IMPROVING LULC CLASSIFICATION FROM SATELLITE IMAGERY USING DEEP LEARNING - EUROSAT DATASET

H.Yassine[1]*, K. Tout[1], and M. Jaber[2]

[1]Faculty of Sciences, Lebanese University, [2]Geoconsult-International, Lebanon

KEY WORDS: Deep learning, CNN, Remote Sensing, LULC, EuroSat, Satellite imagery, Sentinel-2, Calculated indices

ABSTRACT:

Machine learning (ML) has proven useful for a very large number of applications in several domains. It has realized a remarkable growth in remote-sensing image analysis over the past few years. Deep Learning (DL) a subset of machine learning were applied in this work to achieve a better classification of Land Use Land Cover (LULC) in satellite imagery using Convolutional Neural Networks (CNNs). EuroSAT benchmarking data set is used as training data set which uses Sentinel-2 satellite images. Sentinel-2 provides images with 13 spectral feature bands, but surprisingly little attention has been paid to these features in deep learning models. The majority of applications focused only on using RGB due to high availability of the RGB models in computer vision. While RGB gives an accuracy of 96.83% using CNN, we are presenting two approaches to improve the classification performance of Sentinel-2 images. In the first approach, features are extracted from 13 spectral feature bands of Sentinel-2 instead of RGB which leads to accuracy of 98.78%. In the second approach features are extracted from 13 spectral bands of Sentinel-2 in addition to calculated indices used in LULC like Blue Ratio (BR), Vegetation index based on Red Edge (VIRE) and Normalized Near Infrared (NNIR), etc. which gives a better accuracy of 99.58%.

1. INTRODUCTION

Remote sensing plays a significant role in the modern world. It supports and assists the society in several fields not limited to agriculture, environmental monitoring, geology, hydrology, and LULC. Remote sensing facilitates the country's development as to classify and monitor land use, as well as to detect problems in dangerous, or inaccessible areas.

Satellite technology can generate regular updates on urban areas, desertification, agriculture land monitoring, crop area estimation, soil mapping and monitoring, water resources monitoring, identification of the characteristics of soil, water, crop, etc.

This study uses the Sentinel-2 images for analysis because these satellite data are free to use, easy to obtain, enough revisit time, and are capable of supporting LULC analysis.

In this paper, we aim to build a model able to classify and indicate the land use and cover situation using satellite imagery. This work looks into improving a LULC classification method using remote sensing technology. In addition, it explores new proposed methods for supervised LULC classifications using deep learning methods, specifically CNN.

This research proposes two approaches to improve performance.

The first approach investigates the us[1]e of all 13 spectral bands which have shown success in other image classification models. While the second approach involves the use of LULC calculated indices. These two approaches can be formulated in the following research sub-questions:

1. Can using the 13 bands of Sentinel-2 spectral features improve the LULC classification performance?

2. Can involving the LULC calculated indices into the 13 bands of Sentinel-2 spectral features improve the LULC classification performance?

The overall objective is to evaluate the new model, which suggests to use the major 13 bands of Sentinel-2 in addition to LULC calculated indices.

(Chong, 2020) implemented increasingly complex deep learning models to identify LULC classifications on the EuroSAT dataset [16].

The best overall model uses all 13 bands and a 50% training set with 4 convolution-max-pooling layer pairs before a dropout layer and a dense layer. Indeed, training data has been augmented through random shearing, rotating, and flipping. It was able to accurately predict the classification for 94.9% of the testing set.

While his best RGB model based on VGG16 with image augmentation through random shearing, rotating, and flipping for the training dataset, this model accurately classified 94.5% of the testing set images.

Knowledge distillation (KD) is one kind of teacher-student training (TST) method first defined in (Hinton et al., 2015) [1], in which they distill knowledge from an ensemble of models into a single smaller model via high-temperature softmax training.

(Chen et al., 2018) introduced the KD into remote sensing scene classification for the first time to improve the performance of small and shallow network models [15]. They performed experiments on several public datasets including EuroSat, and make quantitative analysis to verify the effectiveness of KD. (Chen et al., 2018) were able to achieve a total accuracy of 94.74% in their proposed model.

---

* Corresponding author

(Sonune, 2020) tried to fine-tune 3 models using RGB bands [17]. The first model is VGG_19 which achieved a classification accuracy of 97.66%. While the second model ResNet_50 model had a classification accuracy of 94.25%, and for the RandomForest model classification the accuracy was 61.46%.

(Helber et al., 2019), the creators of the novel dataset EuroSAT [12], experimented two approaches. First they compared GoogLeNet and ResNet-50 models using RGB bands and achieved consecutively an accuracy of 98.18%, and 98.57%. In the second approach, they tried to evaluate different band combinations, RGB, CI, and SWIR and achieved the following accuracies respectively 98.57%, 98.30%, and 97.05%.

