Available online at www.sciencedirect.com

Journal of Applied Research and Technology

Journal of Applied Research and Technology 15 (2017) 250–258

www.jart.ccadet.unam.mx

Original

# A comparative study of the use of local directional pattern for texture-based informal settlement classification

Abuobayda M. Shabat, Jules-Raymond Tapamo *

*University of KwaZulu-Natal Durban, South Africa*

Received 7 August 2015; accepted 9 December 2016
Available online 25 May 2017

**Abstract**

In developing and emerging countries progression of informal settlements has been a fast growing phenomenon since the mid-1990s. Half of the world's population is housed in urban settlements. For instance, the growth of informal settlements in South Africa has amplified after the end of apartheid. In order to transform informal settlements to improve the living conditions in these areas, a lot of spatial information is required. There are many traditional methods used to collect these data, such as statistical analysis and fieldwork; but these methods are limited to capture urban processes, particularly informal settlements are very dynamic in nature with respect to time and space. Remote sensing has been proven to provide more efficient techniques to study and monitor spatial patterns of settlements structures with high spatial resolution. Recently, a new feature method, local directional pattern (LDP), based on kirsch masks, has been proposed and widely used in biometrics feature extraction. In this study, we investigate the use of LDP for the classification of informal settlements. Performance of LDP in characterizing informal settlements is then evaluated and compared to the popular gray level co-occurrence matrix (GLCM) using four classifiers (Naive-Bayes, Multilayer perceptron, Support Vector Machines, *k*-nearest Neighbor). The experimental results show that LDP outperforms GLCM in classifying informal settlements.


*Keywords:* Informal settlements; Texture features; Gray level co-occurrence matrix; Local directional pattern; Classification

## 1. Introduction

In developing countries, informal settlements have become a phenomenon which grow very fast specially in the 21st Century. Half of the world's population is housed in urban settlements. The reason for this phenomenon is the immigration of people from the rural areas to the cities. Many previous studies aimed at extracting houses outline to quantify shape-based features of informal settlements. Object-based image analysis (OBIA) method estimates the size, spacing and shape of the houses by extracting the houses footprint (Blaschke & Lang, 2006). OBIA partitions remote sensing (RS) imagery into meaningful image-objects and assesses their characteristics through spatial, spectral and temporal scale (Hay & Castilla, 2008). Previous studies on geospatial methods have been used to estimate populations and to distinguish the human settlements. For example, a study by Aminipouri, Sliuzas, and Kuffer (2009) estimates the population by creating an accurate inventory of buildings. The computer vision community is facing a very complex and challenging task extracting the spatial data from informal settlements. Constructions in informal settlements are built using various materials and are very close to each other and have no suitable organization. It makes the classification of informal settlements images an uphill task (McLaren, Coleman, & Mayunga, 2005). A number of researchers have tried to develop tools and techniques to characterize the informal settlements areas from remotely sensed data. Mayunga, Coleman, and Zhang (2007) present a new semi-automatic approach to extract buildings from informal settlements images obtained using Quick Bird. Snackes and radial casting algorithm were used to map the informal settlements images. The main limitation in this study is the difficulty of characterizing small houses. Khumalo, Tapamo, and Van Den Bergh (2011) applied two feature methods, Gabor filters and GLCM to distinguish different textural regions in Soweto area (Johannesburg, South Africa). They found Gabor filters more

* Corresponding author.
*E-mail address:* tapamoj@ukzn.ac.za (J.-R. Tapamo).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

http://dx.doi.org/10.1016/j.jart.2016.12.009

accurate than GLCM in classifying informal settlements. A study carried out by

Ella, Van Den Bergh, Van Wyk, and Van Wyk (2008) compared gray level co-occurrence (GLCM) and local binary pattern (LBP) in their ability to classify urban settlement. It is shown that both methods performed very well with a superior performance for LBP. Van Den Bergh (2011) investigates the powers of two features methods, GLCM and LBP, to classify Soweto (Johannesburg, South Africa) areas. It is established that the performance of the gray level co-occurrence matrix is superior to the local binary pattern on a combined spatial and temporal generalization problem, but the LBP features perform better on spatial-only generalization problems. In Owen and Wong (2013) an analysis is conducted on the shape, texture, terrain geomorphology and road networks to characterize the informal settlements and formal neighborhoods in Latin America. The results achieved were promising when finite data were used to recognize informal settlements. Asmat and Zamzami (2012) introduced an automated house detection technique to extract legal and illegal settlements in Pulau Gaya, Saba. The result shows that the edge to edge features can separate between houses that are less than 2 m away from each other. In Graesser et al. (2012), an investigation of nine statistics methods (GLCM PanTex, Histogram of Oriented Gradients, Lacunarity, Line Support Regions, Linear Feature Distribution, Psuedo NDVI, Red-blue NDVI, Scale Invariant Feature Transform, and TEXTONS) is presented with different direction, structure size and shape and tested in four different cities. The GLCM PanTex, LSR, HoG and TEXTON features were found to be the best in characterizing the informal settlements and formal areas. A new feature method, local directional pattern (LDP), based on the known Kirsch kernels was recently proposed by Jabid, Kabir, and Chae (2010a). LDP has mainly been applied in biometrics: face recognition (Jabid, Kabir, & Chae, 2010b), signature verification (Ferrer, Vargas, Travieso, & Alonso, 2010) and facial expression recognition (Jabid, Kabir, & Chae, 2010c). In Shabat and Tapamo (2014) the powers of GLCM and LDP to characterize texture images are compared; the result shows that LDP outperforms GLCM. In this paper, GLCM and LDP are investigated using different numbers of significant bits; the final goal is to identify the most effective amongst them. The computation of the local directional pattern is based on the number of significant bits, and in this work four alternative values are considered: 2, 3, 4, 5 instead of 3 as in the classic LDP.

