

Automated system for the detection of risk in agricultural sugarcane harvesting using digital image processing and deep learning

Alina E. Badillo-Márquez, Jonathan J. Cid-Galio, Rubén Posada-Gómez, Alberto A. Aguilar-Lasserre

Division of Research and Postgraduate Studies, Tecnológico Nacional de México; Instituto Tecnológico de Orizaba, Mexico

Abstract

In recent years, significant changes have been presented in the climatological trends due to climatic change, originating negative impacts on the agricultural production, diminishing mainly the harvest efficiency. The following research proposes the optimization of the agricultural risk identification method for the prediction of the variables: temperature and precipitation; the risk identification method was developed through the digital image processing technique (DIP) and deep learning (DL); subsequently, with the processed images, convolutional neural networks (CNNs) were developed for the detection of areas where there is a potential risk in the sugar cane crop harvest in the southeast of Veracruz in Mexico. The efficiency of CNN detects temperatures over 38°C and the levels of precipitation under 70 millimeters. The efficiency of network detection is 0.9716 and 0.9948 for predicting the temperatures and precipitation variables, which represent a solid basis for detecting zones that depict a risk for the sugarcane harvest.

Introduction

The effects caused by climatic stress represent a possible threat to the ecosystem, principally in the agricultural systems. At this moment, the changes in climatic behavior have caused a

decrease in agricultural production, risking food safety on a global level (Beer, 2018). Under a climatic approach, the danger is a manifestation of a threat in a vulnerable system; when the danger turns into a threat, the impacts that potentially increase the risk of a system become a reality. A risk in a system increases the sensibility, making it more vulnerable (IPCC, 2007).

The sensibility of agriculture beholds climatic conditions reside in alterations in the behavior of factors such as temperature, precipitation, wind speed, and rise in the incident of extreme climatic events that affect the development of cultivation. Inside of extremes phenomena, we find droughts, extremes rains, hails, and cyclones, causing harvest reduction, soil degradation, and infertility on the cultivation's land caused by adverse effects originated by water reduction and hydrological stress, among others (Fowler *et al.*, 2003).

As a result, the efficiency of the agricultural systems is related inherently to climatic variations (Torres Lima *et al.*, 2011), and in recent years it has emerged the necessary to develop investigations that allow analyzing the stress generated by physical impacts that can put in danger the harvests (Eakin *et al.*, 2004).

Previous studies about the vulnerability of the sugarcane before the climatic change developed a model base on agents that allows for emulating the knowledge and production expertise; it determined that the temperature and precipitation variables raise the potential risk of losing the harvest of the sugarcane in the study area (Badillo-Márquez *et al.*, 2021).

To identify meteorological dangers, it is necessary to conduct a diagnosis scene comprising a dynamic valuation of the risk of the exposed zone. The delimitation of the exposed zone to the risk is defined by a spatial scale based on the nature and available information (Cardona, 1992). As opposed to other qualitative scenes, the quantitative risk scenes allow identifying the magnitude of possible future consequences on and off the system through a top-down type approximation where it is proposed to section by region the developed scenes in order to show the impacts on a global level; meaning, it seeks to go from the general to the particular achieving a more focalized understanding of the system (Pielke *et al.*, 2012; Magaña, 2013).

Given the edaphic and climatological conditions, Mexico is also found as one of the most possibly vulnerable countries in Latin America behold the impacts generated by climatic change due to its geographic location rising its exposition before meteorological phenomena which alter mainly the agricultural sector, the hydrological sector, the biodiversity and infrastructure (Monterroso *et al.*, 2015).

Lately, artificial intelligence techniques have been installed in the development of expert systems to find a solution to problems and learning that help during the decision-making process (Díez *et al.*, 2001; Sosa, 2007; Ponce Cruz, 2010). Techniques such as image processing aim to find and draw characteristics of interest (pixels regions) that make a beginning function creating a search algorithm to identify behavioral patterns (Dalal and Triggs, 2005; Montoya Holguin *et al.*, 2014; Kaur, 2016). The objective of the

Correspondence: Alberto A. Aguilar-Lasserre, Division of Research and Postgraduate Studies, Tecnológico Nacional de México/ Instituto Tecnológico de Orizaba, Av. Oriente 9, 852. Col Emiliano Zapata, 94320 Orizaba, México.
E-mail: albertoaal@hotmail.com

Conflict of interest: the authors declare no potential conflict of interest.

Funding: none.

Received: 20 March 2023.

Accepted: 9 August 2023.

©Copyright: the Author(s), 2024

Licensee PAGEPress, Italy

Journal of Agricultural Engineering 2024; LV:1581

doi:10.4081/jae.2024.1581

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Publisher's note: all claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

research is to develop, train and test a convolutional neural network (CNN) capable of processing data from the GOES-East Geostationary Satellite. The CNN will be used as an agricultural risk identification method to predict areas where there is vulnerability for sugarcane cultivation by monitoring the variables: temperature and precipitation. To achieve this objective, a database of satellite images needs to be specifically preprocessed and interpreted to train the CNN to evaluate areas where a thermal stress or aridity condition is predictable.

Firstly, this study revised the investigation literature that has implemented the use of image processing in the agricultural sector is displayed, given that the impact of the climatic change in the agricultural supply chain has intensified by the growing food demand and the increase of population density in the last 20 years; affecting principally on the cultivation's land and the decrease of hydrological resources. On the other hand, uncontrolled agricultural development and biofuel use have originated environmental impacts (greenhouse gas emissions, water pollution, deforestation, loss of biodiversity) and socioeconomics (Wood *et al.*, 2004).

The proposed method is discussed in this study, which consists of two phases. The first stage involved satellite image processing for each variable of interest in the study area using the digital image processing technology (DIP). In the second stage, deep learning (DL) was used to CNN capable of detecting risk zones. In addition, the investigation's contributions to agricultural risk identification in sugarcane agriculture are shown, as is the suggested investigation's contribution to risk estimation.

Background

The effects generated by the climatic change have brought severe environmental consequences that impact agricultural systems (Paloviita *et al.*, 2016). These impacts are funded mainly by variable changes such as temperature and precipitation levels (Mitter *et al.*, 2015).

Currently, some scenarios indicate that for the next fifty years, temperature changes are expected to keep rising; on their behalf, unforeseen precipitation levels regarding magnitude and frequency. As a result, these changes significantly impact cultivation, leaving grains crops one of the most affected because of phenomena such as droughts and heatwaves.

Studies developed concerning thermal stress on cultivation indicate that even moderate growth on temperature reduces the cultivation efficiency. High nighttime temperatures can influence the efficiency negatively because of the acceleration of the cultivation development, altering the reproductive phase, and as a result, its efficiency decrease; among the most affected cultivations by temperature increase are corn, wheat, sorghum, bean, rice, cotton plant, and peanut (Araus *et al.*, 2008; Barnabás *et al.*, 2008; Semenov and Halford, 2009)

Consequently, it generates a more significant impact on blooming, reproductive phase, and cultivation performance by its influence on farming physiological process just as in the pollination process and photosynthesis in a ground level generating crop losses.

In recent years, food systems studies have emerged to strengthen and increase food safety. Around the investigations related to the risk of food system events are model developments to predict climatic changes. To elaborate the model, the authors took as a baseline the species distribution modeling (SDM) (de Carvalho-Pinto *et al.*, 2007; Fourcade, 2016), which is fed by information

systems about the present climatic biodiversity in a determined area, nevertheless; within the main disadvantages of this type of model is the bias of datum whereby is necessary to resort to alternative information resources in order to decrease the bias level. One alternative is dispensing maps of climate types proportionated by the International Union for Conservation of Nature to make predictions about the current and future weather. The model was developed based on expert knowledge in which six categories were defined to describe the climatic diversity based on impact factors: "temperature and precipitation" to minimize the risk of bias during the model process.

