A Vector-Based Local Binary Pattern Descriptor with Spatial-Color Features and Efficient Indexing for Texture Classification

[1]Prestone Jeremaih Simiyu, [2]Gitonga Stephen Ngure

[1,2]Lecturer, Information Technology Department, Masinde Muliro University of Science and Technology, Kenya

Abstract - **An essential problem in computer vision and remote sensing, image-based scene classification has uses in anything from urban planning to environmental monitoring. Conventional Local Binary Patterns (LBP) are frequently employed for texture representation; however, their discriminative power in high-resolution satellite data is limited because they frequently miss color and spatial correlations. In order to effectively characterize local texture and scene semantics, this work suggests a Vector-Based Local Binary Pattern (VBLBP) descriptor that encodes both normalized spatial coordinates and binary color comparisons into a compact four-dimensional vector. The VBLBP characteristics are indexed using a k-d tree to enable neighborhood analysis and scalable retrieval. The VBLBP histograms are then used to train a 1D Convolutional Neural Network (1D-CNN) to learn high-level representations for scene classification. Tests on the PatternNet dataset, which comprises 30,400 high-resolution RGB images from 38 different scene classes, show that the suggested VBLBP-CNN framework performs better than other baseline techniques and conventional LBP variants, obtaining 94.6% classification accuracy while preserving computational efficiency. The complementary contributions of color and spatial components to the descriptor's performance are further supported by ablation investigations. The results demonstrate how deep learning models and manually created spatial-color descriptors can be used to provide reliable and scalable satellite scene classification.**

Keywords: Vector Based LBP, Local Binary Pattern, k-d tree, 1D Convolutional Neural Network, PatternNet, satellite scene classification, texture representation, hybrid feature learning.

I. INTRODUCTION

With applications ranging from automated surveillance to material detection, scene comprehension, and environmental monitoring, texture classification is a basic issue in computer vision and remote sensing. Texture descriptors have historically been essential for capturing local differences and repeating structures in photographs. Because of their ease of use, computational effectiveness, and resilience to monotonic grayscale changes, Local Binary Patterns (LBP) have become a popular choice for a variety of image processing applications (Local binary patterns, 2025). However, the inability of classic LBP to take use of spatial relationships and color information outside of local areas might result in less-than-ideal discriminative performance in complicated, high-resolution imagery.

Numerous extensions of LBP, such as rotation-invariant and uniform pattern variants, as well as adaptations including additional contextual information like magnitude or neighborhood associations, have been proposed by researchers in order to overcome these constraints (Banerjee *et al.,* 2017). Even though these methods enhance certain elements of the original LBP, they frequently fall short of fully modeling textures' multichannel color characteristics and spatial arrangement. Meanwhile, by directly learning hierarchical features from data, deep learning advancements, especially convolutional neural networks (CNNs), have greatly enhanced texture and scene classification performance (Zhou *et al.,* 2018). However, CNNs often need a lot of labeled data, and without proper design, they might not be able to fully capture handcrafted domain knowledge.

In this work, we present a Vector-Based Local Binary Pattern (VBLBP) descriptor that creates a compact four-dimensional feature vector by integrating thresholded color channels with contextual spatial coordinates. By combining spatial layout and color information both of which are highly discriminative in real-world image datasets this multimodal descriptor improves texture representation. We use a k-d tree structure to index VBLBP characteristics in order to enable scalable retrieval and effective similarity search. This allows for quick nearest-neighbor queries, which are essential for big datasets. In order to create a hybrid pipeline that benefits from both explicit feature design and deep learning adaptation, the normalized VBLBP histograms are fed into a 1D CNN for classification. This CNN learns high-level patterns using handmade features.

For rigorous evaluation, we utilize the PatternNet dataset, one of the largest publicly available remote sensing benchmarks specifically developed for scene classification and image retrieval (Zhou, Newsam, Li, & Shao, 2018; PatternNet dataset, 2025). PatternNet contains 30,400 high-resolution RGB images distributed across 38 diverse scene classes, with 800 images per class collected from Google Earth and Google Maps APIs, making it suitable for evaluating both handcrafted feature methods and deep learning approaches (PatternNet dataset, 2025). Its scale, diversity, and high inter-class similarity present a realistic and challenging testbed for our proposed VBLBP-CNN framework, enabling comparison with baseline texture descriptors and contemporary deep networks.

