

## Research Article

# Improvement of Sentinel-1 Remote Sensing Data Classification by DWT and PCA

O. Charfi Marrakchi <sup>1</sup>, C. Masmoudi Charfi <sup>2</sup>, M. Hamzaoui,<sup>3</sup> and H. Habaieb <sup>3</sup>

<sup>1</sup>Carthage University/INSAT/GREEN-TEAM Laboratory (LR17AGR01-INAT), Centre Urbain Nord BP 676, 1080 Tunis, Tunisia

<sup>2</sup>Sfax University/Olive Tree Institute, Tunisia/Integrated Olive Production Laboratory, Avenue Hedi Karray, 1002 Tunis, Tunisia

<sup>3</sup>Carthage University/INAT/LR GREEN-TEAM, 43 Avenue Charles Nicolle, 1082 Tunis, Tunisia

Correspondence should be addressed to O. Charfi Marrakchi; olfa.charfi@gmail.com

Received 5 September 2020; Revised 25 January 2021; Accepted 6 February 2021; Published 28 February 2021

Academic Editor: Roberto Paolesse

Copyright © 2021 O. Charfi Marrakchi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article presents a new alternative for data resource, by applying the proposed methods of Principal Components Analysis (PCA) or of Discrete Wavelet Transformation (DWT) on the VV and VH polarization images of the Sentinel-1 radar satellite, aiming at a better classification of data. The study area concerns the Houareb site located in the city of Kairouan in central Tunisia. In addition to Sentinel-1 data, field truth data and the Euclidian Minimum Distance (EMD) criterion were used for classification and validation. Energy descriptors have been proposed in this study for classifications. Cross validation was used to evaluate the results of the classification. The best classification result was achieved using the DWT method applied on VH and VV images with an Overall Precision (OA) of 0.671 and 0.548, respectively, against an OA value of 0.371 and of 0.449 when the PCA method and the Minimum Distance (MDist) classifier were applied on the dual (VV; VH) polarization, respectively. The DWT transformation gives the highest Kappa Precision Coefficient (KPC) of 0.8.

## 1. Introduction

Sentinel-1 satellite images in Interferometric Wide (IW) swath mode at level 1 with Ground Range Detection (GRD) were widely used in the last years to identify the Land Cover (LC). Makinde and Oyelade [1] established the LC map of the Lagos site, located in Nigeria by using the maximum likelihood method which was applied for both the VV and VH polarization GRD\_IW Sentinel-1 images and achieved an OA of 0.757 and a KPC of 0.719. Abdikan et al. [2] combined VV and VH image polarizations according to eight scenarios and used the Support Vector Machine (SVM) classifier to map the LC in urban area of the megapole site at Istanbul, in Turkey. Results showed higher OA value, of 0.933 instead of 0.739 when the VV polarization is used. Balzter et al. [3] used the full polarizations and the digital elevation data of the SRTM to provide the LC maps of the Thüringen region located in Germany. Four classified images were provided using the Random Forest (RF) method. Result evaluation was made with the 2006-CORINE map. The OA and KPC

values were equal to 0.684 and 0.63, respectively. Various approaches used the synergy between the remotely sensed optical and radar data to elaborate the retrieval LC maps. Recently, Bousbih et al. [4] used the classical RF and SVM classifiers with different indicators, derived from the Sentinel constellations (Sentinel-1 and Sentinel-2) to generate the clay content map of a semiarid LC located in central Tunisia. The used moisture indicator was given from the soil moisture maps over different periods. Maps are obtained by combining the Synthetic Aperture Radar (SAR) and the optic data in the Water Cloud Model (WCM) and in the Integral Equation Model (IEM). Gao et al. [5] used the Sentinel-1 and the sentinel-2 data fusion over a site in Urgell (Catalunya, Spain) to retrieve the soil moisture by using the change detection approach and the backscattered NDVI Sentinel-1 radar index. The synergy use of the optic and the radar data was helpful to LC discrimination and identification. However, the full polarimetric radar data and the optic data were not free at the same time. The almost free IW\_GRD\_Sentinel-1 data are available in VV and VH polarizations, only, like

for the Tunisian country. Sentinel-1 and Sentinel-2 data can be free downloaded from the Copernicus web site (<https://scihub.copernicus.eu/>). Unfortunately, the only use of this dual polarization (VV; VH) provides poor OA and KPC values for Sentinel-1 data classification. For this reason, we propose to use PCA or DWT orthogonal transformations on VV and VH Sentinel-1 images in order to improve the LC classification.

The orthogonal transformations can lead to a spatial or a frequency analysis depending on the nature of the transformation [6]. When the Finite Impulse Response (FIR) filter is used to filter the satellite image, the frequency analysis can be made. Hence, the new components issued from the filtering enclose energy information. The new component issued from the spatial transformation are obtained from the pixel image projection procedure such as the PCA and the Independent Component Analyses (ICA). These methods use interband covariance matrix to build the new space representation in which bands are uncorrelated.

The PCA transformation is widely used in feature extraction methods. Alons and Malpica [7] used the PCA to obtain a pan sharpening remote sensing images from the multispectral bands in order to perform the automatic classification methods, while Wang and Wang [8] elaborated a region-based unsupervised segmentation by adaptively combining the texture and the spectral distributions using the PCA method. Singh and Kaur [9] used the PCA to verify if the reduced feature sets for both the water and the urban area coverage in SAR image database are the same as obtained by the GLCM (Gray Level Cooccurrence Matrix) and the GLRLM (Gray Level Run Length Matrix).

In literature, the DWT was applied on (i) SAR data for noise filtering of the Sentinel-1 data [10], (ii) for preclassification change detection from the Sentinel-1 multipolarized images [11], (iii) to retrieve the wind direction from a series of VV-Sentinel-1 images [12], (iv) for image classification [13], and for parameterizing the feedforward neural networks to improve remote sensing LC identification [14]. In this context, this paper shows the powerful use of the DWT components of the Sentinel-1 data, comparing to the use of the spatial PCA method, for a classification procedure. Classifications accuracies are based on the field measurements acquired over the Kairouan plain study site, located in central Tunisia (Lat.  $35^{\circ} 40.686'$  N, Long.  $10^{\circ} 5.7798'$  E).

