An Efficient Convolutional Neural Network for Fingerprint Pore Detection

MOHAMMED ALI, (Student Member, IEEE), CHUNYAN WANG, (Senior Member, IEEE) and M. OMAIR AHMAD, (Fellow, IEEE)

Abstract—Pore detection for fingerprint recognition has gained much research attention in recent years, in view of the existence of large number of pores in a small fingerprint segment and availability of high-resolution acquisition devices. Current research efforts have focused on developing two-part hybrid schemes comprising a CNN architecture to produce a pore intensity map and then use it to determine the pore centroids in view of the capabilities in extracting automatic and useful pore features by using a CNN architecture and exploiting the knowledge base on pores. However, recent works lack extracting representational features of pores at a reasonable complexity of the CNN architecture and using efficiently the available knowledge base on fingerprint pores. In this paper, a new two-part fingerprint pore detection scheme is proposed, wherein the first part focuses on developing a CNN architecture capable of extracting highly representational pore features and the second part on accurately determining the pore centroids by taking into consideration the inadequacies in fingerprint acquisition and distinguishing the spatial characteristics of true and false pores. Extensive experiments are performed to demonstrate the distinct characteristics to show the superiority of the proposed scheme in performance and complexity over the existing state-of-the-art pore detection schemes.

Index Terms— Biometrics, convolutional neural network-based pore detection, high-resolution fingerprint image, pore feature extraction.

◆

1 INTRODUCTION

Currently, the demand for reliable personal identification systems has increased because of concerns about security, privacy and identity fraud. Fingerprint has become one of the most widely used bio-signatures for personal recognition due to its uniqueness, permanence and convenience of use [1]. Fingerprint recognition has been widely used in both civilian and forensic applications. Fingerprints are identified based on their features, which can be divided into three levels [1]: Level 1 refers to the ridge flow and pattern type. Level 2 is the minutiae points, such as bifurcations and endings of the ridges. Level 3 includes the fine details associated with the ridge, such as ridge edge contour, pores, dots and incipient ridges.

Fingerprint pores are one of the qualitative features that can be detected to improve fingerprint recognition precision [2], given their immutability and uniqueness in terms of their location coordinates. Pores reside on the ridges of the fingertip and they differ in shape, size and their positions on the ridges [3]. It has been stated in [3] that only around 20-40 pores are sufficient to establish human identity. It has also been observed in [3] that in a small segment of the fingerprint image, the number of pores along a 1-cm length of a ridge varies from 9 to 18. Thus, only a small segment of the fingerprint image is needed to have a sufficient number of pores to establish human identity. This contrasts to only a few or even none of the minutia points that may exist in a small segment of a fingerprint image. This fact is a compelling reason for using fingerprint pores as opposed to the other levels of fingerprint features, such as ridge patterns and minutia points, for fingerprint recognition. However, an essential step for carrying out a pore-based fingerprint recognition [3], [4], [5], [6], [7], [8], [9], [10], [11], [12] is to design an accurate and efficient pore detection scheme. Fingerprint images may have pores that are distorted or may not be clearly visible, due to the fully or partially broken ridges, resulting from permanent damage in the finger or an imperfect acquisition of the fingerprint image. Consequently, pores can be classified into two types, either closed pores, i.e., those that entirely lies within a ridge or open pores, i.e., those that partially lies within a ridge [13], as can be seen in Fig. 1 marked with green and yellow circles, respectively. Detecting open pores is more challenging than detecting closed pores. The open pores have more variations in terms of shape, gradient and orientation. As seen from Fig. 1, such pores may be open to the valley from one side (such as those marked as 1), two sides (such as that marked as 2) or may even exist at a location that looks like the beginning and the end of two different ridges (such as that marked as 3). These imperfections make detecting the open pores more challenging and may result in detecting false pores.