## 2. Materials and methods

In the following sections the different feature methods used the in the paper are presented

### 2.1. Gray level co-occurrence matrix (GLCM)

In the early 1970s Haralick, Shanmugam, and Dinstein (1973) proposed the extraction of fourteen features, from the GLCM of a gray level, to characterize the image texture. The computation of GLCM depends on two parameters: the orientation $\theta$ formed by the line-segment connecting the two considered pixels, and the distance ($d$)[number of pixels] between them. The direction $\theta$ is usually quantized in 4 directions (horizntal – 0°, diagonal - 45°, vertical – 90°, anti-diagonal – 135°).

To compute the gray-level co-occurrence matrix of a window in an image, the following parameters are considered:

- The window size, $N_x \times N_y$, where $N_x$ is the number of rows and $N_y$ the number of columns.
- Distance ($d$) and directions $\theta$.
- And the range of gray values to consider in calculations $0, \ldots, G-1$.

We adopt the formulation used in Bastos, Liatsis, and Conci, 2008 and Eleyan and Demirel (2011) to present the calculation of GLCM. The GLCM is defined as the probability of occurrence of two gray levels at a given offset (with respect to given distance and orientation). Given the image $I$, of size $N_x \times N_y$, the value of the co-occurrence for the gray values $i$ and $j$, at the distance ($d$) and direction $\theta$, $P_{d,\theta}(i, j)$ can be defined as

$$P_{d,\theta}(i, j) = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} \delta_{d,\theta,i,j}(x, y) \tag{1}$$

Where

$$\delta_{d,\theta,i,j}(x, y) = \begin{cases} 1 & \text{if } I(x, y) = i \text{ and } I(x + \pi_x(d, \theta), y + \pi_y(d, \theta)) = j \\ 0 & \text{otherwise} \end{cases}$$

The offset $(\pi_x(d, \theta), \pi_y(d, \theta))$ is used to compute the position of $(x, y)$ with respect to its neighbor at the distance ($\underline{d}$) and direction $\theta$. For the 4 directions (0°, 45°, 90°, and 135°) and the offsets are given in Table 1.

#### 2.1.1. Haralick's features

Given an image $I$ with $G$ gray levels, an angle $\theta$ and a distance ($d$), after the gray level co-occurrence matrix, $(P_{d,\theta}(i, j))_{0 \le i,j \le G-1}$, number of features can be extracted, amongst which the most popular are the 14 Haralick features (energy or angular second moment (ENR), contrast (CON), correlation (COV), variance (VAR), inverse different moment (IDM), sum average (SAV), sum variance (SVA), sum entropy (SEN), entropy (ENT), difference variance (DIV), difference entropy (DEN), information measures of correlation (IMC1, IMC2), maximum correlation coefficient (MCC)). The computation of Haralick features is done using a normalized GLCM. The $(i, j)th$ normalized entry,

Table 1
Definition of different offsets.

| $\theta$ | 0° | 45° | 90° | 135° |
|---|---|---|---|---|
| $\pi_x(d, \theta)$ | 0 | $-d$ | $-d$ | $-d$ |
| $\pi_y(d, \theta)$ | $d$ | $d$ | 0 | $-d$ |

$$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad M_0 \text{ (East)} \qquad \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad M_1 \text{ (North East)}$$

$$\begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \quad M_2 \text{ (North)} \qquad \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad M_3 \text{ (North West)}$$

$$\begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad M_4 \text{ (West)} \qquad \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad M_5 \text{ (South West)}$$

$$\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \quad M_6 \text{ (South)} \qquad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \quad M_6 \text{ (South Est)}$$

Fig. 1. The 8 Krisch masks.

$P_{d,\theta}(i,j)$, of $P_{d,\theta}(i,j)$ is defined as

$$P_{d,\theta}(i,\ j) = \frac{P_{d,\theta}(i,\ j)}{||P_{d,\theta}||} \tag{2}$$

where $||P_{d,\theta}|| = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P_{d,\theta}(i,\ j)$. Details on the calculation of all these features can be found in Haralick et al. (1973). For each texture, $T$, a chosen distance ($d$) and a direction $\theta$, 14 Haralick features can be extracted.