The six categories are average annual temperature, maximum temperature, minimum temperature, annual precipitation, precipitation of the most damped month, and precipitation of the driest month. Using a geographic information system (GIS) allows for obtaining a geospatial database, which can be applied in areas such as agriculture and the administration of natural resources. Currently, geospatial datum systems have been proposed to ease the storage and the interchange of geographic information to develop models that help implement and keep improvement actions in the study system (Vázquez *et al.*, 2015).

The studies developed are the modeling and simulating climatological phenomena of precipitation, water evaporation, and groundwater flow obtained by satellite images (Espínola *et al.*, 2016). The objective of the model is to estimate the water supply for a population and to establish future construction and urban planning projects in locations with a high probability of flooding. The satellite images are proportionated by digital elevation model and remote sensing studies. The central meteorological study phenomenon is precipitations; being a virtual model is crucial in estimating various environmental changes, for example, the formation of new rivers or soil erosion.

Maguire (2013) presents a system to supervise and register data from the indicators of climatic change based on ice satellite observations on north Europe zones and the Arctic. It also includes projections of rains, temperature, and average wind speed from Europe. Subsequently, a surveillance satellite of climatic indicators was implemented, obtaining results performed by EuroClim (A European Project about Climatic Change and the Prediction Surveillance System). On their behalf, (Pulighe and Lupia, 2016) proposed a methodology or algorithm to discover cultivation polygons in the urban zone in Rome through a photo interpretation of the images on Google Earth, combined with additional data deviated from auxiliaries sources. The use of Google Earth allows obtaining multiple photos in "real time" to analyze the study phenomenon's spatial pattern, allowing an urban agriculture spatial information system to distinguish the different existing typologies in the area. For all that, the usage of satellite images has contributed to a large extent to the development of techniques that help towards the development of agricultural systems, being precision agriculture the most beneficial modality (Lu and Young, 2020).

Techniques such as Deep Learning allow the construction of more complex models for developing a cultivation support system where the production system presents some difficulty; in this sense, the authors (Fredys *et al.*, 2021) developed a model of a CNN through aerial photos captured by drones. The developed work aimed to identify the four growing stages (seedbed, sowing, develop and cut) by the Robellini palm. One of the main management problems of the crop is the length of the plant, which exceeds 5 meters; therefore, with the development of the model, an increase in the quality of the harvest and a plant length within the standards was obtained.

Hornig *et al.* (2020) proposed a harvest system using the

Internet of Things (IoT) and image recognition to determine the maturity of the crops before being picked with the help of robotic arms. For the development of the system, the authors employ the coordinates of central pixels obtained on the image identification as the outlet of the robotic arm system running the harvest. The obtained results were an average precision of 84% for the training phase, while the robotic arm showed an efficiency of 89%. Within the obtained results by the monitoring systems of climatic change, it has been identified strategies for effective communication from the proportioned information by the climatic satellite images to evaluate how the ergonomic methods can be adapted for a particular application, proving how the early identification through satellite images help to reduce the negatives impacts.

The research analyzed focuses on the development of models and techniques to monitor crop growth stages. However, it is important to further study the relationship between crop growth and climatic factors to reduce crop risk. The optimization method that was developed in this investigation employs image processing under an automatic learning approach to develop CNNs that function as a support system in the decision-making process to improve the traditional agricultural methods by optimizing the estimation zones that could put in danger the harvest cultivation of the sugarcane. The development of the networks focuses on the risk identification zones on temperatures higher than 38°C and precipitation levels below 70 millimeters because values out of that range can affect the cultivation performance.

processing method was performed to identify zones of interest that could put in danger the agricultural cultivation through the development of an algorithm; in stage 2, DL was applied to optimize the identification method through CNN to predict risk zones for the harvest of sugarcane for each variable of interest.

Variable selection: temperature and precipitation

In developing the risk identification method, two variables of interest were selected to be evaluated: temperature and precipitation. The temperature is a factor associated with growing and plant growth, having optimal values between 25-38°C; values out of this range affect the regrowth or formation of the sugarcane. Values above 38°C reduce the photosynthesis process of the cultivation, altering the sucrose concentration (PRONAC, 2009; Stockdale and Watson, 2009). The hydrological requirements in the growing stage of sugarcane cultivation are higher than in other crops. Between 50-70% of water is distributed in the first 30 cm; this zone contains a higher absorption of nutrients, and water exists. During drought periods in which conditions such as low humidity levels prevail, the root of the crop tends to grow more in depth seeking hydrological resources, which originated a higher concentration of water on the root and to a lesser extent in the rest of the plant reducing the production capacity of the cultivation (CONADESUCA, 2018).

Images compilation

In Mexico, the agricultural systems have access to geographic information services that display a detailed way of desert zones and soil vegetation (SEMAR, 2015; Seminis, 2016). For this investigation, it was decided to use the geospatial information that was obtained by Agro-climates (CONAGUA, 2020), inasmuch that it has satellite information of the variables of interest that are evaluated: temperature and precipitation. The compiled images were

Materials and Methods

This section describes the developed methodology of this investigation (Figure 1), which consists of two stages to develop an agricultural risk identification method. In stage 1, the images pro-

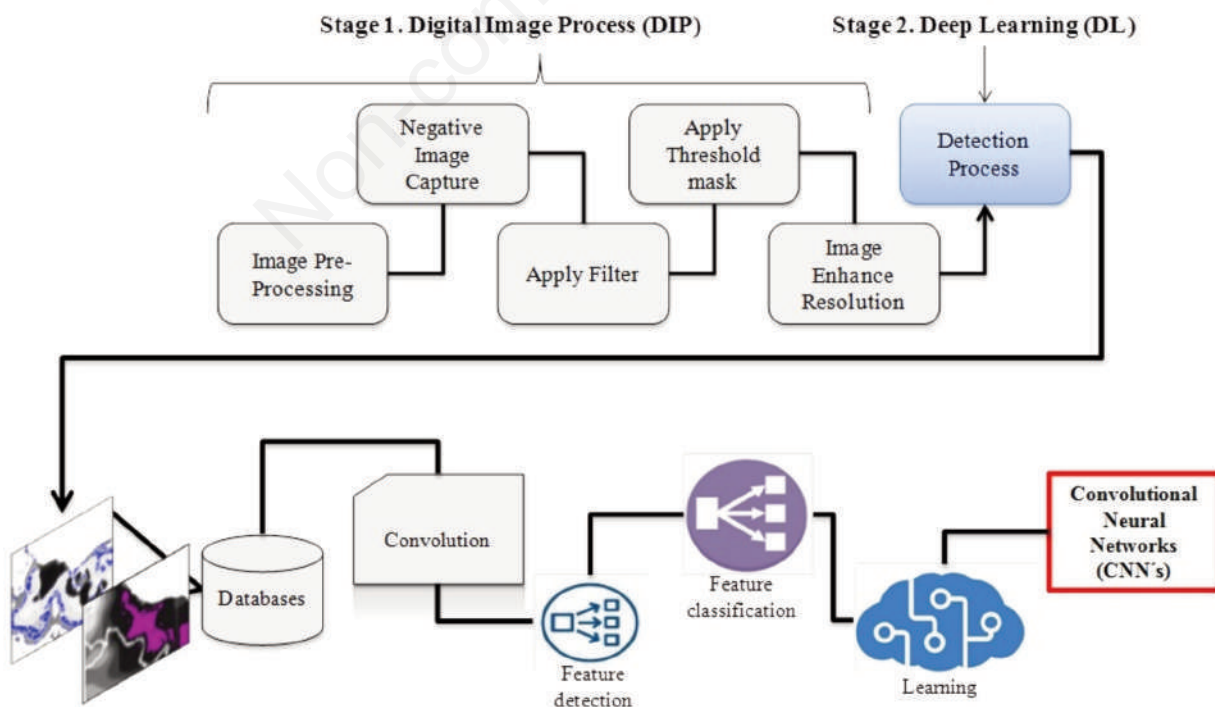


Figure 1. Identification process methodology.

obtained from the Meteorological National System, produced by the Geostationary Satellite GOES-East (Sistema Nacional Meteorológico, 2020). The area of interest corresponds to the southeast region that includes the state of Veracruz and the study area, where the GOES-16 ABI RGB CONVECCION was selected since it shows the highest image sharpness and resolution for analysis purposes.

