A Revisit Histogram of Oriented Descriptor for Facial Color Image Classification Based on Fusion of Color Information

Huy Nguyen-Quoc and Vinh Truong Hoang

Faculty of Information Technology, Ho Chi Minh City Open University, 97 Vo Van Tan, District 3, Ho Chi Minh City, Vietnam

Correspondence should be addressed to Vinh Truong Hoang; vinh.th@ou.edu.vn

Received 30 August 2021; Revised 27 October 2021; Accepted 18 November 2021; Published 30 November 2021

Academic Editor: Ying-Ren Chien

Copyright © 2021 Huy Nguyen-Quoc and Vinh Truong Hoang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Histogram of Oriented Gradient (HOG) is a robust descriptor which is widely used in many real-life applications, including human detection, face recognition, object counting, and video surveillance. In order to extract HOG descriptor from color images whose information is three times more than the grayscale images, researchers currently apply the maximum magnitude selection method. This method makes the information of the resulted image is reduced by selecting the maximum magnitudes. However, after we extract HOG using the unselected magnitudes of the maximum magnitude selection method, we observe that the performance is better than using the maximum magnitudes in several cases. Therefore, in this paper, we propose a novel approach for extracting HOG from color images such as Color Component Selection and Color Component Fusion. We also propose the extended kernels in order to improve the performance of HOG. With our new approaches in the color component analysis, the experimental results of several facial benchmark datasets are enhanced with the increment from 3 to 10% of accuracy. Specifically, a 95.92% of precision is achieved on the Face AR database and 75% on the Georgia Face database. The results are better more than 10 times compared with the original HOG approach.

1. Introduction

Nowadays, image classification is one of the most extensive fields in computer vision which attracts the attention of many researchers because of its wide range of application in real life such as human detection, facial recognition, object classification, and diagnose diseases in medical. Many local and global image descriptors have been proposed in order to handle this task [1–3]. The key step is to find a robust descriptor which can discriminate classes. A large number of descriptors have been proposed to extract feature from image, including Color Local Binary Pattern (LBP) [4], Scale Invariant Feature Transform (SIFT) [5], Histogram of Oriented Gradients (HOGs) [6], and GIST [7]. Among them, HOG is a successful descriptor with various applications in real life, including pedestrian detection, face recognition, object classification, security, and industrial inspection. For example, Ding et al. [8] fuse HOG features and global normalized histogram for human detection task by the AdaBoost classifier. Qi et al. [9] apply HOG for railway track detection by using region-growing methods. Qingbo et al. [10] combine HOG features and Discriminative Multimaniifold Analysis Method (DMAM) for face recognition in few shot learning context. In this approach, facial features were extracted from image patches fusing and then applied DMAM to transform to lower dimensional space. Chowdhury et al. [11] introduced an improved version of HOG for human detection. Nabila et al. [12] present an optimized version of HOG for road car detection. This approach is based on the concatenation of shape features and motion integration. Aytac et al. [13] extract LBP and HOG features from stomach cancer images. Then, multiple reduction techniques are applied to select the most useful attributes. Hmood et al. [14] introduce an improved version of HOG by proposing an approximate windows that can cover whole object, namely, Dynamic-HOG. This method requires less processing time and achieves a higher accuracy for coin classification. Jebril et al. [15] apply HOG for
Various advancement versions of HOG are proposed in recent years by incorporating with convolutional neural networks (CNNs). Zaffar et al. [19] present CoHOG which is based on CNN in order to extract features from ROI for visual place recognition. Xiong et al. [20] introduce a method for handling depth information of images based on HOG, namely, Histogram of Oriented Depth (HOD) features. The proposed approach is applied for pedestrian detection by combining color image edge information and HOD. Wang et al. [21] present an approach based on HOG and multiorientation computation, namely, MO-HOG. This

handwritten character recognition from grayscale images transformed from RGB color space. The authors used both R-HOG and C-HOG with different windows. Uddin et al. create another HOG version named T20-HOG with the ability to extract the textural features from seed varieties for identification [16]. A two-stage classification process is proposed by Chandrakala and Devi in combination with HOG descriptors [17]. With the robustness of deep neural networks and the advantage of HOG, Hung created a hybrid combination of HOG and CNN named HOG-CNN approach with promising results [18].
method combines two-part gradient information to build MO-HOG and applied for facial expression recognition. Pan [22] combines HOG and CNN for facial expression classification. In this approach, HOG features are extracted from CNN’s shallow features. Joshi et al. [23] present a method for selecting optimal parameter for HOG descriptor. A multi-level HOG based on sign language image is applied for several windows size of 30 × 30 and 20 × 20. Wang et al. [24] concatenate HOG and the Zernike moment’s features of SAR images for terrain matching.

