Precise Weed and Maize classification through Convolutional Neuronal Networks

Andrea Concepción Córdova Cruzatty *, Mauricio Daniel Barreno Barreno † and José Misael Jácome Barrionuevo‡
Department of Energy and Mechanics, University of the Army ESPE
Latacunga - Quijano y Ordoñez and Hermanas Paéz
Email: *accordova@espe.edu.ec, †mdbarreno@espe.edu.ec, ‡jmjacome1@espe.edu.ec

***Abstract*—Deep Learning has played a very important role in modern times, being very used in Artificial Vision and especially in Pattern Recognition, which opens many doors in the fields of application among which is precision agriculture. In this paper we present the study of a convolutional neural network with two classes applied to the classification of plants of maize and weeds in real time, focused to fields of maize in its initial stage. The algorithm has two stages, segmentation of images and classification. In the segmentation we wanted to separate the target plant from the original image. While the classification is possible through a convolutional neural network, previously trained with a dataset generated through the segmentation stage. The performance of the network is illustrated by analysis and comparisons with different network architectures, then with the network that presented better results in training, we experimented by modifying its number of filters, to analyze their behavior. Finally, the resulting network was tested on different processors to compare classification times.**

I. Introduction

During the last centuries, significant progress has taken place in science and technology developments. Significant milestones in diverse areas such as communications, numerical computer control and the miniaturization of components have benefited industrial sectors, providing new means to approach problems differently.

The availability of leading edge technologies has considerably increased, thus becoming more accessible to people. This effect of globalization has allowed many countries to have access to state-of-the-art technological products, promoting its application in different areas to generate alternative and more efficient solutions to conventional problems.

Technological advances have caused industry transformations; manufacturing, food processing, and information industries among others, have gone through dramatic changes with industrial revolution. However there are fields that have not been significantly influenced by technological development in Ecuador, as it is the case of agroindustry. Agriculture in Ecuador has undergone some changes since colonial times; while it is true that there are efficient agricultural practices, the lack of technological resources make it impossible for the country to exploit its true potential as an agricultural producer. Also, the excessive use of pesticides have a negative effect on the production, soil, and water quality.

Nowadays, one of agriculture challenges is the development of precision agriculture techniques focused on Weed and Crop segmentation. There are studies that show the impact of Weed in corn crops [1]; its yield is affected by 5000 kg/ha. Currently, growing development of Artificial Vision and Machine Learning algorithms allow researchers to propose solutions for Weed Segmentation in Crops.

One of the first approximations to the algorithms of detection of Crops was developed in 1996 [2], this algorithm permitted to segment crops from weed, it could be possible by the use of IR Images, the image is processed by a hysteresis umbral and the method of Min Neighbouring to identify the row of crops. In recent years the implementation of Machine Learning has opened new possibilities for differentiate the Weed from the crops, recently [3] there had been developed an algorithm by the use of Harris Corner detector, Feature Detector and using the DBSCAN (Density-based spartial clustering of apllications with noise) , it demonstrates an effectiveness of 98% in the identification of Weeds in the Rice, Araguez *et. al* [4] performed their segmentation through the analysis of the green hystogram and performing the segmentation of the crop and weed by classifiers not specified in the document.

Hong and Lei [5] developed their approximation by using an optimal method for detection in various types of luminosity, they achived this by the use an ANN for weed and maize classification, with a precision of 92.5%, also there had been developed an algorithm [6], that through binarization methods of OTSU and Watershed for the segmentation of the images. While its classification was given through an areas analysis to perform a thresholding, although the method is computationally effective when the Weed distribution does not resemble the size to the crop plant, its error increases when there is more density of crop than of weed. Romeo *et. al* [7] propose to use a fuzzy clustering approach to correctly segment the crop green.

The segmentation of weed and crops is not closed to color images, using a multispectral camera [8] permitted to obtain RGB and NIR images, for the segmentation and classification, then the images used a light CNN for the first process and a Deep CNN for classification in Crops, its accuracy is up to 98% in the identification of weeds. One of the most dificults things for the identification of crops is the generation of Datasets due to presence of Weed in the images of training, but there could be correctly generated by a Convolutional Neural Network [9] achieving a acceptable accuracy for the generation of Datasets.

