Image recognition via Local 3-bit Binary Patterns

Abdulaziz Saleh Alraddadi[1] , Essam O. Abdel-Rahman[1,2,*]

[1]Department of Computer Science, Faculty of Science and Computer Engineering, Yanbu Branch, Taibah University, Yanbu, KSA

[2]Department of Computer Science, Faculty of Computers and Information, Minia University, Minia, Egypt

Emails: araddadi@taibahu.edu.sa; eothman@taibahu.edu.sa

Abstract

The current study introduces a trainable object detection model that can be taught to detect an object of a given class within an unconstrained scene. The researchers of the current study use this advanced system in the detection of Relics images, which involves a calculation of Local 3bit Binary Patterns (3bit-LBP). The key highlights of the current work include the integration and analyses of the utilization of the Multi-Support Vector Machine Classification (MSVMC) and Integral image computation analysis. The experimental outcomes of the current study indicate that the method of 3bit-LBP is superior to other methods in accuracy and stability, especially when images of different illumination and object rotation were tested. The researchers further conducted a comparative performance evaluation showing that the presented system gives better detection rates as compared to the conventional strategies, revealing the efficiency in real-world applications. Finally, it is important to note that the implications of the results can be applied to uses beyond just relic detection. To conclude, the current work advances the knowledge of how to improve the functionality of object recognition algorithms further in the context of image recognition systems.

Keywords: Image recognition; Multi-Support Vector Machine Classification; Local Binary Patterns; Binary Classifier

1. Introduction

Image recognition is a fascinating and rapidly evolving sphere within the field of computer vision [1]. This technology is capable of enabling machines they are integrated within to identify and classify diverse objects, which may include people, shapes, and various elements within digital images and videos. Hence, it can be said that image recognition is a technology that holds the potential to mimic the human ability to perceive and interpret visual information [2]. However, this complex task of image detection and handling requires considerable processing power and refined algorithms to examine patterns and regularities present in image data. This ultimately leads to precise classifications based on pixel patterns [3].

The importance of image recognition is apparent in the daily lives of an individual in today's modern era, ranging from facial recognition technology ensuring security on mobile phones, and banking applications, to checking deposits among others [4]. Moreover, it plays a vital role in medical imaging, where it benefits in detecting tumours, fractures, and other anomalies that enables the efficiency of healthcare professionals. Image detection is also a crucial technology in manufacturing environments where it recognizes defective products or categorizes them as they move on assembly lines [5].

The foundation for image recognition technology can be traced back to the late 1950s [6]. The fundamental studies in this landscape aimed to understand how machines could interpret visual data. The field of computer vision started flourishing in the 1970s and 1980s after the formulation of simple algorithms for edge detection and other patterns [4]. Moreover, breakthroughs in computing in the 1990s saw machine learning methods improve the ability of computers to analyse images [7]. Then, the 2000s saw noteworthy innovations with the introduction of

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

deep learning, leading to refined neural networks capable of complex image analysis [8]. Currently, image recognition is a core part of many processes, including facial recognition, relic recognition, and even self-driving cars.

Image recognition typically begins with the processing of simple structures, such as the distinct edges of objects [6]. This principle remains a core part of the deep learning methodologies that support modern image recognition technologies. The primary goal is to 'learn' the image data and identify the patterns of pixels so that the classification of such image data can be efficiently executed. In the realm of computer vision and image processing, texture categorization is especially important. For instance, while visiting sites with archaeological remains, customers are interested in receiving specific information on the artefacts that are on display. The basic static detection system in the current work uses a brute-force search method to find relics in an image. Notably, unlike previous object detection approaches, the system proposed in the current study does not depend on motion or tracking and imposes no assumptions regarding scene structure, object quantity, or camera movement [9, 10, 11, 6].

One of the important concepts that need to be understood before commencing with the study is that of summed-area tables or integral images [12]. Summed-area tables are a technique used in computer graphics for efficient texture mapping. Although it is possible to determine the texture by summing up the texture value from all local pixel values in the texture map for each pixel in a texture map, a summed-area table employs precomputation of the sum of all pixel values in a rectangular area of the resultant map [13]. This precalculated data makes it convenient for sampling the texture during rendering, as the value of a rectangle can be obtained from the table in quite a fast manner, being a sum of just four values. This greatly minimizes the number of times a texture is referenced during rendering; thus, enhancing the performance. Summed-area tables may be also useful for other purposes in computer graphics, for instance, for computing global illumination or shadowing [12]. In general, they are effective means for solving the problem of texture mapping and increasing rendering algorithms' rate [12].