Furthermore, HOG has outperformed other descriptors in the classification stage for grayscale image. However,
when we extract HOG from color image, the performance usually degrades seriously. The reason is because the information that the color image contains is three times more than the grayscale image does. Moreover, most descriptors are first designed to perform on grayscale images only. Therefore, color has been well investigated in recent years for extracting HOG features. For example, Hoang et al. [25] extract local image descriptors, including LBP, HOG, and GIST for an application of rice seed image recognition. In this approach, features extracted from independent color component are then fused to form final feature vector. Aslan et al. [26] compare HOG-SVM and CNN for human tracking based on video in occlusion context. Zhou et al. [27] introduce a method for extracting HOG features based on

\[
\begin{align*}
R & = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & -1 & -1 & -1 & -1 \\
3 & 2 & 1 & & & & & & 
\end{bmatrix} \\
R & = \begin{bmatrix}
1 & -1 & -1 & -1 & 0 & 1 & & & \\
-1 & & & & & & & & \\
-1 & & & & & & & & \\
-1 & & & & & & & & 
\end{bmatrix}
\end{align*}
\]

Figure 3: An illustration of original kernels and extended kernels.

(a) AR dataset

(b) Georgia dataset

(c) CLV dataset

(d) MUCT dataset

Figure 4: The four datasets used to evaluate the proposed approach.
a color filter array. Each Bayer pattern is then applied to convert into three component color images by downsampling operation. A new extension of bin was added to HOG and extracted in HSI color space for removing shadow region. Nhat and Hoang [28] enhance the accuracy of face recognition by the fusion of three popular features (LBP, HOG, and GIST). Duong and Hoang [29] apply to extract rice seed images based on features coded in multiple color spaces using HOG descriptor. Banerji et al. [30] enhance the HOG descriptor by subjecting the image to the Haar wavelet transform to improve the HOG descriptor for image classification. Recently, Z. Zhou and Y. Zhou [31] handle color space information for extracting features, namely, CCS-HOG. They apply a between component color space similarity, combining with HOG for extracting facial features. This approach has been shown its effectiveness in kinship verification. Van and Hoang extract HOG features for coding facial image in different color spaces [32, 33]. The final features are obtained by concatenating all features extracted from each color component. Fekri-Ershad and Tajeripour analyzed the color information of color-texture images for classification using hybrid color LBP [34]. Color-texture can also be analyzed using weighted color order of LBP for classification [35].

Although the current researchers have applied maximum magnitude selection method in order to selectively...
reduce the information of color image so it can meet the requirements of the following stage of HOG extraction process, the performance is not completely optimized. The impact of color component for extracting features is first mentioned in [36] for LBP descriptors. This issue is extensively investigated to incorporate color information by various proposed method in recent years. It can be used by fusing features extracted from each color component independently or jointly [4]. Specifically, better performances are achieved when we use the unselected magnitudes of maximum magnitude selection method. Novel approaches for extracting HOG from color image are Color Component Selection and Color Component Fusion. Furthermore, we also upgrade the kernels in the gradient computation stage by extending it in horizontal and vertical dimensions. The intention is to figure out the connection between surrounding pixels and the computing pixel, whether the surrounding pixels cause any effects on the output performance.

The rest of this paper is organized as follows. Section 2.1 and Section 3 introduce HOG descriptor with Color Component Selection, Color Component Fusion, and the extended kernels. Then, experimental results are presented in Section 4. Finally, the conclusion and future works are discussed in Section 5.

### 2. HOG Descriptor with Color Component Selection and Color Component Fusion

This section briefly introduces HOG computation and the two proposed approaches with Color Component Selection (CCS) and Color Component Fusion (CCF) for HOG descriptor.

