# Intelligent mapping of irrigated areas from Landsat 8 images using transfer learning

**Zouhair Benbahria** *[1] , **Imane Sebari** [2] , **Hicham Hajji** [2] , **Mohamed Faouzi Smiej**[1]

*[1]Royal Center for Remote Sensing, Rabat, Morocco*
*[2]School of Geomatics Sciences and Surveying Engineering, IAV Hassan II, Rabat, Morocco*

**ABSTRACT**

The lack of reliable and up-to-date data in developing countries is a major obstacle to sustainable development. In Morocco, where groundwater withdrawals by farmers are very intensive and informal, maps describing and monitoring the extension of irrigated areas are scarce and labor-intensive to obtain. In this paper a novel transfer learning algorithm is proposed to map irrigated areas at different stages of an agricultural cycle from Landsat 8 images. The results obtained displays satisfactory performance over traditional machine learning algorithms. On a small dataset, we initially tested three well known deep learning architectures (SegNet, DenseNet and Unet). The results obtained were not satisfactory. So, to get high performance, we rely on a transfer learning architecture combining UNet with ResNet50 backbone (trained on 2012 ILSVRC ImageNet dataset) as a baseline after a phase where different configurations were tested. In the first part of this study, we compared the use of three optimization methods: Adam and two variants of Stochastic Gradient Descent (SGD) associated with two techniques (Cyclical Learning Rate and Warm Restart) to find the optimal learning rate and then test the impact of data augmentation on the overall accuracies. Data augmentation had improved the overall accuracy for the three methods. Adam based method from 94% to 97% with mean IoU of 0,79 (for all land cover classes) and 0,86 for irrigated areas class. For SGD based methods, the overall accuracy had increased from 91% to 94% with mean IoU of 0,75 (for all land cover classes) and 0,82 for irrigated areas class. As we are interested in having irrigated areas maps at different key periods of the agricultural cycle, we also explored, in the second part of this study, the temporal generalization of the best model.

## 1. INTRODUCTION

Deep learning (DL) is significantly impacting areas of research, including computer vision, image processing, and remote sensing (Ball et al., 2018) thanks to the increased availability of data and computational resources (Zhu et al. 2017).

Generally, traditional deep networks (DN) are trained using large datasets of imagery. However, in remote sensing the ones available are typically very limited (Ball et al. 2018). In such low to medium learning datasets contexts, some architectures like UNet and SegNet networks are frequently privileged (Younis and Keedwell 2019).

Fortunately, open high-resolution satellite imagery, such as from USGS and Copernicus, is becoming increasingly available. Such imagery can be used to extract useful insights to inform policy decisions in water resources management and feed datasets for training new deep learning architectures.

In this paper, we examine one of the different approaches used to address the lack of training data which are: 1) Unsupervised learning 2) Generative

***** **Corresponding Author**

*(benbahria@crts.gov.ma) ORCID ID 0000 – 0003 – 4326 – 1369
(imane.sebari@gmail.com) ORCID ID 0000 – 0002 – 6754 – 8404
(hajjihi@gmail.com) ORCID ID 0000 – 0001 – 9026 – 8980
(smiej@crts.gov.ma) ORCID ID 0000 - 0002 - 1644 - 5819

adversarial networks (GANs) 3) Transfer learning (Ball et al. 2018) which is still an active research area in remote sensing according to Tuia et al. (2016).

The goal of the next sections is to provide a brief overview of some existing architectures based on DN and the specific objectives of this study.

## 1.1 Deep Learning Approaches for Semantic Image Segmentation

DL architectures have been successfully applied to pixel-based classification of high-resolution satellite images outperforming standard image classifiers. It has been shown that it can achieve far better classification performances (Zhu et al. 2017; Liu et al. 2018).

Despite the lack of training data, Deep networks have proven to outperform at extracting mid- and high-level abstract and discriminative semantic features from images. Recent studies indicate that the feature representations learned by CNNs are greatly effective in semantic segmentation (Long et al. 2015; Khryashchev et al. 2018).

LeNet-5 is the reference structure of a CNN. It was developed by (LeCun et al. 1998). It consists of two convolutional layers followed by three fully connected layers.

Semantic image segmentation is defined as the task of clustering parts of image together which belong to the same object class (Thoma 2016).

So, to produce a land cover map, well known and deep architectures as SegNet, DenseNet and UNet can be used. The latter has received a lot of interest initially for the segmentation of biomedical images using a reduced dataset but then for a lot of applications in remote sensing (Iglovikov et al. 2017; Feng et al. 2019).

UNet architecture (Ronneberger et al. 2015) is like a convolutional autoencoder. It uses skips connections (Figure 1) to reinject the features maps of the encoder part into the decoding phase and also transposed convolutions to reconstruct the original image resolution. These approaches use CNN's pretrained convolutional layers for classification, including VGG-16, as the encoder. The advantage of these symmetric approaches is that they can generate predictions at the same spatial resolution as the input image.
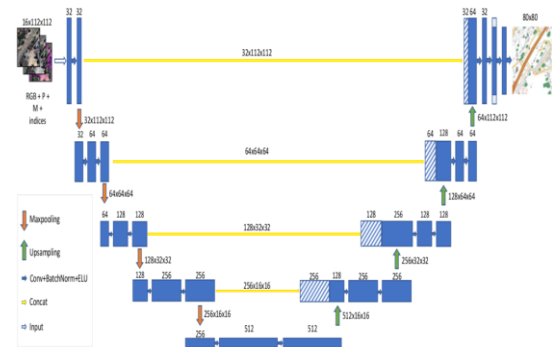


**Figure 1.** Customized UNet architecture for satellite image segmentation (Vooban 2017)

In addition, given the better performance of the ResNet and DenseNet models in object recognition, researchers also tried to adapt these architectures for semantic segmentation. Thanks to the increase in GPU computing capacities (Wu et al. 2016) proposed a first approach for ResNet.

Recent research works have shown that deeper architectures, such as deep residual networks ResNets (He et al. 2106) can gain accuracy from increasing the depth of the network. These residual networks 1) are substantially deeper (Table 1) 2) have fewer parameters 3) are easier to optimize, and 4) can gain accuracy from considerably increased depth (Bilinski and Prisacariu 2018).

