ISSN No:-2456-2165

Automatic Parcelling of Rice Fields based on Sentinel 2 Images and Convolutional Neural Networks in the Valley of the Senegal River

Cheikh Sadibou FAYE[1]
Water and Environmental Sciences and Techniques Laboratory (LaSTEE)
Polytech School of Thies, Senegal

Ababacar FALL[2]
Water and Environmental Sciences and Techniques Laboratory (LaSTEE)
Polytech School of Thies, Senegal

Hyacinth SAMBOU[3]
Institute of Environmental Sciences, Cheikh Anta Diop University of Dakar, Senega

Mbagnick FAYE[4]
Climatology and Environment Laboratory (LCE)
Cheikh Anta DIOP of Dakar, Senegal

El Hadji Bamba DIAW[5]
Water and Environmental Sciences and Techniques Laboratory (LaSTEE)
Polytech School of Thies, Senegal

Abstract:- This study is an application of computer vision techniques to the extraction of the boundaries of rice fields in the middle valley of the Senegal River. The objective of this work is to develop a method for an automatic and reliable parcelization of the contours of irrigated rice plots to practice Sentinel2 images.

The automatic delimitation of plots is a new field of research to allow managers of rice fields in the Senegal River valley to have the exact areas of the sowings. It is based on Sentinel 2 data and semantic segmentation methods using neural networks. The methodological approach consists of using a convolutional neural network to detect parcel boundaries. The results obtained from the extraction of the contours remain entirely satisfactory for a correct calculation of the yields per season. This technique of automation and calculation of rice areas makes it even more reliable and sustainable for updating agricultural databases.

Keywords:- *Sentinel 2, Segmentation, Plot, Irrigated Rice Field, Neural Network.*

I. INTRODUCTION

From natural resource monitoring to precision agriculture, the Earth observation (EO) and teeldectection play a fundamental role in enabling space-based technologies to achieve sustainable solutions in global agri-food systems.

The EO technologies have become a tool to assist agricultural developers, particularly those involved in irrigated agriculture.

In the Senegal River Valley, rice fields are land traps that undergo rapid and permanent spatiotemporal change. They are small areas whose surface areas vary from 1/2 hectare to a few hectares. More specifically, they are village irrigated perimeters that are quite numerous and where it is difficult to have an overall view and the exact limits of the different agricultural crops cultivatedToday, with the mechanization of irrigated agriculture and banking programs, farmers have decided to form cooperatives or groups by pooling their investments (FAYE et al., 2016). Thus, we are witnessing continuous changes in the size of farms, which is leading to the merging of plots into units of several dozen hectares and to a reorganization into groups or cooperatives of producers. These mutations put at stake the reliability of yield calculations. It is to overcome all these shortcomings that we have deemed it useful to set up a method for determining agricultural boundaries and areas. This will rely on automated mechanisms to determine the exploited areas based on the use of satellite imagery and artificial intelligence.

With the multiplication of satellite positioning constellations (GNSS), the miniaturization of on-board electronic systems, and the development of information and communication technologies, certain algorithms make it possible to extract useful parameters for the management of irrigated agricultural plots. Thus, spatial technologies, coupled with deep learning methods, make it possible to automatically update plot boundaries.

This study aims to use space technologies for a reliable and automatic delimitation of the contours of the plots. It proposes to use Sentinel 2 data for the extraction of the limits of the irrigated perimeters. Its realization uses a set of approach, tools, and spatial.

ISSN No:-2456-2165

II. MATERIALS AND METHODS

A. Geographical ituation of the tudy area

The Senegal River Valley is a strip of land that runs along the Senegal River.

It corresponds to the northern part of Senegal. This strip of land is characterized by the presence of the waters of the river (Senegal) favoring irrigated cultivation. The Senegal River valley is located at latitude 14°55 and 17°00'20" North and longitude 15°51'20" and 12°45'20" East (Map 1).

The Senegal River successively crosses the administrative regions of Tambacounda, Matam and Saint-Louis. It flows into the Atlantic Ocean 25 km downstream from Saint-Louis.

From Bakel, the Senegal River serves as a border between the Republic of Senegal and the Islamic Republic of Mauritania. The valley stretches for nearly 650 km.

The study is carried out on hydro-agricultural rice plots in the commune of Ndiayenne Pendao (Middle Valley) in the department of Podor.