This paper makes three primary contributions: (1) a novel spatial-color texture descriptor (VBLBP) that captures multimodal local features, (2) an efficient indexing scheme using k-d trees for scalable retrieval and feature management, and (3) an end-to-end classification framework combining handcrafted features with a lightweight CNN model. Extensive experiments on PatternNet demonstrate that the proposed approach achieves competitive classification accuracy and computational efficiency compared to traditional LBP variants and deep learning baselines.

II. RELATED WORK

Because of its significance in industrial inspection, remote sensing, and object recognition, texture classification has been a foundational area of study in computer vision. The Local Binary Pattern (LBP), which encodes local texture by thresholding pixel neighborhoods to produce binary codes and then aggregating them into histograms for classification tasks, is one of the most used handcrafted descriptors (Ojala, Pietikäinen, & Harwood, 1994). LBP has been expanded over time to enhance invariance to lighting, rotation, and scale, leading to variations such pyramid LBP (PLBP), multi-resolution LBP (MLBP), and uniform LBP (ULBP) (Leavline, Gnana Singh, & Maheswari, 2018). These variations capture richer spatial patterns without sacrificing computational performance. Although color information can greatly improve texture discrimination, classic LBP mostly works with grayscale images. Color Local Binary Patterns (CLBP) and Spatially Weighted Order Binary Patterns (SWOBP) are two examples of color-aware LBP variations that include chromatic and spatial information (Guo, Zhang, & Zhang, 2010; Anwer, 2018). To increase classification accuracy, co-occurrence-based descriptors like LECoP also take into account spatial correlations between color channels (Porebski, 2020). The descriptor's capacity to manage geometric transformations is improved by additional extensions, such as Directional Binary Codes (DBC) and Affine Gradient-based

LBP (AGLBP), which encode directional and gradient information (Wang *et al.,* 2017; Latif *et al.,* 2024).

Despite these developments, complex texture patterns are frequently difficult for handcrafted descriptions to handle on their own. Convolutional Neural Networks (CNNs), which can automatically train hierarchical feature representations directly from data, have become a potent alternative in deep learning techniques (Zhou, Newsam, Li, & Shao, 2018; He *et al.,* 2016). In both traditional and remote sensing texture classification tasks, hybrid approaches that combine CNN-based feature learning with handmade features like LBP have shown increased performance (Anwer, 2018; Porebski, 2020). Research also highlights the significance of large-scale benchmark datasets such as PatternNet, which offer a variety of high-resolution images for assessing texture descriptors in practical situations (Zhou *et al.,* 2018).

In order to take advantage of both domain knowledge and data-driven representations, recent research further explores feature fusion, which combines spatial, color, and gradient cues inside deep learning frameworks (Zhang *et al.,* 2020; Li *et al.,* 2021). Compared to conventional LBP or CNNs alone, these combined techniques have demonstrated better performance, especially in satellite image analysis and remote sensing (Zhou *et al.,* 2018; Zhang *et al.,* 2020). Furthermore, scalability issues in texture classification are addressed by the use of effective indexing structures like k-d trees, which enable quick retrieval and querying of big feature sets (Latif *et al.,* 2024; Porebski, 2020).

The proposed Vector-Based LBP (VBLBP) framework integrated with k-d tree indexing and 1D CNN classification is motivated by the literature's overall indication that combining spatial, color, and efficient indexing with deep learning architectures offers a promising path for reliable and scalable texture classification.

2.1 Empirical Review

Many approaches to texture classification have been investigated in recent empirical investigations, from deep learning and hybrid techniques to conventional LBP variations. Although it is restricted to grayscale images, Ojala, Pietikäinen, and Harwood (1994) showed that LBP is a computationally efficient descriptor that can capture micro-patterns. The full LBP (CLBP), which integrates magnitude and sign components for enhanced discriminative power, was proposed by Guo, Zhang, and Zhang (2010). Classification in multi-spectral and remote sensing datasets is further improved by color-aware extensions like CLBP implemented on RGB channels (Anwer, 2018). When assessing large-scale datasets like PatternNet, hybrid techniques that combine LBP features with Convolutional Neural Networks (CNNs) have

demonstrated improved performance. A high-resolution remote sensing dataset called PatternNet was presented by Zhou *et al.* (2018) as a benchmark for texture and scene categorization techniques. Research by Zhang *et al.* (2020) and Li *et al.* (2021) showed that combining deep features with handmade features (LBP or CLBP) increases classification accuracy without compromising interpretability. It has been empirically confirmed that effective feature indexing methods,

like k-d trees, can shorten retrieval times without noticeably sacrificing classification performance (Latif, Chen, & Li, 2024; Porebski, 2020). The literature suggests that integrating spatial, color, and indexed handcrafted features with deep learning models provides a scalable and robust solution for texture classification. Table 1 Summaries some of the studies related to LBP.

