



PLDANet: Reasonable Combination of PCA and LDA Convolutional Networks

C. C. Zhang, M. Mei, Z. L. Mei, J. K. Zhang, A. Y. Deng, C. L. Lu

Caicai Zhang[&]

School of Modern Information Technology
Zhejiang Institute of Mechanical and Electrical Engineering
528 Binwen Road, Binjiang District, Hangzhou, 310053, Zhejiang, China
zhangcaicai@zime.edu.cn

Mei Mei[&]

Department of Ultrasound, The Second Affiliated Hospital
Zhejiang University School of Medicine
88 Jiefang Road, Shangcheng District, Hangzhou, 310009, Zhejiang, China
meimeizjdx@zju.edu.cn

[&]Co-first authors: these authors contributed equally to this work.

Zhuolin Mei

School of Computer and Big Data Science
Jiujiang University
551 Qianjin East Road, Jiujiang, Jiangxi, 332005, China
6120153@jju.edu.cn

Junkang Zhang

School of Computer and Big Data Science
Jiujiang University
No. 551, Qianjin East Road, Jiujiang, Jiangxi, 332005, China
6120144@jju.edu.cn

Anyuan Deng^{*}

School of Computer and Big Data Science
Jiujiang University
No. 551, Qianjin East Road, Jiujiang, Jiangxi, 332005, China
dengay@jju.edu.cn

^{*}Corresponding author: dengay@jju.edu.cn

Chenglang Lu

School of Modern Information Technology
Zhejiang Institute of Mechanical and Electrical Engineering
528 Binwen Road, Binjiang District, Hangzhou, 310053, Zhejiang, China
luchenglang@zime.edu.cn

Abstract

Integrating deep learning with traditional machine learning methods is an intriguing research direction. For example, PCANet and LDANet adopts Principal Component Analysis (PCA) and Fisher Linear Discriminant Analysis (LDA) to learn convolutional kernels separately. It is not reasonable to adopt LDA to learn filter kernels in each convolutional layer, local features of images from different classes may be similar, such as background areas. Therefore, it is meaningful to adopt LDA to learn filter kernels only when all the patches carry information from the whole image. However, to our knowledge, there are no existing works that study how to combine PCA and LDA to learn convolutional kernels to achieve the best performance. In this paper, we propose the convolutional coverage theory. Furthermore, we propose the PLDANet model which adopts PCA and LDA reasonably in different convolutional layers based on the coverage theory. The experimental study has shown the effectiveness of the proposed PLDANet model.

Keywords: deep learning, principal component analysis, fisher principal component analysis, convolutional coverage;

1 Introduction

Automatic image classification is an important task in many areas, such as face recognition [1, 15, 34], content-based image retrieval [2, 3, 29] and computer-aided medical images classification [4, 28, 30]. The performance of advanced automatic images classification models relies on good features abstracted from images. Currently, deep convolutional neural networks, such as ConvNet [12, 13] and DenseNet [10] are widely used for abstracting features from images, which use the famous back-propagation algorithm to train all the parameters in the networks [20]. However, it lacks of intuitive interpretability, theoretical analysis and results instability. Furthermore, the procedure of parameters training using back-propagation algorithm is time-consuming, and its accuracy highly relies on hyper-parameters optimization [7, 27].

On the contrary, the traditional machine learning algorithms, such as principal components analysis (PCA) [9] and linear discriminant analysis (LDA) [23] have ready-by theoretics, and can calculate perfect features of training data set directly. However, the time complexity of calculating eigen vectors in PCA and LDA is $O(n^3)$, where n is the dimension of input data. The high time complexity prevents them adopting in data sets with large dimension. For example, if the size of an image is 900pt*700pt, then the dimension of the input data is 630000. Therefore, some researchers consider to integrate convolutional neural networks with principal component analysis or linear discriminant analysis. Chan et al. [7] propose PCANet and LDANet. PCANet leverages unsupervised PCA to learn multi-stage filters, while LDANet incorporates supervised LDA to learn convolutional kernels (or filters) for image classification. Sun et al. [27] propose Fisher PCA (FPCA) to learn convolutional kernels based on combination of LDA and PCA. However, to our knowledge, there are no existing works that study how to combine PCA and LDA to learn convolutional kernels to achieve the best performance.

In these models, the labels of patches of images are the same as the labels of images when adopting LDA in calculating each convolutional kernel. The patches of images with different categories may have similar local features. For example, the patches of background areas of different images may be totally the same. Adopting LDA for calculating convolutional kernels may abstract unreasonable local features, and harm the classification performance. Thus, although LDA takes advantage of additional label information rather than PCA, it is suitable for adopting LDA to learn convolutional kernels only when the patches of an image carry information from the whole image.

To our knowledge, there are no existing works that study how to combine PCA and LDA to learn convolutional kernels to achieve the best performance. To abstract better features of images, we propose an improved image classification model called PLDANet, based on a reasonable combination way of PCA and LDA with deep convolutional neural networks.

Our main contributions in this article are summarized as follows.

(1) The reasonable combination schema of PCA and LDA is studied when integrating PCA and LDA in calculating convolutional kernels. The combination schema is achieved based on a proposed convolutional coverage theory, which calculates the number of necessary convolutional layers to guarantee that each local patch of the image covers information from the whole image.

(2) We propose a new model called PLDANet model on the basis of the theory of convolutional coverage. PLDANet adopts a specific number of PCA convolutional layers before the subsequent LDA convolutional layer. The specific number is calculated according to the convolutional coverage theory, so that the LDA convolutional layer takes advantage of labels information completely correctly.

(3) We conduct an extensive experimental study to evaluate PLDANet in different data sets. The experimental results show that the proposed PLDANet have better performance on classification tasks.