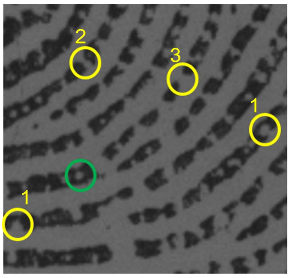


Fig. 1. Example of a fingerprint image patch with incomplete ridges. Open and closed pores are marked by yellow and green circles, respectively.

The process of pore detection used in various schemes can, in general, be divided into two parts. In the first part, a pore intensity map is obtained by extracting pore features from a gray-level input fingerprint image. A pore intensity map is essentially a collection of gray-level blobs characterizing the pores of a

---

The authors are with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC H3G1M8, Canada, e-mails: {Al_ohamm, chunyan, omair}@ece.concordia.ca.

fingerprint image. In the second part, first, the locations of the candidate pores are found from the pore intensity map by locating the coordinates of their centers, and then, by using some prior knowledge about the true fingerprint pores, spurious pores are removed from the set of candidate pores in order to obtain the final set of true pores.

The most critical step in the first part of pore detection is the feature extraction step. Traditionally, in the first part, the pore features generated and extracted are handcrafted features. The works in [14] and [15] are the two earliest methods of generating handcrafted pore features, in which the ridges of a binary skeletonized version of the fingerprint image are tracked and any discontinuity along the ridges are considered to represent the presence of a pore when certain criteria are satisfied. In the first part of other traditional schemes for pore detection, the pore features are first modeled in terms of the orientations of the ridges containing the pores and the pore shapes and sizes, and then a filter, such as Gabor [3], adaptive DOG and DAPM [16], is designed to extract the true pores from the fingerprint images. The pore detection performance of the schemes in which the pore feature extraction in the first part is based on the handcrafted features is very much limited by the modeling capability of the pore features.

In view of the fact that convolutional neural networks (CNN) have provided a capability of automatically generating and extracting features, these networks have been used in a wide variety of applications [17], [18], [19], [20], [21]. This powerful characteristic of CNNs has prompted researchers in the biometric community to develop CNN architectures for the task of human recognition. Recently, in [22], [23], and [24], authors have designed CNN architectures to automatically produce pore intensity maps of fingerprint images. The pore intensity maps so generated have been shown to provide a pore detection performance that is much superior to those provided by the techniques based on handcrafted features. However, in these schemes, little attention has been paid to the design of the network architectures in providing a rich set of pore features, as well as to the number of layers and filters employed by the networks, which are very important considerations in determining the network performance and complexity.

As mentioned earlier, in the second part of a pore detection scheme, the coordinates of the centers of the true pores are determined from the pore intensity map obtained in the first part. It is in this part of the pore detection scheme where our knowledge of some important characteristics of pores and pore fingerprint images can be utilized more efficiently without using a CNN network in order to obtain accurate values for the centroids of true fingerprint pores. In the pore detection schemes of [3], [14], [15], [16] and [22], the pore intensity map is first binarized using a global threshold value, which converts the pore intensity map into a map of white and black regions. Then, the center of each white region in the binary map is computed and it is considered to be the coordinates of the center of a fingerprint pore. In the pore detection schemes of [23] and [24], the pore intensity map is first partitioned into windows. Then, a pixel with a maximum intensity inside a window is considered to be the center of a fingerprint pore if its intensity is found to be larger than a threshold value. In the second part of the schemes of [3] and [16], the pore detection process is further refined by identifying and removing the false pores. A major disadvantage of the second part of all the existing pore detection schemes, traditional or non-traditional, is that all the available knowledge regarding the fingerprint pores are not taken into consideration and even the knowledge that is taken into consideration is not used optimally.

Personal identification or recognition is required in applications such as online banking using mobile devices. Thus, a low-power architecture is necessarily a requirement for the deployment of a fingerprint biometric identification scheme. In other biometric identification applications, such as forensic applications, the algorithm must be run in real-time. These requirements of the biometric applications need the development of biometric identification schemes, including those based on the pores of fingerprint images, with very low complexity.