(Li et al., 2020), proposed the DDRL-AM method for remote sensing scene classification [18]. They addressed the problem of class ambiguity by learning more discriminative features. The approach involves two main tasks:

> (1) Building a two-stream architecture to fuse attention map semantic feature with original image semantic feature;
> (2) Training DDRL-AM that is coupled with a center loss to obtain discriminative feature for remote sensing images.

Extensive experiments were conducted on EuroSAT Dataset and obtained a classification accuracy of 98.74%.

Table.1 summarizes the previous models architectures and the accuracies gained from benchmarking the EuroSAT dataset.

Authors	Model	Bands	Accuracy
(Chen et al., 2018)	Knowledge distillation	RGB	94.74%
(Chong, 2020)	4-convolution max-pooling layer	All 13 sentinel bands	94.9%
(Chong, 2020)	VGG_16	RGB	94.5%
(Sonune, 2020)	VGG_19	RGB	97.66%
(Sonune, 2020)	ResNet_50	RGB	94.25%
(Sonune, 2020)	RandomForest	RGB	61.46%
(Helber et al., 2019)	GoogLeNet	RGB	98.18%
(Helber et al., 2019)	ResNet-50	RGB	98.57%
(Helber et al., 2019)	ResNet-50	CI	98.30%
(Helber et al., 2019)	ResNet-50	SWIR	97.05%
(Li et al., 2020)	DDRL-AM	RGB	98.74%

Table 1. Models architectures and accuracies gained from benchmarking the EuroSAT dataset

2. DATA AND METHODOLOGY

2.1 Data

EuroSAT dataset [12], which this research depends on, is based on Sentinel-2 satellite images covering 13 spectral bands and consisting out of 10 classes within a total of 27,000 labelled and geo-referenced images.

Sentinel-2A is one satellite in the two-satellite constellation of the identical land monitoring satellites Sentinel-2A and Sentinel-2B. The satellites were successfully launched in June 2015 (Sentinel-2A) and March 2017 (Sentinel-2B). Both sun-synchronous satellites capture the global Earth's land surface with a Multispectral Imager (MSI) covering the 13 different spectral bands listed in Table 1.

The three bands B01, B09 and B10 are intended to be used for the correction of atmospheric effects (e.g. aerosols, cirrus or water vapor). The remaining bands are primarily intended to identify and monitor land use and land cover classes. In addition to mainland, large islands as well as inland and coastal waters are covered by these two satellites.

Each satellite will deliver imagery for at least 7 years with a spatial resolution of up to 10 meters per pixel. Both satellites carry fuel for up to 12 years of operation which allows for an extension of the operation. The two-satellite constellation generates a coverage of almost the entire Earth's land surface about every five days, i.e. the satellites capture each point in the covered area about every five days. This short repeat cycle as well as the future availability of the Sentinel satellites allows a continuous monitoring of the Earth's land surface for the next 20 - 30 years. Most importantly, the data is openly and freely accessible and can be used for any application (commercial or non-commercial use).

Figure 1. Sample EuroSAT images with the labelled class

Looking at the preview of the different classes, we can see some similarities and stark differences between the classes (Fig. 1).

Urban environments such as Highway, Residential and Industrial images all contain structures and some roadways. AnnualCrops and PermanentCrops both feature agricultural land cover, with straight lines delineating different crop fields. Finally, HerbaceaousVegetation, Pasture, and Forests feature natural land cover. Rivers also could be categorized as natural land cover as well, but may be easier to distinguish from the other natural classes.

If we consider the content of each image, we might be able to estimate which classes might be confused for each other. For example, an image of a river might be mistaken for a highway, or an image of a highway junction with surrounding buildings could be mistaken for an Industrial site. We'll have to train a classifier powerful enough to differentiate these nuances.

2.2 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a type of Neural Networks [10] which became with the impressive results on image classification challenges [5], [7], [19], the state of-the-art image classification method in computer vision and machine learning.

To classify remotely sensed images, various different feature extraction and classification methods (e.g., Random Forests) were evaluated on the introduced datasets. (Yang et al., 2010) evaluated Bag-of-Visual-Words (BoVW) and spatial extension approaches on the UCM dataset [8]. (Basu et al., 2015) analyzed deep belief networks, basic CNNs and stacked denoising autoencoders on the SAT-6 dataset [14]. (Basu et al., 2015) also presented an own framework for the land cover classes introduced in the SAT-6 dataset.

The framework extracts features from the input images, normalizes the extracted features, and uses the normalized features as input to a deep belief network. Besides low-level color descriptors, (Penatti et al., 2015) also evaluated deep CNNs on the UCM and BCS dataset [4]. In addition to deep CNNs, Castelluccio et al. (2015) intensively evaluated various machine learning methods (e.g., Bag-of-Visual-Words, spatial pyramid match kernels) for the classification of the UCM and BCS dataset [2].