The remainder of this article is organized as follows. Section 2 discusses the related work. Section 3 describes our proposed convolutional coverage theory. Section 4 describes our proposed PLDANet model on the basis of the theory of convolutional coverage. In Section 5, we report the performance evaluation of our techniques over five MNIST variations. Section 6 concludes this paper.

2 Related work

As big data containing images has been continuously generated from many areas[32, 33], automatic image features extraction and classification method is needed in many applications. PCA produces the most distinctive subspace for representation by maximizing the variance ,of the extracted features and discarding noisy dimensions [9]. DPCA model [16] performs a two-layer zero-phase component analysis (ZCA) whitening plus PCA structure to learn, hierarchical features for face recognition. The whole feature representation of images is extracted by concatenating the representations from the two layers.

To further improve the performance of traditional PCA based image classification, researchers have proposed several versions of PCA in the past decade. Pereira et al. [22] proposes a modular PCA algorithm that divides an image into smaller sub-images, which extracts local and global information of images aiming to reduce the effects caused by variable changes. Liu et al. [17] combines the traditional PCA and kernel-PCA to outline the data into high dimension vector space with non-linear transformation.

To achieve a better computational and accuracy rate, some studies combine PCA with different techniques, such as Delaunay Triangulation [14], Linear Discriminant Analysis [24], Deep Neural Networks (DNN) [5] and Convolutional Neural Network (CNN) [21]. Machidon et al. [19] examine the integration of PCA with Delaunay Triangulation. Jing et al. [11] integrate feature extraction of Color 2D-PCA, and Convolutional Neural Network into one decision-level fusion. Generalized discriminant analysis (GerDA) [26] is a generalization of the classical linear discriminant analysis, which uses DNN to extract features. Lu et al. [18] propose a new joint feature learning approach to automatically learn feature representations and stack a deep architecture to exploit hierarchical information. The weighted PCA (WPCA)[31] is used in joint feature learning to map the combined outputs of the first layer to a low-dimensional feature space for reducing the redundancy.

A CNN architecture consists of multiple trainable stages stacked on top of each other, followed by a supervise classifier [25]. Chan et al. [7] propose two deep model variations, PCANet and LDANet. PCANet integrates PCA with CNN and can easily be employed to learn multistage filters. PCANet processes input images by cascaded PCA, binary hashing [6], and block histograms. PCANet is an unsupervised deep learning baseline that mainly leverages PCA to learn multi-stage filters. While LDANet is a supervised deep learning framework that mainly incorporates LDA with CNN to learn filters. Sun et al. [27] propose Fisher PCA (FPCA) to learn each convolutional kernels based on mixing of PCA and LDA. These models adopt LDA in calculating each convolutional kernel, and take the label of an image as that of all patches of the image. However, the patches of images with different categories may have similar local features. Unreasonable local features may be learned by adopting LDA in calculating each convolutional kernel and lead to imperfect performance. Thus, it is suitable for adopting LDA to learn convolutional kernels only when the patches of an image carry information from the whole image. As far as we know, we are the first to study how to combine PCA and LDA reasonably to learn convolutional filters.

3 The convolutional coverage

In the LDA convolutional layer, the labels of all patches of a image are the same as the label of the image. However, it is not reasonable to adopt LDA to learn filter kernels for every convolutional layer, local features of images from different classes may be similar, such as background areas. It is meaningful to adopt LDA to learn filter kernels only when all the patches carry information from the whole image.

In this section, we propose a theory of convolutional coverage and analyze when a patch will carry information from the whole image. Then a new convolutional network model based on the convolutional coverage theory will be introduced in Section 4.

First, the definition of the convolutional coverage is given in Definition 1. The symbols in the paper are summarized in Table 1.

Table 1: The summarization of symbols

Symbols	Description
$I \in R^{m \times n}$	an image I with $m * n$ pixels
$C_{[w, h]}$	the convolutional coverage of the pixel $[w, h]$ in I
$C_{[w, h]}^i$	the convolutional coverage of the pixel $[w, h]$ in I in the i^{th} convolutional layer
$\lceil n \rceil$	round a float point number to the next integer

Definition 1 (The convolutional coverage). *Given an image $I \in R^{m \times n}$ and a pixel $[w, h]$ in I , if $k_1 \times k_2$ is the kernel size of the convolutional layer, then we say that $C_{[w, h]} = [w_{min} : w_{max}, h_{min} : h_{max}]$ is the convolutional coverage of the pixel $[w, h]$ in I , where*

$$\begin{aligned}
 C_{[w, h]} \cdot w_{min} &= \max\{0, w - k_1/2\}, \\
 C_{[w, h]} \cdot w_{max} &= \min\{m - 1, w + k_1/2\}, \\
 C_{[w, h]} \cdot h_{min} &= \max\{0, h - k_2/2\}, \\
 C_{[w, h]} \cdot h_{max} &= \min\{n - 1, h + k_2/2\}.
 \end{aligned}$$

The convolutional coverage defines the range of the image carrying by a pixel after the convolutional operation. For example, given a image with size 28×28 , if the kernel size is 7×7 , then $C_{[6, 6]} = [3 : 9, 3 : 9]$

Then the method of computing the convolutional coverage of a pixel of an image in the i^{th} convolutional layer is given in Theorem 1.