The image compilation period was about five months, obtaining a sample of 280 images since a duplicate sample was performed for each variable: temperature and precipitation. The database consisted of a sample of 280 valid images, representing 45% of the images captured, of which 343 images were discarded due to the presence of clouds. The sampling period was related to July and November because the sugarcane is blooming before being picked in the Zafra season in the state of Veracruz (CONADESU-CA, 2018).

Figure 2 indicates the corresponding interest zone to be evaluated for each color for the temperature [R:255, G:216, B:123] and precipitation [R:157, G:249, B:184] values. The corresponding scales for each image point to the range of interest for the temperature value that varies from the orange color range onwards ($>32^{\circ}\text{C}$) since the values above that temperature could risk the cultivation. For the precipitation variable, the values below 70 mL (in orange color) caused water concentration mainly on the root and in minor proportion in other parts of the plant, decreasing the productive capacity of the cultivation.

Stage 1

Image processing

A set of n images needs to be interpreted with the aim to have a dataset useful to train the neural network. In order to improve the interpretation some common techniques of visualization and were applied. To reduce the noise of the image an average filter with a kernel of 9×9 pixel was applied to all the images. Also, a low pass filter was applied to limit the investigation to the zones of interest. At the last a segmentation of each image was carried on. The methodology applied (Figure 1) for stage 1 features five steps, which are described below.

Image pre-processing

Image pre-processing removes the background interfering with

the RGB channels of each image (Wang, 2014). The image preprocessing was performed in the MatLab[®] software student version, in the “Apps” section with the threshold tool on the $L^*a^*b^*$ channel which allows to represent the color of an object in a uniform way improving the visual perception.

Subsequently, image loading preprocessing was performed in MatLab[®]. Figure 3 a,b presents the background conversion in the $L^*a^*b^*$ channel of the study area for the temperature and precipitation variables.

Negative image capture

In negative image capture, the image is binarized by detecting the RGB color scale in which the pixels in the area of interest oscillate. The objective of the application of the negative filter is to avoid the presence of points saturated by brightness that interfere in the image processing allowing a better interpretation of the image values.

$$Neg(i) = MVP(i) - APV(i) \quad (1)$$

where:

Neg is the negative channel; I is the color channel; MVP is the maximum pixel value; APV is the actual pixel value.

Later, a white balance correction was applied to the negative image to perceive the real color better and increase the difference between the values of the pixels and the monolayer and background.

For the pixel identification in the area of interest on the RGB channel, the function *imtool* by MatLab[®] was used. Afterward, the RGB scale was applied to determine the threshold at which the pixels vary in the risk zone in both variables. Figure 4 a,b displays the obtained results on RGB channels for the temperature and precipitation variables, respectively.

The obtained values on the RGB channel for the temperature variable are: [R=247; G=217; B=131], whereas the precipitation variable values are [R=156; G=246; B=194]. In Figure 4 a,b, for the temperature variable, it can be observed that in the red channel (R), the target area approaches quicker to white, indicating that this is mostly composed of red, whereas the precipitation value, the target area is the green channel (G), where it approaches quicker the white color. Therefore, for image processing, the red channel is

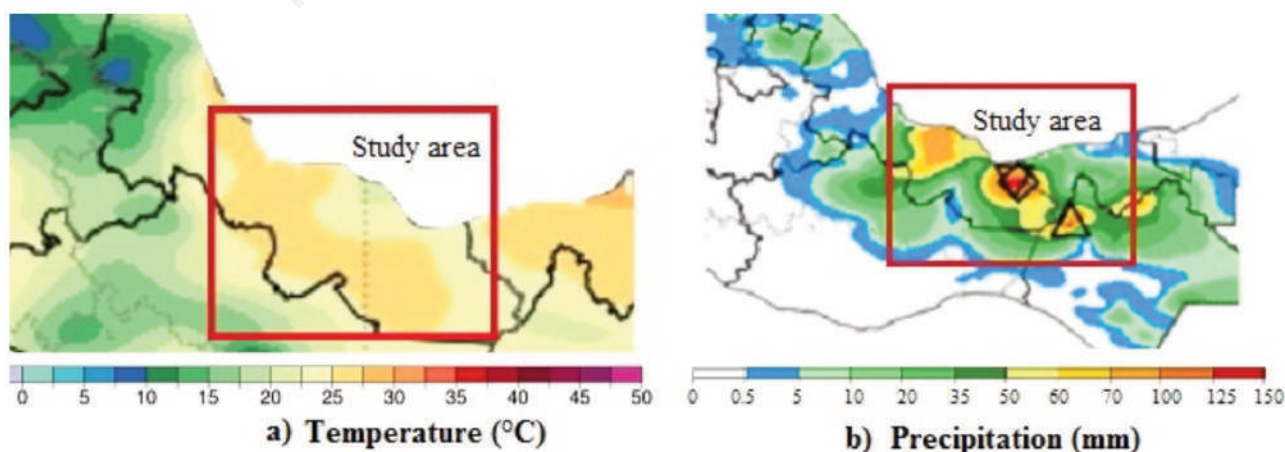


Figure 2. Satellite images of the variables of interest.

selected for temperature and the green for precipitation.

In Figure 5, the evaluation of the specific interception point is observed between the risk zone and the other areas in their RGB values after the negative capture of each variable was applied, obtaining an increased definition of the study zone separation regarding the surrounding regions, allowing the software to

achieve a better comprehension of the colors to section and identify the risk zones. Later, achieve a more optimal training of the neural network.

The negative image capture analysis allows a better interpretation of the image, where the target increases its definition, easing a better interpretation of the contained colors in the image.

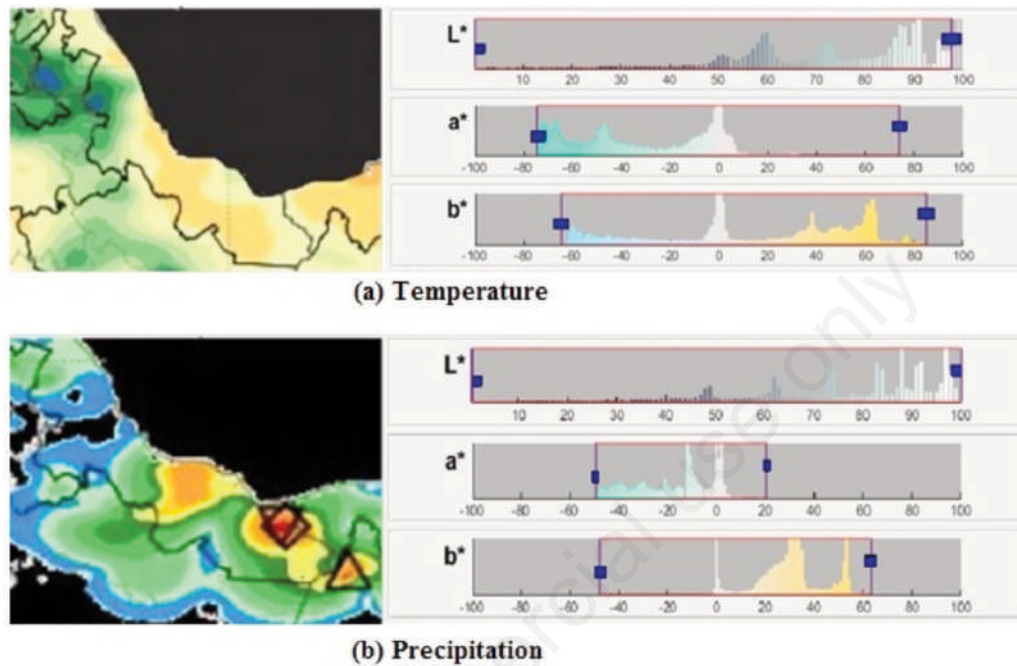


Figure 3. L*a*b* channel in MatLab® for the variables of interest.