Moreover, another crucial term is Local Binary Patterns (LBP), which is a strong texture descriptor presented by Ojala et al in 1994, that describes local image textures based on the comparison of a pixel with 3x3 neighbourhood pixels [14]. The LBP method encodes each pixel by comparing its intensity value with the intensity values of its neighbours to arrive at a string of binary numbers that correspond to the texture of a given region. In the subsequent years, several advancements of the original LBP have been proposed especially concerning background subtraction methods [15]. However, since LBP is fast and quite resistant to noise, there are other methods, for example, the 2-bit Binary Pattern (2BitBP). The 2BitBP produces only four different codes compared with the 256 codes of the LBP has the benefit of a smaller number of codes for representation and therefore is useful in some applications where computational time is critical.

Finally, SVM stands for Support Vector Machines, which is another supervised machine-learning algorithm for classification, regression, and outlier detection [16]. They are used for various tasks, including face detection, text and hypertext categorization, customer segmentation, bioinformatics, handwriting recognition, stock market prediction, disease diagnosis, and image classification. An SVM looks for the best line or the hyperplane that separates different classes in an N-dimensional space with the largest margin [17]. Most algorithms fail in high dimensional space or when several features are larger than several samples, but it is not sensitive to these. SVMs utilize a part of the training data only, the so-called support vectors in the decision function, and thus they are memory efficient [18]. There are three main SVM classifiers including SVC (Support Vector Classification) for binary and multi-class classification, NuSVC (Nu-Support Vector Classification) which is similar but with different parameter sets and mathematical formulation, and finally LSVC (Linear Support Vector Classification) that is used for faster implementation for linear kernel classification [19].

Hence, these concepts are the key to understanding the methodology proposed in the current study. the current study aims to introduce a trainable model for object detection through the utilization of Local 3-bit Binary Patterns (3-bit LBP) along with the integration of Multi-Support Vector Machine Classification (MSVMC) for the use of detecting ancient relics. The researchers of the current study show that their 3bit-LBP is much more accurate and robust than other methods, especially in cases of changes in illumination and object orientation. The system demonstrates superior detection rates than the conventional techniques and has proven useful in practical applications. The importance of these results goes beyond the context of relics' identification since the work can be considered as the further development of the field of image recognition and improvement of the object detection systems' accuracy and performance. The main research question of the study is given by the following question: Does the proposed 3bit-LBP and MSVMC approach accurately detect relics in images and provide higher accuracy and robustness than the previous methods?

2. Literature Review

Fardo and Rodrigues (2024) present a novel approach for object classification for event-based sensors in a study where a sparse LBP extracted algorithm was proposed. The authors demonstrate in their research that the described

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

algorithm is useful for processing asynchronous event streams from new active pixel sensors, which output not only entire images but also their local change in brightness. The study shows that their LBP-based methodology can significantly reduce the feature vector size by up to 99.73% compared to state-of-the-art methods while maintaining comparable accuracy [20].

However, as a downside, the study does not present a detailed overview of real-life scenarios and different objects, where the supposed algorithm could be applied. The authors also fail to provide an adequate explanation of the computational complexity of the algorithm they developed and the performance of their algorithm on different hardware platforms, which may hinder the feasibility of their solution. The literature gap here ranges from a lack of empirical studies that would demonstrate the algorithm's effectiveness in various operational environments and a detailed analysis of the trade- offs between accuracy and computational efficiency when applied to large-scale datasets [20].

On the other hand, the authors in the study Rashed and Popescu (2022) systemically reviewed the developments in medical image processing technologies in the last five years concerning diagnosis in medical imaging. The authors categorized and reviewed forty articles published within the last decade as per PRISMA guidelines, and discussed its application of multiple imaging techniques, preprocessing & segmentation techniques, feature extraction, and classification algorithms, and the emergence of machine learning and deep learning techniques [21]. However, there is a potential gap in exploring how specific feature extraction techniques, like LBP, can be further optimized or adapted within contemporary deep learning frameworks for enhanced image recognition in medical contexts.

