Weed Recognition in Coffee Crops Using Texture Analysis and Machine Learning

Reconocimiento de Malezas en Cultivos de Café por Medio de Análisis de Textura y Aprendizaje Automático

M. J. Muñoz-Neira ; J. A. Ortiz-Delgado
DOI: 10.22517/23447214.24589
Scientific and technological research paper

Abstract— **The present work presents the results of research elaborated to recognize two classes of leaves of weed present in the coffee crops through machine learning techniques, a topic few have explored in the coffee agroindustry, and that can significantly impact the management of herbicides in this important crop. The study involved twenty-four experiments, utilizing a database of 210 images, 70 for each weed class and 70 for coffee leaf samples. All images were processed and transformed into HSV color format. From each image, 33 texture patterns were extracted and reduced to four through principal component analysis. The fractal dimension was added as a fifth pattern. The recognition used three machine learning techniques: support vector machine (SVM), k-near neighbors (KNN), and artificial neuronal networks. The machine learning techniques permitted classification with precision and recall upper or equal to 95%, on average, when the fractal dimension was not used and upper or equal to 97% on average when the fractal dimension was used. SVM and ANN were methods with better outcomes. These experiments constitute a first step towards implementing an automatic system for selective weed eradication in a coffee crop, with promising implications for future developments.**

Index Terms—**Coffee Crop; Machine Learning; Texture Analysis; Weed Recognition.**

Resumen— **El presente trabajo presenta los resultados de una investigación elaborada para reconocer dos clases de hojas de maleza presentes en los cultivos de café mediante técnicas de aprendizaje automático, un tema poco explorado en la agroindustria cafetalera, y que puede impactar significativamente el manejo de herbicidas en este importante cultivo. El estudio involucró veinticuatro experimentos, utilizando una base de datos de 210 imágenes, 70 para cada clase de maleza y 70 para muestras de hojas de café. Todas las imágenes se procesaron y transformaron en formato de color HSV. De cada imagen se extrajeron 33 patrones de textura, que se redujeron a cuatro mediante un análisis de componentes principales. La dimensión fractal se añadió como quinto patrón. Para el reconocimiento se utilizaron tres técnicas de aprendizaje automático: máquina de soporte vectorial (SVM), k-vecinos más cercanos (KNN) y redes neuronales artificiales. Las técnicas de aprendizaje automático permitieron la clasificación con una precisión y exhaustividad superiores o iguales al 95% en promedio, cuando no se utilizó la dimensión fractal, y superiores o iguales al 97%, en promedio, cuando se utilizó la dimensión fractal. SVM y ANN fueron los métodos con mejores resultados. Estos experimentos constituyen un primer paso hacia la implementación de un sistema automático para la erradicación selectiva de malezas en un cultivo de café, con implicaciones prometedoras para desarrollos futuro**.

Palabras claves— **Análisis de Textura; Aprendizaje Automático. Cultivo de Café; Reconocimiento de Malezas.**

This manuscript was submitted on October 03, 2023. Accepted on November 13, 2024. And published on December 19, 2024.

Milton Muñoz: Electronic Engineer with a Máster of Science degree in this field who graduated from the Industrial University of Santander (UIS), Colombia. (emeil: mmunoz@unisangil.edu.co).

Jaime Alberto Ortíz Delgado: an-Electronic Engineer who graduated from the University Foundation of San Gil, UNISANGIL, Colombia. (emeil: jamiejaimeortiz@unisangil.edu.co).

I. INTRODUCTION

WEED control is crucial to ensure adequate growth and crop performance. Indeed, weeds compete with crops for water, nutrients, light, CO2, and space [1], with a special affectation in the first plant years [2]. For coffee, one of the most representative Colombian crops, this control must be realized regularly throughout the year and imply considerable spending of time and money. Anzalone and Silva ([3]) estimate this spending to be around 35% when nothing is done and between 16% and 27% when partial control is realized. The conventional form of address weed eradication involves applying chemically prepared herbicides for this purpose [4]. However, the regular or indiscriminate use of herbicides can lead to severe environmental affectation, health problems, and resistance of weeds to its application ([5], [6], [7], [8]). This forces, in many cases, to alternate the use of glyphosate, which is perhaps the most common herbicide used in coffee farms, with other chemical herbicides ([9], [10], [11], [12], [13]).

Some possible secondary effects on the environment and health are erosion, ground degradation, crop contamination, water contamination, fauna, and human habitat contamination, among others [14]. Roundup, for example, a widely used herbicide in Colombia that contains glyphosate and polyoxyethylene amine surfactant (POEA), is mentioned in [15] as an herbicide capable of altering the hormonal balance of the organisms that are exposed to it; its use is related to cases of cancer and births with malformations. The affectation of flora in coffee crops by herbicides is commentated by [16]. Damages to earthworm biomasses, coleopterous, beneficial fungi, or edaphic mites are researched in [17], [18], [19], [20], [21], [22], [23], [24]. Herbicides can even alter coffee crops themselves [25], [26], [27].

The previous considerations show that the selective application of herbicides or alternative methods for weed eradication is fundamental. In both cases, the use of technology for the automatic recognition of weeds is essential. Many works have been developed for this purpose using different techniques, like edge detection [28], image filtering [29], hyperspectral sensing [30], crop signaling (making crops distinguishable from weeds) [31], deep learning ([32], [33], [34], [35], [36]), between others. Some research experiments with systems for weed recognition in real-time [37], [38], too. However, despite these works, there is no specific research addressing the recognition of weeds in coffee crops.