#### 2.1. HOG Descriptor

Before extracting HOG feature, an image \( I \) is split into three subimages \( I_{C1}, I_{C2}, \) and \( I_{C3} \) which are three color components of \( I \). Next, several image processing algorithms are applied on these images in order to reduce noisy for enhancing the performance. After preprocessing step, on each image, gradient magnitude and direction of each pixel are computed by using the horizontal and vertical gradients. The gradient computation of pixel located at coordinate \((x, y)\) is defined as follows [6]:

\[
\Delta_x = |G_{(x-1,y)} - G_{(x+1,y)}|, \quad (1)
\]

\[
\Delta_y = |G_{(x,y-1)} - G_{(x,y+1)}|, \quad (2)
\]

\[
M_{(x,y)} = \sqrt{\Delta_x^2 + \Delta_y^2}, \quad (3)
\]

**Figure 6:** Different values of extended kernels for HOG by CCS and CCF on Georgia dataset.
\[ \alpha_{(x,y)} = \arctan \left( \frac{\Delta_y}{\Delta_x} \right) \]  

(4)

where \( G \) is grayscale value of the computing pixel, \( \Delta_x \) and \( \Delta_y \) represent horizontal and vertical gradients, and \( M_{(x,y)} \) and \( \alpha_{(x,y)} \) sequentially define gradient magnitude and gradient direction.

2.2. Proposed Approaches. Each image then results in a pair of matrices which contains one gradient magnitude matrix and one the gradient direction matrix. So there are three achieved pairs of matrices in total which are \( M_{C1} \) and \( \alpha_{C1} \), \( M_{C2} \) and \( \alpha_{C2} \), and \( M_{C3} \) and \( \alpha_{C3} \). As the following step requires only one pair of matrices, the current task is to figure out which pair of matrices should be selected in order to produce an optimal feature vector. The most popular solution is maximum magnitude selection. This method compares three gradient magnitudes in each pixel. Then, the maximum is chosen for the final magnitude of this pixel. The final direction is also extracted from the same color component with the selected magnitude. The detail of maximum magnitude selection method is illustrated in Figure 1.

After selection step, two final matrices are obtained to meet the requirement of the orientation binning step. In this step, the original image is divided into cells (8 x 8 per cell) and a 9-bin histogram is built in each cell based on the gradient features of the inner pixels. The bins are ranged from 0 to 180 degrees for unsigned gradient \((\alpha_{\text{unsigned}})\) and from 0 to 360 degrees in case of signed gradient \((\alpha_{\text{signed}})\). Gradient magnitude of each pixel is added into the corresponding index number of a bin in the histogram. The index number \(B_{\text{idx}(x,y)}\) is then computed by Equation (5) or (6) and used the ceiling value.

\[
B_{\text{idx}(x,y)} = \left\lceil \frac{\alpha_{(x,y)} \alpha_{\text{unsigned}}}{B_{\text{num}}} \right\rceil \quad (\alpha_{\text{unsigned}} \in (0, 180], B_{\text{num}} = 9),
\]

(5)

or

\[
B_{\text{idx}(x,y)} = \left\lceil \frac{\alpha_{(x,y)} \alpha_{\text{signed}}}{B_{\text{num}}} \right\rceil \quad (\alpha_{\text{signed}} \in (0, 360], B_{\text{num}} = 9),
\]

(6)

where \(B_{\text{num}}\) stands for the default number of orientation bins of the histogram and is usually set to 9 by default. \(\alpha_{(x,y)}\) is computed above by Equation (4).
For the normalization stage, the original image is then divided into blocks (each block contains 2×2 cells which equal to 256×256 pixels). An effective normalization is able to reduce noise and cancel the ill effects. In each block, 50% of the surrounding blocks are overlapped so each cell can be normalized more than once except the cells locate in the corner. Histograms of each block are concatenated and then normalized by using L1-norm, L2-norm, or L1-sqrt. Finally, all normalized histograms are combined together as a feature vector.

In practice, we observe that the maximum magnitude selection method is not optimal enough as it may cause information loss. When we use the unselected values of the maximum magnitude selection for HOG extraction, the achieved results are higher than use the maximum magnitude in several experiments. Therefore, we apply Color Component Selection and Color Component Fusion methods to improve the performance. For the Color Component Selection method, instead of considering which magnitude or direction should be selected for the next stage, we choose the final pair of matrices based on the color components. There are three matrices of each kind so nine different pairs of matrices are achieved in total. The orientation binning stage sequentially applies these pairs to proceed. In the end, nine feature vectors are obtained which are corresponding to nine selected pairs. These vectors are then evaluated to find out the best performance. The process of the Color Component Selection method is presented in Figure 2. The Color Component Fusion simply takes the achieved vectors and concatenates all of them into a fusion vector.

