Real-world multiclass land cover analysis and classification using a novel lightweight ConvNet

Abraham Sánchez, Alexander Quevedo, Raúl Nanclares, Jorge Martínez,
Juan Pacho, Paola Montoya and E. Ulises Moya*
Coordinación General de Innovación Gubernamental del Estado de Jalisco
*Corresponding author: eduardo.moya@jalisco.gob.mx

Abstract

The understanding of global climate change, agriculture resilience, and deforestation control rely on the timely observations of the Land Use and Land Cover Change (LULCC). Recently, some deep learning (DL) methods have been adapted to make an automatic classification of Land Cover (LC) for one season-homogeneous data. However, most of these DL models can not apply effectively to real-world data. i.e. a large number of classes, multi-seasonal data, diverse climate regions, high imbalance label dataset, and medium-spatial resolution. In this work, we present a novel lightweight (only 89k parameters) Convolution Neural Network (ConvNet) to handle these problems in annual LC classification and analysis for the Jalisco region which is located in central-western Mexico. Our embedded analysis anticipates the limited performance in some classes and gives us the opportunity to group the most similar, as a result, the test accuracy performance increase from 73% to 83%. We hope that this research helps other regional groups with limited data or computational resources to attain the United Nations Sustainable Development Goal (SDG) concerning "Life on Land" and plan effectively the land use, conservation areas, or ecosystem services.

1. Introduction

The diversity in climatic conditions and vegetation pose different obstacles to the annual monitoring land cover using remote sensing [22]. Mexico is considered a megadiverse country [17] due to its location in a transition zone between Nearctic and Neotropic regions making it more difficult for land use classification and monitoring. In recent years, Deep Learning (DL) outpacing the other machine learning techniques remote sensing for the LULC classification [23]. However, there are still big challenges to solve for real-world data, i.e. highly imbalanced and heterogeneous datasets, big-size areas for classification (don't take into account the Minimum Mapping Unit (MMU)), low spatial-resolution images, noisy labels and multi-channel images. In this scenario, it is difficult to handle the limited GPU memory for multiple input channels or use the most common classification architectures such as ResNet, Xception or MobileNet due to the tiny input size (3×3 pixels) required to detect changes on the Minimum Mapping Unit. In this context, we present a novel lightweight (only 89k parameters) Convolution Neural Network (ConvNet) . The numerical results in the test set confirm that this strategy achieves an overall accuracy of 83%.

2. Related work

Many previous works collected in [2, 8, 16, 21] have addressed the problem on LC classification only for 2-4 classes with outstanding results up to 90% overall accuracy. However, is not easy to compare those works with or work due to the following: i) Jalisco data is more diverse, not only with more classes but also, with more diverse types of LC and LU over the region affecting the class separability [22]. ii) We use 30 m ground sampling distance. iii) The base map was semiautomatic-labeled using [15] with an overall accuracy of 89.48%. iv) We use very small input size based on the MMU 3×3 with a large number of channels. v) The multi-seasonal data from Landsat 8. vi) Most LC datasets were created with big homogeneous generic areas [1, 22]. A fair comparison to our data, would be [5], due to their 12 of classes, study area, and use of multi-season images with an overall accuracy of 76 %. According to [13] the accuracy assessment, the tropical evergreen forest was very accurately classified with an accuracy of > 75 %.

3. Study area and data

The study area is Jalisco which is located in central-western Mexico (see Figure 1), with an area of $80,200 km^2$. The climates are temperate, tropical rainy and dry [19] with altitudes from 0 to 4,300 meters above sea level.