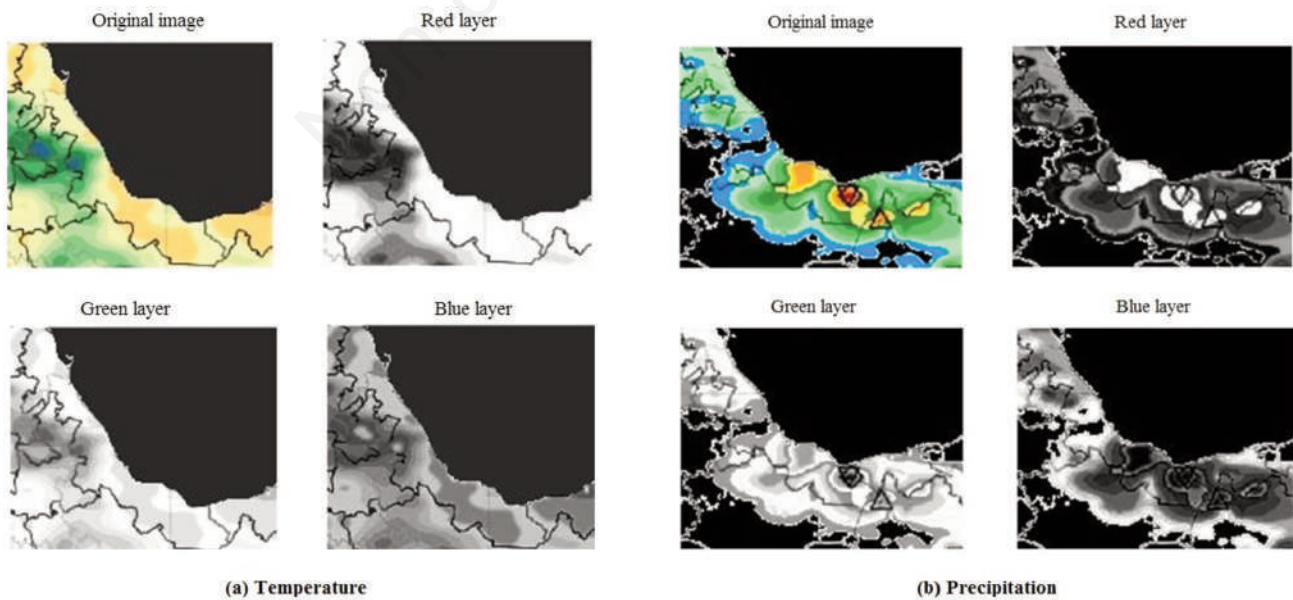


Figure 4. RGB channel for variables of interest.

Filter application

The satellite images have a certain amount of noise, which makes it difficult for the process of acquisition of information. For noise reduction, filters are applied to improve image quality. The average filter method is a grouping method where the average of a vector is calculated [i, j]; this process seeks a grouping of a dominant color pattern within a matrix, creating a two-dimensional filter to standardize the colors levels of a given region, obtaining the average of the pixels and the noise is standardized to the closest absolute value of the average.

Different sized vectors were tested by “trial and error,” in which it was determined that the average within a vector [10 10] proportion the optimal value reducing a large portion of image noise. This average calculation gets repeated in the entire image to homogenize the colors. Hereunder the employed equation for the average filter is displayed.

$$h = \left(n(1), \frac{n(2)}{n(1)+n(2)} \right) \tag{2}$$

where:

h is the Kernel correlation.

Figure 6 mainly compares the applied filters for the images of the temperature and precipitation variables. Figure 6 shows that resolution and detail decrease concerning the original images after applying filters due to the standardizing of the regions in a group of pixels, where the boundaries that define the image are not dominant and are absorbed by their surroundings. Therefore, even though the image definition was reduced, the acquisition of information is more accessible since, simultaneously, the colors have been homogenized, reducing the noise.

Threshold mask application

Color spaces are considered a method capable of interpreting and expressing the color of an object by numerical notation. The International Commission on Illumination (CIE), known for its approach to illumination and color studies, has defined and studied different color spaces that help to describe the perception of millions of perceptible colors by the human eye. Some of the spaces defined by the CIE are CIEI XYZ, CIE L*C*h*, and CIE L*a*b*. The most employed color space is the CIE L*a*b* because, as was mentioned before in the paragraph “image pre-processing”, it is the most consistent space to represent the color of an object, which uses correlations of numerical color values, making it the most consistent space for the visual human perception.

In Figure 3 a,b the levels of color are displayed, where channel L* (brightness) is almost used ultimately, while for temperature and precipitation the dominant colors are red and green, respectively. To extract the information to identify the risk zones, 280 trials were performed with different images for each evaluation criteria (temperature and precipitation) and, in addition to that, determined standard color level in the color space L*a*b*; these values on the channels L*a*b* will be the boundary mask that allows identifying the color pixel regions (risk zones) contained in the monolayer applying the following equation:

$$TM(g) = \begin{cases} 0 & \text{if } g \notin APV \\ 1 & \text{if } g \in APV \end{cases} \tag{3}$$

where:

TM is the threshold mask; g is the L*a*b* channel value; APV is the actual pixel values.