It is in a hybrid (non-traditional) pore detection scheme where one can more efficiently exploit the strength of CNNs in extracting automatic pore features and also use the knowledge base on the characteristics of the pores and fingerprint images. The objective of this paper is to develop a two-part pore detection scheme in which for the first part a low complexity deep neural network is designed to produce a pore intensity map by focusing on the connectivity of its various hierarchical parts that result in the extraction of meaningful features leading to a highly representational pore intensity map, and for the second part, a method is developed for the pore detection by making efficient use of the knowledge base on fingerprint pores.

Since different kinds of features are very important in detecting the pores, the strategy in designing the network is to generate hierarchically low, middle and high level features and to concatenate them to produce a very rich set of pore features. The features at each of the three levels are learnt locally in a residual framework, which helps in curtailing the gradient vanishing problem [25], [26], [27]. By placing emphasis on the architecture design of the network for generating an enriched set of features, it is possible to keep the number of filters and layers low, which results in a low complexity network.

The second part of the pore detection scheme aims at eliminating the false pores and determining the coordinates of all the true pores accurately from the pore intensity map obtained in the first part in conjunction with the fingerprint image itself. In this part of the pore detection scheme, the knowledge on fingerprint images, such as the pores intensity, their variation from one region to another region of the fingerprint image and the minimum distance between two neighboring pores, is used to accurately detect the pores of a fingerprint image.

The accuracy of a neural network-based pore detection scheme is highly dependent on its ability to extract a very rich set of pore features and augmenting this ability with the known attributes of fingerprint pores in determining their true locations in the image. In this regard, the main contributions of our proposed pore detection scheme are as follows. (i) A new CNN architecture is developed to extract low, middle and high level features in a residual framework and then concatenate them to generate a very rich set of pore features, which leads to a highly representational pore intensity map. (ii) The first part of the proposed pore detection scheme is essentially a pixel classification problem to determine whether or not a fingerprint pixel belongs to a pore. The use of depthwise convolution which is known to provide a superior classification accuracy [28] is used in the design of the CNN architecture of the proposed scheme.

In addition, the use of depthwise convolution has led to the design of an ultralightweight network. (iii) Since the intensity of a pore pixel relative to that of a non-pore pixel varies from region to region of a fingerprint image, a locally adaptive threshold rather than a global threshold is used to binarize the pore intensity map. (iv) Since the center of a pore is affected by the intensities of the pixels belonging to the pore, the center of a pore is obtained as the mean of the intensity-weighted coordinates of the pixels belonging to the pore region. (v) In order to distinguish between a true pore and a false pore, a locally adaptive intensity threshold that is determined based on the local mean and the variance of the intensities of the pixels belonging to the local pore regions is used. (vi) Since a pair of true pores observe a certain minimum distance between them, this minimum distance is used to distinguish between true and false pores in a pair of neighboring pixels.

The paper is organized as follows. In Section 2, a literature review of the existing works on the CNN-based pore detection schemes is presented. In Section 3, the proposed two-part scheme for pore detection are systematically developed, and the motivation and rationale behind ideas used in the two parts are provided. Section 4 provides experimental results of the proposed and other pore detection schemes when applying it to a bench-mark high-resolution fingerprint database. In this section, the contribution and the impact of each part of the proposed scheme on the overall performance are investigated and the results are compared with those of the existing pore detection schemes. Finally, some concluding remarks on the study undertaken in this paper are given in Section 5.

2 RELATED WORKS IN CNN-BASED PORE DETECTION

As mentioned earlier, all the pore detection schemes have two parts, in which the second part is always based on some knowledge of the pore characteristics in fingerprint images. Based on the first part in which the pore features are extracted using a traditional method or a CNN architecture, the pore detection schemes can be categorized as traditional or CNN based schemes. Since our proposed scheme belongs to the second category, in this section, we briefly review all the existing CNN-based pore detection schemes so as to have a better appreciation of the ideas and rationale behind the proposed pore detection scheme to be presented in the next section and to understand the differences between the proposed and these existing schemes.