Theorem 1. *Suppose that $I \in R^{m \times n}$ is an image, $I^i \in R^{m \times n}$ is the output in the i^{th} convolutional layer of I , and the kernel size is $k_1 \times k_2$ in each convolutional layer. If $C_{[w, h]}^i = [w_{min} : w_{max}, h_{min} : h_{max}]$ is the i^{th} convolutional coverage of the pixel $[w, h]$ in I , then*

$$\begin{aligned}
 w_{min} &= \max\{0, w - i \times k_1/2\}, \\
 w_{max} &= \min\{m - 1, w + i \times k_1/2\}, \\
 h_{min} &= \max\{0, h - i \times k_2/2\}, \\
 h_{max} &= \min\{n - 1, h + i \times k_2/2\}.
 \end{aligned}$$

Proof. Let I^1 be the output in the first convolutional layer, then we have,

$$\begin{aligned}
 C_{[w, h]}^1 \cdot w_{min} &= \max\{0, w - k_1/2\}, \\
 C_{[w, h]}^1 \cdot w_{max} &= \min\{m - 1, w + k_1/2\}, \\
 C_{[w, h]}^1 \cdot h_{min} &= \max\{0, h - k_2/2\}, \\
 C_{[w, h]}^1 \cdot h_{max} &= \min\{n - 1, h + k_1/2\}.
 \end{aligned}$$

Let $I^{(i-1)} \in R^{m \times n}$ be the output in the $(i-1)^{th}$ convolutional layer, and $C_{[w, h]}^{(i-1)} = [w_{min} : w_{max}, h_{min} :$

h_{max}] be the $(i - 1)^{th}$ convolutional coverage in I for the pixel $[w, h]$, then we have,

$$\begin{aligned} C^i_{[w, h]}.w_{min} &= \max\{0, C^{(i-1)}_{[w-k_1/2, h-k_2/2]}.w_{min}\} \\ &= \max\{0, C^{(i-j)}_{[w-j \times k_1/2, h-j \times k_2/2]}.w_{min}\} \\ &= \max\{0, C^1_{[(w-(i-1)) \times k_1/2, h-(i-1) \times k_2/2]}.w_{min}\} \\ &= \max\{0, w - (i - 1) \times k_1/2 - k_1/2\} \\ &= \max\{0, w - i \times k_1/2\} \\ C^i(w, h).w_{max} &= \min\{m - 1, C^{(i-1)}(w + k_1/2, h + k_2/2).w_{max}\} \\ &= \min\{m - 1, C^{(i-j)}(w + j \times k_1/2, h + j \times k_2/2).w_{max}\} \\ &= \min\{m - 1, C^1(w + (i - 1) \times k_1/2, h + (i - 1) \times k_2/2).w_{max}\} \\ &= \min\{m - 1, w + (i - 1) \times k_1/2 + k_1/2\} \\ &= \min\{m - 1, w + i \times k_1/2\} \end{aligned}$$

Similarly,

$$\begin{aligned} C^i(w, h).h_{min} &= \max\{0, h - i \times k_2/2\}, \\ C^i(w, h).h_{max} &= \min\{n - 1, h + i \times k_2/2\} \end{aligned}$$

□

Example 1. Given an image with size 28×28 , if the kernel size in each convolutional layer is 7×7 , then

$$\begin{aligned} C^1_{[6, 6]} &= [3 : 9, 3 : 9] \\ C^2_{[6, 6]} &= [0 : 12, 0 : 12] \end{aligned}$$

Next we give the Theorem 2 to obtain the number of the convolutional layer that each patch of the image carries information from the whole image.

Theorem 2. Given an image $I \in R^{m \times n}$, and the kernel size is $k_1 \times k_2$ in each convolutional layer, after $G = \lceil \max\{(m - k_1)/(k_1/2), (n - k_2)/(k_2/2)\} \rceil$ convolutional layers, each patch carries information from the whole image.

Proof. If the left upper patch $[0 : (k_1 - 1), 0 : (k_2 - 1)]$ covers information from the whole image, then each patch does. If the pixel $[(k_1 - 1), (k_2 - 1)]$ covers information from the pixel $[(m - 1), (n - 1)]$, then the left upper patch covers information from the whole image.

Based on Theorem 1,

$$C^G_{[(k_1-1), (k_2-1)].w_{max}} = \min\{(m - 1), k_1 - 1 + G * k_1/2\}$$

When $k_1 - 1 + G * k_1/2 = m - 1$, we have

$$G = (m - k_1)/(k_1/2)$$

Similarly,

$$C^G_{[(k_1-1), (k_2-1)].h_{max}} = \min\{(n - 1), k_2 - 1 + G * k_2/2\}$$

When $k_2 - 1 + G * k_2/2 = n - 1$, we have

$$G = (n - k_2)/(k_2/2)$$

Since G have to be an interger, we round up the value to the greater one of $(m - k_1)/(k_1/2)$ and $(n - k_2)/(k_2/2)$.

Therefore, after $G = \lceil \max\{(m - k_1)/(k_1/2), (n - k_2)/(k_2/2)\} \rceil$ convolutional layers, each patch carries information from the whole image. □

Example 2. Given an image with size 28×28 , if the kernel size in each convolutional layer is 7×7 , then $G = \lceil (28 - 7)/3 \rceil = 7$.

4 The PLDANet model

The convolutional coverage theory calculates the necessary number of the convolutional layers, which is important for designing reasonable combination of PCA and LDA convolutional networks. Section 4.1 gives the definitions of PCA convolutional layer and LDA convolutional layer. Section 4.2 introduces the proposed PLDANet model.

4.1 Definitions

PCA extracts main features to maximum the variance of data, which is an unsupervised learning method. Definition 2 defines the PCA convolutional layer.

Definition 2 (PCA convolutional layer). *A convolutional layer is called PCA Convolutional layer if the filter kernel parameters are learned by PCA.*

Let $\{I_i \in R^{m \times n \times c}\}_{i=1}^N$ be the input images of the PCA convolutional layer, where N is the number of input images, $m \times n \times c$ is the size of each image I_i . A patch with the filter kernel size $k_1 \times k_2$ is taken around each pixel of each image. Let $\{X_i\}_{i=1}^N \in R^{k_1 k_2 \times cmnN}$ be the patch set.