In the context of deep learning, the used deep CNNs have been trained from scratch or fine-tuned by using a pretrained network [2], [3]. The networks were mainly pretrained on the ILSVRC-2012 image classification challenge dataset [11]. Even though these pretrained networks were trained on images from a totally different domain, the features generalized well. Therefore, the pretrained networks proved to be suitable for the classification of remotely sensed images [9]. The previous works extensively evaluated proposed machine learning methods and concluded that that deep CNNs outperform non-deep learning approaches on the considered datasets [2], [6], [11], [9].

2.3 Satellite Remote Sensing

Remote sensing is the process of identifying the physical characteristics of an object or an area by measuring its reflected and emitted radiation at a distance.

Each material has a unique spectral signature which becomes the basic criterion for material identification. The graph below in Fig. 2 depicts the typical reflectance characteristics of water, vegetation, and soil. The typical curves of water, vegetation and soils are close in the visible region and quite different in the infrared spectrum.
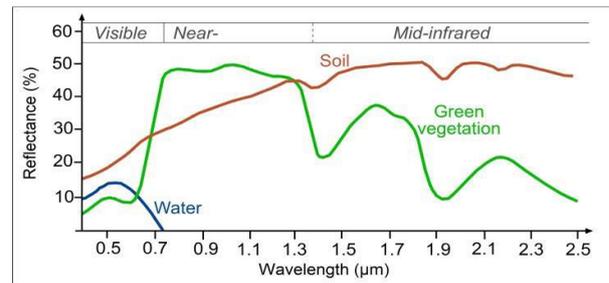

Figure 2. Material reflectance at different wavelengths

The basic information for developing classification models is discriminative reflectance characteristics of materials at different wavelengths [13]. The spectral reflectance characteristics of Sentinel -2 image data for different crops are presented in Fig. 3. In the visible region, the spectral reflectance values of different crops are close but in the infrared region, crops are discriminable.
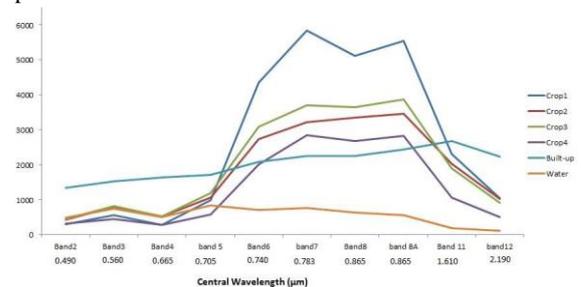

Figure 3. Characteristics of different crops in different bands

2.4 Dataset Preparation

As mentioned before, the EuroSAT dataset is split into 10 classes of land cover. Each class varies in size (Fig. 4), thus we have to split the data into training, validation, and testing sets respectively 70%, 20%, and 10% per class.
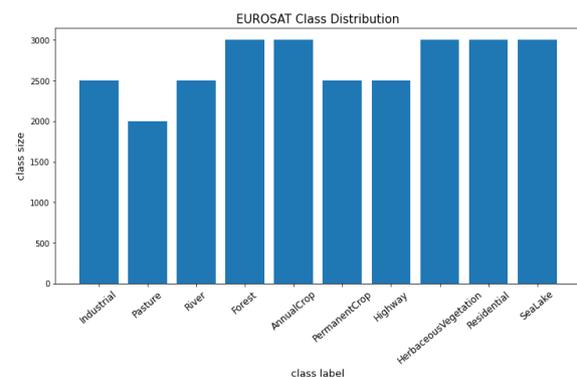

Figure 4. EuroSAT Class Distribution

In the ImageDataGenerator, the batch size is 64. For the training dataset, we applied rotation, horizontal, and vertical flip for the

images to generate and handle more data. Moreover, we calculated the mean and standard deviation(std) for the bands and indices as shown in Table 2. These values were used to normalize the inputs as recommended in deep learning models by subtracting the mean per channel and then divide by the std value.

In the RGB model, we read the bands 2, 3, and 4 respectively, while in All 13 bands model, we read all the 13 bands, while in the new proposed model, we read all the 13 bands and calculated the following indices (MSI, NDSI, NDWI, NDVI, GNDVI, BNDVI, NDRE, BareSoil, NDCI, NBR, DSWI, CVI) and appended them to the data.

Number	Bands and Indices	Mean	Std
1	Aerosols	1353.439	65.571
2	Blue	1117.253	154.376
3	Green	1042.253	188.262
4	Red	947.128	278.926
5	Red edge 1	1199.404	228.244
6	Red edge 2	2002.936	355.633
7	Red edge 3	2373.488	454.901
8	NIR	2300.642	530.549
9	Red edge 4	732.159	98.716
10	Water vapor	12.113	1.187
11	Cirrus	1820.932	378.496
12	SWIR 1	1119.173	304.439
13	SWIR 2	2598.82	501.747
14	MSI	0.348	0.136
15	NDSI	0.282	0.115
16	NDWI	0.245	0.122
17	NDVI	0.251	0.107
18	GNDVI	-0.095	0.123
19	BNDVI	0.082	0.104
20	NDRE	-0.282	0.115
21	BareSoil	-0.731	0.251
22	NDCI	0.138	0.074
23	NBR	-0.055	0.046
24	DSWI	1.283	0.172
25	CVI	4.322	116.297

Table 2. Mean and Standard Deviation for the training images of the EuroSat Dataset Bands and Calculated Indices.