We use the land use and land cover (LULC) map labels

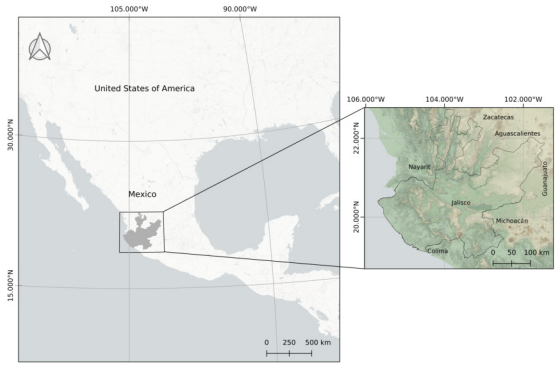

Figure 1. Regional study area: Jalisco, Mexico.

from [3] which is public. This map was developed jointly by the National Forestry Commission of Mexico (CONAFOR) and the Secretaría de Medio Ambiente y Desarrollo Territorial (SEMADET) and Coordinación de Innovación Gubernamental (CGIG). The overall accuracy of the map labels is 89.48 $\pm$ 0.2%. In addition, we used the USGS Landsat 8 Surface Reflectance Tier 1 collection from January to Dec 2016, to generate a cloud-free median composite image, in addition, we compute the normalized Difference Vegetation Index (NDVI) [18] and Normalized Difference Water Index (NDWI) [14]. Moreover, a digital elevation model was included based on the Shuttle Radar Topography Mission (USGS) [4] using the GRASS software [6]. Finally, Landsat 8, relief and LULC map 2016 (Labels) data were stacked into a single dataset with 13 channels: Lansat8 (Band 2 (BLUE), Band 3 (GREEN), Band 4 (RED), Band 5 (NIR near infrared), Band 6 (SWIR Shortwave infrared), (SWIR 2), NDVI = (Band 5 – Band 4) / (Band 5 + Band 4), NDWI = (Band 6 – Band 5) / (Band 6 + Band 5), Digital elevation model, Relief Slope, Relief Aspect, Relief Tangencial Curvature, and Relief Profile Curvature. The name of each 17 LULC map 2016 classes is presented in Table 1. We generate 21,046 patches with ($3 \times 3 \times 13$), divided into 70% for training, 15% for validation and 15% for test.

4. LUCC Lightweight ConvNet

The proposed ConvNet architectures is defined as: (input shape of $3 \times 3 \times 13$), (2DConv,128,1, ReLU), (2DBatch-normalization) (2DConv,64,1, ReLU), (2DBatch-normalization), (Flatten), (Fc,128,ReLU), (Fc,32,ReLU), (Fc,16,ReLu), (Fc,N,sofmax). We designed this model taking into account, the following: i) Classify the land cover using MMU (input size 3×3). ii) The possibility to handle 2D signal with 13 channels (1×1 kernel 2D convolution). iii) Low number of weights to reduce the inference and training time. To solve these requirements we combine our experience with expert-knowledge, tests and image analysis to select the most appropriate architecture. The 1×1 convolution is used as channel-wise pooling to promote learning across channels such as [7, 12, 20]. The Batch-normalization layer was added [10] to reduce the *internal covariate shift* problem, in fact, we expect to have this problem due to the study area having small unevenly distributed examples across the territory. We decide to add three dense (fully connected = Fc) layers [128,32,16] with ReLU, each one having a Gaussian dropout (30%) to reduce overfitting [11].

5. Experimental setup

We conducted three trainings with the proposed architecture: i) a base-line training using a dataset with 17 balanced-classes, presented at Table 1 (Column Class), ii) coarse-grain training, due to embedded analysis, showed that eighth classes could be grouped into four, resulting in 13 classes iii) fine-grain training for binary classification in each group (with similar classes). The hyperparameters values in all the training are: learning rate= 0.0001, epochs= 150 and batch size= 32. In addition, we used 0-90^o random rotations, horizontal and vertical flips as data augmentation. The embedded analysis is based on a *embedded layer* replacing the last two layers in our CNN architecture for a embedded layer consisting of an output dimension of 17 and an L2 normalization. Once the model is trained, we apply a t-distributed stochastic neighbor embedding (t-SNE) [9].