The pore detection scheme of [22] is the first CNN-based scheme, in which the first part is based on a simple CNN architecture. In this architecture, two convolutional layers are employed using, respectively, 30 filters and one filter, each of kernel size 5×5, and a max-pooling layer with a 3×3 kernel in between the two layers. For determining the pore centroids in the second part of this scheme, a global binary threshold using Otsu's algorithm [29] is first used to convert the gray-level pore intensity map into a binary map and then each white region in this map is considered to be a pore region if its area is larger than a threshold area, which is determined empirically. Finally, the coordinates of a pore center are obtained as the mean of the coordinates of all the pixels in the respective pore region. In this pore detection method, there are a number of factors that lead to its performance, which has little improvement over the traditional schemes. First of all, this scheme uses only a shallow network, which relies only on its width to extract the pore features. Hence, it is not able to learn the deep high-level features of the pores. Since the intensities of the pores vary from one region to another region of a fingerprint image, the use of a single global binary threshold is not appropriate for the binarization of a pore intensity map, since it will result in inaccurate estimation of the pore regions, and therefore, incorrect determination of the pore centroids. Another factor that affects the accuracy of the pore centroids is due to the fact that a pore centroid is determined using only the coordinates of the pixels in the pore region and ignoring the intensity values of the pixels.

The scheme of [23] is another CNN-based pore detection scheme in which a deep network, referred to as DeepPore, consisting of 10 convolution layers is used. Each of the layers in this network with the exception of the last one (the reconstruction layer) uses 64 filters of kernel size 3×3. In the second part of the scheme, a spatial filtering-based approach is used to detect the coordinates of the pores directly from the gray-level pore intensity map. A pore intensity map is first divided into 11×11 windows and then the coordinates of the pixels within a window with maximum intensity are considered to be the centers of the pores within that window provided this maximum intensity is equal to or larger than a threshold intensity value. The same threshold is used for all the windows and its value is empirically determined to be 0.25. In view of the fact that this scheme uses a deep network, it is capable of extracting high-level pore features and provide performance superior to that provided by the scheme of [22]. However, this scheme is not able to improve the performance further by making the network deeper, since the network cannot handle the gradient vanishing problem associated with a very deep network. The use of a single global threshold for all the blocks and the maximum intensities of the individual blocks for determining the pore centroids should result in ignoring the true pores in fingerprint blocks with very low intensity. Also, the use of a single maximum intensity for a given block may result in ignoring some of the true pores in that block that do not have their intensities equal to that maximum intensity of the block.

The pore detection scheme of [24] uses a very deep CNN, referred to as DeepResPore, in its first part. This network employs a total of 18 convolutional layers, with the first one using 64 filters each of kernel size 7×7 and the last one (the reconstruction layer) employing a single filter with kernel size 3×3. The main body of this network consists of 8 residual blocks each containing two convolutional layers. The kernel size of each of the filters used in the residual blocks is also 3×3. The number of filters employed in each of the convolutional layers of the first two residual blocks is 64 and these numbers are doubled in the convolutional layers of the succeeding pairs of the residual blocks. The second part of this scheme is the same as the one used in the scheme of [23] with the only difference of using a smaller size window for the partitioning the pore intensity map. The network of this scheme is able to handle the gradient vanishing problem in view of its residual architecture. Thus, by using a larger number of filters for each layer, this network is able to provide a performance much superior to that provided by DeepPore. However, this improvement in performance is achieved at the expense of very large complexity in terms of the number of parameters and multiply-add operations. The disadvantages of this scheme resulting from the second part are the same as in the scheme of [23] since they both used the same methods in their second parts.