**Table 1.** ResNet (50 and 152 layers) architectures for ImageNet (He et al. 2106)

| Layer name | Output size | 50-Layer | 152-layer |
|---|---|---|---|
| Con1 | 112*112 | 7x7, 64, stride 2 | |
| | | 3 x 3 max pool, stride 2 | |
| Conv2.x | 56*56 | $\begin{pmatrix} 1x1, & 64 \\ 3x3, & 64 \\ 1x1, & 256 \end{pmatrix} x3$ | $\begin{pmatrix} 1x1, & 64 \\ 3x3, & 64 \\ 1x1, & 256 \end{pmatrix} x3$ |
| Conv3.x | 28*28 | $\begin{pmatrix} 1x1, & 128 \\ 3x3, & 128 \\ 1x1, & 512 \end{pmatrix} x4$ | $\begin{pmatrix} 1x1, & 128 \\ 3x3, & 128 \\ 1x1, & 512 \end{pmatrix} x8$ |
| Conv4.x | 14*14 | $\begin{pmatrix} 1x1, & 256 \\ 3x3, & 256 \\ 1x1, & 1024 \end{pmatrix} x6$ | $\begin{pmatrix} 1x1, & 256 \\ 3x3, & 256 \\ 1x1, & 1024 \end{pmatrix} x36$ |
| Conv5.x | 7*7 | $\begin{pmatrix} 1x1, & 512 \\ 3x3, & 512 \\ 1x1, & 2048 \end{pmatrix} x3$ | $\begin{pmatrix} 1x1, & 512 \\ 3x3, & 512 \\ 1x1, & 2048 \end{pmatrix} x$ |
| | 1*1 | Average pool, 1000-d fc, softmax | |
| FLOPs | $3.8*10^9$ | $7.6*10^9$ | $11.3*10^9$ |

Also, a fully convolutional version of the DenseNet (Jégou et al. 2017) has also been proposed for semantic segmentation by combining an encoder-decoder approach with UNet inspired activation passing.

In SegNet, (Badrinarayanan et al. 2017) the authors proposed a convolutional encoder-decoder architecture for image segmentation. Similar to the deconvolution network. It consists of an encoder network, which is topologically identical to the 13

convolutional layers in the VGG16 network, and a corresponding decoder network followed by a pixel-wise classification layer.

Fully convolutional network (FCN) approaches for semantic segmentation of remotely sensed images have become much more popular. FCNs infer a pixel prediction for the entire image in a single pass, avoiding as well the problem of the classification per patch. This drastically reduces computation times without requiring unsupervised pre-segmentation.

Generally, deep learning methods using FCN have emerged in few years as the new state of the art for many remote sensing image interpretation tasks (Liu et al. 2019) and UAV images as well (Figure 2).
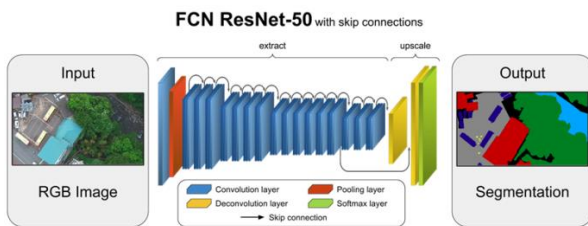


**Figure 2.** UAV image segmentation using ResNet 50 architecture

## 1.2 Related Works

The first applications of FCN on optical aerial data appear many years ago (Paisitkriangkrai et al. 2015; Sherrah 2016). Since (Mnih 2013) where authors tried using FCN for the extraction of roads and buildings in aerial images from image patches, these approaches have been successfully used on many Very High-Resolution satellite data (Lagrange et al. 2015).

In (Papadomanolaki et al. 2016), the authors used Convolutional Neural Networks (CNN) for the classification of SAT-4/SAT-6 dataset given by US National Agriculture Imagery Program. They compared different deep architectures (AlexNet, VGG. etc).

Some researches (Xu et al. 2018; Iglovikov et al. 2017) start from UNet based and adapted architectures to extract buildings, urban patterns and other land cover classes from satellite images.

Also, in (Audebert et al. 2016; Audebert et al. 2017) the authors train variants of the SegNet architecture on remotely sensed imagery over an urban area. The goal was to study different strategies to have an accurate semantic segmentation.

For hyperspectral image (HSI), Lin et al. (2013) introduced, for the first time, the concept of deep learning in a new framework of spectral-spatial feature extraction.

(Pirotti et al. 2016), by using mainly ESA and USGS free images, benchmarked 9 machine learning algorithms (Random Forest, SVM and Neural Networks...etc.). These models were tested for accuracy and speed in training and classification of land-cover classes in a Sentinel-2 dataset.

Research on semantic segmentation includes some works that deal with training data scarcity as in our paper. (Acquarelli et al. 2018) proposed, in the case of lacking training data, a convolutional neural network with a single hidden layer that can achieve state-of-the-art performance by using three tricks: a spectral-locality-aware regularization term, smoothing and label-based data augmentation.

(Vooban 2017; Jiang 2017) used only 25 labelled satellite images for training. While (Younis and Keedwell 2019) modified the structure of SegNet architecture (Figure 3) and train it using 6 RGB images. The results of this study were also promising.
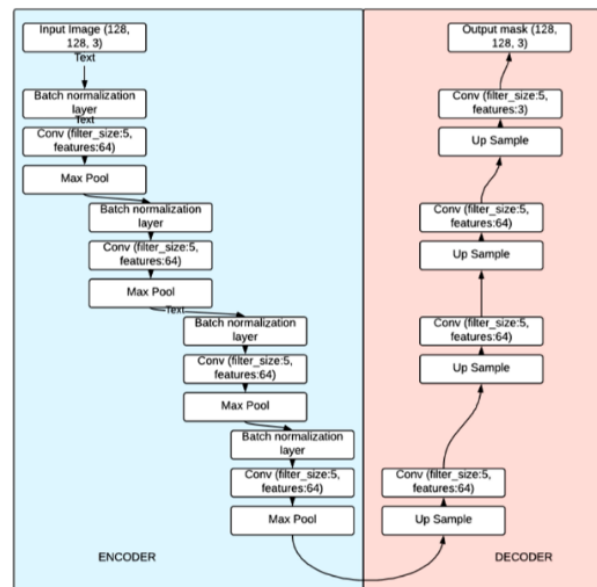


**Figure 3**. SegNet adapted architecture to semantically segment satellite data (Younis and Keedwel 2019)

As we mention before, transfer learning can be used as solution to fine-tune pretrained networks based on a small training dataset. This is possible thanks to transfer learning which seeks to learn from one area to another (Tuia et al. 2016). It can improve the learning process of a target predictive function from a knowledge-based source predictive function (Pan and Yang 2010).