The filter kernel parameters in a PCA convolutional layer are learned by PCA on the patch set. Each eigen vector calculated by PCA corresponds to a filter kernel.

LDA extracts discriminative features to maximum the variance of data in different classes and minimum the variance of data in the same class, which is a supervised learning method. Definition 2 defines the LDA convolutional layer.

Definition 3 (LDA convolutional layer). *A convolutional layer is called LDA Convolutional layer if the filter kernel parameters are learned by LDA.*

The filter kernel parameters in a LDA convolutional layer is learned by LDA on the patch set, and the labels of all patch of an image are the same as the label of the image. Each eigen vector calculated by LDA corresponds to a filter kernel.

4.2 The model architecture

Figure 1 shows the detailed diagram of the proposed PLDANet model. First, we evaluate the number of PCA convolutional layers in the PLDANet model. $G = \max\{(m - k_1)/(k_1/2), (n - k_2)/(k_2/2)\}$, where (m, n) is the size of the input image, and (k_1, k_2) is the size of each filter kernel. Then the Deep PLDANet model is constructed by the number of G PCA convolutional layers and a subsequent LDA convolutional layer. Then the Hashing, Histograms and the linear SVM stages are the same as the PCANet. The training process of PLDANet for classification is demonstrated in Algorithm 1.

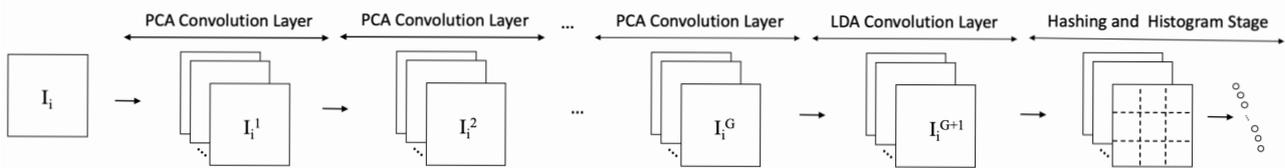


Figure 1: The detailed diagram of the proposed PLDANet model

Example 3. *Given an image with size 28×28 , if the kernel size in each convolutional layer is 7×7 , then $G = \lceil (28 - 7)/3 \rceil = 7$.*

In the PLDANet model, the convolutional network includes 7 PCA convolutional layers and a subsequent LDA convolutional layer.

5 Experiments

Since the training time of PCANet and LDANet is much less than that of CNN, and the superior performance of PCANet and LDANet over CNN has been studied in [7, 27], the comparison of our proposed PLDANet and CNN has not been conducted in our work. In this section, we compare PLDANet with PCANet and LDANet.

We conduct experiments on MNIST and its four variations. The MNIST dataset is one of the most popular datasets in image classification. It contains 60,000 training and 10,000 test images

Algorithm 1 The training process of PLDANet for image classification**Input:**

- (1) The training image set $\{(I_i, y_i)\}_{i=1}^N$;
- (2) $I_i \in R^{m \times n}$;
- (3) The size of filter kernel k_1, k_2 ;

Output: The results of classification.

- 1: $(m, n) \leftarrow$ the size of the input image
- 2: $G = \max\{(m - k_1)/(k_1/2), (n - k_2)/(k_2/2)\}$
- 3: **for** each $l = 1 : G$ **do**
- 4: $I_i^l \leftarrow \text{PCA_Convolutional_Layer}(I_i), i = 1, 2, \dots, N$
- 5: **end for**
- 6: $I_i^{(G+1)} \leftarrow \text{LDA_Convolutional_Layer}(I_i), i = 1, 2, \dots, N$
- 7: Hashing and Histograms stage: take $I^{(G+1)}$ as input;
- 8: Train linear SVM: take Histograms codes as input;
- 9: **return** The results of image classification.

of 10 hand-written digits. MNIST variations [8] introduce addition of rotation (rot), addition of a background composed of random pixels (bg-rand) or patches extracted from a set of images (bg-img), or combinations of these factors (rot-bg-img) to MNIST digit classification. These variations make the problem become particularly challenging. The five classification datasets are summarized in Table 2. The size of all the images are 28×28 . An instance of MNIST variations is shown in Figure 2.

Table 2: Details of the 5 classification tasks on MNIST variations.

Data Sets	Description	Train-Valid-Test
basic	Smaller subset of MNIST	10,000-2000-10,000
rot	MNIST with rotation	10,000-2000-50,000
bg-rand	MNIST with noise background	10,000-2000-50,000
bg-img	MNIST with image background	10,000-2000-50,000
bg-img-rot	MNIST with rotation and image background	10000-2000-50000