### 3. Extended Kernels for HOG Extraction

In order to speed the gradient computation process up, researchers used kernels which are illustrated in Figure 3. By filtering the image with these kernels, the horizontal gradient \( \Delta_x \) and vertical gradient \( \Delta_y \) are computed faster than calculating based on Equations (1) and (2). We define these original kernels are kernels with \( R = 1 \). As we sequentially extend the kernels through horizontal and vertical dimension, the parameter \( R \) also increases. Figure 3 describes specifically the kernels which parameter \( R \) in range from 1 to 4. Then, the achieved horizontal and vertical gradients are divided by parameter \( R \).

### 4. Experiments

#### 4.1. Dataset Description.

The proposed approaches are evaluated on four common facial databases, including AR,
Georgia, CLV, and MUCT (see Figure 4). In order to experiment the proposed approaches, these databases are split into 50% of training and 50% of testing. However, databases, such as Georgia, CLV, and MUCT, have an odd number of images per class. For instance, the MUCT dataset includes 3 images in each class so we decide to randomly split one image for training and the other two for testing in order to make the training more challenging. The 1-NN classifier is employed to evaluate the classification performance, and accuracy metric is employed to measure the performance on testing set. The summary of those databases is presented in Table 1.

4.2. Experimental Setup. As our proposed approaches are designed to fit in every three-component color images, we also experiment them in several well-known color spaces, including HSV, ISH, I1I2I3, and YCbCr. These spaces are frequently applied for pattern recognition [4]. According to Section 2, each image after applying Color Component Selection results in 10 different feature vectors, including 9 different vectors and one fusion vector which is extracted by the Color Component Fusion method. These results are then deployed to a 1-NN classifier for evaluating. Moreover, we set the kernels with R from 1 to 5, respectively, for comparison. The experiments are implemented by Matlab 2017b and conducted on a PC with a configuration of a CPU I3 8100 3.60 GHz and 8 GBs of RAM.

4.3. Results. The experimental results are presented by charts in Figures 5–8. By converting in many color spaces and increasing parameter R, the performance of maximum magnitude selection method has been upgraded. Furthermore, the Color Component Selection and the Color Component Fusion have also outperformed the maximum magnitude selection method according to the charts. For the AR database, the highest achieved accuracy is 95.92% which is 0.54% more than the highest accuracy can be achieved by using the maximum magnitude selection method, and this result is obtained by the Color Component Fusion with R = 5 kernels, and the used color space is ISH. For the Georgia database, when we apply the Color Component Fusion method combined with R = 1 kernels and YCbCr color space, the achieved result is 75.00%. This is the highest accuracy in all experiments on the Georgia database. Better performance is also obtained in experiment for the MUCT database and even in experiment with the CLV database.

| Color space | AR | Max magnitude | C1 | C2 | C3 | Fusion |
|-------------|----|---------------|----|----|----|--------|
| RGB         | 1  | 92.69         | 87.00 | 91.23 | 89.77 | 85.46 | 90.62 |
|             | 2  | 93.46         | 88.46 | 93.15 | 92.92 | 87.85 | 92.77 |
|             | 4  | 94.54         | 87.77 | 93.85 | 93.69 | 88.85 | 93.69 |
|             | 5  | 94.92         | 98.15 | 94.38 | 94.38 | 90.08 | 94.54 |
|             | 1  | 94.38         | 97.30 | 93.69 | 94.46 | 89.23 | 93.77 |
|             | 2  | 92.62         | 84.69 | 94.00 | 74.85 | 82.77 | 86.31 |
| HSV         | 3  | 94.46         | 88.15 | 92.46 | 86.38 | 87.54 | 92.15 |
|             | 4  | 94.54         | 97.77 | 93.69 | 91.31 | 87.54 | 93.69 |
|             | 5  | 94.69         | 92.00 | 94.77 | 87.92 | 87.62 | 94.15 |
|             | 1  | 92.08         | 92.77 | 82.08 | 92.23 | 88.69 |
|             | 2  | 94.77         | 94.54 | 88.15 | 81.77 | 92.27 | 92.85 |
| ISH         | 3  | 94.85         | 97.77 | 92.15 | 85.85 | 93.85 | 94.23 |
|             | 4  | 95.38         | 94.46 | 93.69 | 87.23 | 94.00 | 95.08 |
|             | 5  | 94.69         | 94.62 | 92.43 | 87.69 | 94.08 | 95.92 |
|             | 1  | 92.77         | 93.28 | 86.15 | 85.54 | 93.15 | 91.31 |
|             | 2  | 94.00         | 94.15 | 90.15 | 91.00 | 94.85 | 93.77 |
| YCbCr       | 3  | 94.54         | 94.15 | 91.54 | 92.31 | 95.15 | 94.31 |
|             | 4  | 94.77         | 94.38 | 92.32 | 92.54 | 94.92 | 94.69 |
|             | 5  | 94.54         | 94.08 | 90.77 | 91.69 | 94.92 | 94.62 |
|             | 1  | 93.08         | 93.15 | 89.62 | 78.23 | 93.31 | 92.00 |
|             | 2  | 93.77         | 94.46 | 93.08 | 87.31 | 95.69 | 94.23 |
| I1I2I3      | 3  | 94.46         | 94.69 | 93.31 | 89.38 | 95.08 | 95.38 |
|             | 4  | 94.54         | 94.62 | 93.00 | 90.31 | 95.31 | 95.62 |
|             | 5  | 94.62         | 94.85 | 93.00 | 91.54 | 95.31 | 95.69 |