In the paper by Anwer et al. (2018), the authors propose a new two-stream deep architecture referred to as TEX-Nets where texture-coded mapped images are used as the second stream in addition to the conventional RGB image stream. For fusing the texture and colour information, two fusion techniques; the early and the late fusion methods are explored [22]. The study demonstrates that the late fusion TEX-Net architecture always improves performance compared to standard RGB deep networks without requiring fine-tuning or ensemble methods. However, the paper also has some limitations as the transformation of the LBP codes into a 3D metric space using the MDS, the EMD is not well described, and the effect on the performance is not thoroughly investigated.

Here, one of the research gaps that can be stated is the absence of literature where other hand - constructed features, aside from LBP, are integrated into deep learning frameworks for texture recognition and the classification of remote sensing scenes. Although LBP is a very effective descriptor, there can be other joint features, which, if fused with deep features, can provide more information. For example, features based on Gabor filters, fractal analysis, or co-occurrence matrices have shown promise in texture analysis and could be explored in a similar two-stream framework. In addition, the efficacy of the proposed approach for large-scale datasets such as ImageNet is not assessed. This would inform on the generalization capacity of the two-stream architecture on a data set that is as diverse as ImageNet.

The literature review displays a lack of empirical studies in the new object detection algorithms especially in the recognition of cultural relics. Several research studies have investigated different feature extraction methods and classification techniques, including LBP and SVM and other methods but few comprehensive works were done combining such methodologies under a single framework to reveal cultural relics under different conditions. The majority of the related work has been done with an emphasis placed only on one particular technique without proper regard to their efficiency when applied in real cases or evaluated for their flexibility when it comes to artefact type. This is an area particularly lacking in previous research; therefore, this study seeks to address this gap using a new method of Local 3-bit Binary Patterns (3-bit LBP) with Multi-Support Vector Machine Classification (MSVMC) to enhance the accuracy and robustness of relic detection.

3. Methodology

The methodology employed for this study involves creating a trainable object detection model employing 3bit-LBP and MSVMC. The first stage of data preparation includes the acquisition and pre-processing of an image dataset of artefacts. The images are also scaled down and the colour is changed to grey to help in the analysis of the images. The researcher in the current study used ‘Integral Images’ to compute the sum of pixel values rapidly in rectangular areas for fast feature extraction to be used in the object detection stage.

3.1. Feature Extraction with 3bit-LBP

The images were segmented into cells such as 32*32 pixels and for each pixel within such a cell; comparisons were made with 15 other neighboring pixels. A binary code is produced depending on whether the center pixel’s intensity is higher, equal, or lower than the neighbours’, this yields a 16-digit binary number. These binary codes are histogrammed and the histograms are concatenated to form a feature vector for the cell. Then the feature vectors of these cells are concatenated to form a feature vector for the image window.

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

3.2. Multi-Support Vector Machine Classification

Projections of the feature vectors are then utilized to train the Multi-Support Vector Machine (MSVM) classifier. The MSVM is capable of solving multi-class classification problems and, thus, can be used for finding different classes of relics in the images.

3.3. Performance Evaluation

A comparison with the classical detection methods is made to analyze the effectiveness of the proposed method in performance. Metrics such as accuracy, detection rates, and robustness under varying conditions (illumination and object orientation) are assessed to authenticate the usefulness of the 3bit-LBP and MSVMC approach.

4. Algorithm Steps to calculate Local 3bit Binary Patterns (3bit-LBP)

i. Divide the examined window into cells (e.g. 32x32 pixels for each cell).
ii. For each pixel in a cell, compare the pixel to each of its 16 neighbours (on its left top, left middle, left-bottom, right top, etc.). Follow the pixels along a circle, i.e. clockwise or counterclockwise.
iii. Where the centre pixel's value is greater than the neighbour's value, write "1". Otherwise, write "0". This gives a 16-digit binary number (which is usually converted to decimal for convenience).
iv. Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller, and which are greater than the centre).
v. Optionally normalize the histogram.
vi. Concatenate (normalized) histograms of all cells. This gives the feature vector for the window.