This territorial community is located 177 km to the northeast of Saint-Louis.

In this part of Senegal, rice is one of the most widely grown crops. It is located in the Steppe region, and the climate is of the continental Sahelian type, characterized by the alternation of two seasons: a rainy season, which extends from July to September, and a dry season, from October to June.

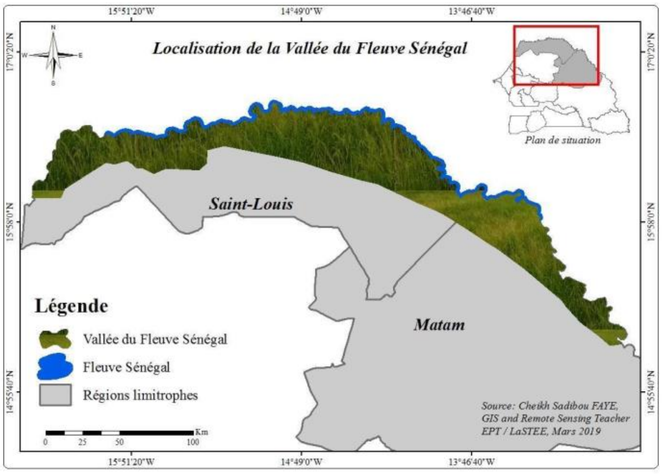


Fig 1: location of the eco-geographical area of the Senegal River valley

B. Data

Remote sensing is the technique which through the acquisition of images, makes it possible to obtain information on the earth's surface without direct contact with it. According to the Inter-ministerial Commission for the teledectection Terminology (1988), teledectection is the "set of knowledge and techniques used to determine the physical and biological characteristics of objects by measurements carried out at a distance, without material contact with them" (Paul S. and al., 1985).

It encompasses any process of capturing and recording the electromagnetic signal emitted or reflected by a target, the processing as well as the analysis and interpretation of data.

The acquisition of teledectection data can take different forms and come from different vectors: satellites, planes, helicopters, balloon drones, etc. These vectors are "vehicles" that carry sensors. In this research, we are interested in geospatial information captured by earth observation satellites. Knowing that there are several satellite constellations for earth observation, our choice focused on Sentinel 2 images. The availability, accessibility of the images and their spatial resolution from 10 to 100 m as well as their revisit time (5 days) were decisive in this choice.

In this study, it is a question of exploring the spectral properties of cultures and contours. The Sentinel-2 data used was derived using the Sentinel-2A L1C. Sentinel-2 A and B images are level 1C, which means that they have undergone geometric and radiometric correction, but no atmospheric correction (**Lillesand et al., 2015; Lopez** et al., 2020; Nguyen et al, 2020, Acharki et al., 2021). They give the advantage of being already rectified by radiometric and geometric corrections. But it is always useful to scientifically validate all image corrections.

A Sentinel 2A image was downloaded from the "scihub Copernicus" platform for automatic parcelization

Table 1: Satellite images acquired during the dry off-season in the middle valley

satellites Images	spatial Resolution (m)	Date
Sentinel-2A L1C	10	April 2020

Fig. 2: Hydro-agricultural developments in the west of the Commune of Richard Toll in Senegal. False color representation (PIR, R, V) of an image taken by the Sentinel-2A satellite in April 2020.*Sentinel2A images*

ISSN No:-2456-2165

Sentinel 2A images are data from the Copernicus mission of 2A on June 23, 2015 which bears the same name. This mission includes a constellation of two polar-orbiting satellites placed in the same sun-synchronous orbit, phased at 180° in relation to each other. The Sentinel 2 satellites have a swath of 290 km and a revisit time ranging from 10 to 5 days depending on the geographical location. Images captured by Sentinel 2 are cloud-free and enable monitoring of land surface changes.

Sentinel 2 data is downloaded during the first half of May 2020 from the archive platform (https://scihub.copernicus.eu/).

Sentinel 2A data were preprocessed using ENVI software. The preprocessing steps included (1) radiometric calibration, orthorectification using a Digital Terrain Model (DTM) of the Shuttle Radar Topography Mission (SRTM) and Ground Control Points (GCP) for geometric corrections.