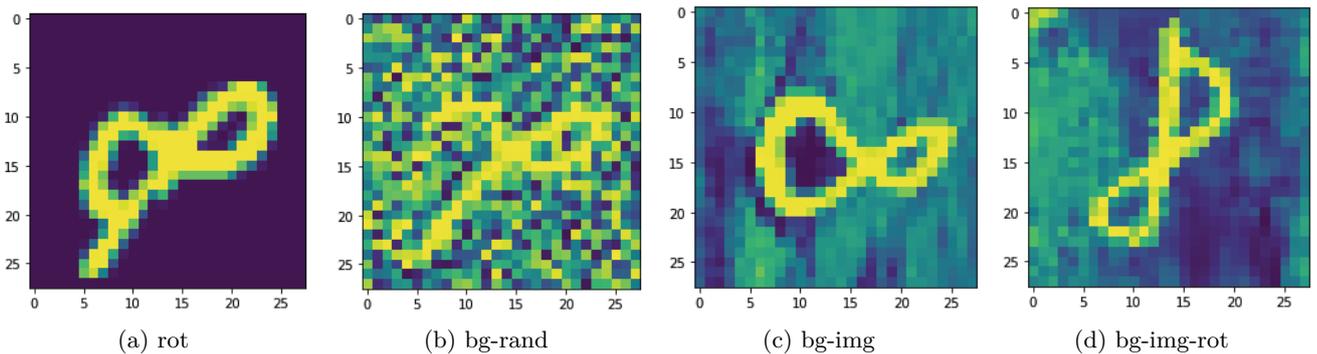


Figure 2: The number eight in the MNIST variations

To evaluate the performance of PCA convolutional layer and LDA convolutional layer in different convolutional layers, we compare PCANet- i , PLDANet- i and PLDANet2- i , where PCANet- i means there are number of i PCA convolutional layers, PLDANet- i means there are PCA convolutional layers in the first $(i-1)$ layers and LDA convolutional layer in the last layer, and PLDANet2- i means there are PCA convolutional layers in the first 6 layers and LDA convolutional layer in the rest layers.

Limited to the memory size in our experimental environment, the hyperparameters of PCANet- i and PLDANet- i are set as follows: (1) the number of filters are $L_1 = L_i = 8, L_j = 1, 1 < j < i$; (2) the patch size is $k_1 = k_2 = 7$ in all convolutional layers; (3) the block size for local histograms in the

output layer is set to be 7×7 , and the block overlap ratio is 0.5.

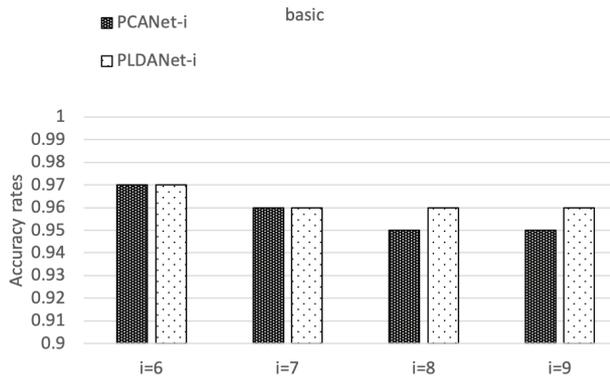


Figure 3: Comparison of PCANet-i and PLDANet-i on basic MNIST

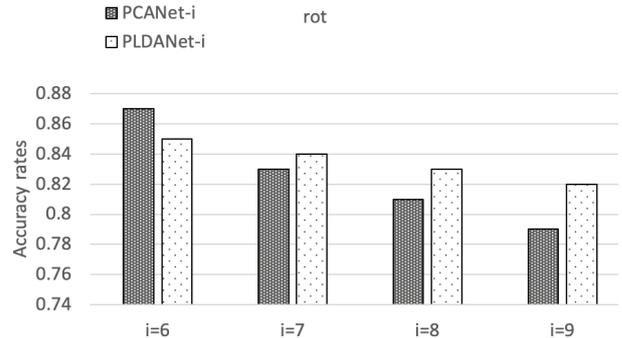


Figure 4: Comparison of PCANet-i and PLDANet-i on rot MNIST

We compare PCANet-i and PLDANet-i on five MNIST variations when varying i from 6 to 9. Figure 3 shows the accuracy rates of PCANet-i and PLDANet-i on the basic MNIST. Figure 4 shows the accuracy rates of PCANet-i and PLDANet-i on the rot MNIST. Figure 5 shows the accuracy rates of PCANet-i and PLDANet-i on the bg-rand MNIST. Figure 6 shows the accuracy rates of PCANet-i and PLDANet-i on the bg-img MNIST. Figure 7 shows the accuracy rates of PCANet-i and PLDANet-i on the bg-img-rot MNIST.

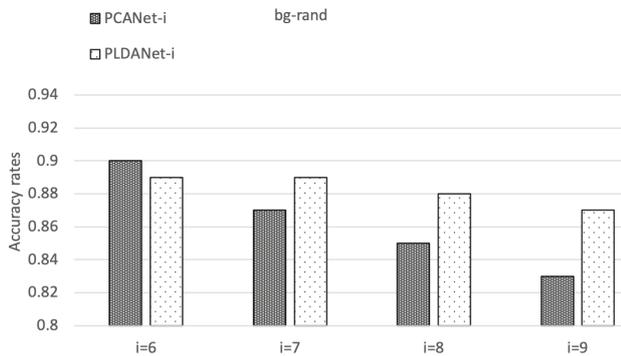


Figure 5: Comparison of PCANet-i and PLDANet-i on bg-rand MNIST

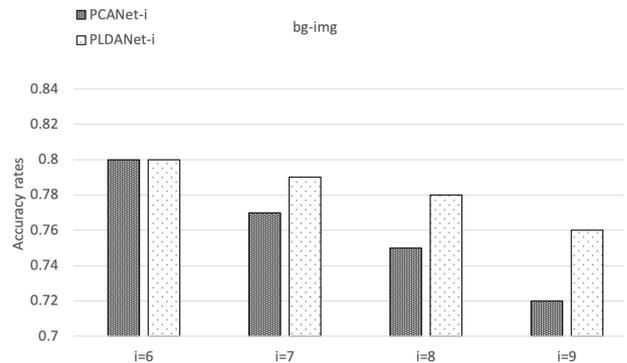


Figure 6: Comparison of PCANet-i and PLDANet-i on bg-img MNIST

The experimental results show that (1) PLDANet-6 have the same accuracy rates as PCANet-6 on basic MNIST and bg-img MNIST, and have lower accuracy rates than PCANet-6 on bg-rand, rot MNIST and rot-bg-img MNIST; (2) PLDANet-7 have higher accuracy rates than PCANet-7 on bg-img MNIST, bg-rand, rot MNIST and rot-bg-img MNIST, and have the same accuracy rate as PCANet-6 on the basic MNIST; (3) PLDANet-i ($i=8, 9$) have higher accuracy rates than PCANet-i ($i=8, 9$) on the five MNIST variations.

