A CONVOLUTION-FREE LBP-HOG DESCRIPTOR FOR MAMMOGRAM CLASSIFICATION

Zainab Alhakeem and Se-In Jang

School of Electrical & Electronic Engineering, Yonsei University, Seoul, Korea 03722
Email: {zrkhakim and sijang}@yonsei.ac.kr

ABSTRACT
In image based feature descriptor design, an iterative scanning process utilizing the convolution operation is often adopted to extract local information of the image pixels. In this paper, we propose a convolution-free Local Binary Pattern (CF-LBP) and a convolution-free Histogram of Oriented Gradients (CF-HOG) descriptors in matrix form for mammogram classification. An integrated form of CF-LBP and CF-HOG, CF-LBP-HOG, is subsequently constructed in a single matrix formulation. The proposed descriptors are evaluated using a publicly available mammogram database. The results show promising performance in terms of classification accuracy and computational efficiency.

Index Terms— Convolution Free, Local Descriptor, Local Binary Pattern, Histogram of Oriented Gradients, Mammogram Classification

1. INTRODUCTION
Breast cancer is the most common and leading cause of cancer death in women [1]. For breast cancer diagnosis, mammography, which uses X-rays to examine the human breast, is usually used for early diagnosis. However, even for skilled radiologists, the mammogram diagnosis between benign and malignant is difficult due to their similarity in the target shape and texture. In order to reduce the human bias as well as to automate the process, ongoing effort has been paid to developing Computer-Aided Diagnosis (CAD) systems for mammogram classification. Typically, these systems rely on extraction of the texture and the shape information from the radio images [2], [3]. The mammogram classification performance depends heavily on precise segmentations of the target tumors (i.e., complex lesion boundary of the tumors) for reliable features. To avoid this dependency, the texture information obtained from local descriptors such as the LBP [4], the Local Binary Convolution (LBC) [5] and the HOG [6] are adopted [7], [8].

Our motivation for mammogram classification consists of the following: (i) The shape based feature extraction can be significantly affected by the irregular lesion boundaries. (ii) The texture based feature extraction such as LBP, HOG and LBC typically use the convolution based iterative scanning process that is time-consuming and often produces a large feature dimension. (iii) To utilize the diversity of different local descriptors for performance improvement.

The main contributions of this work are as follows: (i) Matrix formulations of two descriptors, namely the LBP and the HOG, to avoid the iterative convolution process in local computation and to get away from the high dimensional feature size. (ii) An integrated formulation of the LBP and the HOG under linear matrix products.

The remainder of the paper is organized as follows: Section 2 includes a brief review of the LBC descriptor, the DMP (Difference Matrix Projection) method and the total-error-rate based classifier. In Section 3, we propose two convolution-free descriptors in matrix formulations. Then we construct an integrated matrix formulation of these two descriptors. Section 4 presents our experiments, observations and evaluations of the proposed method using a public mammogram database. Concluding remarks are given in Section 5.

2. PRELIMINARIES

2.1. Local Binary Convolution network (LBC)
The Local Binary Pattern (LBP) [4] is a simple and popular descriptor adopted in many applications (e.g., face and palm print recognitions [9] [10] [11]). The LBP scans each central pixel of an image and its local neighborhood pixels (P) within an odd size window determined by R (e.g., R = 1 indicates a 3 x 3 window and R = 2 indicates a 5 x 5 window). The computed output of the LBP can be expressed as $LBP_{p,r} = \sum_{i=0}^{P-1} s(x_i - x_c)2^i$, where $x_c$ and $x_i$ are respectively the intensity values of the center pixel and the neighborhood pixels. $s(\cdot)$ is the thresholding operation given by $s(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0 \end{cases}$. Frequently, for images with high enough resolution, the, P and R are set at $P = 8$ and $R = 1$.

Based on the LBP, a Local Binary Convolution network (LBC) was proposed in [5] to encode the LBP efficiently utilizing the sparse convolutional filters. The LBC was developed as an alternative to the standard Convolutional Neural Networks (CNN) layer to reduce its complexity by reducing the number of learnable parameters. For a $3 \times 3$ window size, the LBC uses a weighted sum of eight sparse convolutional filters.
filters to reformulate in [5] as the basis of LBC as follows:

\[
LBC_{PR} = \sum_{i=1}^{P} \sigma(b_i \ast x) \cdot v_i, \quad i = \{1, 2, \ldots, P\},
\]

where \( \ast \) is the convolution operation, and \( b_i \) is the sparse convolutional filter with two non-zero values \( \{+1, -1\} \) which convolves with the vectorized input image \( x \), \( \sigma \) is the Heaviside step function. This function \( \sigma \) can be replaced by a Sigmoid or a ReLU activation function for differentiability. \( v = [2^{p-1}, 2^{p-2}, \ldots, 2^0]^T \) contains predefined weight values. \( P \) is the number of neighbor pixels (e.g., \( P = 8 \) with \( 3 \times 3 \) window).