Working on new target domain, two options for transfer learning can be explored (Ghazi et al. 2017). Using completely a pre-trained network to learn new features or fine-tune its weights.

Generally, most remote sensing-based transfer learning works are focused on updating the weights of a DL solution from another context to the current task based on available training data.

In (Yang et al. 2016) the authors used dual CNNs and transfer learning as inputs to a fully connected layer for classification. Lower and middle layers were trained on external dataset whereas the top layers were trained on the available training samples.

(Othman et al. 2016) used transfer learning architecture trained on the ILSVRC-12 dataset. The

trained system was next applied to the UC Merced Land Use and Banja-Luka datasets.

The nature of used mapping approaches by machine learning made it necessary to invest a lot of effort in training the models. In (Wurm et al. 2019) the study seeks to analyze transfer learning capabilities of FCNs to slum mapping in various satellite images. A model trained on very high-resolution optical satellite imagery from QuickBird is transferred to Sentinel-2 and TerraSAR-X data.

Recently, advanced methods based on domain-specific transfer learning are proposed for semantic segmentation of remote sensing data. Panboonyuen et al. (2019) proposed a novel CNN called global convolutional network (GCN) which can capture different resolutions by extracting multi-scale features from different stages of the network.

In summary, automated land cover mapping based on satellite image is a great source of information for many fields such as land management, forestry, agriculture and so on. In Morocco, where groundwater withdrawals by farmers are very numerous and informal, the need for information on the location of irrigated areas rises up as a strategic objective.

Studies on the use of remote sensing for mapping irrigated areas in Morocco are uncommon. In (Merdas et al. 2015) we used low-resolution data (MODIS) and a time series of NDVIs to map irrigated areas using a pixel-based approach (86.29% of well-classified pixels). This study was based on previous studies (Ozdogan et al. 2010). In (Benbahria et al. 2018), a new automatic mapping framework was proposed based on Landsat 8 (L8) time series images and using pixelwise classification Random Forest algorithm.

(Zhang et al. 2018) tested an approach based on well-known image classification convolutional neural networks to automatically detect only center pivot irrigation systems from Landsat 5 TM images.

Our main objective, in this paper, is to evaluate recent Transfer Learning approaches and Semantic Segmentation to monitor the extension of irrigated areas at different stages of an agricultural cycle.

As a preliminary step, we experimented known deep learning architectures as SegNet, DenseNet and UNet with our train dataset and then a transfer learning architecture combining UNet with ResNet50 as backbone. Three specific objectives are set:

- ✓ To compare the use of three optimization methods (Adam and two variants of Stochastic Gradient Descent (SGD)).
- ✓ To evaluate the impact of data augmentation on the overall accuracies of the three methods.
- ✓ To assess the temporal generalization of the model to imagery collected at different times and under different atmospheric conditions (in the same agricultural cycle).

## 2. STUDY AREA and DATA

### 2.1 Study Area

Experiments are conducted in Gharb site which is located in the north-west of Morocco (Figure 4). It currently counts 190 000 ha irrigated area. About 80 % of the rains are concentrated between November and April. The dry period is usually between June and September.
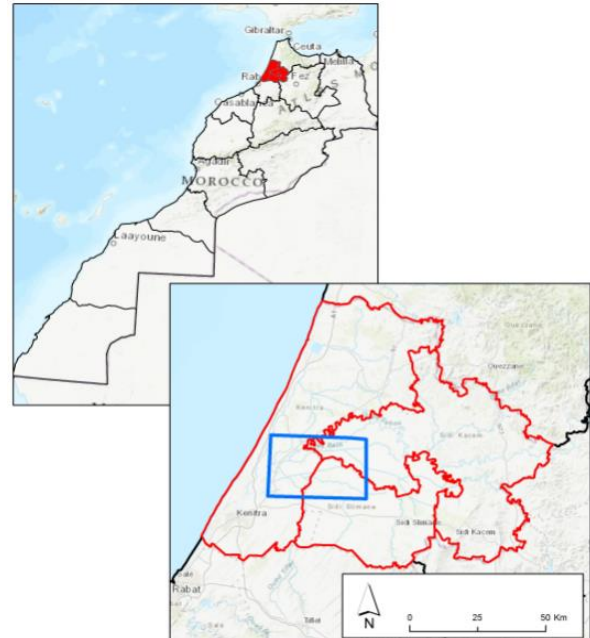


**Figure 4.** Study area localization (blue rectangle) (© Open Street Map).

### 2.2 Used Data

Three L8 cloud free scenes (October 2015, May and August 2016) were acquired and pre-processed. The choice is based on three key periods in the agricultural cycle (Autumn, Spring and Summer).

In addition to RGB (R = NIR, G = R, B = G) L8 images, ground truth data are collected and used for learning and validation.

For collecting reference data, Corine Land Cover (CLC) based classification scheme is used by adapting it to the Moroccan context. Seven (7) classes were selected (Table 2) to appear in the final land cover maps.

**Table 2.** Classification scheme for Gharb region

| Code | Land cover |
|------|------------|
| 0 | Impervious |
| 1 | Forest |
| 2 | Water |
| 3 | Arable land |
| 4 | Irrigated land |
| 5 | Greenhouse |
| 6 | Orchard |

The training image dataset includes 50 images (128x128 pixels each with 3 channels (NIR, R, G))

with corresponding ground truth masks. The ground truth was collected by photointerpretation of the images and existing recent data (Ortho-images SPOT6 and 7 with 1.5m resolution and Google Earth).