C. Pre processings and images procesing

Pre-processing is a preparatory phase to improve the visual quality of images for later analysis. It starts from a raw format of Sentinel 2A images as produced by an acquisition station to build a usable image. This process includes, among other things, optical calibration, orthorectification, and sampling.

1) Geometric corrections

Raw satellite images, as recorded by satellite sensors, often exhibit so-called “geometric” errors. These errors can be semantic (related to recording systems) or accidental. They lead thus to errors in the positioning of the image but also in the area estimates due to the deformations of the image. These errors on the geometry of the satellite images must be corrected before the processing phase. A correction of the geometry of the image is applied to semantic errors and an orthorectification to accidental ones. Geometric corrections are key algorithms in the preprocessing of satellite images. The objective of geometric corrections is to reposition the image in its geographical area. They often require the use of a Digital Terrain Model (DTM) to take into account the topography of the area where the image is captured and the ground control points (GCP) which give the references of the geodetic system used.

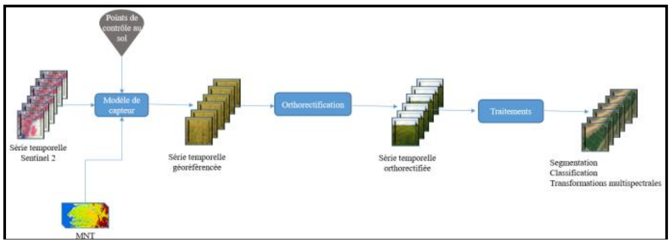


Fig 3: process *modeling for processing geometric corrections applied to the time series of Sentinel 2A images acquired between January and July 2020 in the Commune of Ndiayenne Pendao*

2) Atmospheric corrections

Atmospheric effects and other mechanisms can alter the radiometry of the images such as topographic and environmental effects caused by atmospheric diffusion or even directional effects related to the solar angle, the viewing angle and to the nature of the surface. The light energy recorded by the sensor is the result of multiple interactions with the atmosphere and the earth's surface. Solar radiation is attenuated by scattering and absorption mechanisms in the atmosphere, then its energy is absorbed, transmitted and/or reflected by the earth's surface.

3) Data

In the Senegal River valley, the cultivation methods and the spatial organization of the plots do not facilitate the large-scale monitoring of cultivated developments. From one season to another, the plots are fragmented, expanded or merged. These practices require a recalculation of the areas of the developments in order to ensure the establishment of reliable statistics on crop yields. Thus, there is a strong need to automate the delimitation of plots in each campaign.

This part deals with the automation of the delimitation of hydro-agricultural developments with a view to having the areas actually exploited by campaign. The extraction of plot limits can be performed by different processing techniques on satellite images. This section uses segmentation methods.

D. Satelites images segmentation

There are several algorithm-based methods for object segmentation. These algorithms use information such as the color, the shape of the details visible on the satellite images.

Since these algorithms are hard-coded or learned from selected features, they have certain limitations in terms of generalization. Moreover, the accuracies of these algorithms are low and the results are dependent on parameter selection due to manual feature extraction and the lack of large datasets and high-performance computers (source).

In doing so, new in-depth techniques for the automatic extraction of object characteristics have emerged: this is called segmentation.

Segmentation methods are new approaches for automatically extracting features from an object.

In the field of irrigated agriculture, they can provide relief for the determination of sown areas. They offer the possibility of automatically extracting each pixel from a detail of the image or even of defining the contours of the plots used.

This part shows the potential of using high spatial resolution satellite images such as Sentinel 2A for the automatic delimitation of irrigated plots in the Senegal River valley.

1) Segmentation methods

Instance segmentation consists of assigning a unique label to each instance of the same object within the image. Different segmentation methods are proposed depending on the available field data and the research objectives. It is called an instance when one can differentiate between a parcel of rice and a parcel of tomato and semantics when it does not differentiate between the parcels.

The objective of this work not being the differentiation of plots according to speculations, we therefore use semantic segmentation or Deep Learning in our processing.

a) Semantic segmentation

Semantic Segmentation is an active learning approach using artificial intelligence. It is applied in several applications, robotics, medical medical diagnostics, multimedia, satellite mapping and precision agriculture.