2.5 Calculated Indices

Today many different remote sensing indices exists. It was successfully used for the identification of drifting sand areas, vegetation cover, loss of wetlands, and urban land use mapping. The mentioned calculated indices used in the new model are mainly related to LULC classification, we present an overview for these indices as followed:

2.5.1 MSI: Moisture index

$MSI = (B8A - B11) / (B8A + B11)$

The index is inverted relative to the other water vegetation indices; higher values indicate greater water stress and less water content. The values of this index range from 0 to more than 3. The common range for green vegetation is 0.4 to 2.

2.5.2 NDSI: Normalized difference snow index

Normalized difference snow index is a ratio of two bands: one in the VIR (Band 3) and one in the SWIR (Band 11). Values above 0.42 are usually snow.

$NDSI = (B03 - B11) / (B03 + B11)$

2.5.3 NDWI: Normalized difference water index

NDWI was proposed by (McFeeters 1996). It is used to monitor changes related to water content in water bodies. As water bodies strongly absorb light in visible to infrared electromagnetic spectrum, NDWI uses green and near infrared bands to highlight water bodies. It is sensitive to built-up land and can result in over-estimation of water bodies.

$NDWI = (B03 - B08) / (B03 + B08)$

Index values greater than 0.5 usually correspond to water bodies. Vegetation usually corresponds to much smaller values and built-up areas to values between zero and 0.2.

2.5.4 NDVI: Normalized difference vegetation index

NDVI is simple, but effective index for quantifying green vegetation. It normalizes green leaf scattering in Near Infra-red wavelengths with chlorophyll absorption in red wavelengths.

$NDVI = (B08 - B04) / (B08 + B04)$

The value range of the NDVI is -1 to 1. Negative values of NDVI (values approaching -1) correspond to water. Values close to zero (-0.1 to 0.1) generally correspond to barren areas of rock, sand, or snow. Low, positive values represent shrub and grassland (approximately 0.2 to 0.4), while high values indicate temperate and tropical rainforests (values approaching 1). It is a good proxy for live green vegetation.

2.5.5 GNDVI: Green normalized difference vegetation index

GNDVI is similar to NDVI, but it uses visible green instead of visible red and near infrared. Useful for measuring rates of photosynthesis and monitoring the plant stress.

$GNDVI = (B08 - B03) / (B08 + B03)$

2.5.6 BNDVI: Blue Normalized Difference Vegetation Index

BNDVI is an index similar to NDVI, but without red channel availability it uses the visible blue, this index is useful for areas sensitive to chlorophyll content.

$BNDVI = (B08 - B02) / (B08 + B02)$

2.5.7 NDRE: Normalized Difference Red Edge

This index formulated with NIR and red edge band and it is useful for areas sensitive to chlorophyll content in leaves against soil background effects.

$NDVI\_RedEdge = (B08 - B05) / (B08 + B05)$

2.5.8 BareSoil

This index identifies all observations in which the feature of interest (FOI) is bare soil which is a result of ploughing or covered with non-photosynthetic vegetation as a consequence of harvest or vegetation drying up on the field.

$BareSoil = 2.5 * ((B11 + B04) - (B08 + B02)) / ((B11 + B04) - (B08 + B02))$

2.5.9 NDCI: Normalized difference chlorophyll index

NDCI is an index that aims to predict the plant chlorophyll content which plays a critical role in plant growth and helps predicting the plant type in addition to its health condition. It is calculated using the red spectral band B04 with the red edge spectral band B05.

$NDCI = (B05 - B04) / (B05 + B04)$

2.5.10 NBR: Normalized burn ratio

To detect burned areas, the NBR index is the most appropriate choice. Using bands 8 and 12 it highlights burnt areas in large fire zones greater than 500 acres, with Darker pixels indicate burned areas.

$$NBR = (B08 - B12) / (B08 + B12)$$

To observe burn severity, you may subtract the post-fire NBR image from the pre-fire NBR image.