For testing the models, initially three datasets of ten manually labelled images were prepared and then augmented to make tests on 100 images for each period of the agricultural cycle.

## 3. METHODOLOGY

### 3.1 Data Preparation

L8 images were pre-processed through the following steps: 1) Converting DN of bands (B3 = Green, B4=Red, B5=NIR and B8=Panchromatic) into reflectance (ToC) by applying DOS atmospheric correction, 2) Apply a Pan-sharpening using Brovey Transform (Johnson et al. 2012), and 3) Stacking the three pansharpened bands (B3, B4 and B5).

From the first image, we extracted well distributed 50 patches (128x128 pixels each with 3 channels (NIR, R, G)) covering all land cover classes. These patches were photo-interpreted and the vector results were converted to raster masks. Finally, all the generated images were resized to 224x224 to be suited for the training and validation of the experimented DL architecture (UNet with ResNet50 as backbone).

### 3.2 Evaluating Deep Learning Architectures for Irrigated Areas Mapping

Although it has not been designed specifically for satellite images, UNet architecture is increasingly applicable in remote sensing. As we will explain later, we use this structure as a base reference but combined with ResNet.

In this paper, due to the limited training dataset, we tested a UNet based transfer learning architecture (Yakubovskiy 2018) to perform land cover semantic segmentation of L8 images.

Our multi class segmentation is based on UNet with ResNet50 backbone (which has weights trained on 2012 ILSVRC ImageNet).

In the first part of this study, this architecture was experimentally assessed through three different use cases depending on the used optimization methods: Adam (Diederik and Jimmy 2014) and two variants of Stochastic Gradient Descent (SGD).

As we know, an important part of developing a DL architecture is the selection of hyperparameters. Different methods exist for choosing such values:

1. Manual: hyperparameters are set through trial-and-error until a usable set of parameters are found.

2. Search algorithms: A grid search, or random search algorithm can be deployed. The network is then trained on multiple models by using all combinations of parameters made available in these ranges (Bergstra and Bengio, 2012).

The learning rate (LR) range (base_lr, max_lr) was firstly determined through trial-and-error test by observing the variations of loss value. Then two techniques (associated with SGD) were experimented: Cyclical Learning Rate (CLR) (Kenstler 2018) and Warm Restart (SGDR) (Loshchilov and Hutter 2016).

According to (Smith 2017), letting the learning rate cyclically vary between reasonable bounds (Figure 5) can increase the accuracy of the model in fewer steps and escape the saddle points more efficiently.
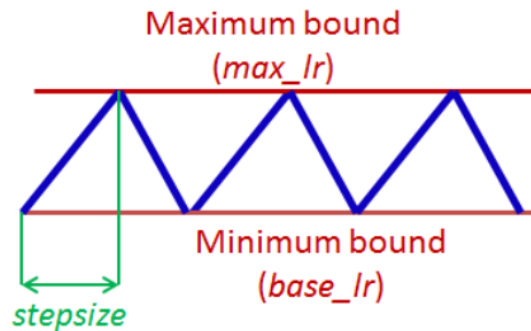


**Figure 5.** Cyclical learning rate (Smith 2017)

SGDR (Figure 6) is similar to CLR. It applies an aggressive annealing strategy. So, the learning rate is varied during training deep neural networks and performance is improved (Loshchilov and Hutter 2016).
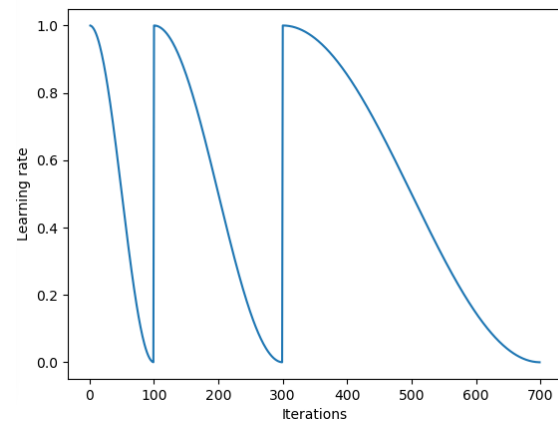


**Figure 6.** SGD with restart and cosine annealing

In the second part of this study, we assess (for the three adopted architectures) the impact of data augmentation. We considerably augment the initial training dataset (outside training process) combining 90 degrees rotations with top-bottom and left-right flip (Bloice, 2017) which increased the training dataset 10 times (500 images with associated ground truth).

In addition, we also assess the temporal generalization of the best model (from the three architectures) learned from the first Landsat 8 image acquired at the beginning of the agricultural cycle (Autumn period). The idea is to explore the feasibility of predicting (using the same model) the

location of irrigated areas at some specific times in the same agricultural cycle (Spring and Summer periods).

As evaluation metrics (Liu et al. 2019) for image segmentation, we use Pixel Accuracy (Pacc) (Eq. 1) and Intersection over Union (IoU) (Eq. 2):

$$P_{acc} = \frac{\sum_i n_{ii}}{\sum_i t_i} \qquad (1)$$

$$IoU = \frac{1}{n_{cl}} \sum \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}} \qquad (2)$$

Where,
nij: number of pixels of class i predicted correctly to belong to class j
ncl: number of classes
ti: number of pixels of class i (ti = $\sum_j$ nij)

Our architectures were implemented based on Keras API with TensorFlow backend. All models are trained and tested using Google Colaboratory.

## 4.  RESULTS and DISCUSSIONS

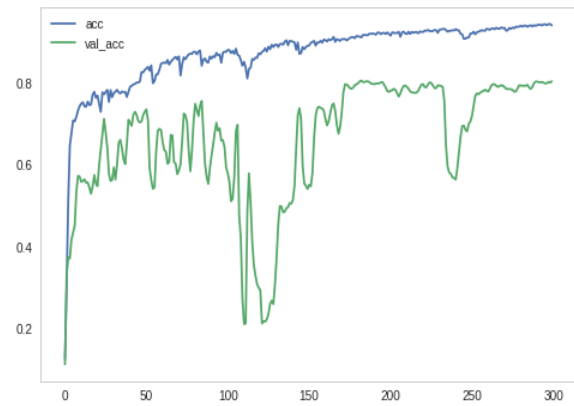### 4.1 Generating Land Cover Maps without Data Augmentation

On the limited train dataset, we experiment three of the known DL architectures: SegNet, DenseNet and UNet. The overall accuracy obtained doesn't exceed 47%.