3 PROPOSED METHODOLOGY

Fig. 2 shows the block diagram of a neural network-based pore detection scheme (hybrid scheme) depicting its two-parts. The first part is the pore feature extraction part using a convolutional neural network, and the second part is a postprocessing part that determines the centroids of the true pores. The input to the pore detection system is a grey-level fingerprint image $\boldsymbol{X_i}$, and the output is a map $\boldsymbol{X_t}$ showing the centroids of the detected pores. The first part of the pore detection scheme extracts the features at various levels using a neural network from the input fingerprint image $\boldsymbol{X_i}$ and then uses them to constructs a single grey-level pore intensity map $\boldsymbol{X_o}$. The postprocessing part of the pore detection scheme is divided into two steps, Step 1 and Step 2. Step 1 is a pore centroid detection step, which receives the pore intensity map $\boldsymbol{X_o}$ as input and first transforms it into a binary pore map in which a value of 1 indicates that the pixel in question belongs to one of the pore regions, whereas a value of zero indicates that it does not. Using this binary pore map, the centroid of each candidate pore is determined. Thus, the output of step 1 of the second part of the scheme is the map $\boldsymbol{X_c}$ giving the coordinates of the centroids of each candidate pores. Step 2 is a pore refinement scheme, in which some of the candidates pores falsely detected as pores are discarded to yield a pore map $\boldsymbol{X_t}$ representing the coordinates of the centroids of the detected true pores. We next develop the proposed neural network-based scheme in the framework of the block diagram of Fig. 2.

3.1 Part 1: Pore Feature Extraction

In this subsection, a CNN model for pore extraction is proposed. The design of this model aims at developing a feature extraction process that is robust enough to extract a very rich set of pore features in a computationally efficient manner. This is achieved through hierarchical feature extraction locally using convolutional layers in a residual framework and through their concatenative fusion globally.

Our objective of developing a very lightweight network with the capability of providing a rich set of pore features leading to a highly representational pore intensity map is based on the following principles.

1. The network is designed to have three stages, stages 1 to 3, that produce, respectively, low, middle and high levels of hierarchal pore features.
2. A concatenative fusion of the three levels of features must produce a very rich set of pore features for the construction of a highly representational pore intensity map.
3. Recognition that the generation of a pore intensity map is essentially a classification problem, in which each pixel is classified to belong or not to belong to a pore region. A classification problem can be better served in a convolutional network by using depthwise convolution rather than the standard convolution.
4. A major consequential advantage of depthwise convolution is that its use leads to the development of a CNN architecture with a dramatically reduced complexity.

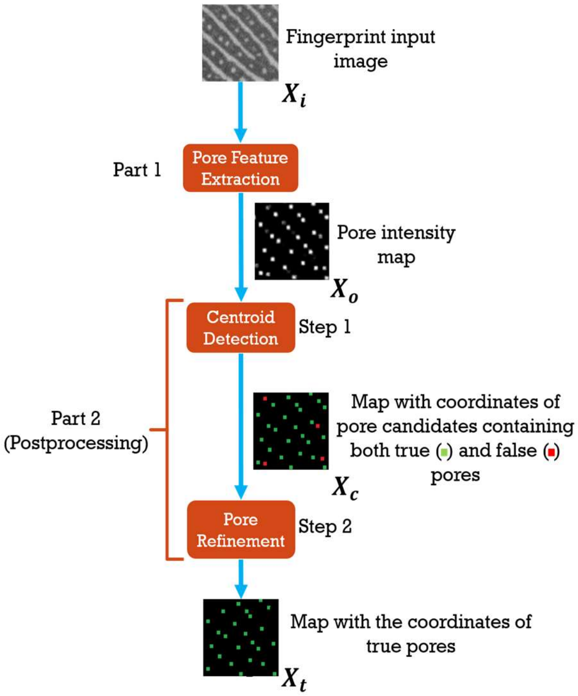


Fig. 2. The general structure of a neural network-based pore detection scheme. The approach takes a fingerprint image as an input and detects the true pore coordinates.