II. Matherials and Methods

El desarrollo del modelo para la identificacion de mala hierba y maiz usando Redes Neuronales Convoluciones se describe abajo:

The development of the model for the weed and maize identification algorithm using convolutional neural networks is described below.

A. Hardware

Para la etapa de entranamiento y captura de imagenes, los siguientes elementos fueron utilizados: Una Raspberry Pi 3 con su modulo de camara V2.1 para la adquisicion de imagenes, fue configurado para obtener video en resolucion de 1280x720 pixeles con el fin de obtener mayor detalle en la segmentacion de imagenes. Debido a la popularidad de la computacion por GPU, fue necesario consider hardware compatible con computacion paralela como las tarjetas graficas fabricadas por NVIDIA, estas tienen la ventaja de poseer mejores tiempos de procesamiento comparados con la CPU, con lo cual se logra optimizar los tiempos de procesamiento y entrenamiento de Modelos de Deep Learning, entonces, utilizamos un computador con una tarjeta Grafica NVIDIA GTX950M. Tambien fue necesario considerar el uso de un CPU Core i7 2.7Ghz, 8 Nucles and el CPU de la propia Raspberry Pi 3, un ARM 3, a ARM Cortex-A53, 1.2GHz, 4 nucles para probar los tiemposde clasificacion de imagenes en las arquitecturas de red.

For the training and image capturing stages, the following elements were used: a Raspberry Pi 3 with Pi Camera V2.1 for image acquisition, it was configured to obtain a video at a resolution of 1280x720 pixels in order to get more detail in the segmented images. Due to the popularity of the GPU computing, it was neccesary to consider hardware compatible with Parallel Computing like the Graphic Cards manufactured by NVIDIA, they have the advantage of better times of processing compared to CPU, in order to optimize the time of image processing and training of Deep Learning Models, so we used a Computer with a Nvidia Graphics Card GTX950M. Also it was neccesary to consider the use of a CPU Core i7 2.7 Ghz, 8 cores and the CPU of the Raspberry Pi 3, a ARM Cortex-A53, 1.2GHz, 4 cores for testing the classification of images of the Network.

B. Software

Para la aplicacion en este documento, se requiere el uso de Poderosas librerias de desarrollo para el procesamiento de imagenes y deep Learning, consideramos el uso de software libre para extender su uso en futuros prototipos, debido a esta consideracion hemos escogido la libreria de OpenCV debido a su eficiencia computacional y su enfoque a aplicaciones en tiempo real, para este caso fue utilizada(en referencia a OpenCV) para la adquisicion y segmentacion de las imagenes de muestra. Tambien hemos considerado Caffe debido a que esta desstinado a el entrenamiento y desarrollo de Redes Neuronales Convolucionales de Proposito General y otros modelos de Deep Learning con soporte completo de la Libreria de Computacion GPU CUDA y una generosa base de datos de recursos para aprendizaje

For the application in this documents it requires the use of powerful libraries of Image Processing and Deep Learning Development, we considered to use Open Software for extend its use in future prototypes, due to this considerations we had chosen the OpenCV library for its computational efficiency and focus on real-time application [10] in this case, it was used for acquisition and segmentation of the sample images. Also we had considered Caffe because it is intended to use in training and developing general-purpose Convolutionals Neural Networks and other Deep Lerning models [11] with fully support of CUDA GPU Computation Library and a generous database of resources for learning,

Tambien el la etapa de desarrollo fue importante el suo del Sistema Operativo Linux, sus caracteristicas son: Sistema Operativo Ligero y Libre tambien completamente compatible con el software mencionado anteriormente. En la computadora principal fue usado la Distribucion Ubuntu Linux 16.04 por su soporte oficial a CUDA, Caffe y OpenCV, tambien en la Raspberry Pi fue usado el Sistema Operativo Pixel para adquirir imagenes y probar el rendimiento de clasificacion la Red Neuronal Convolucional

Also in the development stage was important the use of Linux OS , its features are: lightweight and free OS also fully compatible with the software mentioned above. For the main computer was used Ubuntu Linux 16.04 for its official support to CUDA, Caffe and OpenCV, also in the Raspberry Pi we used the Pixel OS for acquire images and test the classification performace of the CNN.