To improve the performance, we rely on UNet architecture with ResNet50 as backbone and used a loss based on the categorical cross entropy. For all models, we use a minibatch size of 32 images (this choice was a result of trial and error testing for four values: 8, 16, 24 and 32) and fixed learning rates between base_lr=0.001 and max_lr=0.01. The results for the three use cases based on the variation of the optimization method are as follow.
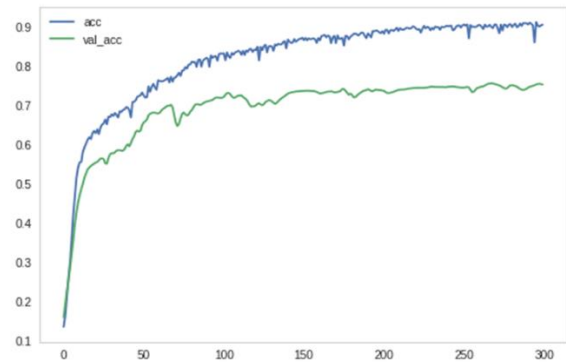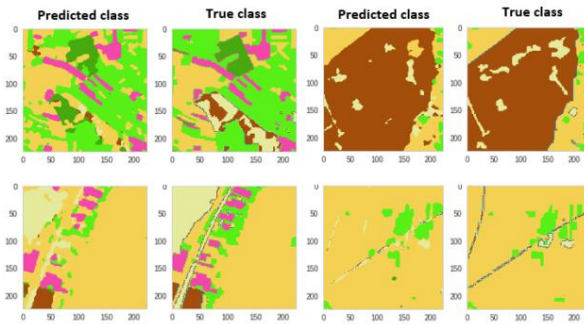


**Figure 7.** Accuracies on training and validation datasets based on 300 epochs (Adam method)



**Figure 8.** Accuracies on training and validation datasets based on 300 epochs (SGD method with CLR and Restart)

The results show a clear improvement with an overall accuracy fluctuating around 94% (76% on validation dataset for Adam method) (Figure 7). The two variants of SGD lead to relatively smooth increasing to reach 91% as overall accuracy and 72% on validation dataset (Figure 8) but with a value of loss (on validation dataset) that doesn't decrease under the value of 1.

**Table 3** Pixel accuracy and IoU metrics on test dataset

|  | ADAM | | SGD_CLR | | SGDR | |
|---|---|---|---|---|---|---|
| Land Cover | Acc (Pixel) | IoU | Acc (Pixel) | IoU | Acc (Pixel) | IoU |
| Impervious | 75% | 43% | 54% | 32% | 36% | 26% |
| Forest | 56% | 39% | 81% | 72% | 74% | 67% |
| Water | 45% | 37% | 6% | 6% | 2% | 2% |
| Arable land | 85% | 73% | 83% | 69% | 90% | 72% |
| Irrigated land | 74% | 63% | 78% | 62% | 81% | 62% |
| Greenhouse | 82% | 63% | 6% | 6% | 5% | 5% |
| Orchard | 42% | 40% | 18% | 9% | 15% | 14% |
| *Mean* | **66%** | **51%** | **47%** | **37%** | **43%** | **35%** |

With mean IoU of 51% (63% for Irrigated land class) and with less confusion among classes, Adam gives the best result (Table 3). Hereafter some predictions outputted on some test images (Figure 9) and the whole first image (Figure 10).

**Figure 9**. Predictions on some test images by Adam based model

From this, it can be seen that Adam based network is particularly good at predicting irrigated areas and green houses.
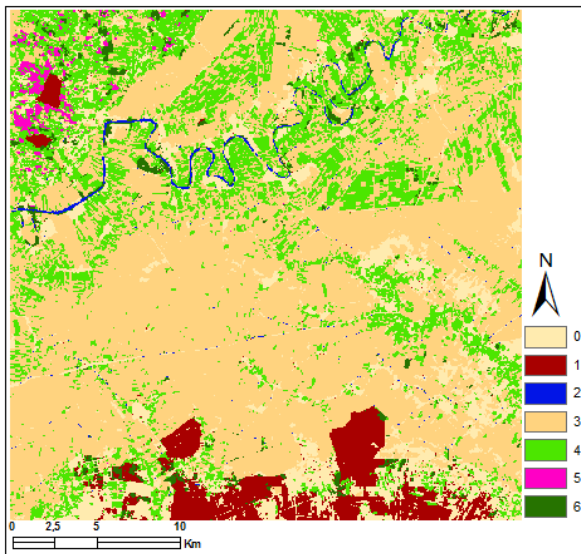


**Figure 10.** Land cover map generated by Adam based model

## 4.2 Generating Land Cover Maps with Data Augmentation

In the deep learning field, it is commonly known that a large amount of data is required to properly train a network. Unfortunately, accessing a suitable amount of data is not possible (or time consuming) for everyone along with data ground truth information.

In this second part of the study, we experimented the same three models on the augmented training dataset to assess the impact on performance and accuracies.