The proposed network model consists of four stages, as shown in Fig. 3. The first three stages carry out the processes of feature extraction and their fusion using a local skip connection. The last stage performs a global feature fusion of the features extracted locally at three hierarchal levels by the first three stages. Each of the three feature extraction stages consists of two convolution layers, and a module to residually learn the features locally. The skip connections used in each of the three stages curtail the gradient vanishing problem of a deep network [25], [26], [27] leading to better training of the network parameters, and hence, generation and extraction of more accurate features. The features produced by stages 1, 2 and 3 can be considered to be, respectively, the low, middle and high level pore features in view of the hierarchical levels of the convolutional layers used in these stages.

In order to achieve sufficient diversity among the extracted features of different channels and yet to keep the complexity low, we apply eight different 3×3 kernels on the input image $\boldsymbol{X_i}$ to obtain eight distinct channels in the first layer of stage 1. Each of the eight resulting feature maps of the first layer of this stage is operated on by another 3×3 kernel. Thus, the second layer of this stage also employs eight different kernels to produce eight distinct maps. The additive residual fusion of the corresponding maps of the two layers then yield the output feature maps $\boldsymbol{X_1}$ of the first stage.

In order to provide further diversity in the extracted features while keeping the complexity of the network low, we double the number of filters, in each of the other two feature extraction stages from that in the previous stage. Specifically, in the first layer of stage 2, we apply two 3×3 kernels on each of the eight maps of $\boldsymbol{X_1}$ resulting from stage 1 and one 3×3 kernel in the second layer to each of the extracted maps from the first layer of this stage. Thus, each of the two layers of stage 2 uses 16 3×3 different kernels. Similarly, in stage 3, two kernels are ap-

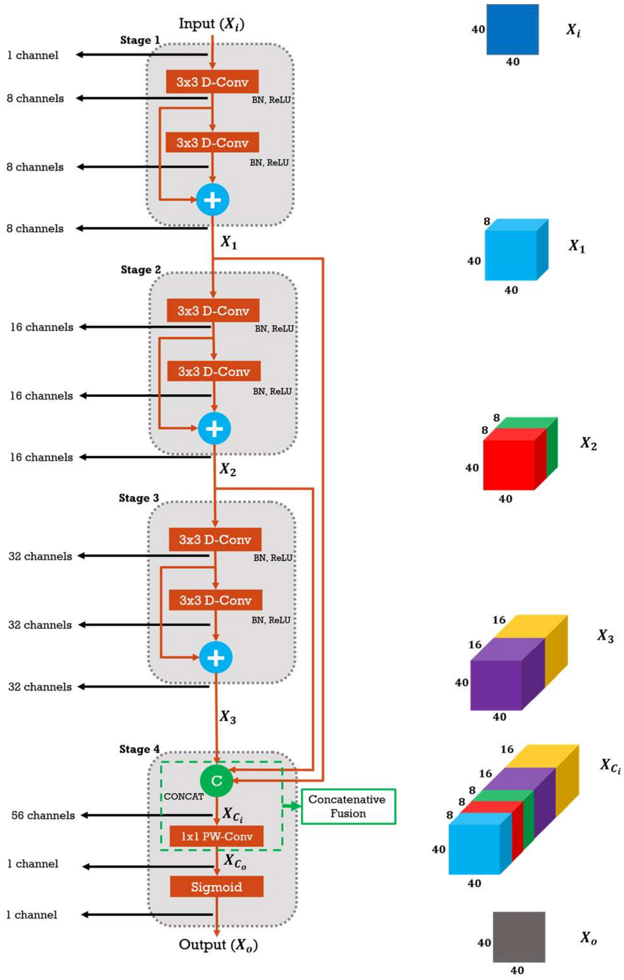


Fig. 3. The proposed network architecture.

plied to each of the 16 maps of X_2 and one 3×3 kernel to each of the extracted maps from its first layer. Hence, each of the two layers of stage 3 employs a total of 32 3×3 different kernels, a number that is twice that used by the layers in stage 2. In the first three stages, the batch normalization (BN) [30] and the non-linear activation function (ReLU) [31] are performed after each convolution operation in order to normalize the distribution of the features produced by the previous layer.