$$\left(glcm^{i}_{T,d,\theta}\right)_{i=1,\ldots,14} \tag{3}$$

### 2.2. Local directional pattern (LDP)

The local binary pattern (LBP) operator depends on the change of the intensity around the pixel to encode the micro-level information of spot, edges and other local features in the image (Jabid et al., 2010a). The gradient is known to be more stable than the gray level; that is why some researches have replaced the intensity value at a pixel position with its gradient magnitude and calculated the LBP (Ferrer et al., 2010). The local directional pattern (LDP) was proposed by Jabid et al. (2010b) to resolve the problem with LBP, mentioned earlier. Since the LBP depends on the neighboring pixels' intensity which makes it unstable. Instead, LDP considers the edge response value in different direction. LDP features are based on eight bit binary codes assigned to each pixel of an input image. It is composed of three steps (Kabir, Jabid, & Chae, 2010):

- **Calculation of eight directional responses** of particular pixels using the Kirsch compass edge detector in eight orientations ($M_7, \ldots, M_0$) centered on its own position as shown in Figure 1. Given a pixel $(x, y)$ of an image, $I$, for each direction $i$, and using the corresponding mask $M_i$ the *ith* directional response $m_i$ can be computed as

$$m_i = \sum_{k=-1}^{1}\sum_{l=-1}^{1} M_i(k+1, l+1) \times I(x+k,\ y+l) \tag{4}$$

  For the 8 directions a vector $(m_7, \ldots, m_0)$ is obtained. Figure 2 shows the Kirsch directional responses of a pixel $(x, y)$.
- **LDP code generation of the directional** responses obtained in the previous step. It is based on the selection of $k$ most significant responses and set the corresponding bit to 1 leaving other $(8-k)$ bits to 0. Finally, the LDP code, $LDP_{x,y}(m_0, \ldots, m_7)$, of the pixel $(x, y)$ with directional response $(m_0, \ldots, m_7)$, is derived using Eq. (5).

$$LDP_{x,y}(m_0, \ldots, m_7) = \sum_{i=0}^{7} s(m_i - m_k) \times 2^i \tag{5}$$

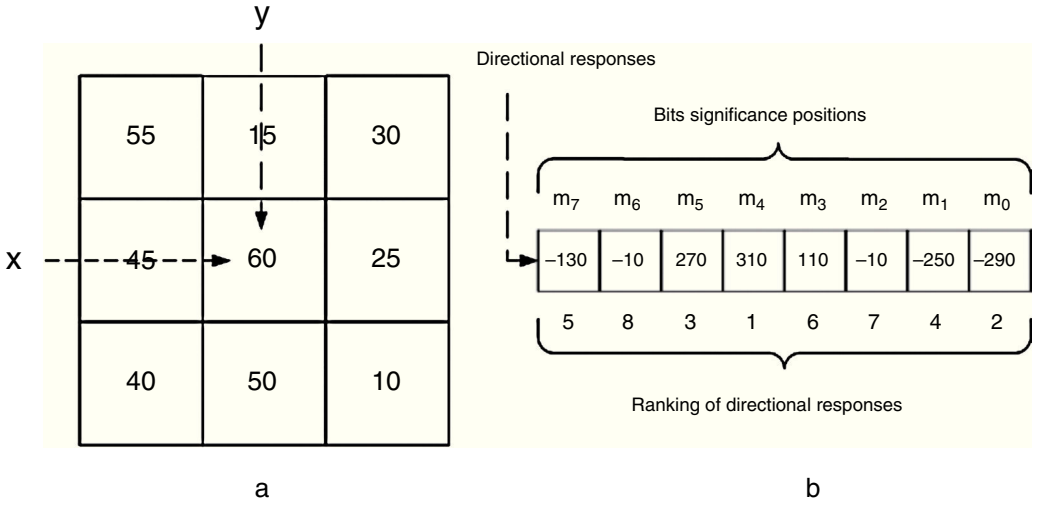


Fig. 2. Krisch directional response. (a) This figure shows a pixel $(x,y)$ that has a gray level 60. (b) Directional responses, together with the ranking of those responses, and the associated bit significance, with $m_0$ being at the less significant position and $m_7$ at most significant position. Note that the ranking of responses is done on absolute values.

where, $m_k$ is the *kth* most significant response and $s(x)$ is defined as

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{6}$$

Given the directional responses generated by the Kirsch convolution on pixel $(x, y)$ presented in Figure 2, the LDP code for $k = 3$ is computed as follows:

- $m_5 = 270$ is the 3*rd* most significant directional response.
- The code of the LDP code of the pixel $(x, y)$ is then

$$\begin{aligned} LDP_{x,y}(m_0, \ldots, m_7) &= \sum_{i=0}^{7} s(m_i - m_5) \times 2^i \\ &= 0 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 \\ &\quad + 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 \\ &\quad + 0 \times 2^1 + 1 \times 2^0 \end{aligned}$$

- The LDP code, $LDP_{x,y}$, of the pixel $(x, y)$ is then 49.