2.5.11 DSWI: Disease-Water Stress Index

This index is used to classify the water stress diseases and its health condition.

$$DSWI = (B03) / (B04)$$

2.5.12 CVI: Chlorophyll vegetation index

CVI is an index that aims to predict the plant chlorophyll vegetation index. Chlorophyll plays a critical role in plant growth and helps predicting the plant type in addition to its health condition.

$$CVI = (B08 * B04) / (B03)2$$

2.6 Model Architecture

In this research we use **DenseNet201** architecture excluding the fully-connected layer at the top of the network and without loading its defined weights (random initialization) with three inputs: width, height, and number of channels respectively (W, H, C#). Width and height is fixed 64 by 64 while number of channels varies based on the model as specified below:

- RGB Model (64, 64, 3)
- All Bands Model (64, 64, 13)
- All Bands with Calculated Indices Model (64, 64, 22)

Moreover, we added a global spatial average pooling layer using **GlobalAveragePooling2D**, and the **Softmax** activation function is applied to the very last layer in the model since it is useful as the activation for the last layer of a classification network.
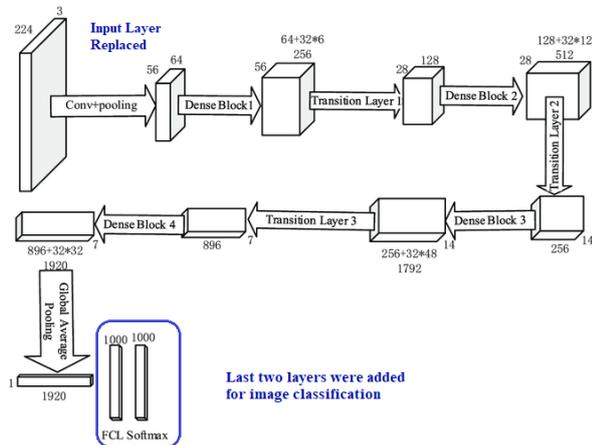

Figure 5. Updated Dense201 Model Architecture

This model was compiled using Adam optimizer, this optimizer is an extension to stochastic gradient descent that has recently seen broader adoption for deep learning applications in computer vision and natural language processing. The compile parameters configured the *'categorical_crossentropy'* as the loss function, and *'categorical_accuracy'* as model metric.

After compiling the model, we run the training on the training datasets using fit function, with two callback functions, *'Checkpointer'* to monitor the accuracies and save the best weights for the model which are saved on Google Cloud storage buckets, and *'EarlyStopping'* to stop the training when a monitored metric has stopped improving, which specified in the training with patience equal to 50 epochs. The number of

epochs specified to train the model is maximum 10000 which is the number of maximum iterations over the entire data provided. Moreover, each epoch trains 1000 batch of samples which is 64, i.e. 64000 image to be trained every epoch.

3. EXPERIMENTAL RESULTS

This section presents the results of the various models including Confusion Matrix, Classification Report, and mislabeled images.

3.1 Models Benchmarking

We are presenting and comparing the results of the new models supported by analyzing the Confusion Matrix, Classification Report, and mislabeled images of each model.

3.1.1 RGB Model

After 60 epochs the training process for the RGB model finished with accuracy 96.83%, with an unstable learning as seen in the accuracy graph (Fig. 6) and the loss graph (Fig. 7).
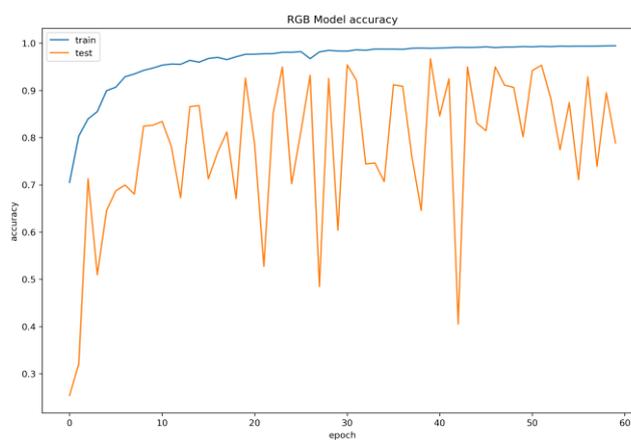

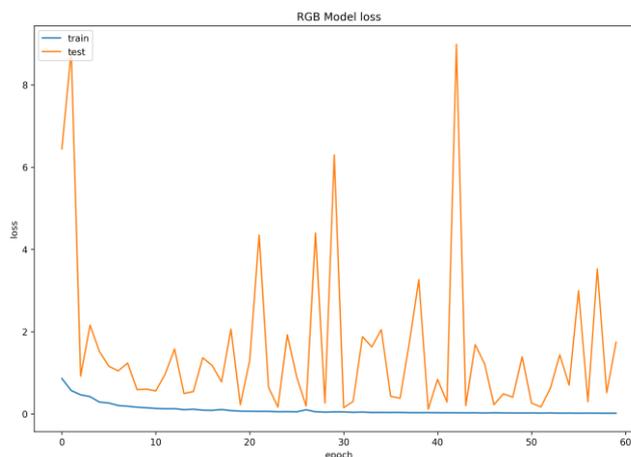**Figure 6.** Training Accuracy graph for the RGB Model

Figure 7. Training Loss graph for the RGB Model.