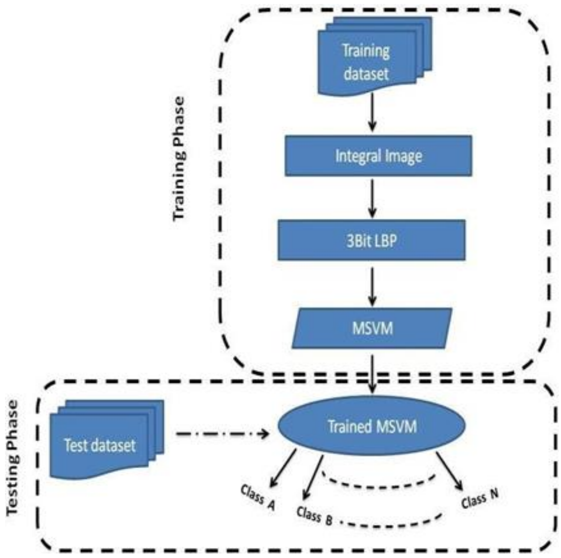


Figure 1. An internal dataset is typically split into two sets: training (Train) and testing (Test).

5. Results

The experimental outcomes of the current study indicate that the method of 3bit-LBP is superior to other methods in accuracy and stability, especially when images of different illumination and object rotation were tested. From the evaluation, the following features of the object detector were revealed.

The object detector developed in this study was based on a collection of features defined by the 2Bit Binary Pattern (2bitBP). Integral images were utilized to represent these patterns. The features contained information regarding the gradient orientation within specific areas. These features were quantized, resulting in the creation of four possible codes, as illustrated in Figure 2.

The bounding box for the object to be tracked was divided into a collection of features, referred to as sub-boxes. In each sub-box, the gradient orientation was measured using integral images. When

the gradient orientation moved vertically from top to bottom and horizontally from left to right, the corresponding binary pattern was identified as the upper left pattern in the 3-bit Binary Pattern section of Figure 2. This procedure was consistently applied to the remaining patterns. The logical operators NOT and AND were used in the

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

expressions presented in Table 1, where I(r1) represented the intensity gradient at region 1. Only two bits were required to code the represented patterns. The findings indicated that the 2Bit binary pattern effectively accommodated changes in object orientation, illumination, and noise present in the scene.

The expressions in Table 1 give the same meaning, where NOT and AND are the logical operators, I (r1) is the intensity gradient at region 1. Two bits only are needed to code the represented patterns.

The results also showed that the 2Bit binary pattern was significant in accommodating changes in object orientation, illumination, and the addition of noise across the scene. For example, an increase in scene illumination resulted in higher contrast. However, the intensity orientation remained unchanged, as the differences in intensities across each region were consistent.

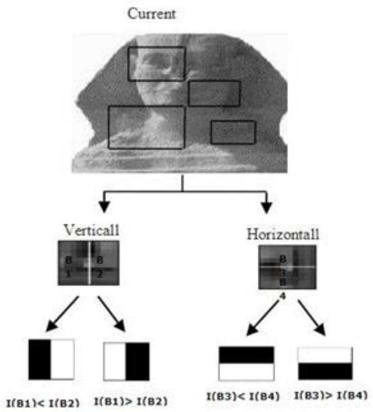
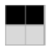
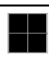


Figure 2. Local 3bit Binary Pattern used in our feature extracting, the output has only eight possible codes representing the gradient orientation in every patch. The vertical/horizontal combinations give eight possibilities, which construct the pattern.

Table 1: binary representation of the 3bit Binary Pattern

Binary Representation	**Binary equivalent**	**Binary pattern**
(I(B1) < I(B2)) AND (I(B3) < I(B4))	000	
(I(B1) < I(B2)) AND (I(B3) > I(B4))	001	
(I(B1) > I(B2)) AND (I(B3) > I(B4))	010	
(I(B1) > I(B2)) AND (I(B3) < I(B4))	011	
(I(B1) = I(B2)) AND (I(B3) < I(B4))	100	
(I(B1) < I(B2)) AND (I(B3) = I(B4))	101	
(I(B1) = I(B2)) AND (I(B3) > I(B4))	110	
(I(B1) > I(B2)) AND (I(B3) = I(B4))	111	
(I(B1) = I(B2)) AND (I(B3) = I(B4))	1000	

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

6. Discussion

As demonstrated by the results, the proposed 3bit-LBP method performs well in object detection especially concerning accuracy before illumination changes and object rotations. The ability to transform the 2Bit Binary Pattern (2bitBP) features into the integral images enables the computation of the proper gradient orientation within a certain region of interest of the bounding box [23]. By partitioning the bounding box into sub-boxes and using the integral images to calculate the gradient orientation in each sub-box the 3bit-LBP method encodes the patterns with only two bits and as a result, encumbers the least computational complexity while retaining the necessary amount of information.