| Color space | Georgia | Max magnitude | C1 | C2 | C3 | Fusion |
|-------------|----------|---------------|----|----|----|--------|
| RGB         | 1  | 65.75         | 67.00 | 65.75 | 66.50 | 66.25 | 66.50 |
|             | 2  | 66.00         | 67.50 | 66.50 | 68.25 | 68.25 | 67.75 |
|             | 4  | 67.00         | 68.75 | 69.25 | 68.75 | 68.50 | 69.00 |
|             | 5  | 67.00         | 68.50 | 67.50 | 68.25 | 69.25 | 68.00 |
| HSV         | 3  | 64.50         | 65.75 | 65.00 | 64.75 | 65.75 | 69.00 |
|             | 4  | 65.00         | 65.75 | 56.25 | 65.25 | 58.25 | 71.25 |
|             | 5  | 64.25         | 68.25 | 58.75 | 68.00 | 62.50 | 72.75 |
| ISH         | 3  | 64.75         | 66.00 | 58.75 | 65.25 | 54.00 | 69.50 |
|             | 4  | 65.00         | 65.75 | 55.75 | 66.50 | 56.25 | 70.50 |
|             | 5  | 65.75         | 67.75 | 56.75 | 65.50 | 59.25 | 72.50 |
| YCbCr       | 3  | 68.25         | 65.50 | 63.75 | 65.00 | 65.00 | 72.50 |
|             | 4  | 69.00         | 65.00 | 64.50 | 64.25 | 63.50 | 71.00 |
|             | 5  | 69.00         | 65.75 | 63.50 | 63.75 | 64.75 | 70.00 |
| I1I2I3      | 3  | 68.50         | 65.50 | 59.50 | 64.25 | 64.75 | 70.00 |
|             | 4  | 68.50         | 65.00 | 59.75 | 63.75 | 64.00 | 70.50 |
|             | 5  | 68.75         | 65.50 | 60.00 | 64.00 | 63.50 | 69.75 |

Table 2: Several experimental cases on the AR database.

Table 3: Several experimental cases on the Georgia database.
which is the most challenging image set. Generally, we observe that the best results are mostly achieved with Fusion approach in comparing with other approaches, which can tell the promising performance of our proposed approaches.

Several of best cases are reported in Tables 2–5. In these tables, the Max magnitude abbreviation stands for maximum magnitude selection method while Fusion abbreviation stands for Color Component Fusion method. We found that the best accuracy obtained for AR dataset is 95.92% by using Fusion approach. Similarly, this approach achieves 75.00%, 49.77%, and 93.47% for Georgia, CLV, and MUCT datasets, respectively. Note that the performance of each color space is different according to the feature extraction methods. Moreover, the increment of parameter \( R \) in the extended kernels has also improved several experimental results. In most cases, the accuracy when applying \( R=5 \) kernels is higher compared to the others. However, there are also several cases where highest precision is achieved with \( R=1 \) so we cannot yet tell if the higher the parameter \( R \) of kernels, the higher accuracy we can achieve. But we believe that these results are the foundation of the advantageous effectiveness of kernel size in HOG performance, which leads us to study deeper about it in the future.