During the two last decades, this study area was a subject of interest of many researchers. The most recent study is that of Bousbih et al. [4]. Authors generate the clay content map of the plain, using the Sentinel-1 and the Sentinel-2 data. Bousbih et al. [15] used the multirate VV and VH Sentinel-1 measurements in synergy with Landsat-8 optical data, to evaluate the soil characteristics (moisture and roughness) and the vegetation parameters of the same site. Zribi et al. [16] used the ENVISAT ASAR radar data (C-band) in synergy with the SPOT-5 optical images. Zribi et al. [17] elaborated for the same site the moisture model of bare soils by combining the Envisat ASAR and the TerraSAR-X multispectral (L, C, and X) data, which was acquired within in situ measurements of the soil moisture content and the ground

surface roughness. Other studies were carried out to map the Kairouan plain LC by using only the optical sensors [18]. So far, there is no any satisfactory classified map for this site.

This paper aims to provide a Sentinel-1 classified image of the Kairouan plain by using the above-mentioned methods. Our paper is organized as follows: in Section 2, the studied area is described; the ground measurements and the database are presented and the two proposed procedures are described. Section 3 presents the results and the discussion. Finally, conclusions are presented in the last section.

## 2. Materials and Methods

**2.1. Study Site.** The study region is located in the south west of Kairouan as presented in Figure 1 with latitudes  $35^{\circ}$  to  $35^{\circ} 45'$  N and longitudes  $9^{\circ} 30'$  to  $10^{\circ} 15'$  E, in the center of Tunisia. The study area was chosen as an experimental area by the SIDA FAO project. The region is characterized by a semiarid climate, and it is considered as a plain area. In the Kairouan plain, agriculture is the main economic activity. It includes animal husbandry, cereals and vegetable crops, and arboriculture (almond, apricot, citrus, and olive) dominated mainly by olive plantations.

**2.2. GCP Measures.** A campaign of Ground Control Points (GCP) of LCs was carried out from 10 to 12 June 2019 on the study area by using the free ODK brief case android application (<https://odk-demo.readthedocs.io/en/latest/briefcase-install.html>). The accuracy of the GCP is 14.53 meters. The GPS point's data were retrieved on computer and processed on the Qgis 3.4.6 software to obtain the georeferenced LC map which represents the GCPs centered in polygons of each LC species. Seven LC species were identified on this site as presented in Table 1 and Figure 2: fruit trees, cereals, fallow, bare soil, soil covered with straws, olive trees, and vegetable crops. The urban area, daim and wadi were visited in campaign and then localized on the TCI Sentinel-2B subimage of the study area as presented in Figure 1.

**2.3. DEM Map.** The Shuttle Radar Topography Mission (SRTM) data was used to extract the Digital Elevation Model (DEM) at 30 m of the study area as presented in Figure 3(a). The SRTM data can be free downloaded from website <http://edcns17.cr.usgs.gov/EarthExplorer/>.

The study area presents in the north and north-west a mountain chain. The DEM map was been thresholded, so pixel slope values  $>15\%$  were converted to grayscale 255 and the others to 0 as presented in Figure 3(b). From the slope map, we masked some reliefs of the study area with black color (as presented in Figure 3(c)).

**2.4. Sentinel-1 Data and Processing.** Sentinel-1 data were downloaded from the ESA's Copernicus Open Access Hub (<https://scihub.copernicus.eu/>) at the date of 05 June 2019. No LC change was detected from this date to the campaign date. Sentinel-1 satellite provides the level 1 GRD products in IW mode in descending and ascending orbit at an angle of incidence between 39 and 40 degrees and in the C band

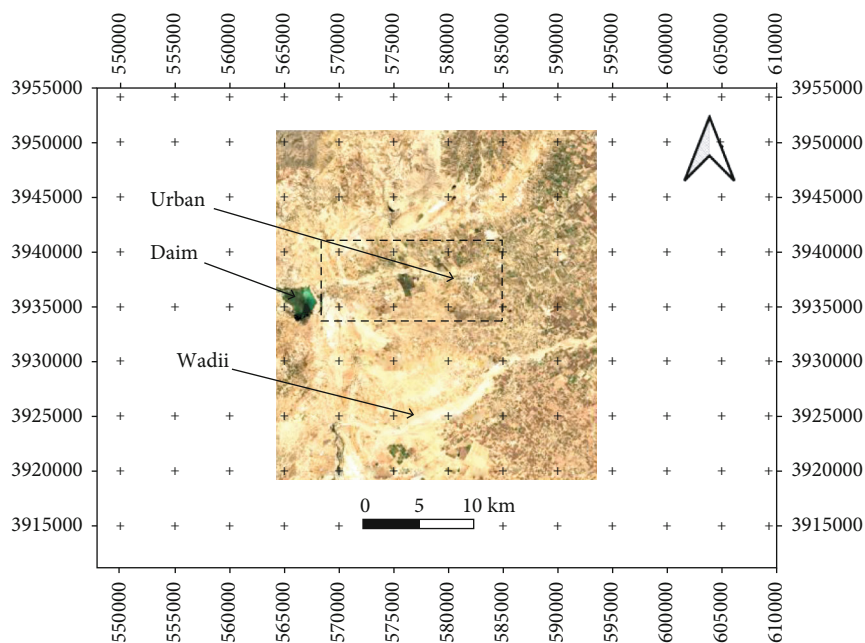


FIGURE 1: TCI Sentinel-2B subimage of the study area—June 5, 2019—TCI image is in true color composite bands (B4: red, B3: green, B2: blue). In image, rectangle delimitates campaign area.

TABLE 1: List of classes identified on the study area.








Classes	Polygon label	Species	LC
C1	None	Urban areas	LC1
C2		Bare soil	LC2
C3		Soil covered with straws	LC3
C4	None	Wadii	LC4
C5		Olive trees	LC5
C6		Fallow	LC6
C7		Cereals	LC7
C8		Vegetable crops	LC8
C9	None	Daim	LC9
C10		Fruit trees	LC10



FIGURE 2: Georeferenced map of the study area (zoom  $\times 4$ ). The LC polygons are centered on the GCPs measures. The polygon label is shown in Table 1.

step permits to correct the geometric distortions by using the Digital Elevation Model (DEM) proposed by the NASA's Shuttle Radar Topography Mission (SRTM); (iii) the thermal noise removal step using the Lee filter.