The target of the threshold mask is to separate the contained

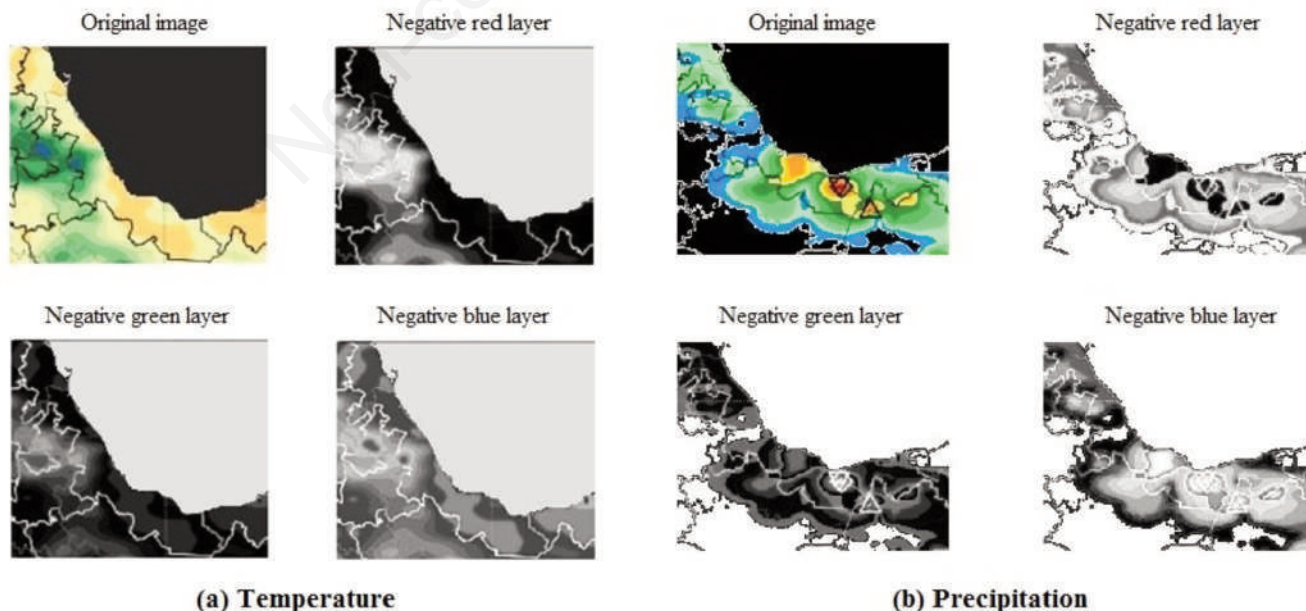


Figure 5. Negative image capture for the variables of interest.

colors in the study area listed as risk, extracting only the required information. Based on the logic “*If... then*”: if the color levels are different from the Threshold mask, then the color will be zero belonging to the black color; otherwise, the value of 1 will be placed belonging to the area of color interest. A threshold mask is a tool located in the MatLab® applications. During the threshold mask execution, an inverted mask was applied to highlight the zones of interest and enable image segmentation.

Image enhancement resolution

In this step, the features of interest were extracted for the acquisition of information through a 280 trials performance for each variable of interest on different images with risk zones to

determine the standard color levels within the L*a*b* threshold. Figure 7a displays the corresponding image to the temperature value after applying the mask. As can be, the shaded area in purple represents the risk zones for the cultivation with temperatures above 38°C. Figure 7b displays the risk zones for the sugarcane concerning the observed precipitation levels. The blue color represents values below 70 mL that could complicate the growing and reproduction cultivation due to the loss of soil fertility due to the presence or increase of arid areas.

Stage 2.

Identification process

The systems that are based on images identification focus on

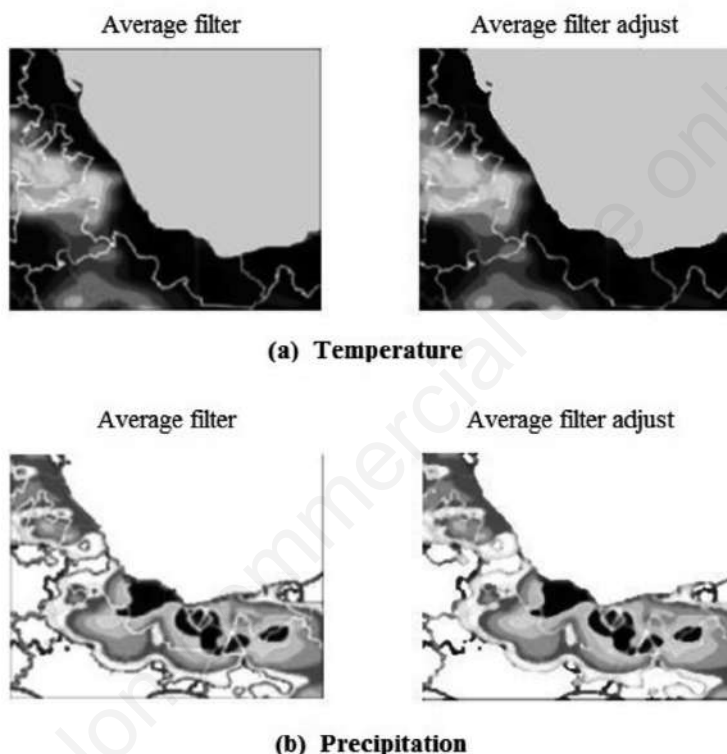


Figure 6. Filter Application on the images of the variables of interest.



Figure 7. Processed images of the variables of interest.

search and navigation within a datum base images, in addition to the recognition of certain features such as color, texture, and spatial location to a local or global level relying upon whether the image information of specific regions is ultimately desired (Aparicio Martín de Loeches, 2015). Even with image processing, the possibility of noise appearing in the image is present; therefore, it is necessary to implement an algorithm capable of reducing it as much as possible.

The DL process is based on CNNs training which extracts information through the interaction of layers learning to detect different features in a group of images. Figure 1 shows the methodology in stage 2.

Development of convolutional neural networks

Convolutional Neural Networks (CNNs) are models that allow identifying the features among thousands of images through object identification and learning.

The construction of the CNN was developed through transfer learning training, using an “AlexNet Network” in the MatLab® environment (student version).

AlexNet Network is formed by eight depth layers, 650,000 neurons with five convolutional layers, followed by layers of maximum grouping, and three fully connected layers with a final 1000-way softmax layer. AlexNet is a pre-training network with 1.2 million high-quality images capable of extracting information from thousands of different object categories (Krizhevsky *et al.*, 2017).

Variable labeling

Image processing was performed on the entire dataset (280 data) allowing the definition and identification of the zone of interest to obtain more accurate and trustworthy labeling for the training fulfillment of each CNN. The variable labeling was performed through the Image Labeler function in MatLab®; 120 images were labeled for the temperature and precipitation, respectively.

Each zone of interest was assigned one label per image to obtain a more substantial training base and provide the CNN training. The label assignment was performed for the sharpest zones. Figure 8 shows the label limit for the variables of interest: temperature and precipitation.

Convolutional neural network test

During the CNN test, the images were entered for the temperature and precipitation variables, which had previous processing. The selected test images were not the ones employed in the labeling and training of the CNN in order to check the efficiency and certainty of the developed networks.

Results and Discussion

The CNN of the temperature variable had two pieces of training to achieve higher efficiency. With the database of images processed of the variables of interest, the CNN training was per-

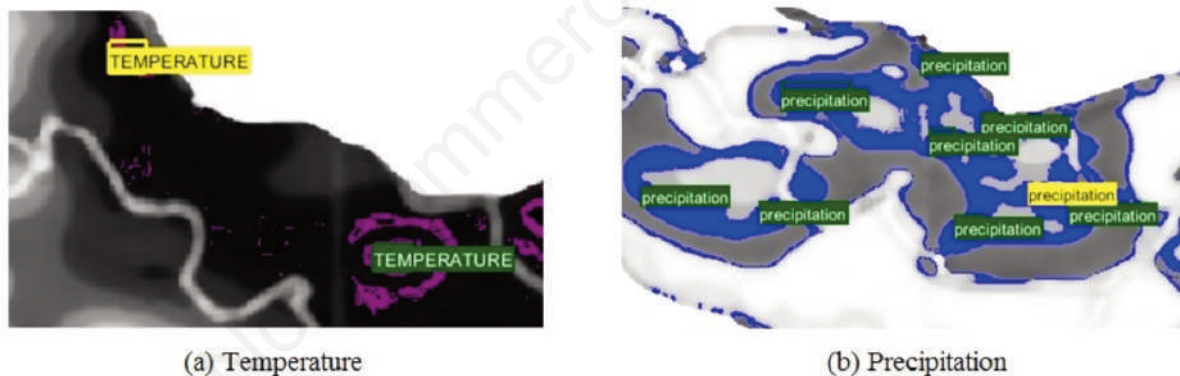


Figure 7. Processed images of the variables of interest.