Although this model is straightforward, it achieves a good accuracy of 96.83%. Below we present the Confusion Matrix (Fig. 8) and Classification Report (Fig. 9) for this model. We can find from both figures that this model perfectly classifies the main classes with minimal wrong classifications, but has limitation in differentiating from similar classes (Table 3)
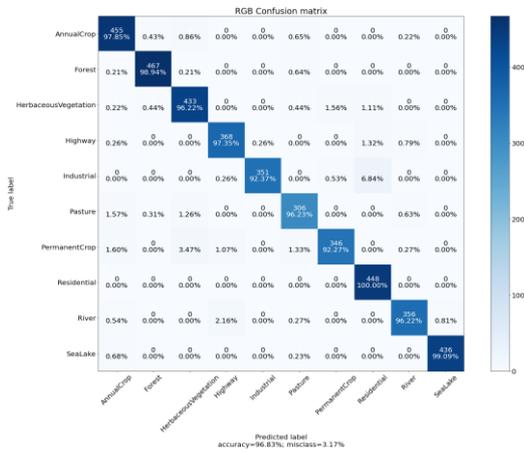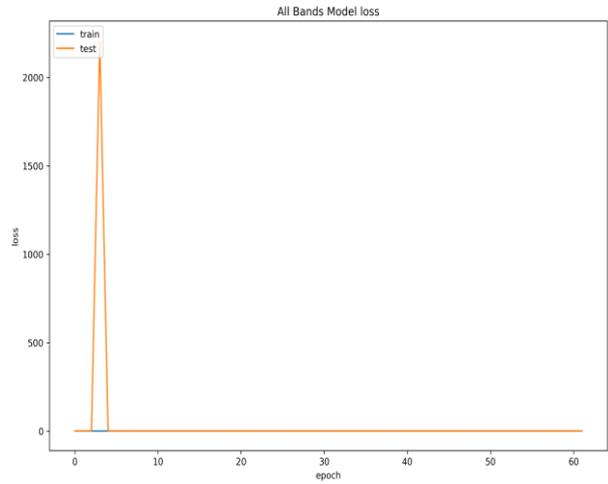
Figure 8. RGB Model Confusion Matrix

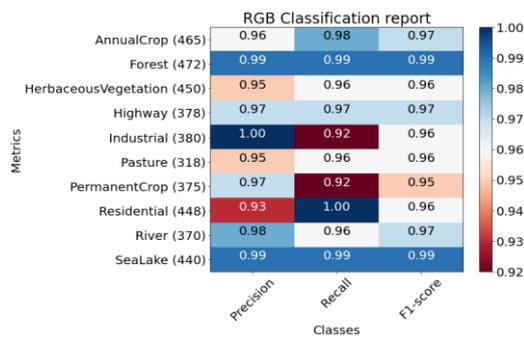

Figure 11. Training Loss graph for the All Bands Model

We can find from the Classification Report (Fig. 12) and Confusion Matrix (Fig. 13) for this model that this approach has better accuracy (98.78%) than the RGB model (96.83%). First it preserves the accuracies for the classes with high classification accuracies. Second it provides good enhancement from the first model for the similar classes (Table 3)

Figure 9. RGB Model Classification Report

3.1.2 All Bands Model

The All Bands model input consists of 13 channels from Sentinel-2 bands, Aerosols, Blue, Green, Red, Red edge 1, Red edge 2, Red edge 3, NIR, Red edge 4, Water vapor, Cirrus, SWIR 1, and SWIR 2.

After 62 epochs the training process finished with accuracy 98.78%, with more stable and better learning as seen in the accuracy graph (Fig. 10) and the loss graph (Fig. 11)
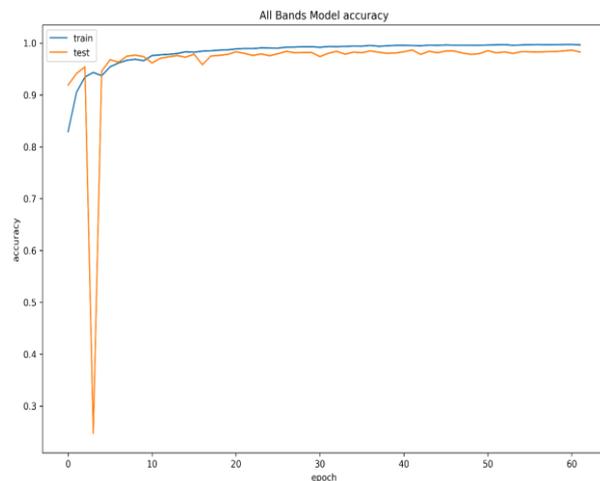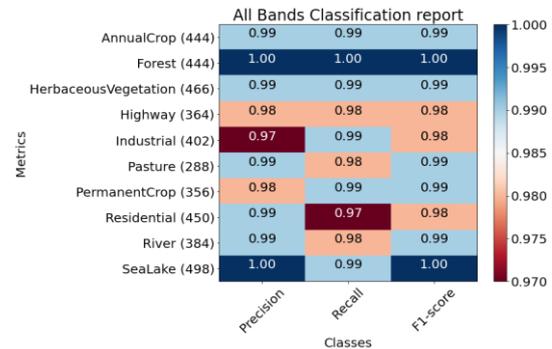