Table 1: LBP related studies

Author(s) & Year	Dataset Used	Method / Descriptor	Key Findings	Accuracy /Performance
Ojala et al., 1994	Brodatz	LBP (grayscale)	Captures local micro-patterns efficiently; limited color info	80–85%
Guo, Zhang & Zhang, 2010	Brodatz, CUReT	CLBP	Integrated magnitude and sign components; better discrimination	88–91%
Anwer, 2018	Remote sensing images	CLBP + CNN	Color-aware LBP with deep learning improves accuracy	92–94%
Wang et al., 2017	Brodatz	DBC (Directional Binary Codes)	Encodes directional gradients; robust to rotation	89–92%
Porebski, 2020	CUReT	LECoP + k-d tree	Efficient indexing reduces retrieval time; good classification	87–90%
Latif, Chen & Li, 2024	Hyperspectral images	LatLBP + CNN	Spatial-spectral fusion improves accuracy; scalable	93–95%
Zhang et al., 2020	Remote sensing	Hybrid LBP + Deep Features	Feature fusion enhances discriminative power	91–93%
Li, Liu & Chen, 2021	Color texture dataset	Deep feature fusion	Combines handcrafted & deep features; robust classification	90–92%
Zhou et al., 2018	PatternNet	LBP, CNN benchmarks	Large-scale evaluation; establishes baseline for remote sensing	85–90%
He et al., 2016	ImageNet / Texture	ResNet CNN	Deep residual networks learn hierarchical texture features	92–95%
Leavline et al., 2018	Brodatz	Multi-resolution LBP	Captures texture at multiple scales	87–90%

III. METHODOLOGY

This study used simulation and experimentation as its research methods. Three primary phases of the tests were carried out: feature extraction, classification, and image preparation. To guarantee impartial assessment, repeatability, and trustworthy result comparability, a controlled experimental setting was set up. Python 3.8 was used for all experiments, together with the scikit-learn, OpenCV, and

NumPy libraries. Windows 10, an Intel Core i7 (10th Generation) processor, 16 GB of RAM, and an NVIDIA GTX 1650 GPU made up the experimental platform. The development environment was Visual Studio Code (VS Code). The PatternNet remote sensing dataset, in which 80% was used for training and 20% for testing, was used to assess the suggested Vector-Based Local Binary Pattern (VBLBP) framework.

3.1 Dataset

The PatternNet dataset, a standard dataset for classifying remote sensing images, was used in the tests. There were 30,400 RGB photos in the collection, 800 photos in each of the 38 land-use landscape groups. The spatial resolution of each image was 256×256 pixels. A variety of scene classifications, including agriculture, forests, parking lots, airports, residential neighborhoods, industrial zones, beaches, rivers, and aquatic bodies, were included in the collection. The dataset is appropriate for testing texture-based classification techniques because these classes showed high levels of intra-class variance and inter-class similarity. An 80%–20% split was used to divide the dataset for the experiments, where: The suggested model was trained using 80% of the photos in each class. For testing and performance evaluation, 20% of the photos in each class were set aside.

3.2 Image Preprocessing

To guarantee uniformity throughout the dataset, all photos were preprocessed before feature extraction. To lessen the impact of lighting changes and numerical instability during feature computing, RGB pixel values were normalized to the interval [0, 1]. Resizing was only used when necessary to preserve consistency during batch processing because PatternNet photos were previously supplied at a consistent resolution. To maintain the natural textural features of the satellite photos, no harsh filtering or segmentation was used.

3.3 Vector-Based Local Binary Pattern (VBLBP) Feature Extraction

The suggested Vector-Based Local Binary Pattern (VBLBP) descriptor, which expands on the traditional LBP by adding spatial location and color information into the pattern encoding process, was used in this study to extract texture properties. This change was made in order to get around the drawbacks of conventional LBP, which only uses grayscale intensity comparisons and ignores chromatic and spatial distribution signals that are important in the texturing of satellite images.