The subimages as shown in Figures 4(a) and 4(b) were extracted from the VH and VV polarization images; they cover the region of interest (as presented in Figure 1). The size of each image is 3227 pixels in line by 2917 pixels in column and is coded on 16 bits. The VH and VV subimages have a covariance of 733.5. Thus, the LC similarities included in the polarization images reduce the accuracy of the classification. We propose to reduce this correlation between bands by the use of PCA or DWT transformations.

frequency with a spatial resolution of 10 meters in dual polarization (VV and VH).

The Sentinel-1 data was preprocessed using the Sentinel Application Platform (SNAP) as follows: (i) the calibration step allows the conversion of the signal recorded by the sensor in the form of digital accounts into backscatter coefficient; (ii) the geometric terrain correction (georeferencing)

2.5. Classifier Algorithms. Classification methods which use pixel values or distribution function give low OA values because of the effect of the cross-pixel's values correlation of the LCs [19]. The PCA and the DWT methods transform

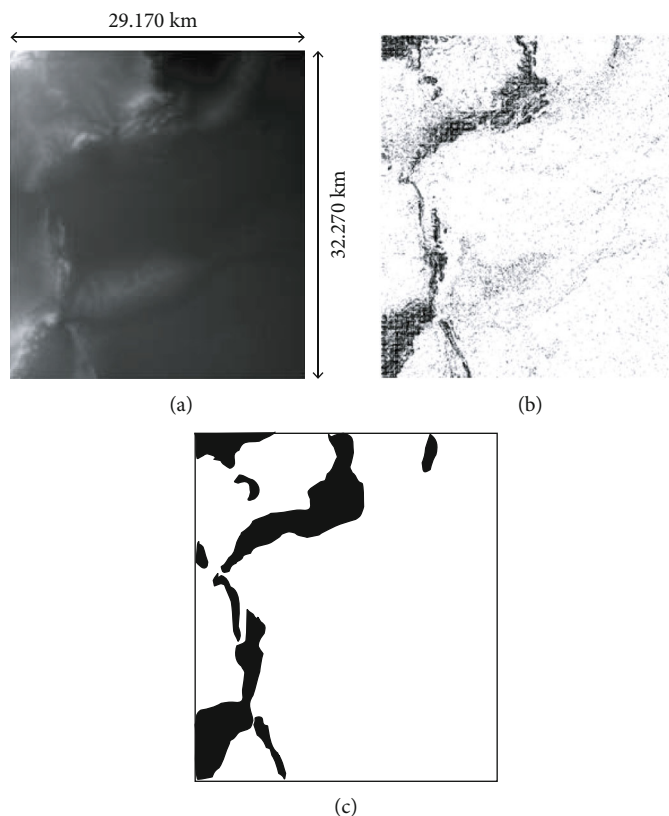


FIGURE 3: (a) Slope map derived from DEM; (b) thresholding DEM map; (c) relief mask.

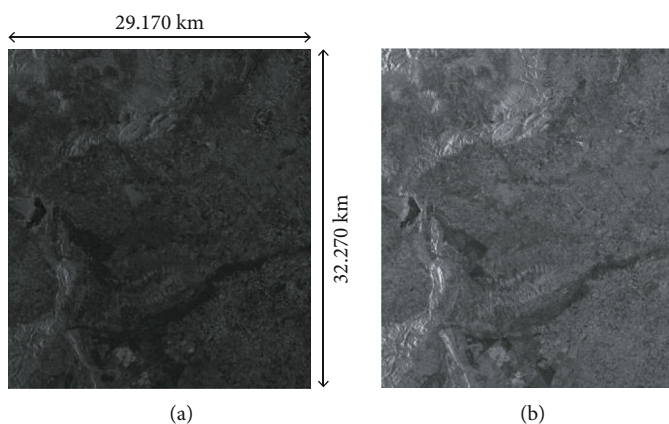


FIGURE 4: Sentinel-1 polarization subimages. (a) VH; (b) VV.

the initial images into a set of new images which are uncorrelated with each from other, and the LCs are uncorrelated in the new space representation. The contribution of these transformations is to reduce the LC correlation between the VV and VH bands and so to increase the LC-OA value of the classified image.

In this paper, DWT and PCA transformations were applied on Sentinel-1 data and the obtained pansharpening images of the study area have been classified according to the EMD criterion (Figure 5) Results were compared to those

obtained with the use of the MDist method in order to see the contribution of orthogonal transformations for the improvement of the classification results.

**2.5.1. DWT Description.** DWT is a multiresolution approach. The bidimensional DWT is achieved by implementing a bank of single-dimensional filters (filters designed from a Mother Function (MF)) which are low-pass,  $h(x)$ , and high-pass,  $g(x)$ , analysis filters. For one level of redundant DWT decomposition as presented in Figure 6, the image is

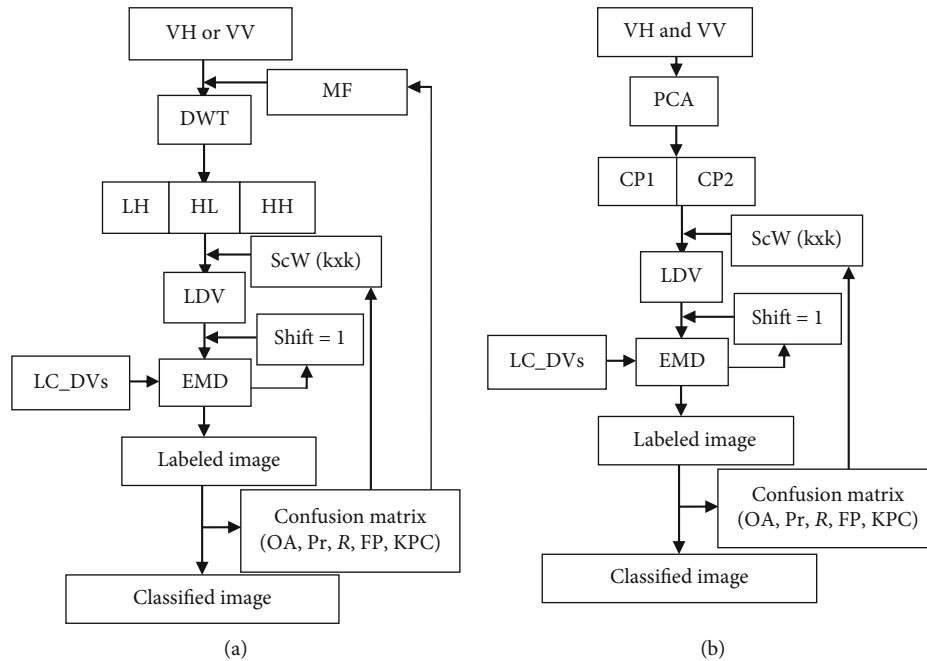


FIGURE 5: Classification flowcharts: (a) DWT method, (b) PCA method.