In recent years, deep learning has led to a remarkable breakthrough in detecting objects in remote sensing images (Kun Fu and al, 2019). With the development of deep learning, neural networks have become powerful and efficient tools in geospatial data processing. Remote sensing applications are many and varied: image classification, change detection, image fusion, image restoration, multi-time series processing, edge detection, etc. All these applications are generated by the development of remote sensing systems which offer a wide variety of optical, multi-spectral, hyper-spectral, radar sensors, etc. The availability of large masses of data favored by the multiplication of sensors are the basis of neural networks.

The challenge today in the use of neural networks lies in the correct and appropriate exploitation of large masses of remote sensing data.

In this section, it will be a question of using convolutional neural networks (CNN) to detect the contours of hydro-agricultural plots. It carries an approach for implementing semantic segmentation based on deep learning in order to detect the limits of irrigation schemes. We use high-resolution Sentinel imagery to automate the delineation of rice developments. This part focuses on the extraction of the limits of the facilities from the semantic segmentation on ENVI. Its objective is to produce a map of the delimitation of developments to facilitate the calculation of yields and the updating of agricultural databases. The automation of the detection of visible land contours from Sentin2 high-resolution images is based on deep learning methods.

b) Artificial neural networks

A biological neural network is a representation of an elementary brain cell. It forms synaptic connections that allow human beings to perform certain tasks.

Unlike biological neuron, the artificial neuron is a human creation like the biological neuron. The latter is the basis of artificial neural networks. The artificial neuron is a simple imitation of the functions of a neuron in the human brain to solve machine learning problems. The first artificial neural network was proposed by Warren McCulloch and Walter Pitts in 1943. It is a simplified representation of a biological neural network. This artificial neuron has one or more inputs and a binary output.

Since 2010, with the power of computers in data processing, the development of neural networks and the availability of very high resolution (VHR) images have led to an interest in the application of deep learning techniques in the geospatial domain.

Deep learning refers to a set of techniques and algorithms that leverage neural networks to solve computer vision or natural language problems. It is a subset of machine learning methods (object detection, semantic segmentation, and instance segmentation).

These deep learning methods and image interpretation algorithms can be grouped into fully convolutional and region-based approaches. The approach used in this work is based on the extraction of the boundaries of rice development from a convolutional neural network.

c) Convolutional neural networks

A convolutional neural network is a neural network that uses a mathematical operation called convolution or convolution product. This is a linear operation. Each convolutional neural network contains at least one convolution layer. Let f and g be two functions defined on R, the convolution product between f and g is generally denoted $f * g$ and it is defined by the following equation:

$$s(x) = (f * g)(x) = \int_{-\infty}^{+\infty} f(t)g(x-t)\,\mathrm{d}t$$

Today, convolutional neural networks, also called CNN or ConvNet for Convolutional Neural Network, are among the most powerful models for image classification.

Several types of deep neural networks exist but we deal in this study with convolutional neural networks (CNN). Indeed, CNNs make it possible both to identify essential characteristics of images and to analyze large masses of data in record time. These are algorithms that are also applied for supervised image classification. Like all other networks, CNNs are characterized by a layer input and layer output relationship useful for understanding or predicting a phenomenon (Belhaouci, 2019).

Convolutional neural networks are a particularly effective tool for semantic segmentation for many types of images: multimedia, aerial, medical or autonomous vehicle (Audebert et al., 2018). They are constituted by a succession of several operations connected to each other.

III. RESULTATS

Extracting information from a Sentinel 2A image requires a CNN composed of several operations: convolution, activation, pooling and connection.

A. Convolution

A neural network is made up of several filters. A simple definition of convolution is applying a mathematical filter to an image. From a more technical point of view, it involves dragging a matrix over an image, and for each pixel, using the sum of the multiplication of this pixel by the value of the matrix. This technique allows us to find parts of the image that could be of interest to us (see fig. 10).

An image is made up of a series of rows and columns forming cells called pixels. Each cell contains a radiometric value between 0 and 255 when the image is coded in 8 bits. To simplify the representation, the values of the pixels of the image are represented by 0 and 1. In reality, the values of the pixels of an image encoded in 8 bits are included between 0 and 255.