C. Dataset Description

Debido a la falta de datasets de plantas de Maize y Mala Hierba, fue necesario la busqueda de cultivos de maiz en su etapa inicial, para ese proposito nos hemos traslatado a Pillaro, esta ciudad es localizada en la Provincia de Tungurahua, en el Centro de la Sierra Ecuatoriana. Pillaro es un productor reconocido de cultivos andinos como: Maiz, Papas y frutas, por lo que podemos encontrar aqui muchos cultivos de maiz que sean utiles para nuestro proposito

Due to the lack of datasets of Weed and Maize plants, it was neccesary to search crops of maize in its initial stage, for that purpose we had traslated to Píllaro, it is a city located in the Tungurahua Province, on the center of the Ecuadorian Highlands region Píllaro is recognized producer of Andean Crops such as: Maize, Potatoes and Fruits so here we could find many crops of Maize useful for our purpose.

Las muestras fueron obtenidas de imagenes capturas en esos campos de maiz, consideramos escoger cultivos de maiz donde podamos visualmente discrimiar las plantas del cultivo de la mala yerba. Las etapas del Maiz (V3-V7) fueron usadas para tomar muestras. Para obtener las muestra, centramos la camara sobre la platna objetivo, asegurandonos que la captura muestre todas las caracteristicas de la planta de maiz, esto fue posible al monitorear el proceso por una pantalla externa. Unas muestras de las imagenes capturadas son mostradas a continuacion:

Samples were obtained from images captured in those maize fields , we considered to choose maize crops where we could visually discriminate the plants of the crop and the weed. Maize in stages V3-V7(3-7 leafs) [12] were used to take samples. To record the samples we centered the camera over the target plant, insuring that the capture shows all features of the maize plant, it was possible by monitoring the get of samples by an external display. An examples of the images captured are shown below.

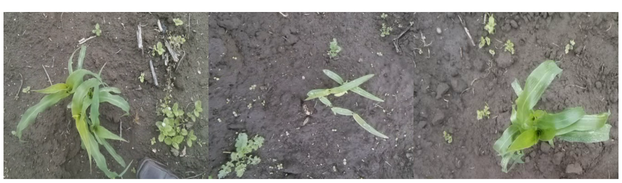

Fig. 1: Crops chosen to obtain samples, notice that it is easy to discramante weed from maize

Tambien fue necesario obtener imagenes de las malas hierbas mas comunes en los campos de maiz de Pillaro, para permitir a la Red discrimiar entre las dos clases definidas en el Dataset.

Also it was neccesary to obtain images of the most common weed plants of the maize crops in Pillaro, to allow the Network to discriminate both classes of the Dataset.

Entonces, a traves de un proceso de procesamiento digital de imagenes, como primera fase, la imagen fue normalizada en su canal verde, esto mejora la deteccion del color verde porque elimina las luces y las sombras de la imagen, despues extrajimos el color verde de la imagen con la ecuacion descrita opr Wang, Men, Luo and Mei, para obtener una imagen final en escala de grises, despues se aplico una Umbralizacion de OTSU para obtener una mascara binaria

Then, through a digital image processing, as first stage the image was normalized in its green channel, it improves the detection of green color due it eliminates lights and shadows of the image, after we extracted the green color to obtain a grayscale image, through the next equation described by Wang, Meng, Luo and Mei [13], then we applied a OTSU Thresholding to grayscale image in order to obtain a binary mask.

$$S = 2 * G - R - B \tag{1}$$

Despues de la imagen fue segmentada, para diferenciarla del suelo y otros elementos que no sean plantas, las imagenes originales fueron enmascaradas con la mascara binaria para proveer al Dataset caracteristicas de Color del Maiz y de la Mala Hierba. Una vez ue las imagenes segmentedas fueron obtenidas, removimos aquellas que tenian una resolucion menor a 64x64 pixeles debido a que son mas pequenhas que el tamanho de entrada de la red, ademas no es conveniente procesarlas debido a que su area total es mucho mas pequenha que la de las muestras de maiz, las cuales son el proposito de este documento, despues las imagenes se almacenan en un formato Png. El proceso se ilustra de mejor manera en la siguiente figura.

After the images were segmented to differentiate them from the soil and other non-plant elements, the original images were masked with the binary mask to provide the dataset features of color of the Maize and Weed. Once the segmented images were obtained, we were removed those that had a resolution lower than 64x64 pixels, because it's smaller than size of input layer of the network, besides it was not convenient its processing since its total area was much smaller than the samples of corn, which is the purpose of this document, later, the images were saved in a png format. The process described before, was illustrated in the next figure.