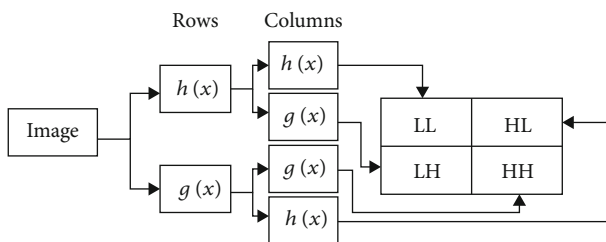


FIGURE 6: One level of DWT decomposition of an image. LH image corresponds to the lines low-pass and then the columns high-pass-filtered version of the image; HL image corresponds to the lines high-pass and then the columns low-pass-filtered version of the image. HH image corresponds to the lines high-pass and then the columns high-pass-filtered version of the image.

decomposed into three detail images (LH, HL, and HH) corresponding to the image distinct frequency bands and into the LL subband which is the low-pass-filtered version of the image. The LL subband is further decomposed with the same manner, in the second level decomposition [20, 21]. The four subbands have the same size as the original image in case of the redundant DWT decomposition. We have applied the redundant DWT on the VV and VH Sentinel-1 images separately at one level of decomposition, and we have obtained three details subbands and an LL band for each image.

**2.5.2. PCA Transform.** To apply the PCA on the Sentinel-1 data (composed of the bibands VV and VH) we have, firstly, to compute the covariance matrix between the VV and VH images and, secondly, to project each pixel of Sentinel-1 image from the VV and VH space representation into the new space CP1 and CP2, respectively, by using a transform

matrix which is composed by the eigenvectors of the covariance matrix [6]. Figure 7 shows the new CP1 and CP2 components of the Sentinel-1 images of the study area.

**2.5.3. LC Descriptors.** The used LC descriptors are described as follows [6, 22]:

$$\text{Energy} : E = \frac{1}{M} \sum_{l,m} (D(l, m))^2, \quad (1)$$

$$\text{L1 norm} : MD = \frac{1}{M} \sum_{l,m} |D(l, m)|. \quad (2)$$

These descriptors are computed in a Region Of Interest (ROI), where  $(l, m)$  are the coordinates of a pixel in the pansharpening image  $D$ ;  $M$  is the ROI size, given in pixels.

**2.5.4. Classification Procedure.** The EMD criterion was used to classify the Sentinel-1 data of the study area by using the DWT or PCA methods as shown in Figures 5(a) and 5(b), respectively. In the classification procedure, the EMD represents the minimum Euclidian distance between the Local Descriptor Vector (LDV) and the LC energy Descriptor Vectors (LC\_DV). The LDV (use (3) in case of DWT, use (4) in case of PCA) was computed for each ROI on the pansharpening images according to formulas (1) and (2), for each position of the ROI-Scanning Window (ScW), which has a size equal to  $k \times k$ . The used scanning shift is equal to 1, and the classification result is a labeled image.

In DWT (PCA, respectively), the LC\_DV (3) (LC\_DV (4), respectively) was computed from each LC feature extracted from DWT subbands (CP1 and CP2, respectively) for each (for both, respectively) polarized Sentinel-1 image

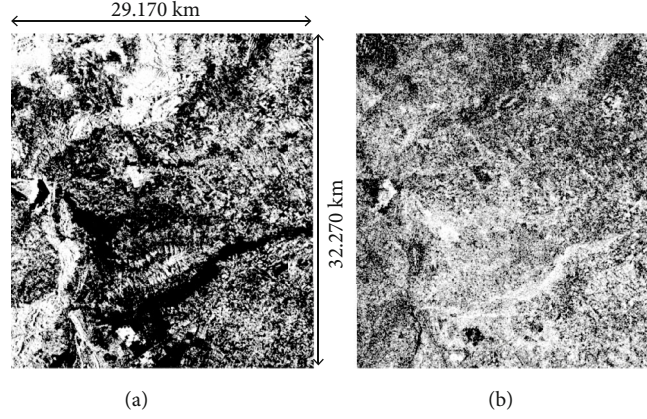


FIGURE 7: PCA components of the Sentinel-1 data of the study area: (a) CP1; (b) CP2.

TABLE 2: Examples of LC\_DV computed on LC ROI located on VH and VV images, PCA components, subbands obtained by the DWT applied to the image of the polarization VH according to MF= 'Coiflet5', subbands obtained by the DWT applied to the image of the polarization VV according to MF= 'sym2'.

Bands	LC_DV	LC1	LC2	LC3	LC4	LC5	LC6	LC7	LC8	LC9	LC10
HL <sub>VH</sub>	<i>E</i> (1.0e-003)	1.2	0.5	0.6	0.3	0.8	0.6	0.4	0.7	0.2	0.9
	MD (1.0e-005)	0.32	0.05	0.06	0.02	0.10	0.09	0.02	0.09	0.01	0.12
LH <sub>VH</sub>	<i>E</i> (1.0e-003)	1.3	0.6	0.6	0.3	0.9	0.6	0.6	0.9	0.2	0.8
	MD (1.0e-005)	0.43	0.06	0.07	0.02	0.13	0.05	0.06	0.15	0.01	0.11
HH <sub>VH</sub>	<i>E</i> (1.0e-003)	0.1	0.1	0.1	0	0.1	0.1	0	0.1	0	0.1
	MD (1.0e-007)	0.28	0.05	0.6	0.02	0.13	0.05	0.03	0.07	0.01	0.09
HL <sub>VV</sub>	<i>E</i> (1.0e-003)	9.5	5.7	5.6	2.8	7.4	5.8	5.0	8.8	1.8	6.3
	MD (1.0e-003)	0.27	0.05	0.05	0.01	0.09	0.05	0.04	0.14	0.01	0.07
LH <sub>VV</sub>	<i>E</i> (1.0e-003)	9.3	5.9	4.8	2.7	6.6	5.4	4.7	6.7	1.9	6.4
	MD (1.0e-003)	0.27	0.06	0.04	0.01	0.07	0.05	0.04	0.07	0.01	0.07
HH <sub>VV</sub>	<i>E</i> (1.0e-003)	1.8	1.1	1.1	0.5	1.3	1.0	0.9	1.2	0.4	1.3
	MD (1.0e-005)	0.9	0.2	0.2	0.1	0.3	0.2	0.1	0.2	0	0.3
CP1	<i>E</i>	35.4	38.8	21.3	53.1	34.5	22.6	26.4	22.9	61.5	27.7
	MD (1.0e+003)	2.3	2.1	0.7	3.3	1.8	0.9	1	0.9	4.6	1.4
CP2	<i>E</i>	15.6	9.1	9.4	8.6	14.9	9	8.2	12.3	12.2	10.6
	MD (1.0e+003)	0.43	0.14	0.14	0.11	0.34	0.13	0.10	0.25	0.25	0.18
VH	<i>E</i>	83	65	61	47	79	66	56	87	60	79
	MD (1.0e+04)	0.75	0.45	0.40	0.24	0.66	0.46	0.32	0.82	0.46	0.59
VV	<i>E</i>	155	146	142	103	161	154	130	173	130	159
	MD (1.0e+04)	2.60	2.20	2.06	1.10	2.69	2.45	1.73	3.17	1.98	2.74