2.2. Difference Matrix Projection (DMP)

The Difference Matrix Projection (DMP) was proposed for pedestrian detection in [12]. Compared with the pixel-wise calculations of the Histogram of Oriented Gradients (HOG) [6] based on the first order gradients, the DMP realized an approximated HOG based on linear matrix products utilizing both the first and the second order gradients with pre-calculated projection matrices. The DMP process uses a cell-based pooling to substitute the histogram construction.

The DMP process can be written as:

\[
F_i = L_i \cdot D_{HOG,i} \cdot R_i, \quad i = \{1, 2, \ldots, 8\},
\]

where

\[
L_i = \begin{bmatrix}
1_{\text{xz}} & 0 & \cdots \\
0 & \ddots & \cdots \\
\vdots & \cdots & 1_{\text{xz}}
\end{bmatrix}, \quad R_i = \begin{bmatrix}
1_{\text{zx}} & 0 & \cdots \\
0 & \ddots & \cdots \\
\vdots & \cdots & 1_{\text{zx}}
\end{bmatrix}^T
\]

are the predefined projection matrices for cell-based non-overlapping pooling, and \( c \) is the cell size. Based on the first and the second order gradients from four orientations (0°, 45°, 90° and 135°), the gradient \( D_{HOG,i} \in \mathbb{R}^{M \times N} \) can be expressed as:

\[
\begin{align*}
D_{HOG,1} &= X(I - H_1), \\
D_{HOG,2} &= X - V_1X_1H_1, \\
D_{HOG,3} &= (I - V_1)X, \\
D_{HOG,4} &= XH_1 - V_1X, \\
D_{HOG,5} &= X(I - H_2), \\
D_{HOG,6} &= X - V_2X_2H_2, \\
D_{HOG,7} &= (I - V_2)X, \\
D_{HOG,8} &= XH_2 - V_2X,
\end{align*}
\]

where \( D_{HOG,1} \) to \( D_{HOG,4} \) are the first-order gradients and \( D_{HOG,5} \) to \( D_{HOG,8} \) are the second-order gradients. \( X \in \mathbb{R}^{M \times N} \) is the input image. The predefined horizontal and vertical shifting matrices are given respectively by

\[
H_l = \begin{bmatrix}
0 & I & 0 \\
I & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} \in \mathbb{R}^{N \times N},
\]

\[
V_l = \begin{bmatrix}
0 & I & 0 \\
0 & 0 & 0 \\
I & 0 & 0
\end{bmatrix} \in \mathbb{R}^{M \times M},
\]

where \( I \) is the identity matrix. \( l \) indicates the number of shifting pixels which determine the first and the second order gradients when \( l = 1 \) and \( l = 2 \) respectively. Similar to HOG in [9], the block normalization is subsequently performed on the cell-based pooling \( F_i \) to obtain the final DMP features.

2.3. Total-Error-Rate (TER) Minimization

The Total-Error-Rate (TER) based classification [13] utilized the sum of the type I and type II errors. The type I error (also known as False Positive Rate (FPR)) is the ratio of falsely recognized positive samples over the negative sample size given by \( \text{FPR} = \frac{P^-}{n} \). The type II error (also known as False Negative Rate (FNR)) is the ratio of falsely recognized negative samples over the positive sample size given by \( \text{FNR} = \frac{FN}{n} \). Then, TER can be written as \( \text{TER} = \text{FNR} + \text{FPR} \). For a classifier which is linear in its parameters, the TER parameters can be optimally determined by

\[
\theta = (P^T W \lambda + \lambda I)^{-1} P^T W y,
\]

where \( P = \begin{bmatrix} P^- \quad P^+ \end{bmatrix} \in \mathbb{R}^{n \times d}, P^- \in \mathbb{R}^{n^- \times d} \) and \( P^+ \in \mathbb{R}^{n^+ \times d} \) are respectively the transformed samples of each category, \( n = n^+ + n^- \) is the total number of samples and \( W = W^- + W^+ \in \mathbb{R}^{n \times n} \) is a class-specific weight matrix where \( W^- = \text{diag}(\frac{1}{n^-}, \ldots, \frac{1}{n^-}, 0, \ldots, 0) \), \( W^+ = \text{diag}(0, \ldots, 0, \frac{1}{n^+}, \ldots, \frac{1}{n^+}) \).