- **Construction of LDP descriptor** which is carried out after the calculation of the LDP code for each pixel $(x, y)$. The input image $I$ of size $M \times N$ is then represented by a LDP histogram using Eq. (7), that is also called LDP descriptor. In this case $k = 3$, is used; It means, ${}^8C_3 = 56$ distinct values are generated and used to encode the image. The histogram $H$ obtained from the transformation has 56 bins and can be defined as

$$H_i = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} p(LDP_{(x,y)}, C_i) \tag{7}$$

where $C_i$ is the *ith* LDP pattern value, $i = 1, \ldots, {}^8C_3$ and the definition of $p$ is given in Eq. (9).

$$p(x, a) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases} \tag{8}$$

Given a texture, $T$, and the number of significant bits $k$, a feature vector $LDP_{k,T}$ can be extracted and represented as

$$LDP_{k,T} = (H_1, \ldots, H_{56}) \tag{9}$$

### 2.3. Classifiers

Four classifiers are used to evaluate the power of LDP to characterize textures and compared Haralick features extracted from GLCM.

#### 2.3.1. Support vector machines

SVM is a learning technique for pattern classification and regression (Cortes & Vapnik, 1995; Vapnik, 2013). It was originally designed as two-class classifier, but many versions have been proposed to perform multi-class classification (Crammer & Singer, 2002; Hsu & Lin, 2002). The principle is, given a labeled set of $M$ training samples $(x_i, y_i)$, where $x_i \in R$ and $y_i$ is the associated label ($y_i \in \{-1, 1\}$), $i = 1, \ldots, M$. A SVM classifier finds the optimal hyperplane that correctly separates the largest fraction of data points while maximizing the distance of either class from the hyperplane. The discriminant hyperplane is defined by the level set function

$$f(x) = \sum_{i=1}^{M} y_i \alpha_i k(x, x_i) + b \tag{10}$$

where $k(\cdot, \cdot)$ a kernel function and the sign of is $f(x)$ indicates the membership of $x$. Constructing an optimal hyperplane is equivalent to finding all nonzero $\alpha_i$. Kernel function $K(x_i, x_j)$ is the inner product of the features space $K(x_i, x_j) \leq \emptyset(x_i), \emptyset(x_j)$. The three following kernel functions are often used:

- Polynomial kernel

$$K(x_1, x_2) = (\langle x_1, x_2 \rangle + d)^d \tag{11}$$

where $c$ is a positive constant and $d$ is the dimension of feature space in question.

- Linear kernel

$$K(x_1, x_2) = \langle x_1, x_2 \rangle \tag{12}$$

- Radial basic function kernel (RBF)

$$K(x_1, x_2) = \exp\left(-\frac{||x_1 - x_2||^2}{2\sigma^2}\right) \tag{13}$$

where $\|x_1 - x_2\|$ is the distance between the vectors $x_1$ and $x_2$, $\sigma \in R$ is the bandwidth of a gaussian curve.

In our case, the $M$ feature vectors $(x_i = 1, \ldots, M)$ are extracted, from texture images that need to be classified, using the local directional pattern, or GLCM.

#### 2.3.2. Naïve Bayes classifier

One of the most popular and most simplest classification models is the naive Bayes classifier (Friedman, Geiger, & Goldszmidt, 1997). The principle of Naive bayes is: given the training data $T$ which contain a set of samples, each sample $X = (x_1, \ldots, x_n)$ and there are $k$ classes $C_1, \ldots, C_k$. Each sample is labeled by one of these classes. Naive Bayes predicts a given sample $X$ belongs to the class that has the highest posterior probability conditioned on $X$. Therefore, sample $X$ is predicted to belong to class $C_i$ if and only if $P(C_i|X) > P(C_j|X)$, for all $j$ such that $0 < j < (m - 1)$ and $j \neq i$. By Bayes' theorem

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \tag{14}$$

If the data set has many attributes, it would be expensive to compute $P(X|C_i)$. To solve this problem, naive assumption assumes that the value of the attributes are conditionally independent of one another. This means that

$$P(X|C_i) \approx \prod_{k=1}^{n} P(x_k|C_i) \tag{15}$$

Fig. 3. Samples FT1 and FT2 are extracts of the formal township, where buildings are placed in a planned manner. This type contains stable structure. The different between T_1 and T_2 is on the houses.

### 2.3.3. k-Nearest neighbor (k-NN)

*k*-NN is a non-parametric classification method introduced in the early 1970s by Fix and Hodges (1951). The process is done by computing the similarity between the sample and the different classes. Let $C_1, \ldots, C_k$ be the classes of our samples. Given a new sample $X=(x_1, \ldots, x_n)$, to find to which class it belongs, the distance $d(x, C_j)$ between $x$ and $C_j$, for $j=1, \ldots, k$, is calculated. Sample $X$ is assigned to class $C_{i_0}$ to which it is closest. Index $i_0$ is calculated as presented in the following equation:

$$i_0 = arg\min_{i=1,\ldots,k} d(x, C_i) \tag{16}$$

### 2.3.4. Multilayer perceptron (MLP)

MLP is well known and widely used in different detection and estimation applications (Burrascano, Fiori, & Mongiardo, 1999; Gati, Wong, Alquie, & Fouad Hanna, 2000; Kasabov, 1996). The principle of MLP is that the input layer in MLP is considered as layer 0. Assume that the total number of the hidden layers is $L$. In the hidden layer $l$ the number of node is $N_l$, $l=1, \ldots, L$. Let $w_{ij}$ be the weight of the connection between the $jth$ nodes of $(l-1)th$ hidden layer and $ith$ nodes of the $l^{th}$ hidden layer, and let $x_i$ be the $ith$ input factor to the MLP. Let $y_i^l$ represent the output of the $ith$ node of the $l^{th}$ hidden layer, which can be calculated by the following equation:

$$y_i^l = f\left(\sum_{i=1}^{N_l-1} w_{ij}^l . y_i^{l-1} + \theta_i^l\right), \qquad i = 1, \cdots, N_l \tag{17}$$

where $\theta_i^l$ represent the bias factor of the $ith$ node of the $lth$ hidden layer, $y_i^0 = x_i$, $i = 1, \ldots, N_0$ and $f(.)$ is the active function. Let $\upsilon_{ki}$ be the weight of the connection between the $kth$ node of the output layer and the $ith$ node of the $Lth$ hidden layer. The MLP output can calculated as

$$y_k = \sum_{i=1}^{N_l} \upsilon_{ki} . y_i^l + \beta_k, \quad k = 1, \ldots, N_y \tag{18}$$

where the $\beta_k$ is the bias factor of the output layer. The MLP algorithm compares between the network output with the desire output which measures the error in the network. To correct the output layer, this algorithm updates the weight until the output of network gets closer to the desire output.

## 3. Experimental results and discussion

### 3.1. Data set

Settlements image categories, shown in Figures 3–8, were identified in Soweto (Gauteng province, South Africa) to work

Fig. 4. Informal squatters (IS type1): the structure of the informal squatters is not stable. The dwellings of this category are shack type (made out of cardboard, wood, tin, etc.). Typically characterized by high building densities.

Fig. 5. Formal township + informal squatter (FTIS Type1, FTIS Type2, and FTIS Type3): In this category we can find any type of any density, of residential unit, but buildings appear in pairs a larger building will be accompanied by a backyard shack.

Fig. 6. Informal Township: the structure of the informal township is recognized as constant or semi-constant structure. The dwelling of this category is shack type and located on serviced and un-serviced sites. The dwelling densities vary from low to high.

Fig. 7. Formal: the structure of the formal residential is a constant structure, located near well-established buildings areas.

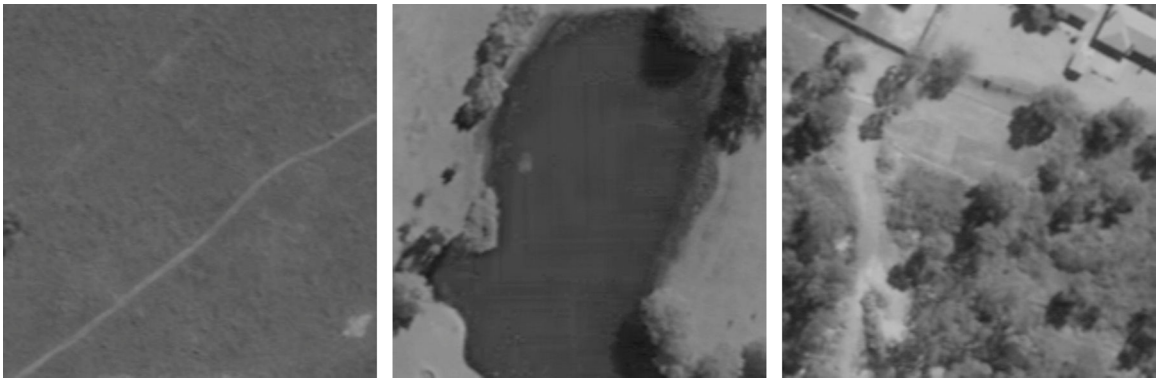

Fig. 8. Non-urban: shows samples of this type. All of them were gray level images with 8 bit per pixel.

Table 2
Number of the sample in each category.

| Samples type | Number |
|---|---|
| Formal | 413 |
| FT1 | 367 |
| FT2 | 637 |
| FTIS1 | 62 |
| FTIS2 | 611 |
| FTIS3 | 100 |
| Informal | 358 |
| IS1 | 312 |
| Non-urban | 479 |

with as target classes. Data used are gray-level images with 8 bit per pixel, and a size of (200 × 200) pixels. Table 2 shows the number of samples in each category.

- Formal township (FT1, FT2): The structure of the formal township building is placed in a planned manner. The difference between FT1 and FT2 relates to the size of the houses.
- Informal squatters (IS): The structure of the informal squatters is not stable. The dwellings of this category are shack type (made out of cardboard, wood, tin, etc.). This is typically characterized by high building density.
- Formal township + informal squatter (FTIS1, FTIS2, and FTIS3): In this category we can establish any type of any density, of residential unit, but buildings appear in pairs and a larger building is accompanied by a backyard shack.
- Informal township: The structure of the informal township is recognized as a constant or semi-constant structure. The dwelling of this category is shack type and located on serviced and un-serviced sites. The dwelling density varies from low to high.
- Formal: The structure of the formal residential is a constant structure, located near well establish buildings area.
- Non-urban area: The areas outside a town or a city. Open swath of land that has few homes or other buildings.