6. Results and analysis

The base-line test-loss and test-accuracy (using 17 classes) are 0.81 and 0.73 respectively. The details of the base-line classification performance per class are presented in Table 1. We observe that class-ID 32 (water) has the best classification performance with and F1-score of 0.96. In contrast, the classes with IDs 2,3,12,15,28,29,34 and 35 has the worst classification results with f1-score ranging from 0.35 to 0.75. The embedding analysis is presented at Figure 2. In this 2D representation, we can observe the presence of clusters formed by the classes [(index = 1: class ID 2 Coniferus forest),(index 3 : class ID 3 Oak forest)], [(index 8 : class ID 34 cultivated grassland), (index 10 : class ID 12 Tropical dry forest)] , [(index 13: class ID 29 rain feed agriculture),(index 14 : class ID 35 cropland irrigated)], [(index 6, class ID 15 crassicaule scrub) (index 9: class ID 28 natural grassland)]. In consequence, we decided to create four new groups: g1=(ID 2, ID 3), g2=(ID 34, ID12), g3=(ID 29, ID 35), and g4=(ID 15, ID 28).

The results of the coarse-grain classification with 13 classes (using the same ConvNet) shows lower test-loss value, 0.53, and higher test-accuracy, 0.83 in comparison to the base-line classification with the 17 classes. The coarse-grain classification details are presented at Table 2. Note that the four groups increase their performance substantially. The fine-grain classification details over the four

P	R	F1	S	ID	Class
0.98	0.93	0.96	207	32	Water
0.60	0.46	0.52	185	2	Coniferous forest
0.89	0.95	0.92	210	1	Upland coniferous forest
0.42	0.30	0.35	180	3	Oak forest and riparian forest
0.69	0.81	0.75	196	7	Cloud forest and low evergreen forest
0.93	0.99	0.96	169	9	Mangrove and peten
0.60	0.85	0.71	183	15	Crassicaule schrub
0.89	0.76	0.82	192	5	Mezquital and submontane shrub
0.45	0.48	0.46	168	34	Cultivated and induced grasslands
0.43	0.36	0.39	185	28	Natural grasslands
0.59	0.58	0.58	183	12	Tropical dry forest
0.74	0.94	0.83	206	13	Tropical semideciduous forest
0.74	0.69	0.72	183	31	Bare land
0.60	0.47	0.52	191	29	Rain fed agriculture
0.76	0.74	0.75	170	35	Cropland irrigated
0.79	0.84	0.81	177	30	Urban areas
0.80	0.88	0.84	172	26	Hydrophilic halophilic vegetation

Table 1. Base-line classification report over the 17 classes. Where P = precision, R = recall , F1 = F1 score S = support.

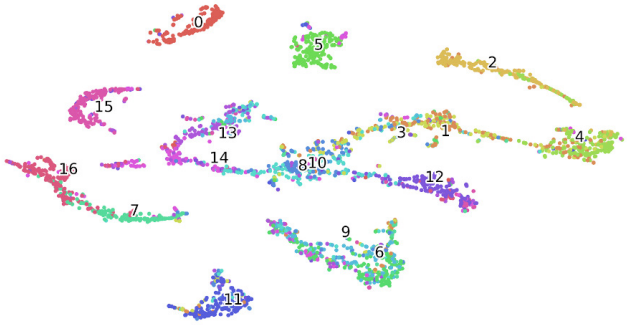

Figure 2. Embedding representation . The closer groups are g1=[(index = 1: class ID =2 Coniferus forest)(index 3 : ID 3 Oak forest)], g2 = [(index 8 : ID 34 cultivated grassland) (index 10 : ID 12 Tropical dry forest)] , g3=[(index 13: ID 29 rain feed agriculture)(index 14 : ID 35 cropland irrigated)], g4=[(index 6, ID 15 crassicaule scrub) (index 9: ID 28 natural grassland)].

groups are presented in Table 3. We observe from this table that the g1 group (2: Coniferous forest, 3:Oak forest) has the lowest classification performance. Figure 3 presents the ground truth map and the prediction in the Guadalajara urban. It is possible to note that the urban area (pink) is relatively well classified. It is possible to see some *salt and paper* artefacts in the natural areas, this artefacts are mainly related with the mask size.