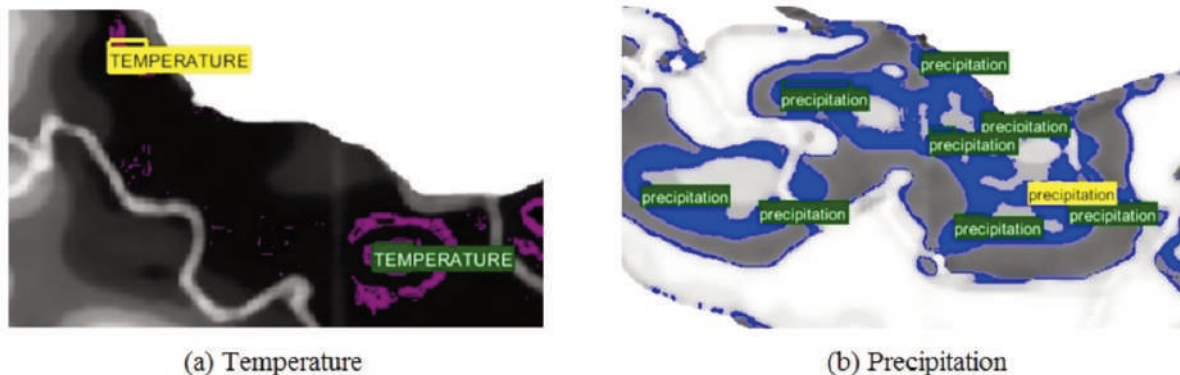


Figure 8. Label assignment for each variable of interest.

formed in MatLab®. In the first training, was obtained an efficiency of 93.7% on the tenth iteration. In the second training, a precision of 100% was obtained on the eighth iteration, 6.25% more effective than the first training. The network optimization times were minimized from 75 seconds to 57 seconds, resulting in 24% reduction of processing time (18 s). Five training sessions were needed for the variable precipitation network (CNN) to achieve high efficiency. It was decided to perform two more training to raise the CNN precision. At the end of the first training, a precision of ≈89.66% was obtained in 39 seconds. In the second and third training, there was an increment of 15% and 87% during the training time compared with the results of the first training achieving an increase up to 93.10% of the CNN precision (+3.44%). After the fourth and fifth training, it was achieved to reduce the time training to 33 seconds (45% less), and the accuracy and efficiency of the network were maximized, obtaining 96.88% and 100%, respectively. All pieces of training were completed in the sixth iteration.

Figure 9 displays the obtained results during the CNN’s trial period or learning phase. A high level of agreement was obtained; therefore, the network development is efficient and robust for identifying risk zones for temperature and precipitation, which could risk the cultivation growth and reproduction in danger. Figure 9b, corresponding to the CNN test for the precipitation variable, displays a higher prediction level than the temperature (Figure 9a) variable, which could result from a more excellent training performance in the CNN of the precipitation variable.

To verify the prediction efficiency of the CNNs through this technique, a group of 20 processed images was selected for each variable, which was not used in the network training to avoid software predisposition. The selected images display different size regions on the identified risk zones by CNNs to verify that the networks can identify risk zones on a lower and higher scale for each variable. Figure 10 obtained an identification level between 0.955 and 0.989 for predicting temperature zones with a risk for the cultivation, while the precipitation variable (Figure 11) obtained an identification level of 0.989 and 0.999. The following figures display the results of the reliable level in identifying the region of interest for the temperature (Figure 10) and precipitation (Figure 11) variable.

The displayed results by networks accomplish the needed features for identifying zones that represent a risk for the sugarcane harvest in the study area.

Discussion

The development activities of the sugarcane industry bring negative impact on the environment; for this reason, it has opted for the introduction of sustainable practices in cultivation development, efficient resources management, and proper waste and emission management. Nevertheless, even with the implementa-

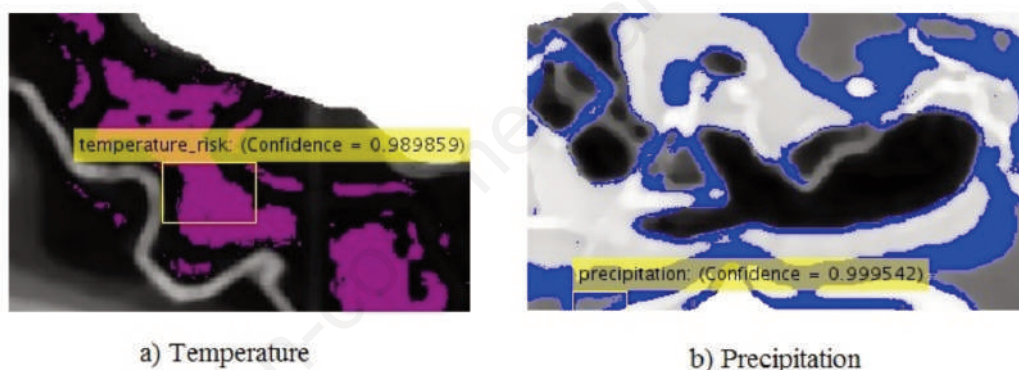


Figure 9. CNN’s prediction for each variable of interest.

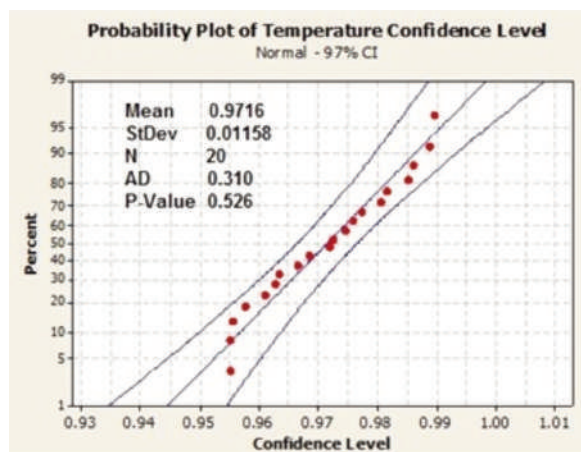


Figure 10. CNN reliable level for the temperature variable.

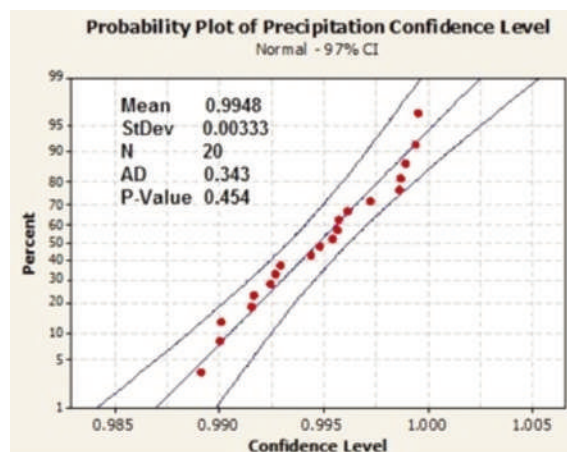


Figure 11. CNN reliable level for the precipitation variable.

tion of these measurements, it is a susceptible sector that is highly influenced by climate conditions. Changes in the behavioral patterns of climatic variables such as temperature and precipitation have caused an efficiency reduction of the sugarcane, originating an increase in production costs behind the growth of harvest requests (PRONAC, 2009; CONADESUCA, 2018). Uncertainty plays an essential role in performing projections and climatic estimations (Rosegrant *et al.*, 2009) Predicting uncertain parameters allows us to make estimations in behavior and how it could impact the system.

With the development of the identification method through image processing, it is possible to estimate the behavior of uncertain variables such as temperature and precipitation to determine the impact on the cultivation of interest in the study area. Optimizing the identification method through Convolutional Neural Networks allows presentation, search, and navigation functionalities through an images repository for the variables of interest: temperature and precipitation; in order to identify climatic stress zones where the sugarcane is cultivated. The system is based on identifying and recognizing certain features such as color and study zone, which allow identification of the harvest risk.