5. Conclusion

In this paper, we propose novel approaches to extract HOG descriptor from color images such as the Color Component Selection method and Color Component Fusion in order to improve the classification performance. In our observation, we recognize that the proposed methods outperform the current maximum magnitude selection method in face classification task, especially the Color Component Fusion method. The color space conversion and extended kernels also efficiently improve the accuracy of the classification. However, in several cases, the extended kernels still cause the accuracy to decrease. The Color Component Selection method requires a long time to finish the extraction from every available case, and the Color Component Fusion may cause computer memory issues due to its deep dimension. Therefore, our future work is to figure out the optimal pair of matrices without experimenting all cases to enhance the inference speed. Feature selection methods are also recommended to reduce the vector dimension. Several CNN-
related works are considered to combine with the proposed approaches for further improvement.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

We have no conflict of interest to declare.

References

[1] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey,” IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 41, no. 6, pp. 765–781, 2011.

[2] A. Humeau-Heurtier, “Texture feature extraction methods: a survey,” IEEE Access, vol. 7, pp. 8975–9000, 2019.

[3] S. Fekri-Ershad, "Gender classification in human face images for smart phone applications based on local texture information and evaluated Kullback-Leibler divergence," Traitement du Signal, vol. 36, no. 6, pp. 507–514, 2019.

[4] V. T. Hoang, Multi Color Space LBP-Based Feature Selection for Texture Classification, [Ph. D thesis], Université du Littoral C’ote d’Opale, 2018.

[5] L. Zhang, J. Chen, L. Yue, and P. Wang, "Face recognition using scale invariant feature transform and support vector machine,” in 2008 The 9th International Conference for Young Computer Scientists, pp. 1766–1770, Hunan, China, 2008.

[6] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection,” in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), vol. 1, pp. 886–893, San Diego, CA, USA, 2005.

[7] M. Douze, H. Jégou, H. Sandhawalia, L. Amsaleg, and C. Schmid, "Evaluation of gist descriptors for web-scale image search,” in Proceedings of the ACM International Conference on Image and Video Retrieval, pp. 1–8, New York, NY, USA, 2009.

[8] J. Ding, Y. Wang, and W. Geng, "An HOG-CT human detector with histogram-based search,” Multimedia Tools and Applications, vol. 63, no. 3, pp. 791–807, 2013.

[9] Z. Qi, Y. Tian, and Y. Shi, "Efficient railway tracks detection and turnouts recognition method using HOG features,” Neural Computing and Applications, vol. 23, no. 1, pp. 245–254, 2013.

[10] J. Qingbo, Z. Enze, Y. Xinqi, X. Yu, and L. Yun, "Face recognition method based on HOG and DDMA from single training sample,” Multimedia Tools and Applications, vol. 75, no. 21, pp. 13163–13177, 2016.

[11] S. A. Chowdhury, M. Mir, S. Kowsar, and K. Deb, "Human detection utilizing adaptive background mixture models and improved histogram of oriented gradients,” ICT Express, vol. 4, no. 4, pp. 216–220, 2018.

[12] M. Nabila, B. J. Yousra, and W. Eric, "Optimized HOG descriptor for on road cars detection,” in Proceedings of the 10th International Conference on Distributed Smart Camera-ICDSC’16, pp. 166–171, Paris, France, 2016.

[13] S. Aytaç Korkmaz and H. Binol, "Classification of molecular structure images by using ANN, RF, LBP, HOG, and size reduction methods for early stomach cancer detection,” Journal of Molecular Structure, vol. 1156, pp. 255–263, 2018.

[14] A. K. Hmood, C. Y. Suen, and L. Lam, "An enhanced histogram of oriented gradient descriptor for numismatic applications,” Pattern Recognition and Image Analysis, vol. 28, no. 4, pp. 569–587, 2018.

[15] N. A. Jebir, H. R. Al-Zoubi, and Q. A. Al-Hajja, "Recognition of handwritten Arabic characters using histograms of oriented gradient (HOG)," Pattern Recognition and Image Analysis, vol. 28, no. 2, pp. 321–345, 2018.

[16] M. Uddin, M. A. Islam, M. Shajalal, M. A. Hossain, M. Yousuf, and S. Iftekhar, "Paddy seed variety identification using t20-hog and haralick textural features,” in Complex & Intelligent Systems, pp. 1–15, Springer, 2021.

[17] M. Chandrakala and P. D. Devi, “Two-stage classifier for face recognition using hog features,” Materials Today: Proceedings, vol. 47, pp. 5771–5775, 2021.