7. Conclusions and Future work

In this paper we have presented a novel lightweight ConVnet to classify the annual LULC images combining real-world open data sources: Lansat8 (multi-seasonal and medium spatial resolution) cloud-free, realief, and 2016 Jalisco's LULC map. In contrast, to most common transfer learning works, the proposed ConvNet can handle 13 channels, and due to 1×1 kernel-size convolution it is possible to take into account the relation over the channels. In addition, this network is designed to work with tiny input size 3×3 conrresponding to the minimum mapeable unit. After an embedded analysis, we grouped the similar classes and increase the test-accuracy from 73% to 83%. The open challenges are the salt and paper noise in the prediction and the lower classification performance with the groped classes. Our research would supports the decision-makers providing evidence, for example, of deforestation or forest degradation.

P	R	F1	S	ID	Class
0.97	0.97	0.97	152	32	Water
0.83	0.98	0.90	199	1	Upland coniferous forest
0.85	0.81	0.83	192	7	Cloud forest and low evergreen forest
0.74	0.70	0.72	184	g1	Coniferus forest and Oak forest
0.64	0.52	0.57	170	g2	Cult. grassland and Trop. dry forest
0.83	0.62	0.71	186	g3	Rain agriculture and Crop. irrigated
0.70	0.86	0.77	184	g4	Crassicaule scrub and Nat. grassland
0.97	1.00	0.98	190	9	Mangrove and peten
0.87	0.87	0.87	193	5	Mezquital and submontane shrub
0.83	0.95	0.88	202	13	Tropical semideciduous forest
0.82	0.73	0.77	173	31	Bare land
0.83	0.87	0.85	190	30	Urban areas
0.89	0.88	0.88	200	26	Hydrophilic halophilic vegetation

Table 2. Classification report with grouped classes. Where P = precision, R = recall , F1 = F1 score and S = support.

P	R	F1	S	ID	Class
0.72	0.64	0.68	184	2	Coniferous forest
0.68	0.76	0.72	188	3	Oak forest
0.83	0.87	0.85	190	34	Cultivated grassland
0.86	0.82	0.84	182	12	Tropical dry forest
0.88	0.82	0.85	182	29	Rain feed agriculture
0.84	0.89	0.86	190	35	Cropland irrigated
0.84	0.75	0.79	194	15	Crassicaule scrub
0.76	0.84	0.80	178	28	Natural grassland

Table 3. Fine-grain classification report per group. Where P = precision, R = recall , F1 = F1 score S = support.

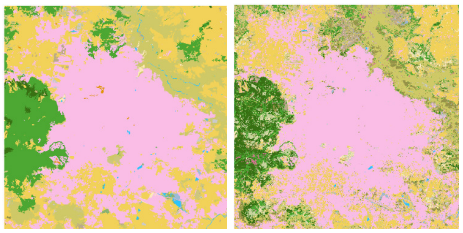

Figure 3. Left: Ground truth map of the urban area of Guadalajara (pink). Right: prediction over the same area.