\[
y = \begin{bmatrix} y^- \\
\eta^+
\end{bmatrix} \in \mathbb{R}^n \text{ is the learning target vector, with } y^- = (\tau - \eta)I \in \mathbb{R}^n \text{ and } y^+ = (\tau + \eta)I^+ \in \mathbb{R}^n \text{ for a given threshold } \tau \text{ and offset } \eta. \lambda = [1, \ldots, 1]^T \text{ is a vector of element ones. } \lambda \text{ is the regularization factor with } I \text{ being the identity matrix that matches the dimension of } P^T W P.$

3. A CONVOLUTION-FREE LBPD-HOG DESCRIPTOR

3.1. Overview

In this section, we propose a Convolution-Free LBP-HOG descriptor (CF-LBP-HOG) in matrix form. Specifically, we propose a Convolution-Free LBP descriptor (CF-LBP) in matrix form in the first step. This is followed by a Convolution-Free HOG descriptor (CF-HOG). The proposed CF-LBP and CF-HOG are then integrated into a single matrix product form.

3.2. Convolution-Free LBP in Matrix Form

The original LBP and the LBC iteratively compute the pixel differences in eight directions at 45° angle difference (0 to 315°). The proposed CF-LBP can be formulated as the weighted sum of eight directional difference matrices as follows:

\[
Z_{PR}^{LBPD}(X) = \sum_{i=1}^{P} \sigma(D_{LBPD,i})2^{i-1}, \quad i = \{1, 2, \ldots, P\},
\]

where

\[
\begin{align*}
D_{LBPD,1} &= V_{LBPD}X_{LBPD,1} - X, \\
D_{LBPD,2} &= V_{LBPD}X - X, \\
D_{LBPD,3} &= V_{LBPD}X_{LBPD,1} - X, \\
D_{LBPD,4} &= X_{LBPD,1} - X, \\
D_{LBPD,5} &= V_{LBPD}X_{LBPD,2} - X, \\
D_{LBPD,6} &= V_{LBPD}X - X, \\
D_{LBPD,7} &= V_{LBPD}X_{LBPD,1} - X, \\
D_{LBPD,8} &= X_{LBPD,1} - X,
\end{align*}
\]

with \( V_{LBPD} \in \mathbb{R}^{M \times M} \) and \( H_{LBPD} = \mathbb{R}^{N \times N} \) being the predefined shifting matrices from equation (5).
3.3. Convolution-Free HOG in Matrix Form

In [12], the DMP method utilized the window based iterative scanning process for block normalization based on the \(L_2\)-norm. The proposed CF-HOG as an approximated HOG is based on the \(L_2\) norm block normalization in matrix form including overlapping. A cell-based overlapping pooling is also proposed to obtain the local connection between the cell groups. The cell-based overlapping pooling can be written as follows:

\[
G_i = L_{c2,v}^{i}F_{i}R_{c2,v}, \quad i = \{1,2,\ldots,8\},
\]

where

\[
L_{c,v} = \begin{bmatrix}
1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \cdots \\
0 & 0 & 0 & \cdots & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\
\end{bmatrix},
\]

and \(F_{i}=L_{c}^{i}D_{HOG},R_{c1}\) is the cell-based non-overlapping pooling from equation (2) using a cell size of \(c_1\) in which \(L_{c2,v}\) and \(R_{c2,v}\) are the predefined projection matrices for the cell-based overlapping pooling from equation (10) and (11) using a cell size of \(c_2\) and an overlapping size of \(v\).

In [6], the \(L_2\) normalization is applied to each block which consisted of a cell group. A block-based overlapping normalization in matrix form can be written as follows:

\[
Z_{i}^{HOG}(X) = G_i \odot \left( L_{b,v}^{i} \left( L_{b,v}^{b} \left( G_{b,v}^{i} R_{b,v}^{i} \right)^{c} \right)^{c} \right),
\]

where \(\odot\) denotes the Hadamard elementwise operation. \(L_{b,v}\) and \(R_{b,v}\) are the predefined projection matrices for the block-based overlapping normalization. \(L_{b,v}^{i}G_{b,v}^{i}R_{b,v}^{i}\) is the sum of the squared elements of each block based on the pooling technique. \(L_{b,v}^{i}\) and \(R_{b,v}^{i}\) are for upsampling to retain the original matrix size. The superscript \((\frac{1}{2})\) is the elementwise inverse of the squared root.