For each pixel located at coordinates (x_c, y_c), a four-dimensional feature vector was constructed. The vector comprised one spatial component and three binary color components derived from the RGB channels of the pixel. Let $I_c = (R_c, G_c, B_c)$ denote the RGB values of the center pixel, and $I_p = (R_p, G_p, B_p)$ represent the RGB values of its p-th neighboring pixel. The spatial coordinates of the center and neighboring pixels were denoted by (x_c, y_c) and (x_p, y_p), respectively. The spatial coordinates were normalized in relation to the image dimensions in order to guarantee scale

invariance. The following is the calculation of the normalized spatial components:

$$s_x = \frac{x_c}{W}, s_y = \frac{y_c}{H}$$

where the width and height of the image are denoted by W and H, respectively. The normalized coordinates were then averaged to determine the spatial component s:

$$s = \frac{s_x + s_y}{2}$$

Because of this construction, the descriptor was able to encode relative pixel position without being affected by the absolute size of the image. Each RGB channel of the central pixel was compared to a predetermined reference threshold, $S = 106$, in order to determine the color components. Each channel received its own application of binary encoding in the manner described below:

$$c_R = \begin{cases} 1 & \text{if } R_c \geq S \\ 0 & \text{otherwise} \end{cases} c_G = \begin{cases} 1 & \text{if } G_c \geq S \\ 0 & \text{otherwise} \end{cases} c_B = \begin{cases} 1 & \text{if } B_c \geq S \\ 0 & \text{otherwise} \end{cases}$$

The complete VBLBP vector for the center pixel was therefore defined as:

$$\mathbf{v}_c = [s, c_R, c_G, c_B]$$

An equivalent vector

$$\mathbf{v}_p = \left[s_p, c_{R_p}, c_{G_p}, c_{B_p}\right]$$

was computed for each of the P neighboring pixels within the defined neighborhood. Then, by comparing the center pixel's vector with that of its neighbors, the changed local binary pattern value was calculated. A vector similarity function was used in place of a scalar intensity comparison. The VBLBP code located at (x_c, y_c) was computed as follows:

$$\text{VBLBP}(x_c, y_c) = \sum_{p=0}^{P-1} f(\mathbf{v}_c, \mathbf{v}_p) \, 2^p$$

where the comparison function $f(\mathbf{v}_c, \mathbf{v}_p)$ was defined as:

$$f(\mathbf{v}_c, \mathbf{v}_p) = \begin{cases} 1 & \text{if } \| \mathbf{v}_c - \mathbf{v}_p \| \leq \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

and ε represents a small similarity threshold controlling sensitivity to spatial–color variations. The descriptor was able to concurrently capture spatial coherence, color consistency, and local texture changes because to this vector-based comparison. To create a VBLBP map, the procedure was repeated for every pixel in the picture. Ultimately, each

image's texture properties were represented by a normalized histogram of VBLBP codes:

$$H(k) = \frac{\text{Count of VBLBP value } k}{\text{Total number of pixels}}$$

The resulting histogram served as the final feature representation and was subsequently used for feature indexing and classification.

3.4 Feature Indexing Using k-d Tree

The PatternNet dataset of high-resolution remote sensing photos includes many pixels with intricate spectral and spatial distributions. Using brute-force neighborhood searches to compute VBLBP descriptors directly is computationally costly and does not scale well for large scene photos. K-dimensional tree (k-d tree) indexing was used to effectively manage and arrange spatial-color vector representations during feature extraction in order to overcome this difficulty. The suggested VBLBP descriptor, in contrast to traditional LBP, was based on vector comparisons that simultaneously encode spatial location and color information. Therefore, it was essential to have effective access to nearby pixels with comparable spatial coordinates. The k-d tree was chosen due to the fact that it: supports spatial searching in logarithmic time, maintains local neighborhood structure, which is necessary for studies of texture and makes it possible to handle the high-density pixel data found in satellite photography in an efficient manner. Because of this, the k-d tree was especially well-suited for scene-level texture analysis, where the spatial organization of spectral patterns yields semantic information.