After artificially augmenting the training dataset and with less epochs (200), the overall accuracy has increased from 94% to 97% for Adam based method (Figure 11) and from 91% to 94% for SGD based methods (Figure 12). Also, the overall accuracy for the three methods on validation dataset has increased and reached respectively 92% and 91%.
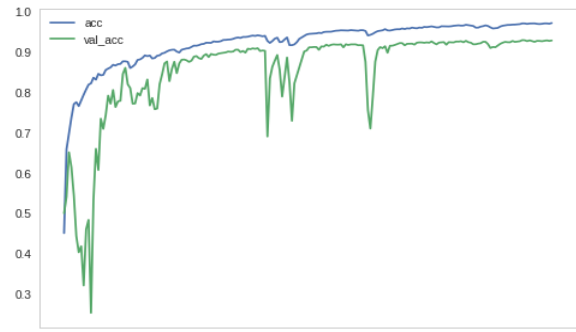


**Figure 11.** Accuracies on augmented training and validation datasets based on 200 epochs (Adam method)
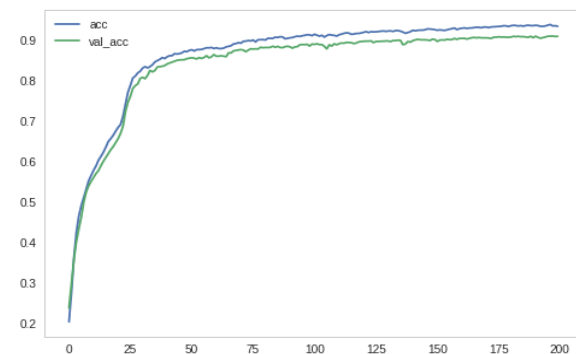


**Figure 12.** Accuracies on augmented training and validation datasets based on 200 epochs (SGD method with CLR and Restart)

All the methods (Table 4) perform well (less confusion among classes). Adam based method outperform the other methods especially for classes water and irrigated land. This latter is mapped with high degree of performance (IoU of 87%).

**Table 4** Pixel accuracy and IoU metrics on test dataset

| | ADAM | | SGD_CLR | | SGDR | |
|---|---|---|---|---|---|---|
| Land Cover | Acc (Pixel) | IoU | Acc (Pixel) | IoU | Acc (Pixel) | IoU |
| Impervious | 89% | 84% | 89% | 74% | 90% | 80% |
| Forest | 93% | 87% | 93% | 89% | 89% | 86% |
| Water | 74% | 56% | 49% | 39% | 44% | 37% |
| Arable land | 96% | 91% | 93% | 87% | 95% | 87% |
| Irrigated land | 93% | 87% | 91% | 82% | 90% | 82% |
| Greenhouse | 78% | 63% | 73% | 63% | 74% | 66% |
| Orchard | 87% | 84% | 96% | 90% | 93% | 84% |
| **Mean** | **87%** | **79%** | **83%** | **75%** | **82%** | **75%** |

Using Google Colaboratory with GPU processing capacities, it has been noticed that training and testing all models show the same speed performance.

Below, some predictions produced on some test images (Figure 13) and the first whole image (Figure 14).
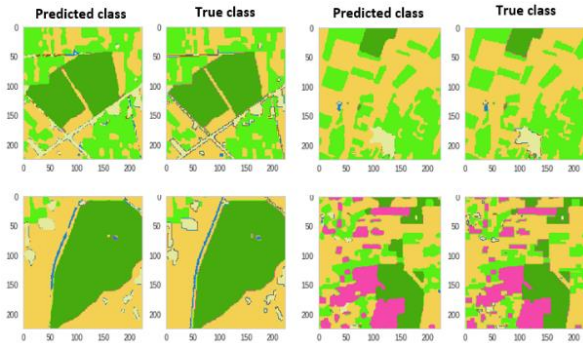


**Figure 13.** Predictions on some test images by Adam based model

### 4.3. Temporal Generalization of the Best Learned Model

The aim here is to assess how robust is the temporal generalization of the best learned model (from the previous phase). Two new test datasets (ground truth masks) are photo interpreted based on two acquisitions corresponding to spring and summer periods in the same agricultural cycle. The results of the predictions evaluations (Table 5) are given hereafter.
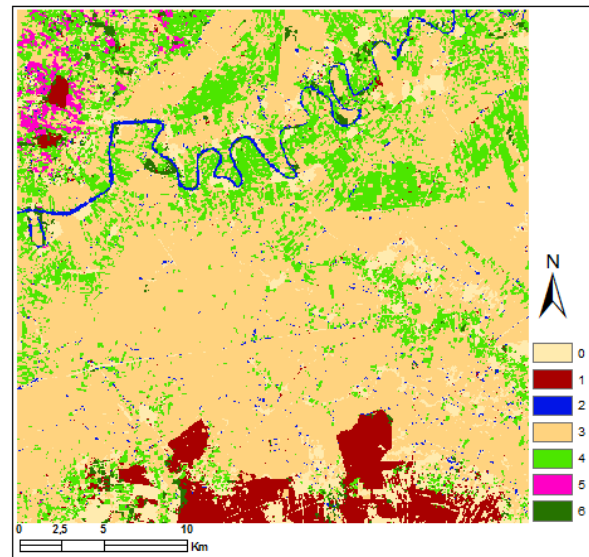


**Figure 14.** Land cover map generated by Adam based model (on augmented training dataset)

**Table 5.** Pixel accuracy and IoU metrics on test datasets for the three image acquisitions
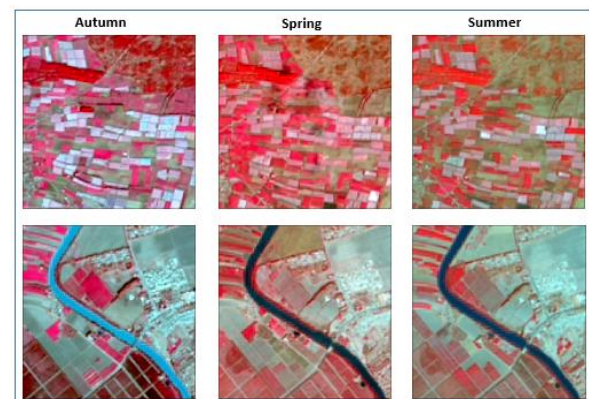
| | AGRICULTURAL CYCLE | | | | | |
| | AUTUMN | | SPRING | | SUMMER | |
| Land Cover | Acc (Pixel) | IoU | Acc (Pixel) | IoU | Acc (Pixel) | IoU |
|---|---|---|---|---|---|---|
| Impervious | 89% | 84% | 79% | 29% | 88% | 9% |
| Forest | 93% | 87% | 65% | 54% | 6% | 6% |
| Water | 74% | 56% | 54% | 23% | 53% | 40% |
| Arable land | 96% | 91% | 58% | 54% | 29% | 25% |
| Irrigated land | 93% | 87% | 52% | 37% | 33% | 28% |
| Greenhouse | 78% | 63% | 21% | 9% | 5% | 5% |
| Orchard | 87% | 84% | 31% | 8% | 0% | 0% |
| **Mean** | **87%** | **79%** | **51%** | **31%** | **31%** | **16%** |
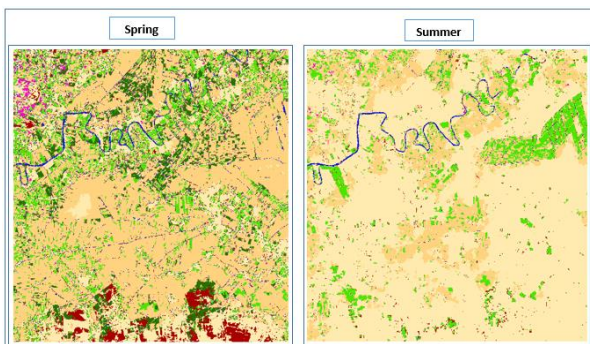