### 3.1.1. Choice of parameters

The computation of GLCM is done using the following parameters:

- Number of gray levels: 256
- Directions: 0°, 45°, 90°, 135°
- Distance: 1

The computation of the local directional pattern is based on the number of significant bits, and in this work four alternative values are considered: 2, 3, 4, 5.

### 3.2. Result and discussion

Performance of the various classifiers using various size of training samples, with various values for $k$ are considered. $k$-NN and MLP achieve reasonably good results, even when only a small proportion of the data is used for training. The superiority of the $k$-NN and MLP over both SVM and NB is easily perceptible. Moreover, it can be noticed that when the number of significant bits ($k$) changes from 2 to 4 the accuracy improves and declines when $k = 5$. Classifiers have the best performance for $k$ equal 4 and the worst performance is registered for $k$ equal to 2. The performance of LDP is evaluated for different values of $k$: The average accuracy of LDP with different $k$ values shows that for LDP when $k$ equal 4 is the highest performance (85.6%) compared to the rest. For LDP, the value 2 for $k$ is the lowest performance (77.7%). The best achievement was obtained by $k$-NN and MLP 98.9%, 98.19% in LDP (4) and LDP (3) when the data proportion is 80% as training 20% testing. NB performances are remarkably low ranging between 55% and 70%. Another observation is the fact that the accuracy of SVM is not very high either. With the gray-level co-occurrence matrix, different combinations of the 14 original Haralick features (entropy, (entropy + IDM), (energy + contrast + correlation), (entropy + energy + contrast), (energy + correlation + entropy + IDM), (entropy + energy + contrast + correlation + IDM), and all 14 features are used to characterize textures. The presentation of different feature combinations to characterize informal settlements reveals that the best performance is achieved by $k$-NN with value (86.04%) using the 14 original Haralick features in 45° direction. With NB, the performance is remarkably low ranging between 29% and 53%. Another observation is that both SVM and MLP are not very effective classifying informal settlements, with accuracies ranging from 30% to77%. Using $k$-NN, the best achievement was achieved with 45° direction at 91%. With NB, the performances are remarkably low ranging between 50% and 55% The accuracy of both SVM and MLP are not very high range between (48%–60%) and (64%–80%). The best performance was achieved in 45° direction with value (80%) using MLP and (60.81%) using SVM. The training set percentage: the average accuracy of the best direction changes from 62% when the training set is 10%–71.17% when the training set is 80%. With the knowledge that the best performance is achieved with GLCM and this when the 14 original Haralick features at 45° direction are chosen, it can be compared to the performance of LDP when $k$ equals 4. The average performances of each classifier in both feature methods are such that, with NB and SVM, the performance is remarkably low in both LDP (4) and GLCM (45°), the best achievement is (83.06%) by SVM using LDP (4). With MLP, LDP (4) achieved 20% more with value (92.77%) compared to GLCM (45°) with value (77.48%). With $k$-NN, The best result was achieved by LDP (4) with value (92.7%), compared to GLCM (45°) with value (86.03%). For GLCM (45), we find the performance from 20% to 80% increase by approximately 1% in each step. The best achievement obtained when the training set is 80% with a value (71.17%). The worst achievement obtained when the training set is 10% with a value (62.6%). For LDP (4), we find the performance almost constant from 50% to 80% with value (86.81%). The best achievement obtained when the training set is 80% with a value (86.81%). Figures 9–10 below summarize the comparison between these two methods (LDP, GLCM). In conclusion, the best combination used in our approach is using the local directional pattern with $k$ equals four applying on

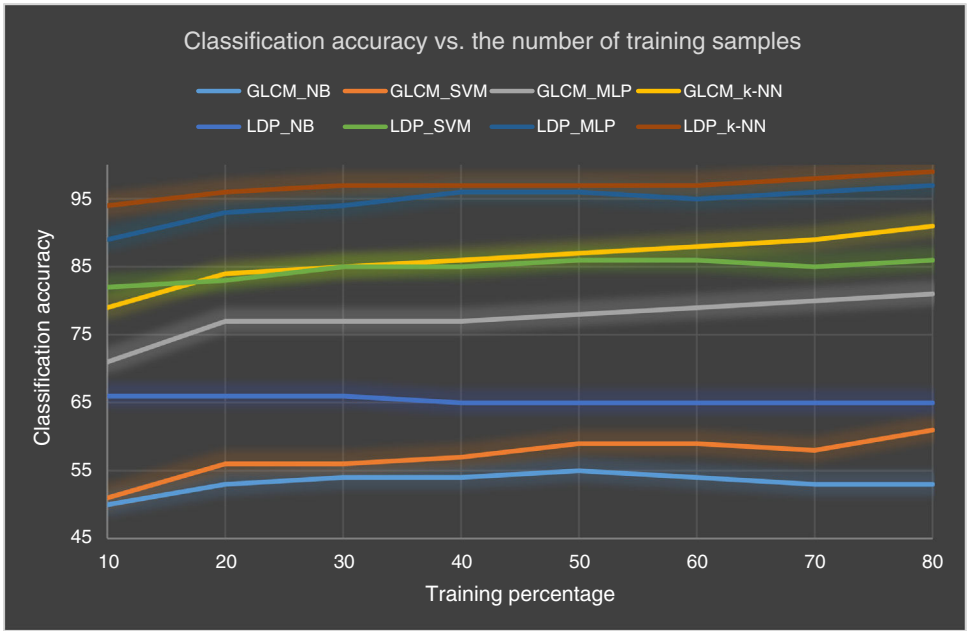


Fig. 9. Classification accuracy vs. the number of training samples using both features method LDP when $k = 4$ and GLCM with distance 1 and 45° direction.

nine categories of image, and the $k$-nearest neighbor being the preferred classifier.