The present paper summarized a series of experiments tending toward recognizing images of two common weeds of coffee crops, *Sida Acuta* and *Paspalum Macrophyllum*. Patterns for recognition were obtained based on texture analysis, an advantageous technique for leaf or fruit recognition ([39], [40]). Three machine learning techniques were implemented: support vector machine, SVM, k-near neighbors, KNN, and artificial neuronal network, ANN, making use of Matlab 2022b®, with the corresponding functions for these machine learning tools, updated by the manufacturer for that version. Despite the existence of more current and robust algorithms and implementation tools, KNN, SVM, and ANNs remain highly relevant and widely used in various applications ([41], [42], [43], [44], [45]). Their enduring relevance is due to their status as widely studied algorithms that are relatively easy to understand and interpret. This makes them particularly suitable for exploratory projects and with a reduced data set, such as the one used in the present research.

The work constitutes a first approximation to the implementation of an automatic system for selective weed eradication. The paper is organized into four sections: introduction, methodology, result analysis, and conclusion.

II. Methodology

Two hundred and ten samples of coffee plants and weed leaves were recollected in an aleatory form from an area of 0.1 km^2, in a coffee crop Castilla variety, in the farm San Carlos from San Gil municipality, at the south of Santander, Colombia. The samples were preserved by placing them into newspaper sheets smeared with alcohol and transported to a laboratory temperature-controlled at 20°C. Seventy leaves corresponded to coffee plants, seventy to leaves of *Sida Acuta* weed, and seventy to *Paspalum Macrophyllum* weed.

Each leaf was photographed with a digital camera in a light cube, obtaining 210 RGB images of approximately 4608 by 3456 pixels, with 24 bits of profundity. This size was adequate for image processing. Experiments with different image sizes are beyond the scope set for investigation and may be the subject of study in future work. The images were processed to eliminate all that was not part of the leaf, resulting in images like those shown in Fig. 1. Afterward, the processed RGB images were transformed into HSI color format (hue, saturation, intensity) through (1) to (3).

$$H = \begin{cases} 0, & \text{if Max} = \text{Min} \\ \dfrac{G - B}{\text{Max} - \text{Min}} + 4, & \text{if Max} = B \\ \dfrac{B - R}{\text{Max} - \text{Min}} + 2, & \text{if Max} = G \\ \dfrac{G - B}{\text{Max} - \text{Min}} & \text{if Max} = R \end{cases} \tag{1}$$

$$I = \frac{1}{2}(\text{Max} + \text{Min}) \tag{2}$$

$$S = \begin{cases} 0, \text{if Max} = \text{Min} \\ \dfrac{\text{Max} - \text{Min}}{\text{Max} + \text{Min}} = \dfrac{\text{Max} - \text{Min}}{2I}, & \text{if Max} = B \end{cases} \tag{3}$$

Fig. 1. Coffee and Weeds Leaves Images

Subsequently, from hue, saturation, and intensity components, three co-occurrence matrixes (one for hue, the other for saturation, and the other for intensity) were calculated, according to (4). Eleven texture patterns were established from each co-occurrence matrix according to (5) to (15).

$$C_{\Delta x \Delta y(i,j)} = \sum_{p=1}^{n}\sum_{p=1}^{n}\begin{cases}1 \; if \; I(p,q) = \; i \; and \; I(p+\Delta x, q+\Delta y) = \; j \\ 0 \qquad otherwise\end{cases} \quad (4)$$

Uniformity

$$I_1 = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}[p(i,j)]^2 \quad (5)$$

Meddle Intensity

$$I_2 = \sum_{i=0}^{N_g-1} i p_x(i) \quad (6)$$

Moment Product

$$\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}(i-I_2)(j-I_2)p(i,j) \quad (7)$$

Inverse Difference

$$I_4 = \frac{\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} p(i,j)^2}{1+(i-j)^2} \quad (8)$$

Entropy

$$I_5 = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1} p(i,j)\ln p(i,j) \quad (9)$$

Entropy Sum

$$I_6 = \sum_{i=0}^{N_g-1} p_{x+y}(k) ln \; p_{x+y}(k) \quad (10)$$

Entropy Difference

$$I_7 = \sum_{i=0}^{N_g-1} p_{x-y}(k) ln \; p_{x-y}(k) \quad (11)$$

Information Correlation 1

$$I_8 = \frac{I_{7-HXY1}}{HX} \quad (12)$$

Information Correlation 2

$$I_9 = [1 - \; e^{-2(HXY2-I_7)}]^{1/2} \quad (13)$$

Contrast

$$I_{10} = \sum_{|i-j|=0}^{N_g-1}(i-j)^2\sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}p(i,j) \quad (14)$$

Mode

$$I_{11} = \sum_{i=0}^{N_g-1}\sum_{j=0}^{N_g-1}\max[\,p(i,j)] \quad (15)$$

Where p is the normalized co-occurrence matrix, p_x is the vector of the sum of columns of p, and p_y is the vector of the sum of rows of p. Values p_{x+y}, p_{x-y}, HX, HXY1 y HXY2, were found conforming (16) to (20).

$$p_{x+y}(k) = \underbrace{\sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1}p(i,j)}_{k=i+j}$$
$$\text{for k} = 0,1,2 \ldots .2\left(N_g - 1\right) \quad (16)$$

$$p_{x-y}(k) = \underbrace{\sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1}p(i,j)}_{k=|i-j|}$$
$$\text{for k} = 0,1,2 \ldots .\left(N_g - 1\right) \quad (17)$$

$$HX = \sum_{i=0}^{Ng-1} p_x(i) * \ln\,[p_x(i)] \quad (18)$$

$$HXY1 = \sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1} p(i,j) * \ln\,[p_x(i)p_y(i)] \quad (19)$$

$$HXY2 = \sum_{i=0}^{Ng-1}\sum_{j=0}^{Ng-1} p_x(i)p_y(i) * \ln\,[p_x(i)p_y(i)] \quad (20)$$

Next, a principal component analysis (PCA) was applied to each group of patterns, including the group comprised of a mix of all. The number of the first components selected (four) was determined in a heuristic form. Consequently, for everyone, the leaf was defined as four groups of patterns (H, S, I, and HSI), with four patterns each.