**Figure 15.** Predicted land cover maps for spring and summer L8 images

The results show low accuracies for the predictions on spring and summer L8 images (Figure 15 and 17). The main cause is the reflectance of the objects which change during the agricultural cycle. Especially for summer acquisition, more confusion between irrigated land and orchard classes (Figure 16) and low performance for the other classes.



**Figure 16.** Samples of two test images acquired during different periods of the agricultural cycle
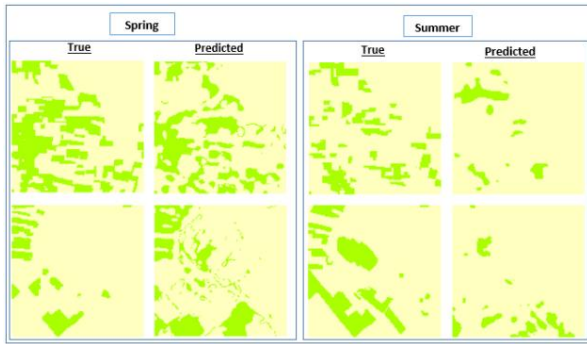
**Figure 17.** Predictions of irrigated areas (Green) on two test images

The aim of this third part experiments is to come up with a prediction model that can map irrigated areas at different times of an agricultural cycle. The reason behind is that our class of interest generally keep the same reflectance during the year. Unfortunately, 2015-2016 was a specific year marked by a severe drought in Morocco which explain the low accuracies and also the confusion among classes.

## 5. CONCLUSIONS

With the aim of irrigated areas automatic mapping from RGB Landsat 8 satellite images, we review three architectures based on UNet with ResNet50 as backbone. Initially, small dataset of 50 images with associated ground truth labels were used in training and validation.

**Table 6.** Irrigated areas mapping accuracies using the three architectures

|  | ADAM | | SGD_CLR | | SGDR | |
|---|---|---|---|---|---|---|
| Model | Acc (Pixel) | IoU | Acc (Pixel) | IoU | Acc (Pixel) | IoU |
| Without AUG | 74% | 63% | 78% | 62% | 81% | 62% |
| With AUG | 93% | 87% | 91% | 82% | 90% | 82% |

AUG: Augmentation

Without data augmentation, Adam based model gives the best result with mean IoU of 51% (63% for Irrigated land class) (Table 6). We believe that better performances can be achieved using more data.

After artificially augmenting the training dataset, the overall accuracy has increased from 94% to 97% for Adam based model and from 91% to 94% for SGD based models. The irrigated areas are mapped with high degree of performance (IoU >= 82%). On the same area of interest and using Random Forest algorithm (Benbahria et al., 2018) we obtained less accuracy. This could confirm the relevance of using these new approaches based on deep learning architectures.

On the other hand, the temporal generalization of the best learned model to spring and summer L8 images (in the same agricultural cycle) leads to low accuracies with 37% and 28% IoUs respectively. Many enhancements can be explored to improve these results. Firstly, we can further augment the

training dataset from different season's images. Adding more spectral bands and indices into training images should also be explored.

## ACKNOWLEDMENT

## REFERENCES

Acquarelli J, Marchiori E, Buydens L M C, Tran T & van Laarhoven T (2018). Spectral-spatial classification of hyperspectral images: Three tricks and a new learning setting. Remote Sensing, 10(7), 1156. DOI: 10.3390/rs10071156

Audebert N, Le Saux, B & Lefèvre S (2017). Fusion of heterogeneous data in convolutional networks for urban semantic labeling. 2017 Joint Urban Remote Sensing Event (JURSE), 1-4.

Audebert, N, Le Saux B & Lefèvre S (2016). Semantic segmentation of earth observation data using multimodal and multi-scale deep networks. Asian Conference on Computer Vision. Springer, Cham, 180-196.

Badrinarayanan V, Kendall A & Cipolla R (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 39(12), 2481-2495.

Ball J E, Anderson D T & Wei P (2018). State-of-the-art and gaps for deep learning on limited training data in remote sensing. IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, 4119-4122.

Benbahria Z, Sebari I, Hajji H & Smiej M F (2018). Automatic mapping of irrigated areas in Mediterranean context using Landsat 8 time series images and random forest algorithm. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 7986-7989. Valencia, Spain, DOI: 10.1109/IGARSS.2018.8517810

Bergstra J & Bengio Y (2012). Random search for hyper-parameter optimization. Journal of Machine Learning Research, 13, 281–305.

Bilinski P & Prisacariu V (2018). Dense decoder shortcut connections for single-pass semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 6596-6605.

Bloice M D (2017). https://github.com/mdbloice/Augmentor Accessed date: January, 27th 2020

Diederik P K & Jimmy L B (2014). Adam: A method for stochastic optimization. 3rd International Conference for Learning Representations, San Diego

Feng W, Sui H, Huang W, Xu C & An K (2019). Water body extraction from very high-resolution remote sensing imagery using deep U-Net and a superpixel-based conditional random field model. IEEE Geoscience and Remote Sensing Letters, 16(4), 618-622.