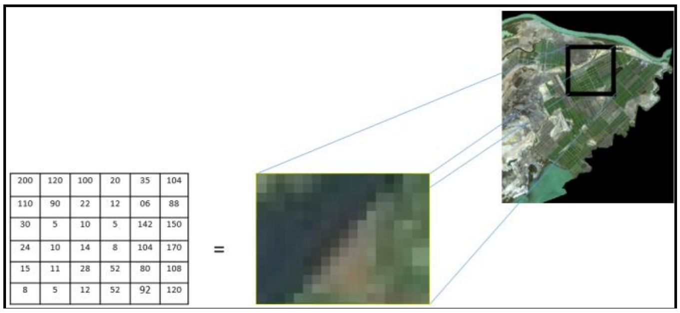


Fig 4: Representation of an image in the form of a matrix of radiometric values

Convolution is an operation that consists of applying a convolutional filter to an image. It involves spatially dragging a convolutional filter (F), on an image (I) by calculating the scalar product at each pixel value. To simplify the calculations, the matrix of the radiometric values of the image pixels are represented by 0s and 1s.

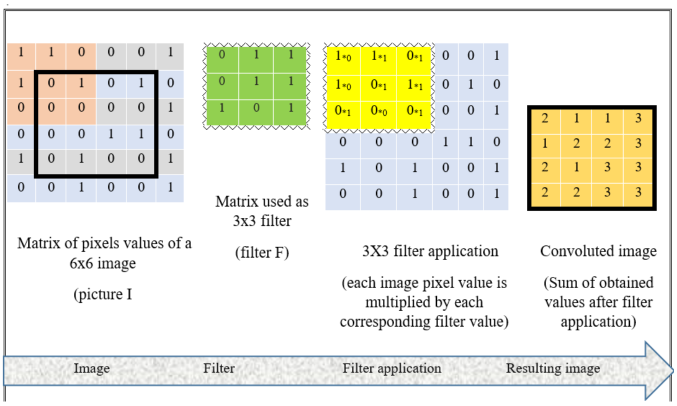


Fig 5: Application of convolution on a Sentinel image

The result of this operation is represented by a so-called "convolution" layer. The dimensioning of the size of the convolution layer depends on the depth of the layer (number of neurons associated with a matrix of pixels, step or stride (The smaller is the size of the matrix of pixels of the image, the more the matrices overlaps and the larger is the output volume) and zero margin or zero panding (controls the output dimensions of the size of the output matrix) of the convolution layer.

As part of this study, the convolution makes it possible to highlight the limits of hydro-agricultural developments. From a Sentinel image We apply 3*3 mathematical filters to identify development boundaries.

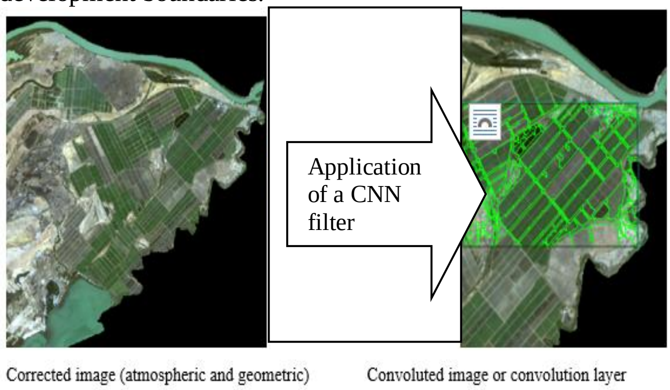


Fig 6: Applying a convolution layer filter

After applying a filter and obtaining the convolution layer, in most cases, the latter is followed by an activation layer in order to modify the output characteristic images called rectified images.

E. Activation layer

It is a mathematical function used in signal processing. The activation function of a neural network attempts to replicate the process of biological neuron activation in humans. It allows the neuron to pass information if a stimulating threshold value is reached. Thus, it decides whether or not a neuron is activated.

There are several activation functions that are used after the convolution operations of a neural network. We distinguish:

- Linear function (units of outputs identical to their input level)
- Hyperbolic tangent function ()
- Sigmoid function
- Gaussian function
- Etc.

The activation function is based on a weighted sum calculation, It allows the activation of the neuron i.e. the production of an output message (ChaireAgroTIC, 2018). The choice of a function depends on the problem studied.

Our choice is the linear rectification function or Rectified Linear Unit (ReLU). Indeed, it is one of the most popular functions currently. It allows a faster training compared to the sigmoid and tanh functions, being lighter. It is widely used for CNNs, RBMs, and multi-perceptron networks (Momotoculteur, 2018).The ReLU is the simplest and most used activation function. It gives x if, x is greater than 0 and 0 if x is less than 0. In other words, it is the maximum between x and 0. The output range of values is between 0 and +∞. Le choix d'une fonction dépend de la problématique étudiée.