Additionally, the fractal dimension of each image was determined, subdividing the binarized image of each leaf in boxes of increasing size, *ri*. The fractal dimension value was obtained employing (21), where *n* is a vector in which each element represents the number of boxes of dimension *ri* that contain white color, and N is the length of the vector.

$$df = \frac{\sum_i \frac{\Delta n_i}{\Delta r_i}}{N} \quad (21)$$

Three classifiers were used: quadratic support vector machine (SVM), K nearest neighbors (KNN), and an artificial neuronal network (ANN). SVM and KNN ran with a cross-validation strategy and seven folds. SVM used a radial basis function as kernel, with box constraint and kernel scale of 1. KNN used the five nearest neighbors classifier and the Minkowski metric. ANN structure had two layers: the first with neurons with sigmoid functions and the hidden layer with softmax functions. Neurons in the hidden layer were five when the number of inputs (patterns) was four and seven when the number of inputs was five (with fractal dimension used as the fifth pattern). All parameters of classifiers were adjusted in a heuristic form.

ANN was trained with a scaled conjugate gradient backpropagation algorithm. 70% of samples (147) were used for training, 10% (21 samples) for validation, and 20% (42 samples) for testing.

With each classifier, eight essays were done, four without fractal dimension (one for H patterns, another for S patterns, another for I patterns, and another for HSI patterns), and four adding the fractal dimension to each group of patterns. In total, 24 experiments were carried out. All codes and machine learning tools were implemented in Matlab 2022b® in a computer with an Intel Core i5 processor and 12GB of DDR4 memory.

III. RESULTS

For each experiment, the evaluation parameters described in Table I were calculated. These are the same components of matrix confusion but are presented in a table because, in this form, it is easier to compare the results of the three systems tested (SVM, KNN, and ANN). The total number of samples determined the percentage of true or false positives. Thus, for example, if the total true positives of coffee samples in SVM classifications was 68, the percentage reported is 32.4% (that corresponds to 68/210). Precision and Recall were estimated according to (22) and (23), respectively, where TP is the total of true positives, FP is the total of false positives, and FN is the total of false negatives.

$$Prec = \frac{TP}{TP + FP} \qquad (22)$$

$$Rec = \frac{TP}{TP + FN} \qquad (23)$$

Precision indicates what percentage of the samples of the systems identified as a positive class are true positives. Precision is, therefore, a quality parameter. Recall indicates what percentage of positive classes the systems could identify.

TABLE I
EVALUATION PARAMETERS

True Positives of Coffee	TPC
False Positives of Coffee by Weed One/False Negatives of Weeds One by Coffee	FPC_W1
False Positives of Coffee by Weed Two/False Negatives of Weed Two by Coffee	FPC_W2
True Positives of Weed One	TPW1
False Positives of Weed One by Coffee/False Negative of Coffee by Weed One	FPW1_C
False Positives of Weed One by Weed Two/False Negative of Weed Two by Weed One	FPW1_W2
True Positives of Weed Two	TPW2
False Positives of Weed Two by Coffee/False Negatives of Coffee by Weed Two	FPW2_C
False Positives of Weed 2 by Weed One/False Negatives of Weed One by Weed Two	FPW2_W1
Coffee Precision	Prec_C
Weed One Precision	Prec_W1
Weed Two Precision	Prec_W2
Coffee Recall	Rec_C
Weed One Recall	Rec_W1
Weed Two Recall	Rec_W2

The following tables (Table I to IX) synthesize the results of the twenty-four experiments. The SVM column is the result of the quadratic Support Machine Vector classifier, the KNN column is the result of the fine K-Near Neighbors classifier, and NNTr, NNV, and NNTe columns are the results of the training, validating, and testing Neuronal Network classifier, respectively.

TABLE II
EVALUATION PARAMETERS FOR HUE (H) TEXTURE PATTERNS

	SVM	KNN	NNTr	NNV	NNTe
TPC	31.9%	31.0%	33.3%	19.0%	38.1%
FPC_W1	0.5%	1.4%	0.7%	0.0%	0.0%
FPC_W2	1.0%	1.0%	0.0%	0,0%	0.0%
TPW1	31.0%	31.4%	30.6%	42.9%	33.3%
FPW1_C	1.0%	0.5%	0.0%	4.8%	0.0%
FPW1_W2	1.4%	1.4%	0.7%	0.0%	0.0%
TPW2	32.4%	32.4%	34.0%	33.3%	28.6%
FPW2_C	0.5%	0.5%	0.0%	0.0%	0.0%
FPW2_W1	0.5%	0.5%	0.7%	0.0%	0.0%
Prec_C	95.7%	92.9%	98.0%	100.0%	100.0%
Prec_W1	92.9%	94.3%	97.8%	90.0%	100.0%
Prec_W2	97.1%	97.1%	98.0%	100.0%	100.0%
Rec_C	95.7%	97.0%	100.0%	80.0%	100.0%
Rec_W1	97.0%	94.3%	95.7%	100.0%	100.0%
Rec_W2	93.2%	93.2%	98.0%	100.0%	100.0%

For hue texture parameters, the three systems had results of *Precision* and *Recall* up to 92%, being the best ANN.