Figure 12. All Bands Model Classification Report

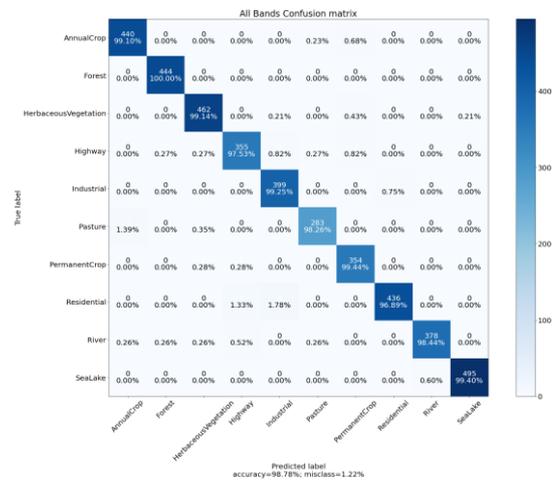**Figure 10.** Training Accuracy graph for the All Bands Model

Figure 13. All Bands Model Confusion Matrix

3.1.3 All Bands with Calculated Indices Model

This model input consists of 25 channels, in which 13 are from the Sentinel-2 bands and 12 are calculated indices as follows: MSI, NDSI, NDWI, NDVI, GNDVI, BNDVI, NDRE, BareSoil, NDCI, NBR, DSWI, and CVI.

After 90 epochs the training process finished with accuracy 99.58%, with better learning as seen in the accuracy graph (Fig. 14) and the loss graph (Fig. 15)
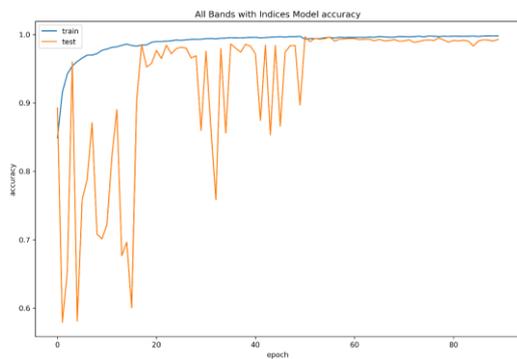

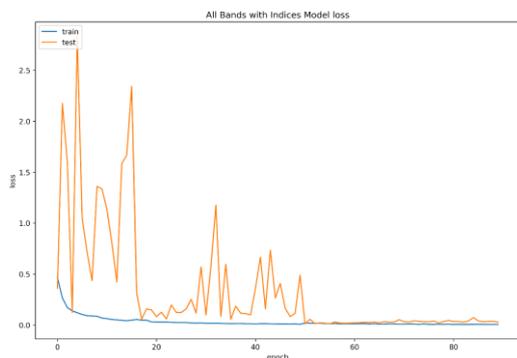**Figure 14.** Training Accuracy graph for the All Bands with Indices Model

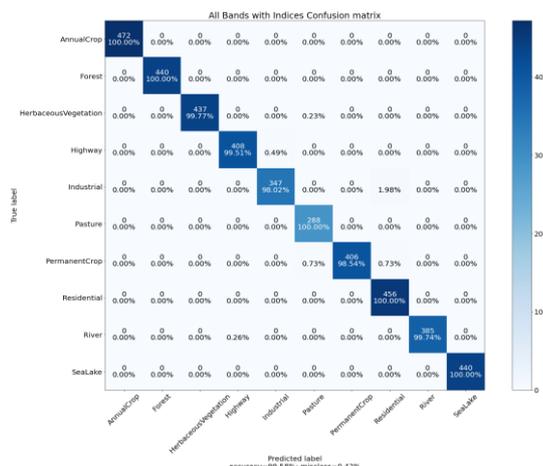**Figure 15.** Training Loss graph for the All Bands with Indices Model

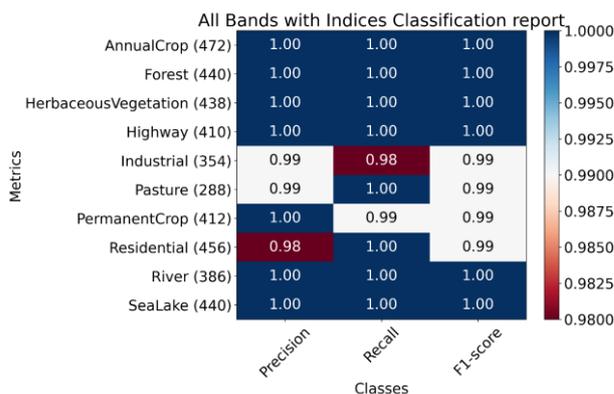**Figure 16.** All Bands with Indices Model Confusion Matrix