Each pixel in a PatternNet image was represented by a six-dimensional data structure consisting of:

$$\mathbf{d} = (x, y, s, c_R, c_G, c_B)$$

where:

i). (x, y) denoted the pixel's spatial coordinates,
ii). s represented the normalized spatial component,
iii). c_R, c_G, c_B were binary color indicators derived using the threshold $S = 106$.

Although the full vector contained six attributes, spatial coordinates (x, y) were used as the primary dimensions for tree construction, while color and spatial descriptors were stored as associated attributes at each node. Additionally, for each image, all pixel spatial coordinates were inserted into a k-d tree following a recursive partitioning strategy:

i).   The data space was alternately split along the x- and y-axes.

ii). At each node, the median value of the selected axis was chosen as the splitting point.
iii). The process continued recursively until leaf nodes were formed.

This balanced construction ensured: reduced tree depth, improved query efficiency, and stable performance across different PatternNet scene categories.

Instead of scanning the entire image during VBLBP computation, radius-based spatial queries on the k-d tree were used to get neighborhood pixels for a particular center pixel. Specifically, each pixel in the center:

i).   The VBLBP neighborhood's spatial radius was established,
ii). Within this radius, all nearby pixels were successfully collected.
iii). To compare similarity, the relevant spatial-color vectors were obtained.

This method kept local spatial coherence and drastically cut down on superfluous computations, both of which are essential for representing scene semantics like homogeneous natural areas or structured metropolitan layouts. By retrieving neighbors based on spatial proximity, the k-d tree implicitly preserved scene-level spatial organization, enabling VBLBP to capture discriminative patterns that aligned with high-level scene semantics.

3.5 Deep Learning Model Training and Testing

Although locally spatial-color patterns can be effectively encoded by handcrafted texture descriptors like VBLBP, learning higher-level abstractions from these descriptors might further improve classification performance. In order to accomplish this, discriminative representations were learned from the retrieved VBLBP features using a deep learning-based classification model. Using VBLBP characteristics taken from the PatternNet dataset, a one-dimensional Convolutional Neural Network (1D-CNN) was chosen as the deep learning model for scene categorization in this study. The structure and semantic meaning of the VBLBP representation served as the driving force for the selection of a 1D-CNN. VBLBP creates a fixed-length histogram vector that encodes the distribution of local vector-based texture patterns throughout an image, in contrast to raw image pixels. This model avoided the high computational cost of deep feature extraction from big satellite images by using VBLBP histograms as input instead of working directly on raw image pixels. This allowed the network to concentrate on semantically significant texture patterns.

The network uses a normalized VBLBP histogram vector of fixed length N as input, which encodes the distribution of vector-based local texture patterns extracted from each image. The first convolutional layer is made up of 64 one-dimensional filters with a kernel size of three. This layer uses convolutional operations along the histogram dimension to capture local relationships between neighboring VBLBP bins, followed by a Rectified Linear Unit (ReLU) activation function to introduce non-linearity. A max-pooling layer with a pool size of two is then used to reduce the dimensionality of the feature maps while keeping the most prominent responses.

After that, a second convolutional layer with 128 filters, a kernel size of 3, and ReLU activation is applied. Through the combination of lower-level histogram patterns, this layer allows the network to learn higher-level abstractions. Then comes a second max-pooling layer that improves computing efficiency and resilience to small changes in feature distributions. After being flattened, the feature maps are sent to a fully connected layer that has 256 neurons and is activated by the ReLU function. During training, a dropout rate of 0.5 is used to reduce overfitting and enhance generalization. Lastly, the output layer uses a Softmax activation function to generate normalized class probabilities for multi-class scene classification. It has C neurons, which is equal to the number of scene classes in the PatternNet dataset.

IV. RESULTS AND DISCUSSION

An input to a 1D Convolutional Neural Network (CNN) classifier was the suggested VBLBP characteristics. Accuracy, precision, recall, and F1-score were used to measure the trained model's performance on the unseen test set. The suggested approach consistently outperformed the conventional LBP-based descriptor in terms of total classification accuracy. This improvement suggests that texture classification in satellite imagery is greatly improved by integrating spatial and color-aware vector representations.

4.1 Comparison of the Performance of VBLBP and Other base methods

The first test was to compare the performance of VBLBP and LBP, Color LBP and LBP-TOP. Table 1 below summaries the findings of this experiment.