The first advantage of integral images is the subdivision of an image into small regions, for which expectation can be computed using only four addresses of the array, regardless of the size of the region. This property allows one to quickly calculate the features that are described as rectangles and is critical when working in real-time object detection. However, arriving at such an image may pose certain problems associated with numerical accuracy because of fixed -size bit data representation. Less accurate computations may bring certain effects to subsequent image processing and analysing procedures. Further discussion on the analysis of the computation of integral images, grey-level processing, and extending SVM to solve multi-class problems is provided in this section.

7. Analysis of Integral Image Computation

Integral images are a widely used and efficient computational tool in many image processing techniques, such as template matching, image thresholding, binarization, pattern recognition, and other general filtering tasks [24]. The primary benefit of integral images is their ability to be reviewed several times for various statistical properties of the input image using only a few simple arithmetic operations, independent of the size of the region of interest after the integral image has been assessed [25]. Nonetheless, using integral pictures presents issues with numerical precision, due to the constrained bit size of the data representation [25]. Conversely, lower precision calculations may have an impact on the outcomes of image processing and analysis assignments.

When features are represented as rectangles, they can be computed easily. Integral Images represent the image intensity by summing the intensity of the pixels above the current pixel. For example, the integral image at location x, y contains the sum of all pixels above and to the left of the pixel at the x, y location including its value. The integral image can be represented by the following equation (1)

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \qquad (1)$$

Where ii is the integral image and i is the original image, the integral image can be calculated by only one pass over the image. The integral image is represented in Figure 3. Using the integral image any rectangular sum can be calculated by four array references (Figure 3). Suppose we need to calculate a feature represented by four adjacent rectangles, nine array references will be needed.

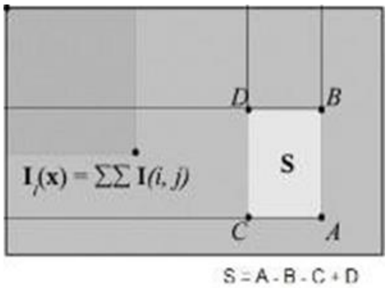


Figure 3. Integral Images: any rectangle in the integral image can be calculated with only four references, for example, the rectangle S=A-B-C+D.

7.1. Grey Level Processing

Colour images are utilized to enable the system to leverage the most visually prominent information across the three colour channels (RGB), which may be lost in grayscale images of the same scene [26]. Both image-processing steps—resize and Haar conversion—are conducted on each colour channel separately. To enhance system speed, a real-time version is implemented to manage image density.

7.2. Extending SVM to solve multi-class problems

A more natural approach to addressing the K-class pattern classification problem involves extending the basic SVM formulation for binary classifiers to encompass the K-class scenario. Weston and Watkins proposed two distinct extensions for this purpose. In the first approach, a decision function is constructed for each class, akin to

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

the K one-versus-rest classifiers method. The optimization problem in the SVM formulation can be generalized to consider all decision functions simultaneously, and the constraints are relaxed. Instead of requiring all decision functions to yield a zero value at the decision boundary, it suffices for the output of the decision function for the correct class to exceed the outputs of the other decision functions by a margin of 2. These concepts are summarized in Equations (2) and (3).

$$L(w;\gamma)=\frac{1}{2}\sum_{m=1}^{K}(w_m w_m)+\mathrm{C}\sum_{i=1}^{l}\sum_{m\neq y}\gamma_i^m \quad (2)$$

$$\begin{aligned}&(w_y.\mathrm{x}_i)+\mathrm{b}_{yi}\geq(\mathrm{w}_m.\mathrm{x}_i)+b_m+2-\gamma_i^m\\&\gamma_i^m\geq 0; i=1,\ldots,\mathrm{l}\,\mathrm{m}\in\{1,\ldots,K\}\setminus y_i\end{aligned} \quad (3)$$

The class label for any testing data is the class whose decision function give the largest output value, Equation (4).