TABLE 3: Nonnormalized confusion matrix.

	C1	C2	C3	Subtotal
LC1	$X_{11}$	$X_{12}$	$X_{13}$	$Y_1$
LC2	$X_{21}$	$X_{22}$	$X_{23}$	$Y_2$
LC3	$X_{31}$	$X_{32}$	$X_{33}$	$Y_3$
Subtotal	$Z_1$	$Z_2$	$Z_3$	$T$

according to the position of the LC polygons (Figure 2). For each LC specie (Table 1), it attributed an LC\_DV (Table 2). We have considering 2/3 of these LC polygons as training samples, and the rest (1/3) were used for validation of the image classification.

$$V_{DWT} = (E_{LH}, E_{HL}, E_{HH}, MD_{LH}, MD_{HL}, MD_{HH}), \quad (3)$$

$$V_{PCA} = (E_{CP1}, E_{CP2}, MD_{CP1}, MD_{CP2}). \quad (4)$$

TABLE 4: Highest values of  $R$  of the classes identified on Sentinel-1 classified data using DWT method.

Classes	MF	Polarization	$k$	OA	Pr	$R$
C1	Sym2	VV	11	0.548	0.574	0.918
C2	Db45	VH	11	0.482	0.48	0.97
C3	Db45	VV	11	0.478	0.41	0.44
C4	Db45	VV	9	0.450	0.64	1.00
		VH	11	0.478	0.78	1.00
C5	Coiflet5	VH	11	0.482	0.35	0.78
C6	Db45	VH	11	0.671	0.77	0.73
C7	Coiflet5	VH	11	0.482	0.58	0.93
C8	Db45	VV	11	0.671	0.60	0.66
C9	Coiflet5	VH	9	0.478	0.36	0.93
C10	Db1	VH	9	0.640	0.90	0.89
				0.474	0.31	0.65

TABLE 5: List of classes identified on Sentinel-1 classified data using DWT method in case of the restriction  $Pr > 0.5$  with ( $R > 0.57$ ,  $FP < 47\%$ ).

Classes	MF	Polarization	$k$	FP (%)	Pr	$R$
C1	Db1	VH	9	16.22	0.83	0.77
		VV	11	15.58	0.82	0.72
	Db45	VH	9	10.91	0.87	0.70
		VH	11	37.58	0.69	0.84
		VH	11	31.26	0.73	0.83
C2	Sym2	VV	11	27.01	0.77	0.74
	Coiflet5	VH	9	46.92	0.55	0.57
	Db45	VV	11	28.33	0.78	1.00
C4	Coiflet5	VH	9	2.61	0.97	0.79
		VH	11	17.06	0.84	0.87
C5	Coiflet5	VH	11	38.48	0.64	0.69
C6	Coiflet5	VH	9	7.62	0.91	0.81
		VH	11	15.86	0.83	0.76
C7	Coiflet5	VH	9	44.45	0.62	0.72
		VH	11	31.36	0.74	0.90
C8	Coiflet5	VH	11	12.85	0.82	0.59
		VH	9	10.14	0.90	0.89
		VH	11	7.29	0.92	0.87
C9	Sym2	VV	11	16.73	0.84	0.93

TABLE 6: The OA and KPC metric values for the best classification results using: the DWT applied on the polarization VH according to MF = 'coiflet5', the DWT applied on the polarization VV according to MF = 'sym2', the classifications by the PCA and the MDist. All for  $k = 11$ .

Methods	Bands	OA	KPC
DWT	VH	0.671	0.8
	VV	0.548	0.735
PCA	VH and VV	0.371	0.336
MDist	VH and VV	0.449	0.441