The main function of the last stage, stage 4, is two-fold. First, it should suitably combine the feature maps produced by the first three stages so that the significance of the individual features is taken into account to produce a pore map. Second, the values of the individual features should provide a reasonable indication of its likelihood of belonging to a pore. The feature maps produced by the different feature extraction stages are comparatively more diverse than the feature maps produced locally by a single stage. Therefore, a weighted combination of the feature maps X_1 , X_2 and X_3 should be carried out. This purpose is effectively served by performing a point-wise convolution of the feature map $X_{C_i} = (X_1 \ X_2 \ X_3)^T$, since this type of convolution provides a mechanism of summation of the corresponding pixels in the maps of X_{C_i} with weights that are learnable. Thus, in our proposed scheme, a point-wise convolution is performed on X_{C_i} using a single filter of kernel size 1×1×56 pixels in order to produce a single map X_{C_o}. Now, as mentioned earlier, each pixel value of the network output X_o must be indicative of its likelihood of belonging to a pore. Therefore, a pixel with a large positive value can be regarded to belong to a pore with high certainty, whereas a pixel with a negative value but a large magnitude can be regarded to belong to a pore with very low certainty. On the other hand, the pixel values of X_{C_o} in the neighborhood of zero need some amplification for further consideration of their likelihood of belonging to a pore. Essentially, this last step of stage 4 in going from X_{C_o} to X_o is a classification problem in which the likelihood of a pixel belonging to a pore is determined. In our proposed scheme, we achieve this goal by applying a sigmoid activation function [32] on X_{C_o} to yield X_o. Thus, each pixel value of X_o lies in the range 0 to 1 indicating its likelihood of belonging to a pore.

3.2 Part 2: Postprocessing

In this subsection, a method is presented for accurately determining the coordinates of all the true pores from the pore intensity map X_o obtained from the proposed pore feature extraction network presented in the previous subsection, while attempting to discarding the false pores. This postprocessing part of the proposed pore detection scheme consists of two steps as shown in the block diagram of Fig. 4. Step 1 consists of an image binarization module that transforms the pore intensity map X_o into a binary map X_b and a pore determination module that obtains a candidate pore centroid map X_c from X_b. Step 2 consists of a pore refinement process that identifies the true pores and removes false pores from X_c leading to a final pore map X_t containing the centroids of the true pores in the fingerprint image.

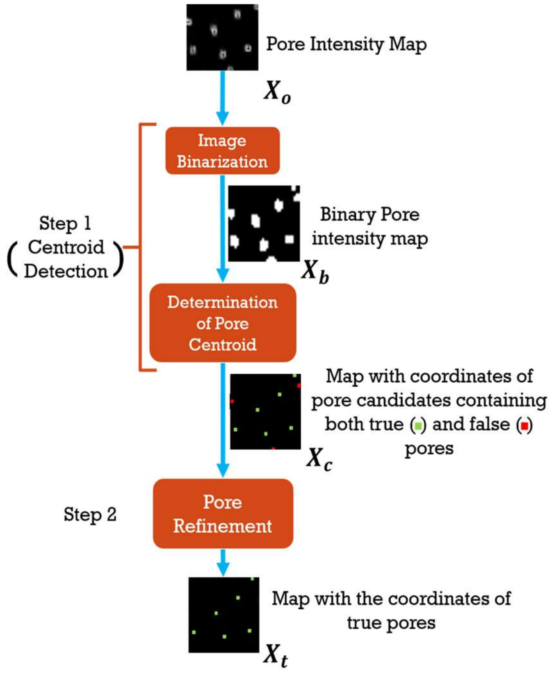


Fig. 4. The structure of the proposed postprocessing part.