Our choice is based on the linear rectification functions or *Rectified Linear Unit* (ReLU). In fact, it is currently one of the most popular functions. They also enable a rapid workout compared to the sigmoid and tanh functions which are lighter. It is widely used for CNN and RBM.

ReLU is the simplest and most widely used activation function. It gives x if x is superior to 0. In other words it is the maximum between x and 0. The output interval between of values is between 0 and +∞.

$$ReLU(x) = max(x, 0) \iff ReLU(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

ReLU is a linear rectification function unit. It allows the transformation of the input from a weighted sum to the output. It is a question of transforming into 0 all the negative input values and of keeping the positive values at the output.

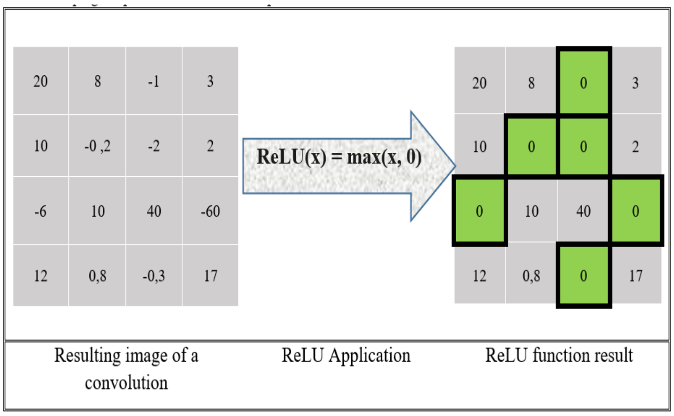


Fig 7: Application of linear rectification function

The ReLU function is applied after each convolution operation. Its objective is the non-linearization of responses.

F. The Pooling or sampling function

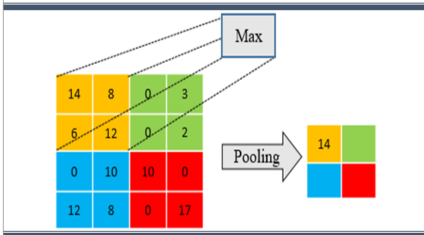
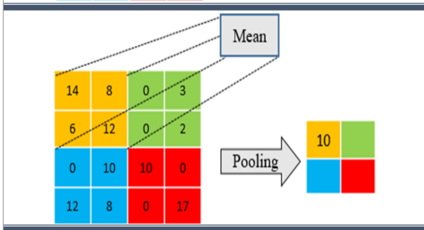
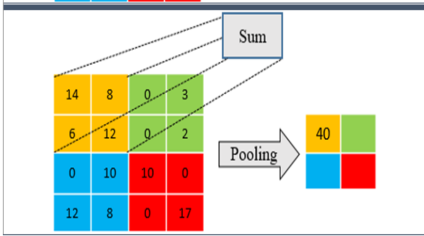
Pooling is an operation of convolutional neural networks. It consists of reducing the size of entered data. However, reducing the data size should preserve the characteristics of the image. There are two methods to apply pooling:

- Extraction of the highest value from an area of the image or a matrix (max-pooling)
- Calculation of the average values of an area of the image or of a matrix (mean-pooling)
- Calculation of the sum of the values of an area of the image or of a matrix (sum-pooling or global pooling)

Pooling is a subsampling operation that helps saving calculation time. It consists of extracting the value of a pixel from an image or calculating the sum or the average of the pixel values. It is said to be global pooling or Sum-pooling when it is based on an approach of pooling all the representations of the nodes into one. Sum-pooling then consists of calculating the sum of all the nodes of the neural network (Itoh et al., 2022). The mean-pooling or Average-pooling is an operation which consists of calculating the average of the values of the pixels of an area or a matrix.

The main advantage of mean pooling is that it is effective when we want to detect weak signals. Max pooling is effective when we want to detect strong signals, such as objects for example it allows the model to be invariant to translations (Pibre et al., 2016).

