: achievements and challenges. Remote Sensing Of Environment, [S.L.], v. 241, p. 1-24, maio 2020. Elsevier BV. http://dx.doi.org/10.1016/j.rse.2020.111716.

WEISS, M., JACOBB, F., DUVEILLERC, G. Remote sensing for agricultural applications: A meta-review. Remote Sensing of Environment, 236. 2020.

ZHANG, Q. et al. Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network. IEEE Transactions on Geoscience and Remote Sensing (TGRS), China, v. 56, n. 8, p. 4274-4288, 14 mar. 2018.

****