$$f(x)=\arg\underbrace{max}_{k}\ (w_k.x+b_k)\ k=1,\ldots,K \quad (4)$$

The same technique as in the basic SVM (i.e. using Lagrangian and prima dual formulation) can be used to solve the optimisation problem in the new formulation above. Using these techniques, the above problem reformulates to a QP problem of maximising Equation. (5) with respect to the linear constraints in Equation. (6). the function for labelling an unknown data point, Equation. (7) is then revised to Equation. (8).

$$W(\lambda)=2\sum_{i,m}\lambda_i^m+\sum_{i,j,m}\left[-\tfrac{1}{2}c_j^{y_i}\sum_{p=1}^{K}\lambda_i^p\sum_{q=1}^{K}\lambda_i^q+\lambda_i^m\lambda_j^{y_i}-\tfrac{1}{2}\lambda_i^m\lambda_j^m\right](\mathrm{x}_i.\mathrm{x}_j) \quad (5)$$

$$\begin{aligned}&\sum_{i=1}^{l}\lambda_i^n=\sum_{i=1}^{l}c_i^n\sum_{m=1}^{K}\lambda_i^m n=1,\ldots K\\&0\leq\lambda_i^m\leq\mathrm{C}\,\mathrm{and}\,\lambda_i^{y_i}=0;\\&i=1,\ldots,\mathrm{l}\,\mathrm{m}\in\{1,\ldots,K\}\setminus y_i\end{aligned} \quad (6)$$

Where

$$c_i^n=\begin{cases}1 & if\, y_i=n\\0 & if\, y_i\neq n\end{cases} \quad (7)$$

$$f(x)=\arg\underbrace{max}_{k}\ \sum_{i=1}^{l}(c_i^k\sum_{m=1}^{K}\lambda_i^m-\lambda_i^k)(x_i.x)+b_k \quad (8)$$

In the method highlighted by Weston and Watkins in 1999, the resolution of the values of bm for m=1,…, K is not specified. From the new formulation, it can be inferred that these values can be obtained by solving K simultaneous equations with K-free variables. The matrix derived from these sets of equations could be singular; hence, least mean squares or another optimization method may be employed to obtain the solution [27].

Analyzing the methodology and applications of SVM reveals that it is a technique with excellent response times for computational execution. Despite being a dichotomous classification method, it also possesses the potential to work with a larger number of classes across different data types. In the standard classification proposed in this work, SVM encountered difficulties in accurately classifying points that are very close to each other due to the generalization of the one- versus-all approach. However, the wide range of classification functions during the machine learning process allows SVM to mitigate this generalization issue by using a greater number of points for classification. Consequently, the higher the number of points representing a class, the greater the accuracy of classification. Overall, the patterns were classified effectively, except the digit "9." The digits "1" and "8" achieved the highest classification rates. The use of mean and variance as characteristics of the data to generate patterns proved to be the most effective method for achieving better separability between points and, therefore, improved classification, as illustrated in Figure 3.

8. Conclusion

To conclude, the existing study indicates that the accuracy and detection effectiveness of cultural relic detection and recognition, when combined with the attention mechanism and pyramid structure, significantly surpass those of other algorithms. The proposed algorithm demonstrates its feasibility and successfully meets the design

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

objectives established in this study. However, it is important to note that this method primarily focuses on identifying museum cultural relic images and incorporates additional neural networks beyond the pyramid convolutional network. Therefore, further experiments are necessary to enhance and validate the findings of this study.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer Nature, 2022.

[2] Y. Li, "Research and application of deep learning in image recognition," in *2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA)*, 2022.

[3] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, "Machine recognition of human activities: A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 11, pp. 1473-1488, 2008.

[4] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," *Electronics*, vol. 9, no. 8, p. 1188, 2020.

[5] G. Rose, *Visual Methodologies: An Introduction to Researching with Visual Materials*, 2022.

[6] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 11, pp. 3212-3232, 2019.

[7] A. Pramod, H. S. Naicker, and A. K. Tyagi, "Machine learning and deep learning: Open issues and future research directions for the next 10 years," in *Computational Analysis and Deep Learning for Medical Care: Principles, Methods, and Applications*, 2021, pp. 463-490.

[8] S. Sengupta, S. Basak, P. Saikia, S. Paul, V. Tsalavoutis, F. Atiah, V. Ravi, and A. Peters, "A review of deep learning with special emphasis on architectures, applications and recent trends," *Knowl.-Based Syst.*, vol. 194, p. 105596, 2020.