TABLE III
EVALUATION PARAMETERS FOR SATURATION (S) TEXTURE PATTERNS

	SVM	KNN	NNTr	NNV	NNTe
TPC	30.0%	25.7%	25.9%	14.3%	19.0%
FPC_W1	0.0%	0.5%	0.0%	0.0%	2.4%
FPC_W2	3.3%	7.1%	23.1%	23.8%	31.0%
TPW1	32.4%	31.9%	32.7%	23.8%	35.7%
FPW1_C	0.5%	1.4%	0.7%	4.8%	0.0%
FPW1_W2	0.5%	0.0%	0.7%	0.0%	0.0%
TPW2	32.9%	29.0%	9.5%	9.5%	2.4%
FPW2_C	0.5%	4.3%	7.5%	19.0%	9.5%
FPW2_W1	0.0%	0.0%	0.0%	4.8%	0.0%
Prec_C	90.0%	77.1%	52.8%	37.5%	36.4%
Prec_W1	97.1%	95.7%	96.0%	83.3%	100.0%
Prec_W2	98.6%	87.1%	56.0%	28.6%	20.0%
Rec_C	96.9%	81.8%	76.0%	37.5%	66.7%
Rec_W1	100.0%	98.5%	100.0%	83.3%	93.8%
Rec_W2	89.6%	80.3%	28.6%	28.6%	7.1%

In the case of saturation parameters, the performance was lower than the results of hue parameters. For example, leaves of coffee and weed number 2 were not recognized. The best results were for the SVM system.

TABLE IV
EVALUATION PARAMETERS FOR INTENSITY (I) TEXTURE PATTERNS

	SVM	KNN	NNTr	NNV	NNTe
TPC	32.4%	31.9%	32.0%	38.1%	31.0%
FPC_W1	0.0%	0.0%	0.0%	0.0%	0.0%
FPC_W2	1.0%	1.4%	0.0%	0.0%	0.0%
TPW1	32.4%	30.0%	22.4%	28.6%	26.2%
FPW1_C	0.0%	0.0%	0.0%	0.0%	2.4%
FPW1_W2	1.0%	3.3%	4.8%	4.8%	7.1%
TPW2	31.9%	28.6%	29.9%	23.8%	23.8%
FPW2_C	0.5%	1.0%	0.7%	0.0%	0.0%
FPW2_W1	1.0%	3.8%	10.2%	4.8%	9.5%
Prec_C	97.1%	95.7%	100.0%	100.0%	100.0%
Prec_W1	97.1%	90.0%	82.5%	85.7%	73.3%
Prec_W2	95.7%	85.7%	73.3%	83.3%	71.4%
Rec_C	98.6%	97.1%	97.9%	100.0%	92.9%
Rec_W1	97.1%	88.7%	68.8%	85.7%	73.3%
Rec_W2	94.4%	85.7%	86.3%	83.3%	76.9%

For intensity parameters, the results were better than those of saturation parameters, globally. Better results were obtained with the SVM method, with Precision and Recall up to 94%.

TABLE V
EVALUATION PARAMETERS FOR HSI TEXTURE PATTERNS

	SVM	KNN	NNTr	NNV	NNTe
TPC	17.6%	17.6%	13.6%	14.3%	19.0%
FPC_W1	10.0%	8.6%	15.0%	4.8%	19.0%
FPC_W2	5.7%	7.1%	8.2%	9.5%	9.5%
TPW1	18.6%	21.9%	8.2%	0.0%	11.9%
FPW1_C	11.0%	6.2%	5.4%	4.8%	0.0%
FPW1_W2	3.8%	5.2%	4.8%	0.0%	2.4%
TPW2	11.0%	19.0%	21.1%	28.6%	16.7%
FPW2_C	15.7%	8.1%	15.0%	19.0%	9.5%
FPW2_W1	6.7%	6.2%	8.8%	19.0%	11.9%
Prec_C	52.9%	52.9%	37.0%	50.0%	40.0%
Prec_W1	55.7%	65.7%	44.4%	0.0%	83.3%
Prec_W2	32.9%	57.1%	47.0%	42.9%	43.8%
Rec_C	39.8%	55.2%	40.0%	37.5%	66.7%
Rec_W1	52.7%	59.7%	25.5%	0.0%	27.8%
Rec_W2	53.5%	60,6%	62.0%	75.0%	58.3%

The machine learning process for the HSI texture parameters was the worst performing, with Precision and Recall, on average, lower than 50% for SVM, KNN, and ANN.

The following tables present the recognition results when a fractal dimension is added. In general, as is evident from the data, this addition improves the quality of machine learning.

TABLE VI
EVALUATION PARAMETERS FOR HUE (H) TEXTURE PATTERNS PLUS FRACTAL DIMENSION

	SVM	KNN	NNTr	NNV	NNTe
TPC	32.4%	31.4%	32.0%	33.3%	35.7%
FPC_W1	0.5%	1.4%	0.0%	4.8%	0.0%
FPC_W2	0.5%	0.5%	0.0%	0.0%	0.0%
TPW1	32.4%	31.9%	33.3%	33.3%	31.0%
FPW1_C	0.5%	1.0%	0.0%	0.0%	2.4%
FPW1_W2	0.5%	0.5%	0.0%	0.0%	0.0%
TPW2	32.4%	32.4%	34.7%	28.6%	31.0%
FPW2_C	0.5%	0.5%	0.0%	0.0%	0.0%
FPW2_W1	0.5%	0.5%	0.0%	0.0%	0.0%
Prec_C	97.1%	94.3%	100.0%	87.5%	100.0%
Prec_W1	97.1%	95.7%	100.0%	100.0%	92.9%
Prec_W2	97.1%	97.1%	100.0%	100.0%	100.0%
Rec_C	97.1%	95.7%	100.0%	100.0%	93.8%
Rec_W1	97.1%	94.4%	100.0%	87.5%	100.0%
Rec_W2	97,1%	97,1%	100,0%	100,0%	100.0%