Based on Theorem 2, $G = (m - k)/(k/2) = (28 - 7)/3 = 7$. Therefore, each patch of the image will cover information from the whole image after 7 convolutional layers. However, in the 6th convolutional layer and 7th convolutional layer, some but not all pathes carry information from the whole image, and LDA convolutional layer have better or worse performance than PCA convolutional layer. In the 8th convolutional layer and 9th convolutional layer, all patches carry information from the whole image, and LDA convolutional layer have better performance than the PCA convolutional layer.

We compare PCANet-i, PLDANet-i and PLDANet2-i on five MNIST variations when varying i from 8 to 9. Figure 8 shows the accuracy rates of PCANet-i, PLDANet-i and PLDANet2-i on the basic MNIST. Figure 9 shows the accuracy rates of PCANet-i, PLDANet-i and PLDANet2-i on the

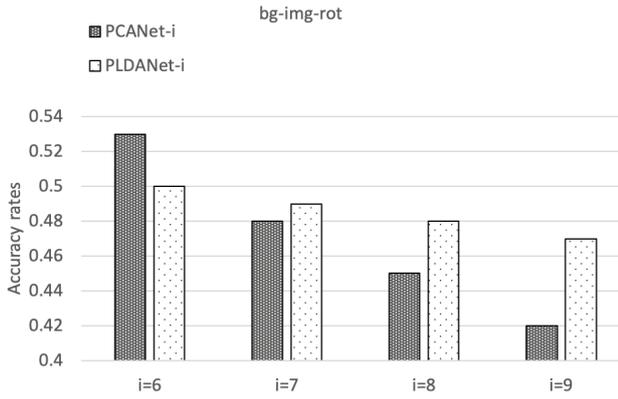


Figure 7: Comparison of PCANet-i and PLDANet-i on bg-img-rot MNIST

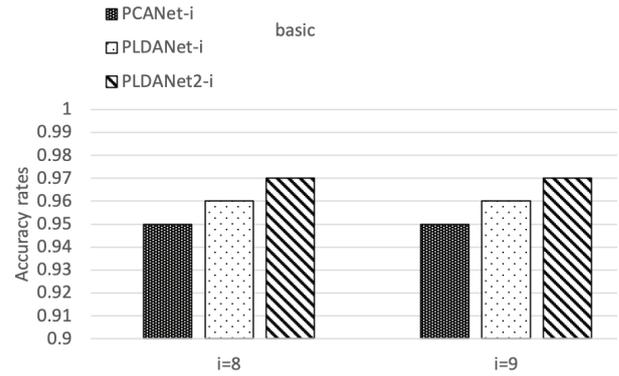


Figure 8: Comparison of PCANet-i, PLDANet-i and PLDANet2-i on basic MNIST

rot MNIST. Figure 10 shows the accuracy rates of PCANet-i, PLDANet-i and PLDANet2-i on the bg-rand MNIST. Figure 11 shows the accuracy rates of PCANet-i, PLDANet-i and PLDANet2-i on the bg-img MNIST. Figure 12 shows the accuracy rates of PCANet-i, PLDANet-i and PLDANet2-i on the bg-img-rot MNIST.

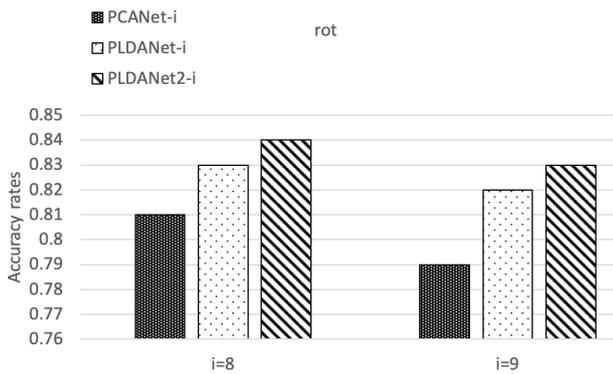


Figure 9: Comparison of PCANet-i, PLDANet-i and PLDANet2-i on rot MNIST

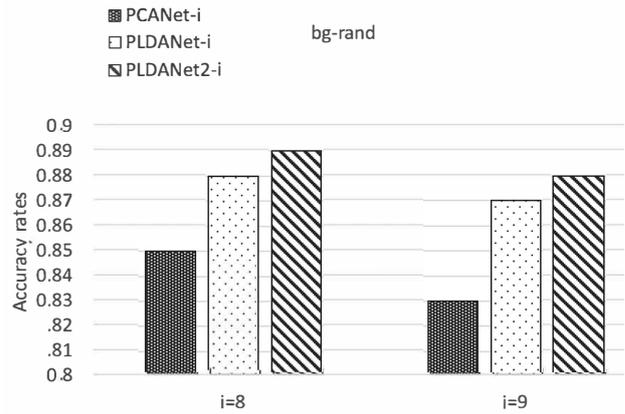


Figure 10: Comparison of PCANet-i, PLDANet-i and PLDANet2-i on bg-rand MNIST

The experimental results show that PLDANet2-i ($i=8, 9$) have higher accuracy rates than PCANet-i and PLDANet-i on all the five MNIST variations. Since PLDANet-7 have higher or the same accuracy rates compared to PCANet-7 on all the five MNIST variations, LDA convolutional layer performs better than PCA convolutional layer in the j^{th} ($j \geq 7$) convolutional layer. PLDANet-9 performs LDA convolutional layer during 7^{th} , 8^{th} and 9^{th} convolutional layer. The phenomenon proves our initial idea that adopting LDA convolutional layer is greatly beneficial for extracting features when each patch of the image covers information from the whole image.