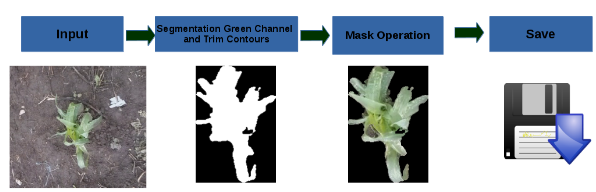

Fig. 2: Preprocessing to obtain samples for the dataset

Debemos acotar que hemos etiquetado manualmente las imagenes en dos clases, maiz y mala hierba, teniendo en cuenta que las imagenes no se repitan(**corregir**). En la presente imagen presentamos los resultados finales de las imagenes procesadas que conforman el dataset.

It should be noted that we had manually labeled the samples in two classes, maize and weed, having no account that the images are not repeated. In the following image we present the final results of the processed images that conform the dataset.

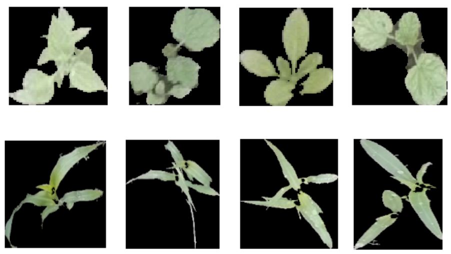

Fig. 3: Segmented and Masked images of dataset Above: Weed, Below: Maize

Una vez completada el procesamiento de imagenes, donde fueron obtenidas 2835 imagenes de maiz y 880 images de Mala Hierba, Lee et. Al dicen que una vez que la CNN ha alcanzado una precision aceptable, el dataset se puede extender por transformaciones geometricas, en adicion sladojevic mencionan que esto permite a reducir el sobreajuste y para mejorar la precision en el entrenamiento, razon por la cual hizimos 11 rotaciones cada 30 grados, obteniendo de esta forma el incremento(corregir) del dataset en 12 veces y tener mejores probabilidades de reconocer als plantas en cualquier orientaciones.

Once we completed the processing of the images, it were obtained 2835 images of Maize and 880 images of Weed, Lee *et. al* [14] said that once CNN has reached an acceptable accuracy, the dataset could extend by Geometric Transformations, in addition [15] they mention that this allows to reduce the overfitting and to improve the precision in the training, reason why we made 11 rotations each 30 degrees, obtaining in this way to increase the dataset in 12 times and have better chances of recognizing plants that are in any orientation.

Para la fase de validacion, hemos escogido aletariamente la quinta parte del total de imagenes de cada clase, asi que tenemos la siguiente tabla con la distribucion del dataset.

For the Validation Phase, we had chosen randomly a fifth part of the total images of each class, so we have the following chart with the distribution of the dataset.

Phase	Train	Validation	Test	Total
Maize	25695	8325	200	34220
Weed	8560	2000	200	10760

TABLE I: Distribution of the Dataset of each class

D. Convolutional Neuronal Network

After the last step have been terminated, it is neccesary classify the images of Maize and Weed, a high accurate method of Image Classification is Convolutional Neuronal Networks(CNN), those models are complex but efficient with a highly rate of discrimination and its demostrated that they have good results in Image Classification, Object Detection and Fine-Grained Classification [16]. Their are also applied in Precision Agriculture [8] for the correct identification of plants.

One of the characteristics of the CNN is that they can have multiple architectures, each architecture can reach different result depending of the application , for the present document we have considered four CNN architectures: Two models were chose of the Caffe Zoo Model(it's a free resource provided by Caffe Developers), LeNet and AlexNet also Potena and Nardi had used two additional types of Models, sNET and cNET, they were sucessfully proved in identification of plants in crops[8], with the Nets chosen, each one was trained with a same Solver of type AdaDelta and with the same dataset, the results of training are the following.