TABLE VII
EVALUATION PARAMETERS FOR SATURATION (S) TEXTURE PATTERNS PLUS FRACTAL DIMENSION

	SVM	KNN	NNTr	NNV	NNTe
TPC	32.4%	31.9%	28.6%	57.1%	35.7%
FPC_W1	1.0%	1.0%	0.0%	0.0%	0.0%
FPC_W2	0.0%	0.5%	0.0%	0.0%	0.0%
TPW1	32.4%	31.9%	36.1%	28.6%	26.2%
FPW1_C	1.0%	1.4%	0.0%	0.0%	2.4%
FPW1_W2	0.0%	0.0%	0.0%	0.0%	0.0%
TPW2	32.4%	32.4%	35.4%	14.3%	0.0%
FPW2_C	1.0%	1.0%	0.0%	0.0%	0.0%
FPW2_W1	0.0%	0.0%	0.0%	0.0%	35.7%
Prec_C	97.1%	95.7%	100.0%	100.0%	100.0%
Prec_W1	97.1%	95.7%	100.0%	100.0%	91.7%
Prec_W2	97.1%	97.1%	100.0%	100.0%	100.0%
Rec_C	94.4%	93.1%	100.0%	100.0%	93.8%
Rec_W1	97.1%	97.1%	100.0%	100.0%	100.0%
Rec_W2	100,0%	98,6%	100,0%	100,0%	100.0%

TABLE VIII
EVALUATION PARAMETERS FOR INTENSITY (I) TEXTURE PATTERNS PLUS FRACTAL DIMENSION

	SVM	KNN	NNTr	NNV	NNTe
TPC	32.4%	31.9%	31.3%	52.4%	28.6%
FPC_W1	0.0%	0.5%	2.7%	4.8%	0.0%
FPC_W2	1.0%	1.0%	0,0%	0.0%	0.0%
TPW1	31.9%	31.4%	32.7%	19.0%	28.6%
FPW1_C	0.5%	0.5%	0.7%	0.0%	0.0%
FPW1_W2	1.0%	1.4%	0.0%	0.0%	0.0%
TPW2	32.4%	31.4%	32.7%	23.8%	40.5%
FPW2_C	0.5%	1.0%	0.0%	0.0%	0.0%
FPW2_W1	0.5%	1.0%	0.0%	0.0%	2.4%
Prec_C	97.1%	95.7%	92.0%	91.7%	100.0%
Prec_W1	95.7%	94.3%	98.0%	100.0%	100.0%
Prec_W2	97.1%	94.3%	100.0%	100.0%	94.0%
Rec_C	97.1%	95.7%	97.9%	100.0%	100.0%
Rec_W1	98.5%	95.7%	92.3%	80.0%	92.3%
Rec_W2	94,4%	93,0%	100.0%	100,0%	100,0%

TABLE IX
EVALUATION PARAMETERS FOR HSI TEXTURE PATTERNS PLUS FRACTAL DIMENSION

	SVM	KNN	NNTr	NNV	NNTe
TPC	25.7%	25.2%	29.3%	38.1%	21.4%
FPC_W1	7.1%	8.1%	8.2%	9.5%	2.4%
FPC_W2	0.5%	0.0%	0.0%	0.0%	0.0%
TPW1	22.4%	23.3%	22.4%	14.3%	40.5%
FPW1_C	9.5%	8.6%	3.4%	0.0%	11.9%
FPW1_W2	1.4%	1.4%	0.0%	0.0%	0.0%
TPW2	32.4%	31.0%	36.1%	38.1%	21.4%
FPW2_C	0.5%	1.0%	0.0%	0.0%	0.0%
FPW2_W1	0.5%	1.4%	0.7%	0.0%	2.4%
Prec_C	77.1%	75.7%	78.2%	80.0%	90.0%
Prec_W1	67.1%	70.0%	86.8%	100.0%	77.3%
Prec_W2	97.1%	92.9%	98.1%	100.0%	90.0%
Rec_C	72.0%	72.6%	89.6%	100.0%	64.3%
Rec_W1	74.6%	71.0%	71.7%	60.0%	89.5%
Rec_W2	94.4%	95,6%	100,0%	100,0%	100,0%

The results obtained are promising compared to those presented by another related research on the subject. Although there are no specific developments for weed recognition in coffee crops in the literature consulted, it is possible to compare the results obtained with those of works oriented, in general, to weed recognition through different techniques. In [46], for example, using convolutional neural networks, CNN, to recognize different types of weeds, recognition percentages of 97.78% are achieved in validation. On the other hand, [47], an exhaustive review of works on machine learning and deep learning, highlights research results for crop and weed discrimination with accuracy percentages of up to 95.1% using ANN and up to 98.2% using SVM. In [48], for weed recognition in strawberry and pea crops, recognition accuracies of 95.3% are achieved with CNN and 63.7% and 84.9% for SVM and KNN, respectively.

In synthesis, in this work, the experiments with only texture patterns, hue (H), saturation (S), and intensity (I), without considering fractal dimension, showed that quadratic SVM reliably classifies coffee and weed samples, with Precision and Recall near or upper to 95%, on average. KNN, for its part, classifies with Precision and recall upper to 90% for Hue and Intensity patterns and upper to 86% for S patterns, on average, too; nevertheless, classification is better for SVM.