6 Conclusion

Adopting PCA in learning convolutional kernels takes much less time than training a CNN, and avoids the laborious fine-tuning process. The supervision of the learned kernels can be further enhanced by incorporating the labels information based on LDA. However, it is meaningful to assign class labels to the patches and adopt LDA only when all the patches carry information from the whole image. In this paper, we propose a convolutional coverage theory, which indicate that all patches

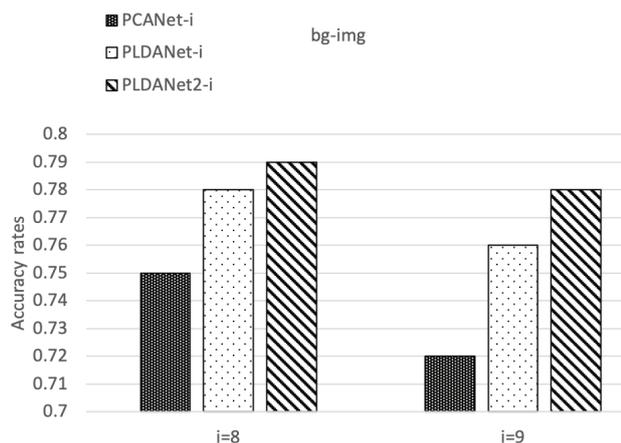


Figure 11: Comparison of PCANet-i, PLDANet-i and PLDANet2-i on bg-img MNIST

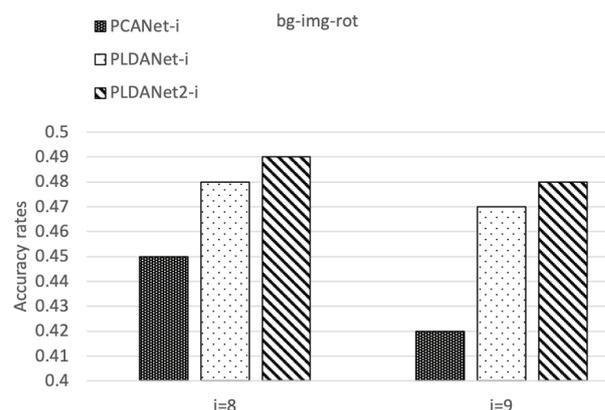


Figure 12: Comparison of PCANet-i, PLDANet-i and PLDANet2-i on bg-img-rot MNIST variations

cover information from whole images after a specific number of convolutional layers. Based on the convolutional coverage theory, we propose PLDANet model, which includes a specific number of PCA convolutional layers and a subsequent LDA convolutional layer. Benefiting from the reasonable combination of PCA and LDA, PLDANet enhances the representation ability of PCANet and LDANet. Finally, a variety of experiments are conducted to demonstrate the effectiveness of the proposed PLDANet.

Funding

This work was funded by the National Natural Science Foundation of China (No. 61962029); the Natural Science Foundation of Jiangxi Province (No. 20202ACBL202005 and No. 20202BAB212006); MOE (Ministry of Education in China), Youth Fund Project of Humanities and Social Sciences Research (No.21YJ CZH096); Science and Technology Foundation of Jiangxi Educational Committee (No. GJJ201832).

Author contributions

Caicai Zhang and Mei Mei contributed significantly to the conception of the study and wrote the manuscript; Zhuolin Mei contributed to analysis. Junkang Zhang performed the experiment; Anyuan Deng and Chenglang Lu helped perform the analysis with constructive discussions.

Conflict of interest

The authors declare no conflict of interest.

References

- [1] Ahmed, S.B.; Ali, S.F.; Ahmad, J.; Adnan, M.; and Fraz, M.M. (2020). On the frontiers of pose invariant face recognition: a review. *Artificial Intelligence Review*, 53(4), 2571–2634, 2020.
- [2] Alzyoud, F.Y.; Maqableh, W.; and Faiz Al, S. (2021). A semi smart adaptive approach for trash classification. *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*, 16(4172), 1–13, 2021.