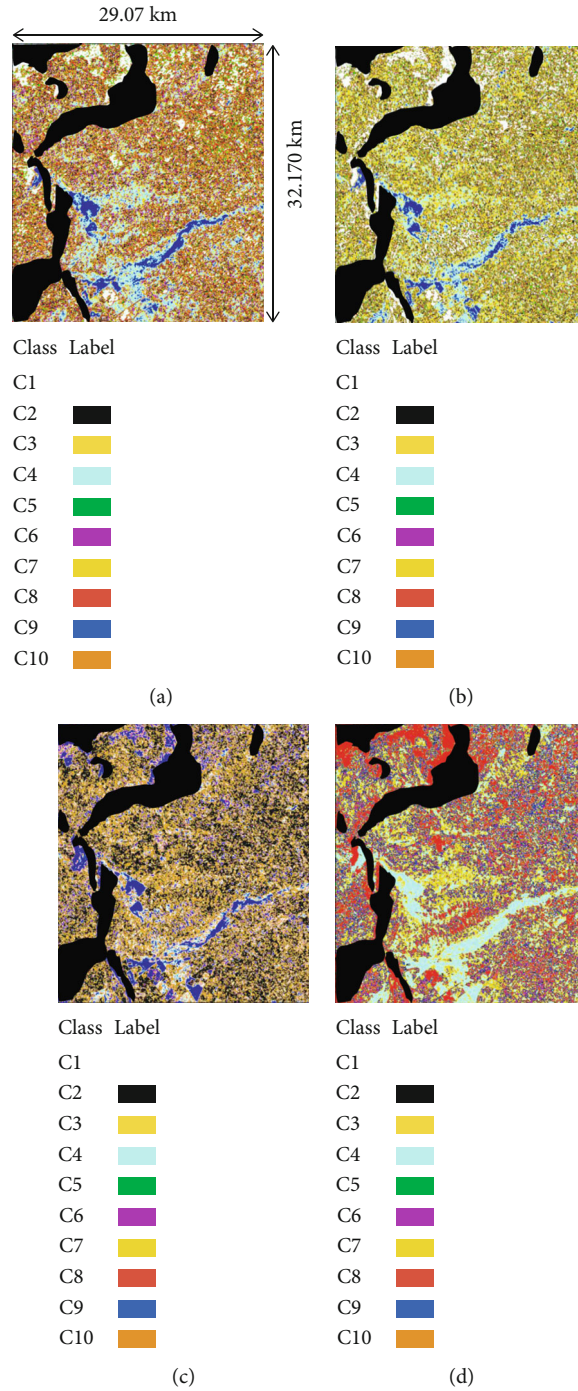


FIGURE 8: (a) VH image classified by the DWT procedure considering MF = 'coiflet5' and  $k = 11$ , (b) VV image classified by the DWT procedure considering MF = 'sym2' and  $k = 11$ , (c) Sentinel-1 image classified by PCA procedure for  $k = 11$ , (d) Sentinel-1 image classified by MDist method for  $k = 11$ .

**2.6. Evaluation Criterion.** The classification accuracy assessment is provided by the following metrics: Overall Accuracy (OA), Precision (Pr) (5), and true-positive rate of the considered class ( $R$ ) (6). Metrics are computed from the confusion matrix. The confusion matrix is a table that shows the correspondence between the classification result and a reference data (validation polygons). The coefficients of the matrix are given in terms of probability (normalized by the total

number of test examples of any class). OA is calculated as the sum of correct classifications divided by the total number of classifications.

$$\text{Pr} = \frac{\text{TP}}{(\text{TP} + \text{FP})}, \quad (5)$$



TABLE 7: PCA classification results given by the (FP, Pr, and R) metrics values, in case of the restriction (FP &lt; 50%).

$k$	OA	Classes	FP (%)	Pr	R
3	0.260		2.55	0.28	0.21
5	0.302	C4	49.44	0.34	0.25
		C3	32.24	0.42	0.23
7	0.327	C7	33.37	0.44	0.26
		C10	11.28	0.15	0.26
9	0.326	C4	43.00	0.40	0.28
		C7	38.32	0.39	0.25
		C4	27.48	0.54	0.32
11	0.372	C3	37.60	0.47	0.33
		C4	27.48	0.54	0.32
		C5	14.19	0.00	0.00
		C6	43.88	0.35	0.24
		C7	25.67	0.60	0.39
		C8	8.21	0.18	0.12

$$R = \frac{TP}{TP + FN}, \quad (6)$$

where TP and FN are the numbers of the true-positive and false-positive predictions for a considered class. FP is the sum of values in the corresponding column (excluding the TP), and FN for a class is the sum of values in the corresponding row (excluding the TP).

The Kappa coefficient (KPC) is also a measure used in remote sensing classification data assessment. KPC is computed from the nonnormalized confusion matrix (Table 3). In this case, coefficients ( $X_{ij}$ ) of the confusion matrix represent the number of pixels in  $LC_i$  which is associated to the class  $j$ . KPC is defined using equation (7), where  $p_e$  is the expected agreement ratio (8), and  $oa$  is the observed agreement (9).

$$KPC = \frac{(oa - pe)}{(1 + pe)}, \quad (7)$$

$$pe = \frac{(\sum_i (Y_i \cdot Z_i))}{T} \text{ with } Y_i = \sum_j X_{ij} Z_j = \sum_i X_{ij} T = \sum_{ij} X_{ij}, \quad (8)$$

$$oa = \sum_i X_{ii}. \quad (9)$$

### 3. Results and Discussions

The results show that by using the DWT method, just one decomposition level is enough to achieve satisfying for all LC classes' accuracy assessment. We could identify nine LC classes among the ten observed classes with an  $R$  value (6) above than 0.65 as presented in Table 4. However, the best  $R$  value achieved for the C3-class (soil covered with straws), not well identified by DWT procedure, is equal to 0.44. As seen in Table 4, these best results, for each LC classes, depend on a specific choice of the DWT MF, the polarization, and the ScW size (kxk). So, with only the use of VH polarization (resp. VV), we can well identify classes C2, C4, C5, C6, C7,

C9, and C10 (resp., C1, C4, and C8). The ScW size is equal to 9 or 11, and the OA values are beyond from 0.474 to 0.671, in these cases. The OA and  $R$  gives good accuracy assessments, but, for some classes such as C2, C6, C8, and C10 (resp., C1 and C8) for which  $R > 0.6$  and  $Pr < 0.58$  (5), they present a risk of confusion with other classes in the event of use of the polarization VH (resp., polarization VV). To avoid this risk,  $Pr$  must converge to 1 (high  $R$  value and low FP value). Table 5 illustrates the eight classes for which  $Pr > 0.5$  with ( $R > 0.57$ ,  $FP < 47\%$ ). With these restrictions, the  $R$  value for some classes (C2, C6, and C8) can be less than those illustrated in Table 4. The classes C3 and C10 are not identified in the case of this restriction (condition). The classes C1, C4, C5, C6, C7, C8, and C9 give the highest  $R$  values considering the restriction, in case of use of the polarization VH. All the classes, except C3 and C10, can be well identified by the use of the MF named 'coiflet5', the polarization VH and  $k$  equal to 11. The OA and the KPC values are equal to 0.671 and 0.8, respectively, in this case (Table 4, Table 6). Figure 8(a) (resp., Figure 8(b)) illustrates the VH (resp.VV) classified image, in case of  $k$  equal to 11 and MF 'coiflet5' (resp., MF 'sym2').

During the field campaign in June, we noticed that the wadi's soils (C4) were dry and were dotted with straws, so it can be considered as a bare soil. In addition, the farmers begin to harvest grain. In light of these findings, we merged the classes C3 and C4 in classification procedure. The results show that we can achieve an  $R$  value of these merged classes equal to 0.97 by considering MF as 'Db45', the VV polarization and  $k$  equal to 5.