3.4. Integrating LBP and HOG into One Matrix Form

Based on the CF-LBP and CF-HOG, the proposed CF-LBP-HOG in matrix form can be written as follows:

\[
Z_{i}^{LBP-HOG}(X) = \left( C_{c1}^{L} D_{HOG,j}^{c}(Z_{i}^{LBP}(X)) C_{c2}^{R} \right) \left( C_{c1}^{L} D_{HOG,j}^{c}(Z_{i}^{LBP}(X)) C_{c2}^{R} \right)^{c} \left( C_{b}^{L} D_{HOG,j}^{b}(Z_{i}^{LBP}(X)) C_{b}^{R} \right)^{b} \left( C_{b}^{L} D_{HOG,j}^{b}(Z_{i}^{LBP}(X)) C_{b}^{R} \right)^{b}^{2},
\]

where \(C_{c1}^{L} = L_{c2,v}^{c1}, L_{c1,v}^{c1}, C_{c1}^{R} = R_{c2,v}^{c1}, R_{c1}^{c1}, C_{b}^{L} = L_{b2,v}^{b}, L_{b1,v}^{b}, C_{b}^{R} = R_{b2,v}^{b}, R_{b1,v}^{b}\), \(X\) is the input matrix utilized by \(Z_{i}^{LBP}\) in equation (7) and \(D_{HOG,j}\) is obtained by using equation (6) with \(Z_{i}^{LBP}\). For the final CF-LBP-HOG features, each \(Z_{i}^{HOG}\) are concatenated together into one feature vector.

3.5. A Case Study of The LBP Based Descriptors

In this study, we compare the proposed CF-LBP with the original LBP and with the LBC using a mammogram image as the input image. Based on a typical LBP setting at \(R = 1\) and \(P = 8\), the LBP based descriptors, namely, LBP, LBC and CF-LBP, are performed to obtain the LBP based texture image. Fig. 1 shows the same output values within the borders from each descriptor whereas the borders of each descriptor have different output values due to different techniques to perform the LBP.

4. EXPERIMENTS

In this section, we evaluate the proposed descriptor for mammogram classification. The experimental goals are as follows: 1) Observing the effect of overlapping pooling among the HOG, the DMP and the proposed CF-HOG; 2) Performance comparison of the proposed CF-LBP-HOG with state-of-the-arts descriptors.

4.1. Database and Experimental Setup

The most commonly used database in mammography is the Digital Database of Screening Mammography (DDSM) [14]. Recently, in [15], a Curated Breast Imaging Subset of the DDSM (CBIS-DDSM) has been released in view of the segmentation difficulty for a standardized evaluation. In this experiment, we used the CBIS-DDSM database. The database includes 3,568 ROI images which are categorized into two classes (malignant, benign) for 6,671 patients and the images are resized to a 56 x 56 resolution. By following the data split setting in [15], a training set (2,864 images) and a test set (704 images) are obtained.

For state-of-the-art descriptors, the original LBP [4], the LBC [5], the HOG [6] and the DMP [12] are implemented.
4.2. Observing the effect of the overlapping pooling among the HOG based descriptors

To compare the proposed CF-HOG descriptor to the HOG and DMP descriptors, classification test accuracies at different cell and block sizes \((c_1 \text{ and } b)\) were acquired utilizing CBIS-DDSM database. Fig. 2 shows the effect of the proposed overlapping pooling at different cell size \(c_2\) in CF-HOG compared to HOG and DMP where there is no overlapping pooling. The accuracy performances are observed according to increasing the non-overlapping pooling size \(c_1\). According to the increment, the performance of the CF-HOG is increasing while that of HOG and DMP are unstable. The proposed CF-HOG outperformed the other state-of-the-art methods when \(c_1\) is 4 and 6. The best performance is observed in CF-HOG when the overlapping pooling is included.

4.3. Performance Comparison and Summary

In terms of classification performance, Table 1 shows the classification accuracies for the compared descriptors, namely LBP, HOG, DMP, LBP-HOG and CF-LBP-HOG. The proposed CF-LBP-HOG based on the TER classifier showed the best performance compared to the state-of-the-arts while the LBP-HOG based on the SVM classifier and the VGGNet respectively showed the second and the third bests.