ANN does a reasonable classification for hue and intensity patterns, but better for hue, with Precision and Recall equal to 100% in the testing samples of hue patterns and upper to 80% for intensity patterns. Precision and Recall are minor to 60% for saturation patterns, on average.

With Hue, Saturation, and Intensity (HIS) patterns, all classifiers have Precision and recall minor to 60%, on average.

As was commented above, it is important to mention that when the fractal dimension is added as the fifth pattern, Precision and Recall improve in all classifiers, achieving values near or upper to 97%, on average, in most experiments. However, even if the fractal dimension is used, the classifiers function better with individual patterns H, S, and I than with pattern mixed (HIS). For pattern mixed (HIS) plus fractal dimension, Precision and Recall achieve values of 80%, on average.

IV. Conclusions

The extraction of texture patterns from coffee and weeds images, in most experiments done, permits their classification with precision and recall upper or equal to 95%, on average, when the fractal dimension is not used, and upper or equal to 97% on average when the fractal dimension is used as the fifth pattern. From the experimented classifiers, SVM and ANN have better outcomes than the KNN method. The classification has better results for individual texture patterns (hue, saturation, intensity), reduced by PCA, than for the mixed pattern (HSI), also reduced by PCA. In all tests, the fractal dimension improves the performance of classifiers. Experiments suggest that using this technology to identify and classify weeds associated with the coffee crop is viable. Tests with an extended group of weeds and samples, just as classification experiments in real-time, are necessary for classifier validation. Applications of this type of technology are fundamental to improving the efficiency of the weed control system for the benefit of coffee crops and environmental protection.

References

[1] Y. Blanco and Á. Leyva, “Las Arvenses En El Agroecosistema Y Sus Beneficios Naturales,” *Cultivos Tropicales*, vol. 28, no. 2, pp. 21–28, 2007, doi: http://dx.doi.org/10.13140/RG.2.2.10964.19844.

[2] E. Gamboa and L. Sánchez, “Control de malezas con herbicidad y métodos mecánicos en plantaciones jóvenes de café,” *Bioagro*, vol. 16, no. 2, pp. 133–136, 2004.

[3] A. Anzalone and A. Silva, “Evaluación de herbicidas sulfonilureas para el control de malezas en cafetales,” *Bioagro*, vol. 22, no. 2, pp. 95–104, 2010.

[4] CENICAFE, “Manejo de las malezas o arvenses en los cafetales,” 2004, *CENICAFÉ*.

[5] C. P. de Souza, T. de A. Guedes, and C. S. Fontanetti, “Evaluation of herbicides action on plant bioindicators by genetic biomarkers: a review,” *Environmental Monitoring and Assessment*, vol. 188, no. 12, 2016, doi: 10.1007/s10661-016-5702-8.

[6] R. Hoagland and D. B. Clyde, “Controlling herbicide-susceptible, -tolerant and -resistant weeds with microbial bioherbicides,” *Outlooks on Pest Management*, vol. 27, pp. 196–197, 2016, doi: 10.1564/v27_dec_04.

[7] W. T. Tsai, “Status of herbicide use, regulatory management and case study of paraquat in Taiwan,” *Environment, Development and Sustainability*, vol. 22, no. 3, pp. 2673–2683, 2018, doi: 10.1007/s10668-018-0293-x.

[8] M. Bilal, H. M. N. Iqbal, and D. Barceló, “Persistence of pesticides-based contaminants in the environment and their effective degradation using laccase-assisted biocatalytic systems,” *Science of the Total Environment*, vol. 695, 2019, doi: 10.1016/j.scitotenv.2019.133896.

[9] N. M. Freitas, F. C. L. Freitas, I. F. Furtado, M. F. F. Teixeira, and V. F. Silva, “Herbicide mixtures to control dayflowers and drift effect on coffee cultures,” *Planta Daninha*, vol. 36, 2018, doi: 10.1590/S0100-83582018360100047.

[10] P. V. da Silva, G. C. Barbosa, A. Ferrari, S. M. Tronquini, and P. A. Monquero, “Chemical control strategies of commelina benghalensis in coffee crop,” *Coffee Science*, vol. 14, no. 2, pp. 231–239, 2019, doi: 10.25186/cs.v14i2.1576.

[11] X. M. de S. Vilela, F. C. L. de Medeiros, A. H. Gonçalves, and L. C. da Silva, “Quizalofop-p-ethyl controlling sourgrass (Digitariavilela insularis)x. m.andde s. egoosegrass (eleusine indica) in infested coffee areas,” *Coffee Science*, vol. 14, no. 4, pp. 530–537, 2019, doi: 10.25186/cs.v14i4.1646.

[12] J. A. López, D. Villalba, L. F. Salazar, and O. A. Cárdenas, “Manejo integrado de arvenses en el cultivo de café Nueva alternativa de control químico,” *Avances Técnicos Cenicafé*, pp. 1–8, 2012.

[13] A. Anzalone, M. Arizaleta, and J. Vargas, “Respuesta del cafeto (Coffea arabica) ‘catuaí’ a los herbicidas glifosato, clomazone, linuron, 2,4-D, metsulfuron-metil, rimsulfuron y clorimuron-etil,” *Bioagro*, vol. 26, no. 1, pp. 3–12, 2014.

[14] H. Menza and L. F. Salazar, “Estudios de resistencia al glifosato en tres arvenses de la zona cafetera colombiana y alternativas para su manejo,” *Avances Técnicos Cenicafé*, no. 350, 2006.