References

[1] Felix Abramovich and Marianna Pensky. Classification with many classes: challenges and pluses. *Journal of Multivariate Analysis*, 174:104536, 2019. 1

[2] Hamed Alemohammad and Kevin Booth. Landcovernet: A global benchmark land cover classification training dataset. *arXiv preprint arXiv:2012.03111*, 2020. 1

[3] CONAFOR and SEMADET. Mapa de Cobertura del Suelo del Estado de Jalisco al año base 2016 [Vector]. Escala 1:75,000. Versión 1.3. México, 2020. 2

[4] Tom G Farr, Paul A Rosen, Edward Caro, Robert Crippen, Riley Duren, Scott Hensley, Michael Kobrick, Mimi Paller, Ernesto Rodriguez, Ladislav Roth, et al. The shuttle radar topography mission. *Reviews of geophysics*, 45(2), 2007. 2

[5] Steffen Gebhardt, Thilo Wehrmann, Miguel Angel Muñoz Ruiz, Pedro Maeda, Jesse Bishop, Matthias Schramm, Rene Kopeinig, Oliver Cartus, Josef Kellndorfer, Rainer Ressl, et al. Mad-mex: Automatic wall-to-wall land cover monitoring for the mexican redd-mrv program using all landsat data. *Remote Sensing*, 6(5):3923–3943, 2014. 1

[6] GRASS Development Team. *Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.2*. Open Source Geospatial Foundation, 2017. 2

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. 2

[8] Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2217–2226, 2019. 1

[9] Geoffrey Hinton and Sam T Roweis. Stochastic neighbor embedding. In *NIPS*, volume 15, pages 833–840. Citeseer, 2002. 2

[10] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167, 2015. 2

[11] Alex Labach, Hojjat Salehinejad, and Shahrokh Valaee. Survey of dropout methods for deep neural networks. *arXiv preprint arXiv:1904.13310*, 2019. 2

[12] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. *arXiv preprint arXiv:1312.4400*, 2013. 2

[13] Jean-François Mas, Stéphane Couturier, Jaime Paneque-Gálvez, Margaret Skutsch, Azucena Pérez-Vega, Miguel Angel Castillo-Santiago, and Gerardo Bocco. Comment on gebhardt et al. mad-mex: automatic wall-to-wall land cover monitoring for the mexican redd-mrv program using all landsat data. remote sens. 2014, 6, 3923–3943. *Remote Sensing*, 8(7):533, 2016. 1

[14] Stuart K McFeeters. The use of the normalized difference water index (ndwi) in the delineation of open water features. *International journal of remote sensing*, 17(7):1425–1432, 1996. 2

[15] Pontus Olofsson, Giles M Foody, Martin Herold, Stephen V Stehman, Curtis E Woodcock, and Michael A Wulder. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148:42–57, 2014. 1

[16] Darius Phiri and Justin Morgenroth. Developments in landsat land cover classification methods: A review. *Remote Sensing*, 9(9):967, 2017. 1

[17] Jorge Enrique Ramírez Albores, Ernesto Ivan Badano, Joel David Flores Rivas, José Luis Flores Flores, and Laura Yáñez Espinosa. Scientific literature on invasive alien species in a megadiverse country: advances and challenges in mexico. 2019. 1

[18] John W Rouse Jr, R Hect Haas, JA Schell, and DW Deering. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Technical report, 1973. 2

[19] Hermes Ulises Ramírez Sánchez, Ángel Reinaldo Meulenert Peña, and José Antonio Gómez Reyna. Actualización del atlas bioclimático del estado de jalisco. *Investigaciones Geográficas, Boletín del Instituto de Geografía*, 2013(82):66–92, 2013. 1

[20] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions, 2014. 2

[21] Swapan Talukdar, Pankaj Singha, Susanta Mahato, Shahfahad, Swades Pal, Yuei-An Liou, and Atiqur Rahman. Land-use land-cover classification by machine learning classifiers for satellite observations—a review. *Remote Sensing*, 12(7), 2020. 1

[22] Mirela G Tulbure, Patrick Hostert, Tobias Kuemmerle, and Mark Broich. Regional matters: On the usefulness of regional land-cover datasets in times of global change. *Remote Sensing in Ecology and Conservation*, 2021. 1

[23] Ava Vali, Sara Comai, and Matteo Matteucci. Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review. *Remote Sensing*, 12(15):2495, 2020. 1