The results obtained by the PCA classification procedure (Figure 5(b)) do not give satisfaction (Table 7). The OA and KPC values do not exceed 0.371 and 0.336, respectively (Table 6). Results show only the classes C2 and C9 can present a  $TP > 50\%$  but with an  $FP > 50\%$  (87.57% and 89.12%, respectively). These classes have a high-class confusion risk (Table 8). So, DWT method permits to give best results than PCA and MDist method for all classes with lower values of FP.

TABLE 8: Confusion matrices: DWT (VH 'coiflet5'; VV 'sym2'), PCA and MDist.  $k = 11$ .

	Methods	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
LC1	PCA	26.51	8.58	5.85	11.40	2.0	8.38	6.04	5.46	0	25.73
	DWT_VV	91.90	0.30	0	0	3.15	0	0	1.62	0	3.03
	DWT_VH	86.55	0	0	0	5.48	0	0	2.19	0	5.78
	MDist	33.46	14.62	3.08	0	10	16.15	0	6.15	8.46	8.08
LC2	PCA	0.00	92.73	7.27	0.00	0.00	0	0	0.00	0.00	0.00
	DWT_VV	0	30	28.18	0	2.73	18.18	3.64	0	0	17.27
	DWT_VH	0	51.43	0	0	0	22.86	25.71	0	0	0
	MDist	0	40	32.73	0	0	14.55	2.73	0	10	0
LC3	PCA	0	38.24	32.84	0	2.45	13.24	3.43	0	0	9.80
	DWT_VV	0	3.65	54.17	0	0	11.46	30.73	0	0	0
	DWT_VH	5.88	4.90	43.63	0	3.43	0.98	0	33.33	0	7.84
	MDist	0	43.18	52.27	0	0	1.14	0	0	3.41	0
LC4	PCA	29.55	1.41	0.63	31.97	0.17	0.72	3.05	0.29	12.67	19.53
	DWT_VV	0	0.60	0	72.81	0	0.17	9.68	0	16.73	0
	DWT_VH	0	0	0	86.90	0	0	8.04	0	5.07	0
	MDist	0	0	0	100	0	0	0	0	0	0
LC5	PCA	26.42	0	.0	4.69	0	0	0	38.52	29.63	0.74
	DWT_VV	0	7.79	8.44	0	37.01	10.39	3.90	22.73	0	9.74
	DWT_VH	0	0	0	0	73.44	0	0	26.56	0	0
	MDist	32.25	28.13	9.38	0	4.69	4.69	0	0	21.88	0
LC6	PCA	3.23	46.33	12.46	0.15	2.35	23.61	4.84	0.88	0.73	5.43
	DWT_VV	0	9.09	28.03	0	1.52	19.70	37.88	0	0	3.79
	DWT_VH	0	46.75	5.20	0	0	38.31	5.84	1.30	0	2.60
	MDist	0	52.5	17.5	0	0	30	0	0	0	0
LC7	PCA	1.09	15.49	10.87	0	0.82	15.22	38.59	0	0	17.93
	DWT_VV	0	0	0	35.71	0	0	64.29	0	0	0
	DWT_VH	0	6.52	2.72	15.76	0	9.24	65.76	0	0	0
	MDist	0	8.6	6.72	5.38	0	1.08	76.34	0	1.88	0
LC8	PCA	17.28	1.23	0	2.47	2.47	3.70	1.23	12.35	54.32	4.94
	DWT_VV	28.57	0	0	0	0	0	0	71.43	0	0
	DWT_VH	32.14	0	0	0	0	0	0	64.29	0	3.57
	MDist	3.7	0	0	0	7.93	0	0	91.36	0	0
LC9	PCA	6.41	0.03	0.16	7.50	0.13	0.26	0.38	1.28	80.93	2.92
	DWT_VV	0.07	0.90	0	5.1	0.10	0	0.53	0	93.08	0.17
	DWT_VH	0	0	0	1.30	0	0	4.65	0.43	93.62	0
	MDist	3.21	4.04	2.21	3.61	5.51	4.14	8.22	61.12	4.34	3.58
LC10	PCA	40.58	0.54	0.36	1.27	3.80	2.36	6.70	11.78	0.54	32.07
	DWT_VV	38.67	16	0.67	0	13.33	4.67	2	10.67	0	14
	DWT_VH	17.33	0	0	0	12.67	0	0	2.67	0	67.33
	MDist	5.33	0	0	0	32	2	0	42	0	18.67

Figure 8 illustrates the best results of image classifications in case of use of ScW size equal to 11 for: DWT method applied on the polarization VH with the use of MF = 'coiflet5', for DWT method applied on the polarization VV with the use of MF = 'sym2'; the PCA method and the MDist classification. These configurations give the highest OA and KPC values.

#### 4. Conclusion

In this article, we present the potential of the Discrete Wavelet Transform (DWT) to retrieve efficient subbands for classification of Sentinel-1 images (VH and VV). The proposed classification approach was tested on the Kairouan plain,

located in central Tunisia. The Land Cover (LC) of this semi-arid region presents a diversity of classes: fruit trees, cereals, fallow, bare soil, soil covered with straws, olive trees, vegetable crop, urban area, daim, and wadi. The proposed LC descriptors (energy, norm L1) and the Euclidian Minimum Distance criterion were used to classify the Sentinel-1 data of the study area. Classification results were assessed by the use of the ground truth data.

As shown in our study, DWT approach gives better classification results than the Principal Components Analysis approach and the minimum distance classification. High representation of data with DWT improves OA from 0.371 to 0.671 and the KPC from 0.336 to 0.8.

In the future, we plan to use features extracted from the DWT pansharpening images as inputs for machine learning to automate the classification process in order to optimize time for the LC mapping.