Parameters	LeNet	AlexNet	cNET	sNET
Input size of images	32x32	64x64	64x64	64x64
Layers numbers	9	11	8	4
Number of parameters	652500	20166688	6421568	135872
Iterations	3000	3000	3000	3000
Accuracy(%)	86.48	93.86	95.33	80.4
Loss(%)	32.80	15.32	13.72	15.32

TABLE II: Distribution of the Dataset of each class

Inside of the shown parameters in the Table II, we considered the net who achieved the highest rate of accuracy and lowest rate of loss, clearly the cNET is the chosen due it shows a precission near to human classification.

Although results given for the net are excellents, one of the main purposes of this paper is the implementation in real time, so the number of parameters to be computed must be a problem, so we considered to modify the net in order to improve the processing time. To achive this, it was neccesary to reduce the number of filters of cNET from 64 to 16.

Parameters	cNET 16 filters	cNET 64 filters
Number of parameters	1651376	6421568
Iterations	9000	9000
Accuracy(%)	97.26	96.40
Loss(%)	8.39	14.48

TABLE III: Distribution of the Dataset of each class

This permitted to reduce the number of parameters to 1651376(it's a 25% of parameters than the original net), and what about the precision of the net, it was slighty affected but it reach a 97.2% of accuracy.

The complete network model is formed as follows: the input layer where an RGB image 64x64 pixels, followed by a convolution layer of 16 filters, a kernel of 5 and a stride of 1 pixel, this layer obtaining only its main characteristics, then a pooling layer is applied, which has the function to find the maximum values of the previous layer with a kernel size and stride of 2, then again a second layer of convolution is applied that has 16 filters, 5 kernels and a stride of 1, followed by a new layer of pooling with a kernel size and stride of 2 again obtained a size of 16x16x16 and finally the layers of type fully-connected with a total of 478 neurons, where the last two neurons that is the output represent the two classes established in this network.

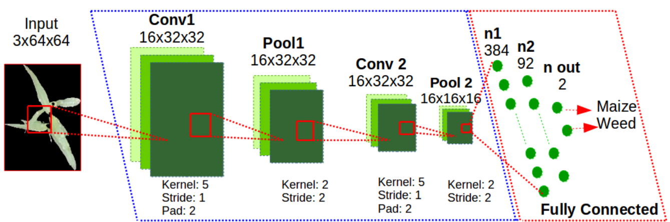

Fig. 4: The schematic of CNet

The follow picture shows the schematic of the cNET of 16 filters and its process of training.Note that by reducing the number of filters the network gained accuracy because it does not have so much data to process.

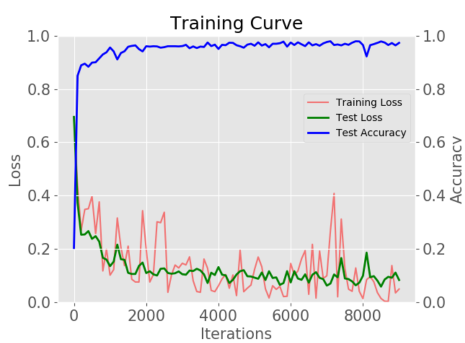

Fig. 5: Graph of the Training Process

III. Test

For this stage, we considered an extra dataset with 200 images of maize plants in stage of V3-V7 and 203 images of weed, obtained in maize crops, the results were compared between cNET with 64 and 16 filters in the Table IV.

Like the table shown the results were achieved by the GPU processor, it is the best option to classificate a large amount of images however this project is thought to be portable, so

	GPU		
cNET	**Accuracy Weed**	**Accuracy Maize**	**Average time of classification**
64 Filters	95.05%	74.75%	2.34 ms
16 Filters	94.55%	87.13%	951 μs
	CPU		
cNET	**Accuracy Weed**	**Accuracy Maize**	**Average time of classification**
64 Filters	95.05%	74.75%	51.6 ms
16 Filters	94.55%	87.13%	17.23 ms
	CPU Raspberry Pi 3		
cNET	**Accuracy Weed**	**Accuracy Maize**	**Average time of classification**
64 Filters	95.05 %	74.75%	615 ms
16 Filters	94.55%	87.13%	161 ms

TABLE IV: Test

we consider the RaspBerry Pi for the analysis of performance. With cNET of 64 filters the results fo time of classification was 615 miliseconds, so it time isn´t the optimum for our objective that it is the processing in a real time.

IV. Conclusion

In this paper we have studied the application of Convolutional Neuronal Networks for Classification between Maize and Weed Plants, for that we have used some CNN Architectures, the Network with best results in precision was cNET, with a 99.33 % of accuracy of training, using a dataset of both classes of 44580 segmented images.