[15] C. Martínez, “Taller sobre : “ Impactos toxicológicos y ecotoxicológicos asociados ‘ Efectos tóxicos del glifosato y sus formulaciones comerciales ,’” 2011, *Facultad de Agronomía Universidad de la República de Uruguay*.

[16] M. Arizaleta, A. Anzalone, and A. Silva, “Efecto del uso de metsulfuron-metil y glifosato sobre malezas asociadas a cafetales en venezuela,” *Bioagro*, vol. 20, no. 2, pp. 79–88, 2008.

[17] J. A. García, E. Alarcón, Y. Hernández, and C. Hernández, “Impact of litter contaminated with glyphosate-based herbicide on the performance of Pontoscolex corethrurus, soil phosphatase activities and soil pH,” *Applied Soil Ecology*, vol. 104, no. 2015, pp. 31–41, 2016, doi: 10.1016/j.apsoil.2016.03.007.

[18] R. Melloni, E. M. Silve, M. I. N. Alvarenga, E. G. P. Melloni, and E. N. de Alcântara, “Impact of weedcontrol on amf propagules and mycorrhization of coffee trees,” *Coffee Science*, vol. 12, no. 2, pp. 207–215, 2017.

[19] J. C. Niemeyer, F. B. de Santo, N. Guerra, A. M. Ricardo Filho, and T. M. Pech, “Do recommended doses of glyphosate-based herbicides affect soil invertebrates? Field and laboratory screening tests to

risk assessment," *Chemosphere*, vol. 198, pp. 154–160, 2018, doi: 10.1016/j.chemosphere.2018.01.127.

[20] P. De Pádua Marafeli, P. R. Reis, L. F. De Oliveira Bernardi, E. N. De Alcântara, and P. A. Martinez, "Effects of weed management on soil mites in coffee plantations in a neotropical environment," *Neotropical Biology and Conservation*, vol. 14, no. 2, pp. 275–289, 2019, doi: 10.3897/neotropical.14.e38094.

[21] C. B. de Novais, L. Avio, M. Giovannetti, S. M. de Faria, J. O. Siqueira, and C. Sbrana, "Interconnectedness, length and viability of arbuscular mycorrhizal mycelium as affected by selected herbicides and fungicides," *Applied Soil Ecology*, vol. 143, no. June, pp. 144–152, 2019, doi: 10.1016/j.apsoil.2019.06.013.

[22] S. Pochron *et al.*, "Temperature and body mass drive earthworm (Eisenia fetida) sensitivity to a popular glyphosate-based herbicide," *Applied Soil Ecology*, vol. 139, no. March, pp. 32–39, 2019, doi: 10.1016/j.apsoil.2019.03.015.

[23] J. A. García-Pérez, E. Alarcón-Gutiérrez, and F. Díaz-Fleischer, "Interactive effect of glyphosate-based herbicides and organic soil layer thickness on growth and reproduction of the tropical earthworm Pontoscolex corethrurus (Müller, 1857)," *Applied Soil Ecology*, vol. 155, no. July 2019, p. 103648, 2020, doi: 10.1016/j.apsoil.2020.103648.

[24] Y. P. Setyawan, M. Naim, A. D. Advento, and J. P. Caliman, "The effect of pesticide residue on mortality and fecundity of Elaeidobius kamerunicus (Coleoptera: Curculionidae)," *IOP Conference Series: Earth and Environmental Science*, vol. 468, no. 1, 2020, doi: 10.1088/1755-1315/468/1/012020.

[25] G. Belutti *et al.*, "Phytotoxicity and Growth of Coffee Plants As a Function of The Application of Herbicide 2,4-D," *Coffee Science*, vol. 14, no. 4, pp. 438–445, 2019.

[26] D. Toledo, A. de Oliveira, G. Belutti, T. Teruel, P. Menicucci, and R. J. Guimarães, "Growth, anatomy and physiology of coffee plants intoxicated by the herbicide glyphosate," *Coffee Science*, vol. 14, no. 1, pp. 76–82, 2019, doi: 10.25186/cs.v14i1.1530.

[27] D. Toledo *et al.*, "Selectivity of the herbicide chlorimuron ethylcastanhon young coffee plants," *Coffee Science*, vol. 14, no. 4, pp. 467–472, 2019, doi: 10.25186/cs.v14i4.1615.

[28] L. Parra, J. Marin, S. Yousfi, G. Rincón, P. V. Mauri, and J. Lloret, "Edge detection for weed recognition in lawns," *Computers and Electronics in Agriculture*, vol. 176, no. July, 2020, doi: 10.1016/j.compag.2020.105684.

[29] C. A. Pulido-Rojas, M. A. Molina-Villa, and L. E. Solaque-Guzmán, "Machine vision system for weed detection using image filtering in vegetables crops," *Revista Facultad de Ingenieria*, vol. 2016, no. 80, pp. 124–130, 2016, doi: 10.17533/udea.redin.n80a13.

[30] X. E. Pantazi, D. Moshou, and C. Bravo, "Active learning system for weed species recognition based on hyperspectral sensing," *Biosystems Engineering*, vol. 146, pp. 193–202, 2016, doi: 10.1016/j.biosystemseng.2016.01.014.

[31] R. Raja *et al.*, "Crop signalling: A novel crop recognition technique for robotic weed control," *Biosystems Engineering*, vol. 187, pp. 278–291, 2019, doi: 10.1016/j.biosystemseng.2019.09.011.

[32] A. dos Santos Ferreira, D. M. Freitas, G. G. da Silva, H. Pistori, and M. T. Folhes, "Unsupervised deep learning and semi-automatic data labeling in weed discrimination," *Computers and Electronics in Agriculture*, vol. 165, no. July, 2019, doi: 10.1016/j.compag.2019.104963.