The main advantage of implementing the identification method is identifying the harvest risk during the plant's blooming stage. In the blooming stage, the plant reaches the reproductive phase allowing a greater cultivation efficiency.

Previously analyzed research have implemented the geospatial information used to ease the management and geographic exchange information for database creation. The development identification method developed in this research, on their behalf, allows the creation of an information repository that displays the existing climatic biodiversity in the study area in order to implement improvement actions and agricultural practices that allow to increase the cultivation performance of the sugarcane even under climatic stress conditions and with that reduce the harvest loss. The information of the developed identification method will help support the decision of the climatic conditions in cultivation growth for other Zafra seasons.

Currently, methods exist to classify and recover information in certain desired regions: systems that assign interest zones automatically (SDR) and systems with interest regions designated by the user (UDR) (Temniranrat *et al.*, 2021). On their behalf, the automated systems divide the study zones through algorithms to identify peaks, while the manual systems (UDR) are defined to offer an alternative in order to select the region of interest. The features were evaluated manually on a local scale to obtain more reliable and accurate results.

For object identification and learning, hundreds of images are necessary models with excellent learning capacity; CNN constitute these models. The CNN capacities can be controlled by changing its depth and extent, and it makes solid assumptions, and the majority of them are correct according to the nature of the images (in other words, the statistic seasonality and the locality of pixels dependencies). Hence, compared with standard feed-forward neural networks with similar-sized layers, the CNN has far fewer connections and parameters, making them easier to train (Krizhevsky *et al.*, 2017).

This algorithm performs the information extraction more quickly than training from scratch, which would require thousands of images to train a new image recognition network. The development of CNN allows emulating the human learning process by training a repository of previously processed images. The CNNs rely on DIP and DL to extract information effectively from the desired features. For the training of each CNN, a database of 30

processed images was performed for each variable.

In image pre-processing, in filter application, it was determined that the vector [10 10] provides the optimal value decreasing noise image in significant proportion for both variables of interest (temperature and precipitation). Using this vector size (optimal kernel) effectively reduces the noise in the image without affecting the image information. In vectors < [10 10] the background noise is almost not reduced; whereas with vectors > [10 10] the background pixel values invade the monolayer. For the creation of the threshold mask, the space CIE L*a*b* was selected; because it is oriented to describe color perception behold a change in some dimension of the image and facilitate the identification of risk zones during CNN's training and development; in addition, it is less sensitive to mistakes because the CIE L*a*b* ignores completely image brightness in channels a and b.

With the network efficiency test phase (CNNs), an identification of 0.989859 and 0.999542 was obtained to evaluate the parameters risk of temperature and precipitation, respectively. Afterward, during the networks validation process (CNNs), 20 tests were performed for each one, in which average efficiency identification values for temperatures of 0.9716 and precipitation of 0.9948 were obtained.

In recent years, statistic models and techniques to estimate the cultivation efficiency were implemented (Nagageetha and Ramesh, 2021); however, most of these are limited to cultivation characterization (seed, sowing, growth), leaving aside the physical environment in which the crop was developed.

The launch of monitored systems based on image processing and DL allows the development of decision support systems for the improvement of traditional agricultural practices, with the development and optimization of the method for timely risk identification in the harvest of the sugarcane. Timely predicting climatological risk on cultivation allows continuous follow-up to improve techniques and agricultural practices to increase harvest efficiency. The development of CNNs allows automatic learning to locate risk through identifying and classifying features in the image (color) listed as the risk zone.

Conclusions

The agricultural systems are continuously evolving due to their capacity to respond to climatic, social, and economic changes. The predictions of uncertain parameters allow for performing estimation on the behavior patterns and possible impacts within the system. Although there is research that provides models and techniques to monitor the growth and production stages of crops, it is important to delve deeper into the climatic interaction with respect to crop development in order to detect potential crop risk.

This research proposes an identification method capable of identifying areas of interest that could increase the harvest risk for sugarcane cultivation in the southeast region of Veracruz, Mexico. The identification method is formed by CNNs developed through DIP and DL for the prediction of the variables of interest: temperature and precipitation.

The efficiency of the CNN represents a solid base to identify temperatures above 38°C, where the photosynthesis process of the crop is reduced, altering the sucrose concentration and affecting the regrowth or sugarcane germination. The developed network for the identification of precipitation variables is adequate to identify zones with precipitation levels below 70 millimeters, which cause water to concentrate in the crop root and to a lesser extent in the

plant, reducing the production capacity of the crop; therefore, the networks achieve the needed features for the identification of zones that represent a risk for the harvest of the sugarcane in the study zone.

Due to all of the above, it can be concluded that implementing techniques such as DIP and DL provides the information extraction from desired features. The training network process (CNNs) was performed from a repository of previously processed images. The networks efficiency (CNNs) represents a solid base for identifying zones that depict a risk for the sugarcane harvest in the study zone. The network identification accuracy is 0.9716 for the temperature variable and 0.9948 for the prediction of the precipitation variable.