## Nomenclature

PCA:	Principal Components Analysis
DWT:	Discrete Wavelet Transformation
VH, VV:	Polarizations of the Sentinel-1 image
IW:	Interferometric wide
OA:	Overall Precision
LC:	Land Cover
SVM:	Support Vector Machine
MDist:	Minimum Distance classifier
KPC:	Kappa Precision Coefficient
GCP:	Ground Control Points
DEM:	Digital Elevation Model
SRTM:	Shuttle Radar Topography Mission
EMD:	Euclidian Minimum Distance
MF:	Mother Function
LH, HL, HH, LL DWT:	Energy subbands of the image
CP1 and CP2:	PCA components of the image
E:	Energy
MD:	L1 norm
ROI:	Region Of Interest
ScW:	Scanning Window
LDV:	Local Descriptor Vector
LC_DV:	LCs energy Descriptor Vectors
Pr:	Precision
R:	True-positive rate of the considered class
TP:	True-positive prediction for a considered class
FP:	False-positive prediction for a considered class

## Data Availability

The Shuttle Radar Topography Mission (SRTM) data was used to extract the Digital Elevation Model (DEM) at 30 m of the study area. The SRTM data can be free downloaded from website <http://edcscns17.cr.usgs.gov/EarthExplorer/Sentinel-1data> and from the ESA's Copernicus Open Access Hub (<https://scihub.copernicus.eu/>) The Sentinel-1 data was preprocessed

using the Sentinel Application Platform (SNAP) as follows: (i) the calibration step allows the conversion of the signal recorded by the sensor in the form of digital accounts into backscatter coefficient; (ii) the geometric terrain corrections (georeferencing) step permits to correct the geometric distortions by using the Digital Elevation Model (DEM) proposed by the NASA's Shuttle Radar Topography Mission (SRTM); (iii) the thermal noise removal step using the Lee filter.

## Conflicts of Interest

The authors declare no conflict of interest.

## Acknowledgments

The authors thank the FAO Regional Organization, Tunis, for their support during the ground-truth measurement campaigns. This study was conducted within the LR GREEN-TEAM activities and co-funded by the ERANET-MED CHAAMS and FAO-SIDA projects.

## References

- [1] E. O. Makinde and O. E. Oyelade, "Land cover mapping using Sentinel-1 SAR satellite imagery of Lagos State for 2017," *Proceedings*, vol. 2, pp. 1–5, 2018.
- [2] S. Abdikan, F. B. Sanli, M. Ustuner, and F. Calò, "Land cover mapping using Sentinel-1 SAR data," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 41, pp. 757–762, 2016.
- [3] H. Balzter, B. Cole, C. Thiel, and C. Schullius, "Mapping CORINE land cover from Sentinel1A SAR and SRTM digital elevation model data using random forests," *Remote Sensing*, vol. 7, pp. 14876–14898, 2015.
- [4] S. Bousbih, M. Zribi, C. Pelletier et al., "Soil texture estimation using radar and optical data from Sentinel-1 and Sentinel-2," *Remote Sensing*, vol. 11, no. 13, article 1113120, 2019.
- [5] Q. Gao, M. Zribi, M. J. Escorihuela, and N. Baghdadi, "Synergetic use of Sentinel-1 and Sentinel-2 data for soil moisture mapping at 100 m resolution," *Sensors*, vol. 17, pp. 1–21, 2017.
- [6] O. Charfi Marrakchi, "Elaboration de signatures des textures d'images de télédétection," *HAL Archive ouverte en Sciences de l'Homme et de la Société*, vol. 11, no. 2, 2020 <https://halshs.archives-ouvertes.fr/halshs-01136470>.
- [7] M. C. Alons and J. A. Malpica, "Satellite imagery classification with LIDAR data," *Trees*, vol. 2106, 2010.
- [8] S. Wang and A. Wang, "Segmentation of high-resolution satellite imagery based on feature combination," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, vol. 37, pp. 1223–1228., 2008.
- [9] M. Singh and G. Kaur, "SAR image classification using PCA and texture analysis," in *International Conference on Advances in Information Technology and Mobile Communication*, pp. 435–439, Springer, 2011.
- [10] M. Kumari, C. S. Murthy, V. Pandey, and G. Bairagi, "Soybean cropland mapping using multi-temporal sentinel-1 data," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. XLII-3/W6, 2019.
- [11] A. Schmitt and A. Wendleder, "SAR-sharpening in the Ken-nough framework applied to the fusion of multi-modal SAR

- and optical images,” *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, 2018.
- [12] S. Zecchetto, “Wind direction from Sentinel-1 SAR images in regional seas,” in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 998–1000, Valencia, Spain, 2018.
- [13] M. C. Olfa and M. Jérôme, “Confrontation of filter banks conception procedures for attributes extraction of the airborne SAR image textures,” *International Journal of Engineering Research & Technology*, vol. 8, no. 10, pp. 9–16, 2018.
- [14] O. C. Marrakchi, “Improved remote-sensing land cover identification using parameterized feedforward neural networks,” *Euro-Mediterranean Journal for Environmental Integration*, vol. 6, 2021.
- [15] S. Bousbih, M. Zribi, Z. Lili-Chabaane et al., “Potential of Sentinel-1 radar data for the assessment of soil and cereal cover parameters,” *Sensors*, vol. 17, no. 11, p. 2617, 2017.
- [16] M. Zribi, A. Chahbi, M. Shabou et al., “Soil surface moisture estimation over a semi-arid region using ENVISAT ASAR radar data for soil evaporation evaluation,” *Hydrology and Earth System Sciences*, vol. 15, no. 1, pp. 345–358, 2011.
- [17] M. Zribi, A. Gorrab, N. Baghdadi, Z. L. Chabaane, and B. Mougenot, “Influence of radar frequency on the relationship between bare surface soil moisture vertical profile and radar backscatter,” *HAL Archive ouverte en Sciences de l’Homme et de la Société* May 2020, <https://hal.archives-ouvertes.fr/hal-00907867>.
- [18] A. Chahbi, M. Zribi, Z. Lili-Chabaane et al., “Estimation of the dynamics and yields of cereals in a semi-arid area using remote sensing and the SAFY growth model,” *International Journal of Remote Sensing*, vol. 35, no. 3, pp. 1004–1028, 2014.
- [19] O. M. Charfi, J. Mbainabeye, and W. Jebalia, “Wavelet-based remote sensing heterogeneous textures signatures using generalized Gaussian density model,” *International Review on Computers and Software*, vol. 7, no. 2, pp. 538–545, 2012.
- [20] S. Mallat, “Wavelet for a vision,” *Proceedings of the IEEE*, vol. 84, no. 4, pp. 604–614, 1996.
- [21] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [22] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, “Statistical texture characterization from discrete wavelet representations,” *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 592–598, 1999.