Figure 17. All Bands with Indices Model Classification Report

Based on the accuracies, this model achieves accuracy of 99.58% which is better than the first model 96.83% and the second one 98.78%. From the Confusion Matrix (Fig. 16) and Classification Report (Fig. 17) for this model, we can find that involving the indicated indices that targets LULC classification to the proposed model helps better classify the similar classes as follows (Table 3):

Actual Class	Wrong Predicted Class	RGB Model	All Bands Model	All Bands with Indices Model
Herbaceous Vegetation	Permanent Crop	1.56%	0.43%	0%
Highway	Residential	1.32%	0%	0%
Industrial	Residential	6.84%	0.75%	1.98%
Pasture	Annual Crop	1.57%	1.39%	0%
Permanent Crop	Annual Crop	1.60%	0%	0%
Permanent Crop	Herbaceous Vegetation	3.47%	0.28%	0%
Permanent Crop	Pasture	1.33%	0%	0.73%
River	Highway	2.16%	0.52%	0.26%
Total		19.85%	3.37%	2.97%

Table 3. Comparison of Models wrong predictions between similar classes

4. CONCLUSION

The main focus of this research was to investigate whether we can improve the classification performance of the EuroSAT dataset. Two approaches are presented for this objective. While the RGB model gives an accuracy of 96.83%, in the first approach, features are extracted from 13 spectral feature bands of Sentinel-2 instead of RGB leads to a better accuracy of 98.78%. While in the second approach features are extracted from 13 spectral bands of Sentinel-2 in addition to calculated indices used in LULC like Blue Ratio (BR), Vegetation index based on Red Edge (VIRE) and Normalized Near Infrared (NNIR), etc. which gives a better accuracy of 99.58%.

This significant enhancement in the LULC classification performance for the Sentinel-2 images in the new model can play a major role in our society especially in the environmental and agricultural sectors. These accurate classifications can help us manage, monitor and predict the agricultural areas periodically in a wide range without any efforts in going onsite, as well as to detect problems in dangerous, or inaccessible areas.

Due to time and computational limitations, we are forced in this study to cover only the novel EuroSAT dataset. Future research may focus on the BigEarthNet Dataset using TPU, and benchmarking different sets of calculated indices.

REFERENCES

[1] G. Hinton, O. Vinyals, J. Dean. Distilling the knowledge in a neural network. arXiv:1503.02531, 2015.

[2] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva. Land use classification in remote sensing images by convolutional neural networks. arXiv preprint arXiv:1508.00092, 2015.

[3] K. Nogueira, O. A. Penatti, and J. A. dos Santos. Towards better exploiting convolutional neural networks for remote sensing scene classification. Pattern Recognition, 61:539–556, 2017.

[4] O. A. Penatti, K. Nogueira, and J. A. dos Santos. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains? In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages 44–51, 2015.

[5] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 115(3):211–252, 2015.

[6] G.-S. Xia, J. Hu, F. Hu, B. Shi, X. Bai, Y. Zhong, and L. Zhang. Aid: A benchmark dataset for performance evaluation of aerial scene classification. arXiv preprint arXiv:1608.05167, 2016.

[7] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[8] Y. Yang and S. Newsam. Bag-of-visual-words and spatial extensions for land-use classification. In Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, pages 270–279. ACM, 2010.

[9] D. Marmanis, M. Datcu, T. Esch, and U. Stilla. Deep learning earth observation classification using imagenet pretrained networks. IEEE Geoscience and Remote Sensing Letters, 13(1):105–109, 2016.

[10] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4):541–551, 1989.

[11] F. P. Luus, B. P. Salmon, F. van den Bergh, and B. Maharaj. Multiview deep learning for land-use classification. IEEE Geoscience and Remote Sensing Letters, 12(12):2448–2452, 2015.

[12] P. Helber, B. Bischke, A. Dengel, D. Borth. EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. arXiv preprint arXiv: 1709.00029, 2019.

[13] J.R. Jensen, Remote Sensing of the Environment: An Earth Resource Perspective. Upper Saddle River, New Jersey: Pearson Prentice Hall, 2000, pp-656.

[14] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani. Deepsat: a learning framework for satellite imagery. In Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, page 37. ACM, 2015.

[15] G. Chen, X. Zhang, X. Tan, Y. Cheng, F. Dai, K. Zhu, Y. Gong, Q. Wang. Training Small Networks for Scene Classification of Remote Sensing Images via Knowledge Distillation. Remote Sens. 2018, 10, 719. doi:10.3390/rs10050719, 2018.

[16] E. Chong. EuroSAT Land Use and Land Cover Classification using Deep Learning. https://github.com/e-chong/Remote-Sensing, 2020.

[17] N. Sonune. Land Cover Classification with EuroSAT Dataset. https://www.kaggle.com/nilesh789/land-cover-classification-with-eurosat-dataset, 2020.

[18] J. Li, D. Lin, Y. Wang, G. Xu, Y. Zhang, C. Ding, Y. Zhou. Deep Discriminative Representation Learning with Attention Map for Scene Classification. Remote Sens. 2020, 12, 1366, 2020.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.