References

- Aparicio Martín de Loeches, A, Fernández Guzmán, L. 2015. [Reconocimiento óptico de caracteres en imágenes digitales de contadores de gas]. [in Spanish]. Universidad Complutense de Madrid. Available from: <https://docta.ucm.es/entities/publication/2073b262-9392-4619-88a3-b6ccdeff15e0>
- Araus, J.L., Slafer, G., Royo, C., Serret, M.D. 2008. Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant. Sci.* 27:377-412.
- Badillo-Márquez, A.E., Aguilar-Lasserre, A.A., Miranda-Ackerman, M.A., Sandoval-González, O.O., Villanueva-Vásquez, D., Posada-Gómez, R. 2021. An agent-based model-driven decision support system for assessment of agricultural vulnerability of sugarcane facing climatic change. *Mathematics* 9:3061.
- Barnabás, B., Jäger, K., Fehér, A. 2008. The effect of drought and heat stress on reproductive processes in cereals. *Plant. Cell Environ.* 31:11-38.
- Beer, T. 2018. The impact of extreme weather events on food security, p. 121-33. In: S. Mal, R. Singh and C. Huggel (eds.), *Climate change, extreme events and disaster risk reduction. Sustainable Development Goals Series.* Cham, Springer.
- Cardona Arboleda, O.D., Barbat, Á.H. 1992. [Vulnerabilidad y el riesgo desde una perspectiva holística]. [in Spanish]. Universitat Politècnica de Catalunya.
- CONADESUCA. 2018. [Programa Nacional de La Agroindustria de La Caña de Azúcar]. [in Spanish]. Available from: <https://www.gob.mx/conadesuca/documentos/temas-destacados>
- CONAGUA. 2020. [Agroclimas]. [in Spanish]. Available from: <https://smn.conagua.gob.mx/es/animacion-imagenes-desatelite?satelite=GOESEsteandnombre=Suresteandtipo=RGB>
- Dalal, N., Triggs, B. 2005. Histograms of oriented gradients for human detection. *Proceedings 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 886-93.
- Díez, R.P., Gómez, A.G., de Abajo Martínez, N. 2001. [Introducción a la inteligencia artificial: sistemas expertos, redes neuronales artificiales y computación evolutiva]. [Book in Spanish]. Universidad de Oviedo.
- de Carvalho Pinto, F.D.A., de Queiroz, D.M., Chartuni, E., Ruz, E. 2007. [Agricultura de precisión: nuevas herramientas para mejorar la gestión tecnológica en la empresa agropecuaria]. [Article in Spanish]. *Revista Palmas* 28:29-34.
- Espínola, M., Piedra-Fernández, J.A., Ayala, R., Iribarne, L., Leguizamón, S., Wang, J.Z. 2016. Simulating rainfall, water evaporation and groundwater flow in three-dimensional satellite images with cellular automata. *Simul. Model. Pract. Th.* 100:89-99.
- Fourcade, Y. (2016). Comparing species distributions modelled from occurrence data and from expert-based range maps. Implication for predicting range shifts with climate change. *Ecol. Inform.* 100:8-14.
- Fowler, H.J., Kilsby, C.G., O'Connell, P.E. 2003. Modeling the impacts of climatic change and variability on the reliability, resilience, and vulnerability of a water resource system. *Water Resour. Res.* 39:10-21.
- Fredys, A.S.H., Paez, J.A., Méndez, J.C., Garrido, F.B. (2021). Identifying Robellini palm growth stages through a convolutional neuronal network. *ARPN J. Engin. Appl. Sci.* 16:1402-1411.
- Eakin, H., García, C.G., Estrada, F., Álvarez, A.C.C. 2004. [Impactos potenciales del cambio climático en la agricultura: escenarios de producción de café para el 2050 en Veracruz (México), p. 651-60]. [in Spanish]. In: *Proceedings IV Congr. Asociación Española de Climatología, Santander.* Universidad de Cantabria.
- Hong, G.J., Liu, M.X., Chen, C.C. 2020. The smart image recognition mechanism for crop harvesting system in intelligent agriculture. *IEEE Sensors J.* 20:2766-2781.
- IPCC. 2007. *Climate Change 2007 - The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC.* Available from: <https://www.ipcc.ch/report/ar4/wg1/>
- Kaur, S. 2016. An automatic number plate recognition system under image processing. *Int. J. Intell. Syst. Appl.* 8:14-25.
- Krizhevsky, A., Sutskever, I., Hinton, G.E. 2017. ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60:84-90.
- Lu, Y., Young, S. 2020. A survey of public datasets for computer vision tasks in precision agriculture. *Comput. Electron. Agr.* 178:105760.
- Magaña-Rueda, V.O. 2013. [Guía Metodológica para la Evaluación de la Vulnerabilidad ante Cambio Climático]. [in Spanish]. Mexico City, Instituto Nacional de Ecología y Cambio Climático de México (INECC).
- Maguire, M.C. 2013. An analysis of specialist and non-specialist user requirements for geographic climate change information. *Appl. Ergon.* 44:874-885.
- Mitter, H., Heumesser, C., Schmid, E. 2015. Spatial modeling of robust crop production portfolios to assess agricultural vulnerability and adaptation to climate change. *Land Use Policy* 46:75-90.
- Monterroso Rivas, A., Gómez-Díaz, J., Lluch-Cota, S. E., Sáenz-Romero, C., Pérez-Espejo, R., Salvadeo, C.J., et al. 2015. [Sistemas de producción de alimentos y seguridad alimentaria]. In: C. Gay y Garcia, A. Cos-Gutiérrez and C.T. Peña-Ledón (Eds.), [Reporte Mexicano de Cambio Climático 2]. [Report in Spanish]. Programa de Investigación en Cambio Climático.
- Montoya Holguin, C., Cortés Osorio, J.A., Chaves Osorio, J.A. 2014. [Sistema automático de reconocimiento de frutas basado en visión por computador]. *Ingeniare. Rev. Chil. Ingen.* 224:504-516.
- Nagageetha, M., Ramesh, N.V.K. 2021. A survey of agriculture crop monitoring using iot based image processing and machine learning techniques. *Turk. J. Physiother. Rehabil.* 32:567-571.
- Paloviita, A., Kortetmäki, T., Puupponen, A., Silvasti, T. 2016. Vulnerability matrix of the food system: Operationalizing vulnerability and addressing food security. *J. Clean. Prod.*

- 100:1242-1255.
- Pielke, R.A., Wilby, R., Niyogi, D., Hossain, F., Dairuku, K., Adegoke, J., et al. 2012. Dealing with complexity and extreme events using a bottom-up, resource-based vulnerability perspective. In: A.S. Sharma, A. Bunde, V.P. Dimri and D.N. Baker (Eds.), *Extreme events and natural hazards: The complexity perspective*. American Geophysical Union. pp. 345-359.
- Ponce Cruz, P. 2010. [Inteligencia Artificial con aplicaciones a la ingeniería]. [Book in Spanish]. Alfaomega Grupo Editor, México.
- PRONAC. 2009. [Diagnóstico, Modelaje y Recomendaciones de La Fertilidad de Suelos Del Campo Cañero]. [in Spanish]. Programa Nacional de la Agroindustria de la Caña de Azúcar. Available from: <http://www.cndsca.gob.mx/documentoseficproductiva/8.PRONAC/PRONAC2014-2018.pdf>
- Pulighe, G., Lupia, F. (2016). Mapping spatial patterns of urban agriculture in Rome (Italy) using Google Earth and web-mapping services. *Land Use Policy* 59:49-58.†
- Rosegrant, M., Koo, J., Roberton, R., Sulser, T., Zhu, T., Ringler, C., et al. 2009. [El impacto en la agricultura y los costos de adaptación]. [in Spanish]. Available from: http://www.fao.org/fileadmin/user_upload/AGRO_Noticias/docs/costo%20adaptacion.pdf
- SEMAR. 2015. [Estación Virtual de Imágenes Satelitales]. [in Spanish]. Available from: <https://www.gob.mx/semar/acciones-y-programas/estaciones-satelitales>
- Semenov, M.A., Halford, N.G. 2009. Identifying target traits and molecular mechanisms for wheat breeding under a changing climate. *J. Exp. Bot.* 60:2791-2804.
- Seminis. 2016. [Información Espacial Para El Campo Mexicano]. [in Spanish]. Available from: <https://www.seminis.mx/blog-informacion-geoespacial-para-el-campo-mexicano/Sistema-Nacional-Meteorologico>. 2020. [Agro-Climas]. [in Spanish]. Available from: <https://cmgs.gob.mx:31/agroclimas/>
- Sosa Sierra, M.D.C. 2007. [Inteligencia artificial en la gestión financiera empresarial *Pensamiento and Gestión*]. [in Spanish]. Universidad del Norte Barranquilla, Colombia. *Pensamiento and Gestión* 23:153-186.
- Stockdale, E.A., Watson, C.A. 2009. Biological indicators of soil quality in organic farming systems. *Renew. Agr. Food Syst.* 24:308-318.
- Temniranrat, P., Kiratiratanapruk, K., Kitvimonrat, A., Sinthupinyo, W., Patarapuwadol, S. 2021. A system for automatic rice disease detection from rice paddy images serviced via a Chatbot. *Comput. Electr. Agr.* 185:106156.
- Torres Lima, P., Cruz Castillo, J.G., Acosta Barradas, R. 2011. [Vulnerabilidad agroambiental frente al cambio climático. *Agendas de adaptación y sistemas institucionales*]. [Article in Spanish]. *Política Cultura* 36:205-232.
- Vázquez, H.I.C., Ordóñez, Y.M.F., Ruiz, J.S., Maurice, M.J.E. 2015. [Enfoque metodológico para la construcción de una Geobase como apoyo a la investigación en agricultura y recursos naturales]. [Article in Spanish]. *Invest. Geogr.* 87:39-50.
- Wang, Y.Q. 2014. An analysis of the Viola-Jones face detection algorithm. *Image Process.* On Line 4:128-148.
- Wood, E. C., Tappan, G. G., Hadj, A. 2004. Understanding the drivers of agricultural land use change in south-central Senegal. *J. Arid Environ.* 59:565-582.