[9] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *European Conference on Computer Vision*, 2020.

[10] A. Kuznetsova, H. Rom, N. Alldrin, J. Uijlings, I. Krasin, J. Pont-Tuset, S. Kamali, S. Popov, M. Malloci, and A. Kolesnikov, "The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale," *Int. J. Comput. Vis.*, vol. 128, no. 7, pp. 1956-1981, 2020.

[11] Y. Xiao, Z. Tian, J. Yu, Y. Zhang, S. Liu, S. Du, and X. Lan, "A review of object detection based on deep learning," *Multimedia Tools Appl.*, vol. 79, pp. 23729-23791, 2020.

[12] S. Li, C. Zheng, R. Wang, Y. Huo, W. Zheng, H. Lin, and H. Bao, "Multi-resolution terrain rendering using summed-area tables," *Computers Graphics*, vol. 95, pp. 130-140, 2021.

[13] M.-S. Șerbănescu, "Fractal dimension box-counting algorithm optimization through integral images," in *7th International Conference on Advancements of Medicine and Health Care through Technology: Proceedings of MEDITECH-2020*, 2022, pp. 13-15.

[14] S. Karanwal and M. Diwakar, "EDRM-LBP: Effective directional radial median local binary pattern for face recognition," *Int. J. Embedded Syst.*, vol. 15, no. 6, pp. 475-492, 2022.

[15] Q. Wang and L. Li, "Museum relic image detection and recognition based on deep learning," *Computational Intelligence Neurosci.*, vol. 2022, p. 9670191, 2022.

[16] M. Singla and K. Shukla, "Robust statistics-based support vector machine and its variants: A survey," *Neural Comput. Appl.*, vol. 32, no. 15, pp. 11173-11194, 2020.

[17] S. Ghosh, A. Dasgupta, and A. Swetapadma, "A study on support vector machine based linear and non-linear pattern classification," in *2019 International Conference on Intelligent Sustainable Systems (ICISS)*, 2019.

[18] J. Nalepa and M. Kawulok, "Selecting training sets for support vector machines: A review," *Artif. Intell. Rev.*, vol. 52, no. 2, pp. 857-900, 2019.

[19] S. Saha, "An empirical comparison of linear and non-linear classification using support vector machines," *Int. J. Comput. Sci. Eng.*, vol. 11, pp. 120-126, 2023.

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025

[20] F. A. Fardo and P. S. S. Rodrigues, "A sparse local binary pattern extraction algorithm applied to event sensor data for object classification," *Pattern Recognit.*, vol. 146, p. 110004, 2024.

[21] B. M. Rashed and N. Popescu, "Critical analysis of the current medical image-based processing techniques for automatic disease evaluation: Systematic literature review," *Sensors*, vol. 22, no. 18, p. 7065, 2022.

[22] R. M. Anwer, F. S. Khan, J. Van De Weijer, M. Molinier, and J. Laaksonen, "Binary patterns encoded convolutional neural networks for texture recognition and remote sensing scene classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 138, pp. 74-85, 2018.

[23] C. Ranjeeth Kumar and M. Kalaiarasu, "Multiple vehicles tracking and detection using weight high order singular value decomposition dimensionality reduction and double classifiers," *J. Intell. Fuzzy Syst.*, vol. 44, no. 3, pp. 3943-3958, 2023.

[24] W. A. N. Jasim and R. J. Mohammed, "A survey on segmentation techniques for image processing," *Iraqi J. Electr. Electron. Eng.*, vol. 17, no. 2, 2021.

[25] W. Burger and M. J. Burge, *Digital Image Processing: An Algorithmic Introduction*, Springer Nature, 2022.

[26] N. Elazab, W. Gab Allah, and M. Elmogy, "Computer-aided diagnosis system for grading brain tumor using histopathology images based on color and texture features," *BMC Med. Imaging*, vol. 24, no. 1, p. 177, 2024.

[27] J. Weston and C. Watkins, "Support vector machines for multi-class pattern recognition," in *ESANN*, 1999.

DOI: https://doi.org/10.54216/FPA.200106
Received: January 30, 2025 Revised: March 27, 2025 Accepted: April 20, 2025