In order to search a right processing to work in real time, we have experimented to reduce the number of filters of the Layers of the Net from 64 to 16 filters, achiveiving to reduce the number of parameters to be computed during the classification , we obteined a classification time 2.5 times less than the original Net of 64 filters , this change had a small impact in the training accuray of the net, it value fell 1 % in comparation with the original net.

Also, by using different methods to classify the images in our experiments, the results showed that the best time of classification was achieved by using a Nvidia Graphic Cards, this is due the Parallel Architecture that this GPU has and also it has fully support for the computation platform CUDA, it demostrated to be 18 times faster than a normal CPU and 170 times faster than a Raspberry Pi 3.

The study demostrated that is better to perform some geometrical tranformations to the Dataset images due it benefits to improve the classification of plants because the plants ever have different orientations in the crop field.

Acknowledgment

The authors would like to thank to the "Granja Agroecológica y Demostrativa" of Píllaro City and owners of corn crops of the city which allowed us to take samples of

plants, to obtain to conform the dataset of the project.

This work has been supported by the University of the Army ”ESPE” Latacunga.

References

[1] R. SUÁREZ and J. P. Y. J. VALLADARES, “Distintos sistemas de escarda en maíz forrajero,” *Producciones agroganaderas: Gestión eficiente y conservación del Medio Natural. Actas de la XLV RC de la SEEP. Gijón*, 2005.

[2] R. Brivot and J. Marchant, “Segmentation of plants and weeds for a precision crop protection robot using infrared images,” *IEE Proceedings-Vision, Image and Signal Processing*, vol. 143, no. 2, pp. 118–124, 1996.

[3] B. Cheng and E. T. Matson, “A feature-based machine learning agent for automatic rice and weed discrimination.” in *ICAISC (1)*, 2015, pp. 517–527.

[4] S. Equipo Técnico, “Proceedings of vii seae congress spanish society for organic farming,” 2006.

[5] H. Y. Jeon, L. F. Tian, and H. Zhu, “Robust crop and weed segmentation under uncontrolled outdoor illumination,” *Sensors*, vol. 11, no. 6, pp. 6270–6283, 2011.

[6] S. H. Hlaing and A. S. Khaing, “Weed and crop segmentation and classification using area thresholding,” *IJRET*, vol. 3, pp. 375–382, 2014.

[7] J. Romeo, G. Pajares, M. Montalvo, J. Guerrero, M. Guijarro, and A. Ribeiro, “Crop row detection in maize fields inspired on the human visual perception,” *The Scientific World Journal*, vol. 2012, 2012.

[8] C. Potena, D. Nardi, and A. Pretto, “Fast and accurate crop and weed identification with summarized train sets for precision agriculture,” in *International Conference on Intelligent Autonomous Systems*. Springer, 2016, pp. 105–121.

[9] M. Di Cicco, C. Potena, G. Grisetti, and A. Pretto, “Automatic model based dataset generation for fast and accurate crop and weeds detection,” *arXiv preprint arXiv:1612.03019*, 2016.

[10] G. Bradski, “The opencv library.” *Dr. Dobb’s Journal: Software Tools for the Professional Programmer*, vol. 25, no. 11, pp. 120–123, 2000.

[11] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” *arXiv preprint arXiv:1408.5093*, 2014.

[12] A. Fassio, “Maíz, aspectos sobre fenología, alberto fassio, ana inés carriquiry, cecilia tojo, ricardo romero,” *Serie técnica. 101.*, 1998.

[13] P. Wang, Z. Meng, C. Luo, and H. Mei, “Path recognition for agricultural robot vision navigation under weed environment,” in *7th International Conference on Computer and Computing Technologies in Agriculture (CCTA)*, no. Part I. Springer, 2013, pp. 242–248.

[14] S. H. Lee, C. S. Chan, P. Wilkin, and P. Remagnino, “Deep-plant: Plant identification with convolutional neural networks,” in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 452–456.

[15] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, “Deep neural networks based recognition of plant diseases by leaf image classification,” *Computational intelligence and neuroscience*, vol. 2016, 2016.

[16] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, “Cnn features off-the-shelf: an astounding baseline for recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 806–813.