In terms of computational performance, Table 2 shows the CPU processing time in seconds. The averaged CPU times are reported over 10 runs. The proposed CF-LBP-HOG based on the TER classifier showed the best CPU times in the training and test phases due to the predefined projection matrices under global computation form. In addition, the CF-LBP-HOG produced the smallest feature dimension which is 5 times less than the other descriptors.

In summary, we have shown that 1) the overlapping pooling step of the proposed CF-HOG is effective compared with the HOG and the DMP, 2) the proposed CF-LBP-HOG achieved better performance with a small feature dimension than that of the other state-of-the-art methods in terms of classification accuracy and CPU time.

5. CONCLUSION

Different from the convolution based LBP, we have presented a convolution-free LBP (CF-LBP) in matrix form. In addition, we have shown a convolution-free HOG based on Difference Matrix Projection (DMP). The integrated form of these two proposed descriptors, CF-LBP-HOG, was then proposed in a matrix formulation. The proposed descriptors were evaluated using the CBIS-DDSM database for mammogram classification where the results show promising performance comparing with state-of-the-art descriptors.

6. REFERENCES

[1] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, and Ahmedin Jemal, “Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in...
185 countries,” *CA: A Cancer Journal For Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] Hamid Soltanian-Zadeh, Siamak Pourabdollah-Nezhad, and Farshid Rafiee Rad, “Shape-based and texture-based feature extraction for classification of microcalcifications in mammograms,” in *Medical Imaging: Image Processing*. International Society for Optics and Photonics, 2001, vol. 4322, pp. 301–311.

[3] Alan Joseph Bekker, Moran Shalhon, Hayit Greenspan, and Jacob Goldberger, “Multi-view probabilistic classification of breast microcalcifications,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 2, pp. 645–653, 2016.

[4] Timo Ojala, Matti Pietikainen, and David Harwood, “Performance evaluation of texture measures with classification based on kullback discrimination of distributions,” in *International Conference on Pattern Recognition (ICPR)*. IEEE, 1994, vol. 1, pp. 582–585.

[5] Felix Juefei-Xu, Vishnu Naresh Boddeti, and Marios Savvides, “Local binary convolutional neural networks,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2017, vol. 1.

[6] Navneet Dalal and Bill Triggs, “Histograms of oriented gradients for human detection,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2005, vol. 1, pp. 886–893.

[7] Arnau Oliver, Xavier Lladó, Robert Martí, Jordi Freixenet, and Reyer Zwiggelaar, “Classifying mammograms using texture information,” in *Medical Image Understanding and Analysis*. Citeseer, 2007, vol. 223.

[8] Aditya A. Shastri, Deepti Tamrakar, and Kapil Ahuja, “Density-wise two stage mammogram classification using texture exploiting descriptors,” *Expert Systems with Applications*, vol. 99, pp. 71–82, 2018.

[9] Timo Ahonen, Abdennour Hadid, and Matti Pietikainen, “Face description with local binary patterns: Application to face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[10] Alireza Sepas-Moghaddam, Paulo Lobato Correia, and Fernando Pereira, “Light field local binary patterns description for face recognition,” in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 3815–3819.

[11] Yufei Han, Zhenan Sun, and Tieniu Tan, “Palmprint recognition using coarse-to-fine statistical image representation,” in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2009, pp. 1969–1972.

[12] Xing Liu and Kar-Ann Toh, “A novel set of pixel difference-based features for pedestrian detection,” in *International Conference on Identity, Security, and Behavior Analysis (ISBA)*. IEEE, 2018, pp. 1–6.

[13] Kar-Ann Toh, “Deterministic neural classification,” *Neural Computation*, vol. 20, no. 6, pp. 1565–1595, 2008.

[14] Michael Heath, Kevin Bowyer, Daniel Kopans, Richard Moore, and W. Philip Kegelmeyer, “The digital database for screening mammography,” in *International Workshop on Digital Mammography*. Medical Physics Publishing, 2000, pp. 212–218.

[15] Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, and Daniel L. Rubin, “A curated mammography data set for use in computer-aided detection and diagnosis research,” *Scientific Data*, vol. 4, pp. 170–177, 2017.

[16] Pengcheng Xi, Chang Shu, and Rafik Goubran, “Abnormality detection in mammography using deep convolutional neural networks,” in *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. IEEE, 2018.