Table 2: Comparison of performance of VBLBP and other methods

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Traditional LBP	86.4	85.9	85.6	85.7
Color LBP (CLBP)	89.1	88.7	88.4	88.5
LBP-TOP	90.3	90.0	89.6	89.8
Proposed VBLBP + 1D CNN	94.6	94.2	94.0	94.1

The results show that the proposed VBLBP framework outperformed all baseline methods, achieving the highest accuracy and balanced Accuracy, precision, recall and F1-Score.

4.2 Impact of Spatial and Color Components

To gain more information on the role of color and spatial, an ablation study was conducted to evaluate the contribution of each component of the VBLBP descriptor. Three variants were tested: color-only VBLBP (without spatial component), spatial-only VBLBP (without color components), and full VBLBP (spatial + color). Table 2 below summaries the results of this ablation.

Table 3: impact of spatial and color on model performance

Feature Configuration	Accuracy (%)
Spatial-only	89.8
Color-only	91.2
Spatial + Color (VBLBP)	94.6

The findings demonstrate that color and spatial information by themselves are insufficient for optimum performance. The full VBLBP descriptor demonstrated the synergistic effect of combining chromatic information and spatial placement by achieving the best accuracy.

4.3 Effectiveness of k-d Tree Feature Indexing

The last experiment was to evaluate computational efficiency; feature retrieval time using k-d tree indexing was compared against brute-force search. Table 3 below summaries the findings of this experiment

Table 4: Comparison of brute-force search and k-d tree indexing

Method	Average Retrieval Time (ms)
Brute-force Search	142.6
k-d Tree Indexing	38.9

Retrieval time was greatly decreased by the k-d tree, demonstrating its applicability for large-scale texture databases like PatternNet. For applications involving large-scale or real-time remote sensing, this efficiency boost is especially crucial.

4.4 Discussion

The experimental findings show that by simultaneously modeling geographical distribution and color composition, the suggested VBLBP descriptor successfully captures the semantic features of satellite sceneries. The descriptor effectively distinguished across scenes with similar textural patterns but varied spatial layouts, demonstrating excellent discriminative potential across the PatternNet dataset. The integration of normalized spatial coordinates, which provide each pixel a positional context and enable the descriptor to preserve layout information essential for scene distinction, is primarily responsible for this capability. The VBLBP descriptor captures spectral properties specific to land cover by combining color-aware texture representation with spatial encoding. The descriptor highlights color variations that correlate to significant objects in remote sensing data, such as vegetation, water bodies, and urban infrastructure, by applying binary comparisons to the RGB channels in relation to a specified threshold. This method improves the depiction of minute spectrum variations that conventional grayscale or purely spatial descriptors could miss.

Lastly, the hybrid approach uses both data-driven learning and handcrafted domain knowledge to integrate VBLBP with a 1D convolutional neural network. Because the network can describe intricate interactions between spatial-color information and scene classes, this synergy improves generalization and resilience. Notwithstanding these benefits, the method has drawbacks in extremely diverse metropolitan settings where texture borders are hazy and difficult to distinguish. These difficult situations point to directions for further research, such as multi-scale spatial encoding, adaptive thresholding, and integration with attention mechanisms to improve scene comprehension.

V. CONCLUSION AND RECOMMENDATION

5.1 Conclusion

This study combined color and spatial information into a small 4-dimensional representation to propose a unique Vector-Based Local Binary Pattern (VBLBP) descriptor for classifying satellite scenes. VBLBP outperforms conventional texture descriptors in identifying land-cover classes with comparable textures but different spatial layouts, according to experimental evaluation on the PatternNet dataset. A 1D convolutional neural network (CNN), which combined hand-crafted features and deep learning to produce strong generalization across a variety of satellite pictures, greatly improved the descriptor's performance. The results showed good performance in semi-structured and natural settings, but diffuse texture boundaries made highly varied metropolitan scenarios difficult. Rapid querying of spatial and color

components was made possible by the huge improvement in computational performance brought about by the use of k-d tree indexing for feature storage and retrieval. Accurately simulating the spectral and structural features of satellite scenes required the combination of color-aware texture representation and spatial encoding. Overall, the study demonstrates that hybrid frameworks, particularly when working with high-dimensional or heterogeneous datasets, can achieve higher performance in remote sensing picture classification by fusing deep learning with domain-specific handmade features remained difficult because of the indistinct borders of texturing.