- [3] Arun, K.S.; Govindan, V.K.; and Kumar, S.D.M. (2020). Enhanced bag of visual words representations for content based image retrieval: a comparative study. *Artificial Intelligence Review*, 53(3), 1615–1653, 2020.
- [4] Cao, Z.; Duan, L.; Yang, G.; Yue, T.; and Chen, Q. (2019). An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Medical Imaging*, 19(1), 511–519, 2019.
- [5] Carrara, F.; Falchi, F.; Caldelli, R.; Amato, G.; Fumarola, R.; and Becarelli, R. (2017). Detecting adversarial example attacks to deep neural networks. 1–7, 2017.
- [6] Carreira-Perpián, M. and Raziperchikolaei, R. (2015). Hashing with binary autoencoders. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [7] Chan, T.H.; Jia, K.; and Gao, S. (2015). Pcanet: A simple deep learning baseline for image classification? *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 24, 5017–5032, 2015.
- [8] Dlv, S. and Sor, R. (2007). An empirical evaluation of deep architectures on problems with many factors of variation. In *24th International Conference on Machine Learning*. 473–480, 2007.
- [9] Groth, D.; Hartmann, S.; Klie, S.; and Selbig, J. (2013). Principal components analysis. *Methods Mol Biol*, 930, 527–547, 2013.
- [10] Huang, G.; Liu, Z.; Pleiss, G.; Maaten, L.V.D.; and Weinberger, K. (2019). Convolutional networks with dense connectivity. *IEEE transactions on pattern analysis and machine intelligence*, 1(4), 1–12, 2019.
- [11] Jing, L.; Tao, Q.; Chang, W.; Kai, X.; and Fang-Qing, W. (2018). Robust face recognition using the deep c2d-cnn model based on decision-level fusion. *Sensors*, 18(7), 2080, 2018.
- [12] Krizhevsky, A.; Sutskever, I.; and Hinton, G.E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90, 2017.
- [13] Lecun, Y. and Bottou, L. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324, 1998.
- [14] Lee, D.T. and Lin, A.K. (1986). Generalized delaunay triangulation for planar graphs. *Discrete & Computational Geometry*, 1, 201–217, 1986.
- [15] Lintang, R.A.; Purnawarman, M.; and Wibowo, E.P. (2017). Human face recognition application using pca and eigenface approach. In *2017 Second International Conference on Informatics and Computing (ICIC)*, 2017.
- [16] Liong, V.E.; Lu, J.; and Wang, G. (2013). Face recognition using deep pca. In *2013 9th International Conference on Information, Communications & Signal Processing (ICICS)*. 1–5, 2013.
- [17] Liu, C.; Zhang, T.; Ding, D.; and Lv, C. (2016). Design and application of compound kernel-pca algorithm in face recognition. In *2016 35th Chinese Control Conference (CCC)*. 4122–4126, 2016.
- [18] Lu, J.; Liong, V.E.; Wang, G.; and Moulin, P. (2017). Joint feature learning for face recognition. *IEEE Transactions on Information Forensics and Security*, 10(7), 1371–1383, 2017.
- [19] Machidon, A.L.; Machidon, O.M.; and Ogrutan, P.L. (2019). Face recognition using eigenfaces, geometrical pca approximation and neural networks. In *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*, 2019.
- [20] Ooyen, A.V. and Nienhuis, B. (1992). Improving the convergence of the back-propagation algorithm. *Neural Networks*, 5, 465–471, 1992.

- [21] Oziuddeen, M.A.K.; Poruran, S.; and Caffiyar, M.Y. (2020). A novel deep convolutional neural network architecture based on transfer learning for handwritten urdu character recognition. *Tehnički vjesnik*, 27(4), 1160–1165, 2020.
- [22] Pereira, J.F.; Barreto, R.M.; Cavalcanti, G.D.C.; and Tsang, I.R. (2011). A robust feature extraction algorithm based on class-modular image principal component analysis for face verification. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1469–1472, 2011.
- [23] Riffenburgh, R. and Clunies-Ross, C. (1960). Linear discriminant analysis. *Pacific Science*, 14, 251–256, 1960.
- [24] Riffenburgh, R.H. and Clunies-Ross, C.W. (2013). Linear discriminant analysis. *Pacific Science*, 3(6), 27–33, 2013.
- [25] Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; and Summers, R.M. (2016). Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298, 2016.
- [26] Stuhlsatz, A.; Lippel, J.; and Zielke, T. (2014). Feature extraction with deep neural networks by a generalized discriminant analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 23(4), 596–608, 2014.
- [27] Sun, K.; Zhang, J.; Yong, H.; and Liu, J. (2018). Fpcanet: Fisher discrimination for principal component analysis network. *Knowledge-Based Systems*, 166, 108–117, 2018.
- [28] Swati, Z.N.K.; Zhao, Q.; Kabir, M.; Ali, F.; and Lu, J. (2019). Brain tumor classification for mr images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, 75(7), 34–46, 2019.
- [29] Unar, S.; Wang, X.; Wang, C.; and Wang, Y. (2019). A decisive content based image retrieval approach for feature fusion in visual and textual images. *Knowledge-Based Systems*, 179, 8–20, 2019.
- [30] Xiaofeng; Qi; Lei; Zhang; Yao; Chen; Yong; Pi; Yi; and and, Q. (2018). Automated diagnosis of breast ultrasonography images using deep neural networks. *Medical Image Analysis*, 52, 185–198, 2018.
- [31] Z., L.; M., P.; and S.Z., L. (2014). Learning discriminant face descriptor for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(2), 289–302, 2014.
- [32] Zhang, C.; Mei, Z.; Wu, B.; Yu, J.; and Wang, Q. (2020). Query with assumptions for probabilistic relational databases. *Tehnički vjesnik*, 27(3), 923–932, 2020.
- [33] Zhang, S.; Yang, L.T.; Feng, J.; Wei, W.; Cui, Z.; Xie, X.; and Yan, P. (2021). A tensor-network-based big data fusion framework for cyber-physical-social systems (cpss). *Information Fusion*, (76), 337–354, 2021.
- [34] Zhou, Y.; Wang, Y.; and Wang, X.H. (2018). Face recognition algorithm based on wavelet transform and local linear embedding. *Cluster Computing*, 22, 1529–1540, 2018.



Copyright ©2022 by the authors. Licensee Agora University, Oradea, Romania.

This is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License.

Journal's webpage: <http://univagora.ro/jour/index.php/ijccc/>



This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).

<https://publicationethics.org/members/international-journal-computers-communications-and-control>

Cite this paper as:

Zhang C.C.; Mei, M.; Mei, Z.L.; Zhang, J.K.; Deng A.Y.; Lu, C.L. (2022). PLDANet: Reasonable Combination of PCA and LDA Convolutional Networks, *International Journal of Computers Communications & Control*, 17(2), 4541, 2022.

<https://doi.org/10.15837/ijccc.2022.2.4541>