[18] B. T. Hung, "Face recognition using hybrid HOGCNN approach,” Research in Intelligent and Computing in Engineering: Select Proceedings of RICE 2020, , pp. 715–723, Springer, 2021.

[19] M. Zaffar, S. Ehsan, M. Milford, and K. McDonald-Maier, "CoHOG: a light-weight, compute-efficient, and training-free visual place recognition technique for changing environments,” IEEE Robotics and Automation Letters, vol. 5, no. 2, pp. 1835–1842, 2020.

[20] X. Zhang, H. Shangguan, A. Ning, A. Wang, J. Zhang, and S. Peng, "Pedestrian detection with EDGE features of color image and HOG on depth images,” Automatic Control and Computer Sciences, vol. 54, no. 2, pp. 168–178, 2020.

[21] H. Wang, S. Wei, and B. Fang, "Facial expression recognition using iterative fusion of MO-HOG and deep features,” The Journal of Supercomputing, vol. 76, no. 5, pp. 3211–3221, 2020.

[22] X. Pan, "Fusing HOG and convolutional neural network spatial–temporal features for video-based facial expression recognition,” IET Image Processing, vol. 14, no. 1, pp. 176–182, 2020.

[23] G. Joshi, S. Singh, and R. Vig, "Taguchi-TOPSIS based HOG parameter selection for complex background sign language recognition,” Journal of Visual Communication and Image Representation, vol. 71, 2020.

[24] K. Wang, H. Wang, and J. Wang, "Terrain matching by fusing HOG with Zernike moments,” IEEE Transactions on Aerospace and Electronic Systems, vol. 56, no. 2, pp. 1290–1300, 2020.

[25] V. T. Hoang, D. P. van Hoai, T. Surinwarangkoon, H.-T. Duong, and K. Meethongjan, "A comparative study of rice variety classification based on deep learning and hand-crafted features,” ECTI Transactions on Computer and Information Technology (ECTI-ICT), vol. 14, no. 1, pp. 1–10, 2020.

[26] M. F. Aslan, A. Durdu, K. Sabanci, and M. A. Mutluer, "CNN and HOG based comparison study for complete occlusion handling in human tracking," Measurement, vol. 158, 2020.

[27] W. Zhou, S. Gao, L. Zhang, and X. Lou, "Histogram of oriented gradients feature extraction from raw Bayer pattern images,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 67, no. 5, pp. 946–950, 2020.

[28] H. T. M. Nhat and V. T. Hoang, "Feature fusion by using LBP, HOG, GIST descriptors and Canonical Correlation Analysis for face recognition,” in 2019 26th International Conference on Telecommunications (ICT), pp. 371–375, Hanoi, Vietnam, 2019.
[29] H.-T. Duong and V. T. Hoang, “Dimensionality reduction based on feature selection for rice varieties recognition,” in *2019 4th International Conference on Information Technology (InCIT)*, pp. 199–202, Bangkok, Thailand, 2019.

[30] S. Banerji, A. Sinha, and C. Liu, “Haarhog: Improving the hog descriptor for image classification,” in *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 4276–4281, Manchester, UK, 2013.

[31] Z. Zhou and Y. Zhou, “Cross-channel similarity based histograms of oriented gradients for color images,” in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pp. 1621–1625, Bari, Italy, 2019.

[32] T. N. Van and V. T. Hoang, “Early and late features fusion for kinship verification based on constraint selection,” in *2019 25th Asia-Pacific Conference on Communications (APCC)*, pp. 116–121, Ho Chi Minh City, Vietnam, 2019.

[33] T. N. Van and V. T. Hoang, “Kinship verification based on local binary pattern features coding in different color space,” in *2019 26th International Conference on Telecommunications (ICT)*, pp. 376–380, Hanoi, Vietnam, 2019.

[34] S. Fekri-Ershad and F. Tajeripour, “Impulse-noise resistant color-texture classification approach using hybrid color local binary patterns and Kullback-Leibler divergence,” *The Computer Journal*, vol. 60, no. 11, pp. 1633–1648, 2017.

[35] T. Song, J. Feng, S. Wang, and Y. Xie, “Spatially weighted order binary pattern for color texture classification,” *Expert Systems with Applications*, vol. 147, 2020.

[36] T. Mäenpää and M. Pietikäinen, “Classification with color and texture: jointly or separately?,” *Pattern Recognition*, vol. 37, no. 8, pp. 1629–1640, 2004.