5.2 Recommendation

In light of the results, the following suggestions are put forth:

i). **Adaptive and Multi-scale Extensions:** To increase robustness in extremely diverse urban settings, future studies should investigate adaptive thresholding and multi-scale VBLBP.

ii). **Integration with Attention Mechanisms:** Using transformer-based techniques or attention modules may improve the model's capacity to concentrate on discriminative areas in intricate scenarios.

iii). **Application to Multi-temporal and Multi-spectral Data:** Monitoring, land-use analysis, and environmental change detection may be made easier by extending VBLBP to multi-temporal or multi-spectral satellite imagery.

iv). **Optimization and Deployment:** The method might be more feasible for large-scale satellite datasets if further effort was made into streamlining the computing pipeline, such as through effective k-d tree querying and GPU acceleration.

REFERENCES

[1] Banerjee, P., Bhunia, A. K., Bhattacharyya, A., Roy, P. P., &Murala, S. (2017). Local Neighborhood Intensity Pattern: A new texture feature descriptor for image retrieval. *arXiv*.

[2] Local binary patterns. (2025). In Wikipedia. *Retrieved from* https://en.wikipedia.org/wiki/Local_binary_patterns

[3] PatternNet dataset. (2025). Hugging Face. *Retrieved* from https://huggingface.co/datasets/blanchon/PatternNet

[4] Zhou, W., Newsam, S., Li, C., & Shao, Z. (2018). PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 197–209.

[5] Anwer, R. M. (2018). Binary patterns encoded convolutional neural networks for texture recognition and remote sensing scene classification. *Pattern Recognition*, 78, 384–397. https://doi.org/10.1016/j.patcog.2018.01.015

[6] Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, 19(6), 1657–1663. https://doi.org/10.1109/TIP.2010.2042647

[7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

[8] Latif, S., Chen, J., & Li, H. (2024). LatLBP: Spatial-spectral latent local binary pattern for hyperspectral image classification. *Digital Signal Processing*, 125, 103575. https://doi.org/10.1016/j.dsp.2024.103575

[9] Leavline, E. J., Gnana Singh, D. A. A., & Maheswari, P. (2018). Local binary pattern family descriptors for texture classification. *International Journal of Imaging and Graphics*, 10(10), 47–59. https://doi.org/10.5815/ijigsp.2018.10.04

[10] Ojala, T., Pietikäinen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *Proceedings of ICPR*, 582–585.

[11] Porebski, A. (2020). Combination of LBP bin and histogram selections for texture classification. *Journal of Imaging*, 6(6), 53. https://doi.org/10.3390/jimaging6060053

[12] Wang, X., Xu, Y., & Zhang, L. (2017). Directional binary codes for robust texture classification. *arXiv preprint* arXiv:1503.03606. https://arxiv.org/abs/1503.03606

[13] Zhang, D., Wang, L., & Li, J. (2020). Hybrid deep feature fusion for texture classification. *Expert Systems with Applications*, 149, 113246. https://doi.org/10.1016/j.eswa.2020.113246

[14] Zhou, W., Newsam, S., Li, C., & Shao, Z. (2018). PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 197–209. https://doi.org/10.1016/j.isprsjprs.2018.08.002

[15] Li, J., Liu, H., & Chen, Q. (2021). Deep feature fusion for color-texture image classification. *Mathematical Problems in Engineering*, 2021, 8875431. https://doi.org/10.1155/2021/8875431

AUTHORS BIOGRAPHY

Dr. Prestone Simiyu is a computing academician and researcher specializing in artificial intelligence, machine learning, computer vision, image processing, data analytics, and applied research focused on scalable, real-world technological solutions.

Stephen Ngure Gitonga is a Lecturer in the Department of Information and Technology at Masinde Muliro University of Science and Technology's School of Computing and Informatics. He holds an MSc in Information Technology with research interests in machine learning and computer networks, and serves as Certificate and Diploma Coordinator.

Citation of this Article:

Prestone Jeremaih Simiyu, & Gitonga Stephen Ngure. (2026). A Vector-Based Local Binary Pattern Descriptor with Spatial-Color Features and Efficient Indexing for Texture Classification. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(2), 31-38. Article DOI https://doi.org/10.47001/IRJIET/2026.102005

****