Table 2: Pooling types

Type	Principle
	Each pooling operation selects the maximum value of the satellite image area
	Each pooling operation calculates the average value of the of the satellite image area
	Each pooling operation calculates the sum of the values of the surface of the satellite image

Pooling is a layer usually placed between two convolution layers. It receives feature maps as input and applies the pooling operation to each layer. The latter seeks to cut the image into regular cells and to keep within each cell the maximum value for the max-pooling and the average value for the mean-pooling.

Max Pooling suppresses noise. On the other hand, Mean Pooling simply performs dimensionality reduction as noise suppression mechanism. Therefore, in practice, Max Pooling is used much more than Mean Pooling since it works better (Zhou et al, 1988).

In this part we apply the max pooling that allows to generate a pixel which better reflects the values of the pixels of a window. Max-pooling is similar to convolution. It consists of dragging a window of two (2) or three (3) pixels above an image. To illustrate this process, we use a window of two (2) pixels and steps of two (2) to avoid the overlapping of the windows. The window size is called "kernel size" and the steps are called "strides". For each step, we take the largest value present in the window and this value represents a new pixel in a new image. This is called Max Pooling.

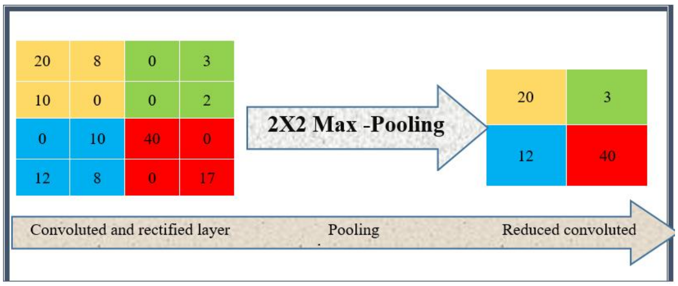


Fig 8: Applying Max pooling function to a convoluted and rectified layer. A 2x2 filter and a step of 2 are used.

The pooling layer helps to reduce the number of parameters and calculations in the network. It improves the efficiency of the neural network and avoids overlearning. Applying mean-pooling reduces the number of segments while providing clear limits around plot developement. The limits of plot developements in the final image provide better contrast with the surrounding features.

G. Connection function

The connection layer (or fully connected layer) is the last layer of the architecture of a neural network. It comes in a flattened form and is connected to all neurons. It can be used to optimize goals such as class scores. It makes it possible to define the limits of plot developments and to classify an image or images entered into the network.

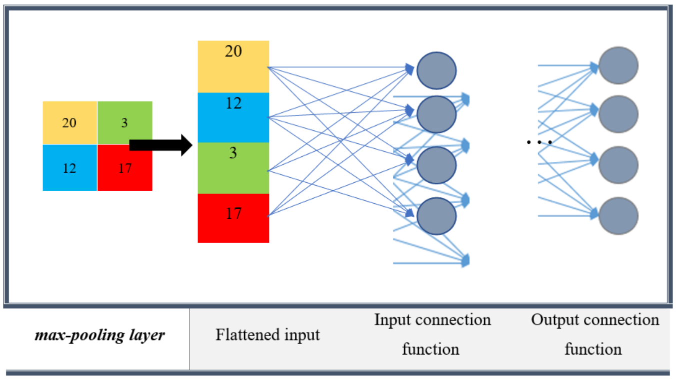


Fig 9: Application of layer connection function

H. CNN product design

A neural network is a collection of several interconnected layers. It is a powerful model for image classification, object detection, etc. It has two distinct parts.

As input, an image is provided as a matrix of pixels. It has 2 dimensions for a grayscale image. Color is represented by a third dimension, of depth 3 to represent the three (3) RGB of a color image. In practice, for an RGB image, the input is not a grid of values but a grid of spectral bands corresponding to the different input channels or to feature maps in hidden layers (Salti, 2018).

As output, several categories of layers, each layer of which represents a neuron.

The convolutional neural network is simply a stack of successive layers of convolution, activation and sampling (pooling) and fully connected layers (multilayer perceptron).