## 4. Running Times for Features Extraction

For an image of size $n \times m$, GLCM running time will be $G(m, n) = O(mn)$. However, the running time for LDP is $L(m, n) = R(m, n) + H(m, n)$ where $R(m, n) = O(mn)$ is the running time to compute the responses and $H(m, n) = O(mn)$ is the running time of the computation of the histogram, then $L(m, n) = O(mn)$. Table 3 shows running times for both GLCM and LDP applied to nine different categories using a computer with a processor Intel Core i5, a CPU of 2.3 GHz and 4G of RAM. With GLCM 14 features were computed and for $k$ equals 4 for LDP; the running time of the GLCM algorithm is remarkably low (less than 13 ms to process an image). Compared to LDP, which takes a very long time to compute, ranging from 4.7 s in average to process an image in FTIS1 to 9 s in FT2. It is worth mentioning that the size of each image is $(200 \times 200)$.

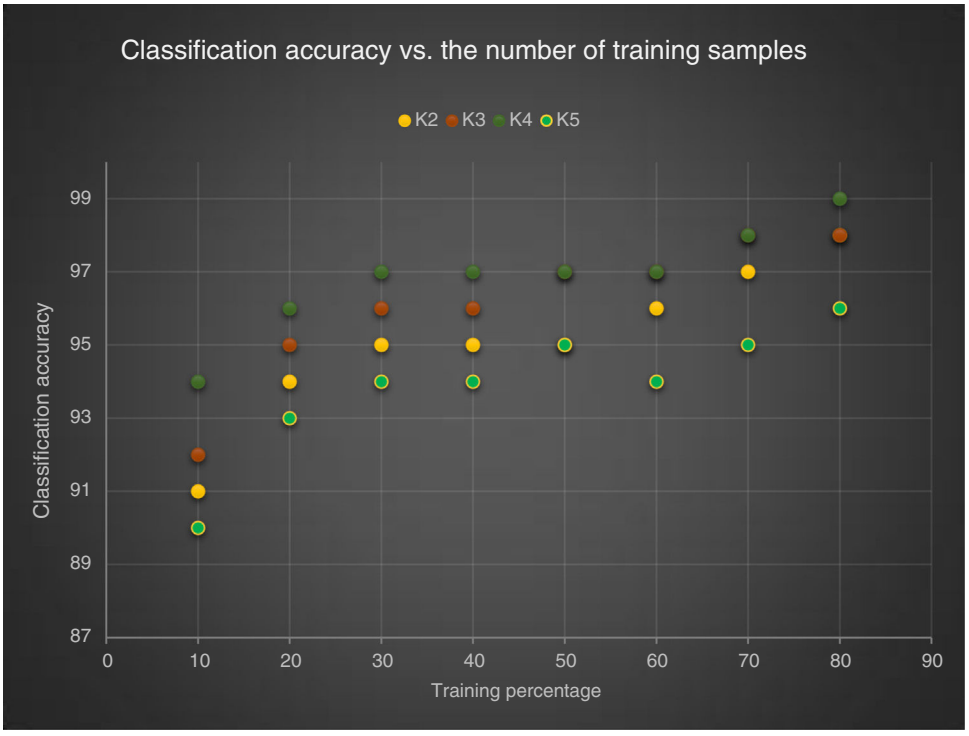


Fig. 10. Classification accuracy vs. the number of training samples using LDP with different number of significant bits ($k$).

Table 3
Average running time for both GLCM and LDP applied to nine category images.

| Image category | GLCM runtime (s) | LDP runtime (s) |
|---|---|---|
| Formal | 0.014 | 7.22 |
| FT1 | 0.013 | 8 |
| FT2 | 0.014 | 9 |
| FTIS1 | 0.016 | 4.7 |
| FTIS2 | 0.013 | 7.4 |
| FTIS3 | 0.01 | 7.93 |
| Informal | 0.013 | 8 |
| IS | 0.012 | 8.7 |
| Non-urban | 0.012 | 6.3 |

## 5. Conclusion

Two feature methods, local directional pattern (LDP) and gray-level co-occurrence matrix (GLCM), have been compared using four different classifiers (Naive-Bayes, multilayer perceptron, support vector machines, and *k*-nearest neighbor). This work has investigated the impact of the number of significant bits considered to code the Kirsch masks application responses. Experiments have shown that the choice of 4 significant bits achieves the best accuracy for texture characterization using LDP. It has also been established that the best texture feature is the local directional pattern when *k*-NN is used as a classifier. On the other hand, it has been demonstrated that the local directional pattern is superior in characterizing informal settlement images. However, the running time for LDP is two orders of magnitude higher than that of the GLCM.