[33] T. Kounalakis, G. A. Triantafyllidis, and L. Nalpantidis, "Deep learning-based visual recognition of rumex for robotic precision farming," *Computers and Electronics in Agriculture*, vol. 165, no. August, 2019, doi: 10.1016/j.compag.2019.104973.

[34] J. Yu, S. M. Sharpe, A. W. Schumann, and N. S. Boyd, "Deep learning for image-based weed detection in turfgrass," *European Journal of Agronomy*, vol. 104, no. November 2018, pp. 78–84, 2019, doi: 10.1016/j.eja.2019.01.004.

[35] K. Hu, G. Coleman, S. Zeng, Z. Wang, and M. Walsh, "Graph weeds net: A graph-based deep learning method for weed recognition," *Computers and Electronics in Agriculture*, vol. 174, no. April, 2020, doi: 10.1016/j.compag.2020.105520.

[36] H. Jiang, C. Zhang, Y. Qiao, Z. Zhang, W. Zhang, and C. Song, "CNN feature based graph convolutional network for weed and crop recognition in smart farming," *Computers and Electronics in Agriculture*, vol. 174, no. April, p. 105450, 2020, doi: 10.1016/j.compag.2020.105450.

[37] R. Raja, T. T. Nguyen, D. C. Slaughter, and S. A. Fennimore, "Real-time weed-crop classification and localisation technique for robotic weed control in lettuce," *Biosystems Engineering*, vol. 192, pp. 257–274, 2020, doi: 10.1016/j.biosystemseng.2020.02.002.

[38] R. Raja, T. T. Nguyen, V. L. Vuong, D. C. Slaughter, and S. A. Fennimore, "RTD-SEPs: Real-time detection of stem emerging points and classification of crop-weed for robotic weed control in producing tomato," *Biosystems Engineering*, vol. 195, pp. 152–171, 2020, doi: 10.1016/j.biosystemseng.2020.05.004.

[39] T. Ashraf and Y. N. Khan, "Weed density classification in rice crop using computer vision," *Computers and Electronics in Agriculture*, vol. 175, no. June, 2020, doi: 10.1016/j.compag.2020.105590.

[40] D. G. Kim, T. F. Burks, J. Qin, and D. M. Bulanon, "Classification of grapefruit peel diseases using color texture feature analysis," *International Journal of Agricultural and Biological Engineering*, vol. 2, no. 3, pp. 41–50, 2009, doi: 10.3965/j.issn.1934-6344.2009.03.041-050.

[41] A. Cravero, S. Pardo, S. Sepúlveda, and L. Muñoz, "Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review," Mar. 01, 2022, *MDPI*. doi: 10.3390/agronomy12030748.

[42] P. Dini and S. Saponara, "Analysis, design, and comparison of machine-learning techniques for

networking intrusion detection," *Designs (Basel)*, vol. 5, no. 1, pp. 1–22, 2021, doi: 10.3390/designs5010009.

[43] C. Ji, T. B. Mudiyanselage, Y. Gao, and Y. Pan, "A review of infant cry analysis and classification," Dec. 01, 2021, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1186/s13636-021-00197-5.

[44] I. Attri, L. K. Awasthi, and T. P. Sharma, "Machine learning in agriculture: a review of crop management applications," *Multimed Tools Appl*, vol. 83, no. 5, pp. 12875–12915, Feb. 2024, doi: 10.1007/s11042-023-16105-2.

[45] M. J. Al Dujaili, A. Ebrahimi-Moghadam, and A. Fatlawi, "Speech emotion recognition based on SVM and KNN classifications fusion," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 2, pp. 1259–1264, Apr. 2021, doi: 10.11591/ijece.v11i2.pp1259-1264.

[46] Y. Mu *et al.*, "DenseNet weed recognition model combining local variance preprocessing and attention mechanism," *Front Plant Sci*, vol. 13, Jan. 2023, doi: 10.3389/fpls.2022.1041510.

[47] M. H. Saleem, J. Potgieter, and K. M. Arif, "Automation in Agriculture by Machine and Deep Learning Techniques: A Review of Recent Developments," Dec. 01, 2021, *Springer*. doi: 10.1007/s11119-021-09806-x.

[48] S. Khan, M. Tufail, M. T. Khan, Z. A. Khan, and S. Anwar, "Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer," *Precis Agric*, vol. 22, no. 6, pp. 1711–1727, Dec. 2021, doi: 10.1007/s11119-021-09808-9.

Biografía:

Milton Muñoz: is an Electronic Engineer with a Master of Science degree in this field who graduated from the Industrial University of Santander (UIS), Colombia. He has over 20 years of experience in research development and engineering education. He was a recipient of a Colciencias scholarship and is currently recognized as a Junior Researcher by Colciencias. Milton has published in international journals, classified as A2, and national journals, categorized as B and C. Additionally, he served as an editor for an international journal, which was also rated A2. Throughout his career, he has developed eight research prototypes and has directed or evaluated more than 30 undergraduate projects. He has also managed, organized, and acted as an academic coordinator for four international engineering and technological development events. His areas of interest include modeling, simulation and control, artificial intelligence, embedded systems, and basic science education. ORCID: https://orcid.org/0000-0001-6724-9888.

Jaime Alberto Ortíz Delgado: is an-Electronic Engineer who graduated from the University Foundation of San Gil, UNISANGIL, Colombia. Since his graduation in 2020, he has worked in different fields, including machine learning, commercial advisor in topics related to computers and industrial electronics, and data center infrastructure management. His areas of interest include pattern recognition, industrial automation, and telecommunication. ORCID: https://orcid.org/0009-0001-9318-2506