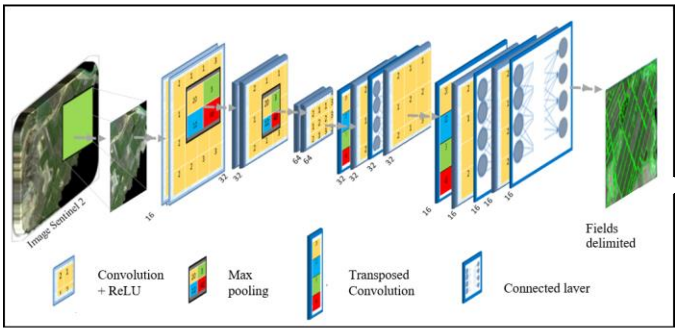
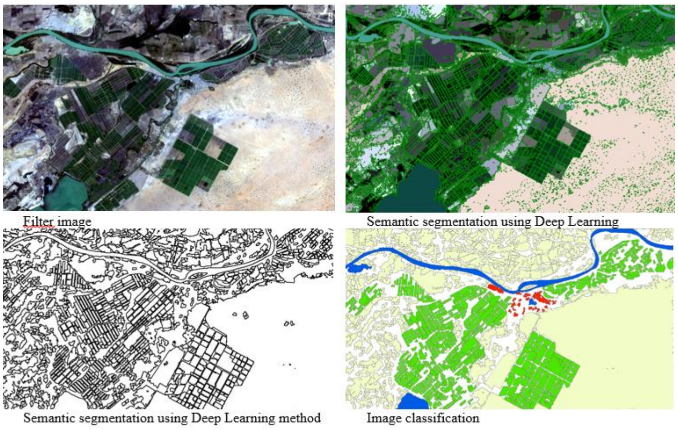


Fig 10: Design of the convolutional neural network used for the automatic parcelling of rice plot developments.

This design represents a multitasking learning model. Convolutional layers are in yellow, pooling layers in green-yellow-red-blue, and connection layers in blue-white.



Map 4: Hydro-agricultural developments in the communes of Fanaye, Ndiayenne Pendao and Guédé Village in 2020

IV. CONCLUSION

This paper has shown an approach to semantic segmentation of Sentinel 2A images using algorithms from artificial intelligence techniques. We have illustrated the automatic parceling process with all the operations of applying a convolutional neural network.

The results obtained show that the method used is well suited for the delimitation of rice paddocks in the Senegal River valley. They provide a database of real areas with a better accuracy of hydroagricultural plots. The process developed constitutes an important step in the periodic updating of the areas exploited.

In addition, the results of the study show that the Sentinel 2A data offer many opportunities for managers of hydroagricultural schemes, including the possibility of having a database useful for calculating rice yields.

In short, operations on large masses of data are made possible with deep learning. Parceling, enlargement and other mutations of agricultural plots are well taken into account in the process of automated delimitation of plot outlines. This automatic delimitation is possible thanks to the use of a convolutional neural network.

REFERENCES

[1]. *Cheikh Sadibou FAYE, Hamath Ndiaye, Papa Meissa Diop, "Rice value chain in the communes of Fanaye, Ndiaye Pendao and Guede Village: Demographic characteristics, Farmers' organizations, partners and infrastructure and equipment", 2016, pp 83.*
[2]. *Lionel Pibre, Marc Chaumont, Dino Ienco, Jérôme Pasquet, "Étude des réseaux de neurones sur la stéganalyse", 2016, pp 7.*
[3]. *Nicolas Audebert, Alexandre Boulch, Bertrand Le Saux, Sébastien Lefèvre, "Segmentation sémantique profonde par régression sur cartes de distances signees", 2018, pp 9.*
[4]. *Takeshi D.Itoh, Takatomi Kubo, Kazushi Ikeda, 2021 "Multi-level attention pooling for graph neural networks: Unifying graph representations with multiple localities", Vol 45, 2022, pp 356-373.*
[5]. *Samuele Salti, "Introduction to deep learning, Lesson 2 - Convolutional Neural Networks", cours, 2018, pp 61.*
[6]. *Djamel Belhaouci, 2019, "Démystifier le Machine Learning, Partie 2: les Réseaux de Neurones artificiels", cours, 2019.*
[7]. *Mathieu Carpentier, "Classification fine par réseau de neurones à convolution", 2019, pp 77.*
[8]. *Yosr Ghanney, "l'apprentissage profond pour la détection des signaux d'interférence radio fréquentiels", 2019, pp 74.*
[9]. S. Paul, A. Louges, M. Bonneval, L. Pontier. *Dictionnaire de Télédétection aérospatiale / Airborne and spaceborne Remote sensing Dictionary* (1982), 1985, pp 241