## Conflict of interest

The authors have no conflicts of interest to declare.

## References


Aminipouri, M., Sliuzas, R., & Kuffer, M. (2009). Object-oriented analysis of very high resolution orthophotos for estimating the population of slum areas, case of Dar-Es-Salaam, Tanzania. In *Paper presented at the Proc. ISPRS XXXVIII Conf.*

Asmat, A., & Zamzami, S. Z. (2012). Automated house detection and delineation using optical remote sensing technology for informal human settlement. *Procedia-Social and Behavioral Sciences*, *36*, 650–658.

Bastos, L. D. O., Liatsis, P., & Conci, A. (2008). Automatic texture segmentation based on *k*-means clustering and efficient calculation of co-occurrence features. In *15th International Conference on Systems, Signals and Image Processing, 2008. IWSSIP 2008* (pp. 141–144). IEEE.

Blaschke, T., & Lang, S. (2006). Object based image analysis for automated information extraction – A synthesis. In *Paper presented at the Measuring the Earth II ASPRS Fall Conference*.

Burrascano, P., Fiori, S., & Mongiardo, M. (1999). A review of artificial neural networks applications in microwave computer-aided design (invited article). *International Journal of RF and Microwave Computer-Aided Engineering*, *9*(3), 158–174.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

Crammer, K., & Singer, Y. (2002). On the algorithmic implementation of multiclass kernel-based vector machines. *The Journal of Machine Learning Research*, *2*, 265–292.

Eleyan, A., & Demirel, H. (2011). Co-occurrence matrix and its statistical features as a new approach for face recognition. *Turkish Journal of Electrical Engineering & Computer Sciences*, *19*(1), 97–107.

Ella, L. A., Van Den Bergh, F., Van Wyk, B. J., & Van Wyk, M. A. (2008). A comparison of texture feature algorithms for urban settlement classification. In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International, Vol. 3, no. 1*. IEEE, pp. III-1308.

Ferrer, M. A., Vargas, F., Travieso, C. M., & Alonso, J. B. (2010). Signature verification using local directional pattern (LDP). In *IEEE International Carnahan Conference on Security Technology (ICCST), 2010* (pp. 336–340). IEEE.

Fix, E., & Hodges, J. L., Jr. (1951). *Discriminatory analysis-nonparametric discrimination: Consistency properties*. DTIC Document.

Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian network classifiers. *Machine Learning*, *29*(2–3), 131–163.

Gati, A., Wong, M., Alquie, G., & Fouad Hanna, V. (2000). Neural networks modeling and parameterization applied to coplanar waveguide components. *International Journal of RF and Microwave Computer-Aided Engineering*, *10*(5), 296–307.

Graesser, J., Cheriyadat, A., Vatsavai, R. R., Chandola, V., Long, J., & Bright, E. (2012). Image based characterization of formal and informal neighborhoods in an urban landscape. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(4), 1164–1176.

Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, *3*(6), 610–621.

Hay, G. J., & Castilla, G. (2008). Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. *Object-based Image Analysis*, 75–89.

Hsu, C.-W., & Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, *13*(2), 415–425.

Jabid, T., Kabir, M. H., & Chae, O. (2010a). Gender classification using local directional pattern (LDP). In *20th International Conference on Pattern Recognition (ICPR), 2010* (pp. 2162–2165). IEEE.

Jabid, T., Kabir, M. H., & Chae, O. (2010b). Local directional pattern (LDP) for face recognition. In *2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE)* (pp. 329–330). IEEE.

Jabid, T., Kabir, M. H., & Chae, O. (2010c). Robust facial expression recognition based on local directional pattern. *ETRI Journal*, *32*(5), 784–794.

Kabir, M. H., Jabid, T., & Chae, O. (2010). A local directional pattern variance (LDPv) based face descriptor for human facial expression recognition. In *2010 Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 526–532). IEEE.

Kasabov, N. K. (1996). *Foundation of neural networks, fuzzy systems, and knowledge engineering*. Cambridge, MA: MIT.

Khumalo, P. P., Tapamo, J. R., & Van Den Bergh, F. (2011). Rotation invariant texture feature algorithms for urban settlement classification. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International* (pp. 511–514).

Mayunga, S., Coleman, D., & Zhang, Y. (2007). A semi-automated approach for extracting buildings from QuickBird imagery applied to informal settlement mapping. *International Journal of Remote Sensing*, *28*(10), 2343–2357.

Owen, K. K., & Wong, D. W. (2013). An approach to differentiate informal settlements using spectral, texture, geomorphology and road accessibility metrics. *Applied Geography*, *38*(1), 107–118.

McLaren, R., Coleman, D., & Mayunga, S. (2005). Sustainable management of mega growth in megacities. In *Paper presented at the International Federation of Surveyors, FIG*.

Shabat, A. M., & Tapamo, J. R. (2014). A comparative study of local directional pattern for texture classification. In *World Symposium on Computer Applications and Research (WSCAR), 2014*.

Van Den Bergh, F. (2011). The effects of viewing-and illumination geometry on settlement type classification of quickbird images. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International* (pp. 1425–1428).

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media.