Ghazi M M, Yanikoglu B & Aptoula E (2017). Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing, 235, 228-235.

He K, Zhang X, Ren S & Sun J (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.

Iglovikov V, Mushinskiy S & Osin V (2017). Satellite imagery feature detection using deep convolutional neural network: A kaggle competition. Computer Vision and Pattern Recognition.

Jégou S, Drozdzal M, Vazquez D, Romero A & Bengio Y (2017). The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 11-19.

Jiang R X (2017). https://towardsdatascience.com/dstl-satellite-imagery-contest-on-kaggle-2f3ef7b8ac40 Accessed date: 1May, 2020)

Johnson B A, Tateishi R & Hoan N T (2012). Satellite image pansharpening using a hybrid approach for object-based image analysis. ISPRS International Journal of Geo-Information, 1(3), 228-241.

Kenstler, B. (2018). https://github.com/bckenstler/CLR Accessed date: January, 27th 2020

Khryashchev V, Ivanovsky L, Pavlov V, Ostrovskaya A & Rubtsov A (2018). Comparison of Different Convolutional Neural Network Architectures for Satellite Image Segmentation. 23rd Conference of Open Innovations Association (FRUCT), 172-179

Lagrange A, Le Saux B, Beaupere A, Boulch A, Chan-Hon-Tong A, Herbin S, Randrianarivo H & Ferecatu M (2015). Benchmarking classification of earth-observation data: From learning explicit features to convolutional networks. International Geoscience and Remote Sensing Symposium (IGARSS), 4173-4176.

LeCun Y, Bottou L, Bengio Y & Haffner P (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

Lin Z, Chen Y, Zhao X &Wang G (2013). Spectral-spatial classification of hyperspectral image using autoencoders. In: 2013 9th International Conference on Information, Communications & Signal Processing.

Liu T, Abd-Elrahman A, Morton J & Wilhelm V L (2018). Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system. GIScience & remote sensing, 55(2), 243-264.

Liu X, Deng Z & Yang Y (2019). Recent progress in semantic image segmentation. Artificial Intelligence Review, 52, p. 1089-1106.

Long J, Shelhamer E & Darrell T (2015). Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431–3440.

Loshchilov I & Hutter F (2016). Sgdr: Stochastic gradient descent with warm restarts. ICLR 2017, 1-16

Merdas M, Ozdogan M, Benbahria Z, Khajour L & Youssoufi E E (2015). Mapping of irrigated crops and monitoring of their evolution using satellite images. GEO OBSERVATEUR, N°22

Mnih V (2013). Machine learning for aerial image labeling. PHD Thesis, University of Toronto, Canada

Othman E, Bazi Y, Alajlan N, Alhichri H & Melgani F (2016). Using convolutional features and a sparse autoencoder for land-use scene classification. International Journal of Remote Sensing, 37(10), 1977–1995

Ozdogan M, Yang Y, Allez G & Cervantes C (2010). Remote sensing of irrigated agriculture: Opportunities and challenges. Remote Sensing, 2, 2274-2304 DOI: 10.3390/rs2092274

Paisitkriangkrai S, Sherrah J, Janney P & Hengel V D A (2015). Effective semantic pixel labelling with convolutional networks and conditional random fields. Conference on Computer Vision and Pattern Recognition Workshops, 36-43

Pan S J & Yang Q (2010). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

Panboonyuen T, Jitkajornwanich K, Lawawirojwong S, Srestasathiern P & Vateekul P (2019). Semantic segmentation on remotely sensed images using an enhanced global convolutional network with channel attention and domain specific transfer learning. Remote Sensing, 11(1), 83.

Papadomanolaki M, Vakalopoulou M, Zagoruyko S & Karantzalos K (2016). Benchmarking deep learning frameworks for the classification of very high resolution satellite multispectral data. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 3(7), 83-88

Pirotti, F. and Sunar, F. and Piragnolo, M. (2016). Benchmark of machine learning methods for classification of a Sentinel-2

image. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 41, 335-340

Ronneberger O, Fischer P & Brox T (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, 9351, 234-241. Springer, Cham.

Sherrah J (2016). Fully convolutional networks for dense semantic labelling of high-resolution aerial imagery. Computer Vision and Pattern Recognition

Smith L N (2017). Cyclical learning rates for training neural networks. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), 464-472

Thoma M (2016). A survey of semantic segmentation. Computer Vision and Pattern Recognition

Tuia D, Persello C & Bruzzone L (2016). Domain adaptation for the classification of remote sensing data: An overview of recent advances. IEEE geoscience and remote sensing magazine, 4(2), 41-57.

Vooban A (2017). https://medium.com/vooban-ai/satellite-image-segmentation-a-workflow-with-u-net-7ff992b2a56e Accessed date 1 May 2020

Wu Z, Shen C & Hengel A V D (2016). High-performance semantic segmentation using very deep fully convolutional networks. Computer Vision and Pattern Recognition

Wurm M, Stark T, Zhu X X, Weigand M & Taubenböck, H (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing, 150, 59-69.

Xu Y Wu L, Xie Z & Chen Z (2018). Building extraction in very high-resolution remote sensing imagery using deep learning and guided filters. Remote Sensing, 10(1), 144. DOI: 10.3390/rs10010144

Yakubovskiy, P. (2018). https://github.com/qubvel/segmentation_models Accessed date :1 May 2020

Yang J, Zhao Y, Chan J C W & Yi C (2016). Hyperspectral image classification using two-channel deep convolutional neural network. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 5079-5082.

Younis M C & Keedwell E (2019). Semantic segmentation on small datasets of satellite images using convolutional neural networks. Journal of Applied Remote Sensing, 13(4) DOI: 10.1117/1.JRS.13.046510

Zhang C, Yue P, Di L & Wu Z (2018). Automatic identification of center pivot irrigation systems from landsat images using convolutional neural networks. Agriculture, 8(10), 147. DOI: 10.3390/agriculture8100147

Zhu X X, Tuia D, Mou L, Xia G S, Zhang L, Xu F & Fraundorfer F (2017). Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8-36.