3.2.1 Step 1: Centroid Detection

A. Image binarization

Since the pore intensity map X_o gives only a measure of the likelihood of its individual pixels to belong to a pore, it is necessary to differentiate more clearly whether or not a pixel of X_o belongs to a pore region by assigning it a distinct binary value. Therefore, in order to obtain a binary map X_b from the pore intensity map X_o, we need a suitable threshold T so that a pixel with a value equal to or greater than T could be classified as a candidate belonging to a pore region by assigning them a value of 1. On the other hand, if a pixel in $\boldsymbol{X_o}$ has a value less than T, it is assigned a value of 0 and, therefore, regarded not to belong to a pore region.

First of all, it is to be noted that the intensity values in $\boldsymbol{X_o}$ are directly related to the intensity distributions of the various pores within a fingerprint image $\boldsymbol{X_i}$. Even though the majority of the pores in the input image $\boldsymbol{X_i}$ have the same shape and intensity pattern, there are always some pores whose shapes and intensity values may differ, and hence, the intensity patterns of such pores may be different from that of a typical normal pore. The distribution of the pixel values in $\boldsymbol{X_o}$ is necessarily affected by such a nonuniformity of the pore shapes and intensity values in $\boldsymbol{X_i}$. Thus, if a fixed threshold T is chosen for the entire map $\boldsymbol{X_o}$ to convert it to a binary map $\boldsymbol{X_b}$, it may result in classifying some of the pixels of $\boldsymbol{X_o}$ to be falsely classified to belong or not to belong to a pore region. Hence, a method for determining the threshold must take into account the non-identically distributed nature of the pore distribution across the pore intensity map $\boldsymbol{X_o}$. In the following, we describe a scheme for determining a spatially adaptive threshold T_l to binarize the pore intensity map $\boldsymbol{X_o}$, that is, it is adaptive from one window to another window of $\boldsymbol{X_o}$. When this threshold T_l is used for the window l of $\boldsymbol{X_o}$, then $\boldsymbol{X_o}$ is converted to a binary map $\boldsymbol{X_b}$.

Window-based locally adaptive binarization methods, which compute a threshold value for all the pixels within a given window, based on the information contained on the pixels value within the window, exist for applications other than pore detection. The Niblack algorithm [33] has been successfully used to determine a locally adaptive threshold T_l to identify the region of an object in document images. In this algorithm, the mean m_l and the standard deviation σ_l in the window l are used to determine a local value for the threshold T_l to be used for all the pixels within the window l, as

$$T_l = m_l + \beta\sigma_l \tag{1}$$

with the parameters m_l and σ_l computed as

$$m_l = \frac{1}{N_l}\sum_{k=1}^{N_l} p_k \tag{2}$$

$$\sigma_l = \sqrt{\sum_{k=1}^{N_l}(p_k - m_l)^2 \Big/ N_l} \tag{3}$$

where p_k is the gray value of the k^{th} pixel in the l^{th} window of $\boldsymbol{X_o}$, N_l is the number of pixels in the l^{th} window, and β, called the Niblack factor, is a parameter that controls the effect of the local standard deviation in determining the boundary of an object. In [33], the value for β was empirically determined to be -0.2. The use of this fixed value of β (i.e. a non-adaptive value) in determining the threshold T_l leads to a successful determination of the object regions in images that are purely text documents. In [34], the Niblack factor was also made locally adaptive by selecting its value using the expression given by

$$\beta_l = -0.3 \times \frac{\left((m_g\sigma_g) - (m_l\sigma_l)\right)}{\max(m_g\sigma_g\,, m_l\sigma_l)} \tag{4}$$

where m_g and σ_g are the global mean and standard deviation of the entire map. Thus, the Niblack factor β_l is made to change from window to window but it remains fixed within a window. In this paper, for the purpose of transforming the pore intensity map $\boldsymbol{X_o}$ to a binary map $\boldsymbol{X_b}$, we employ the adaptive threshold T_l given by (1) with the adaptive Niblack factor β_l of (4). For the sake of simplicity, we use windows of fixed size $N \times N$ for the entire pore intensity map $\boldsymbol{X_o}$, i.e., $N_l = N$. The intensity and the pore shapes in a typical fingerprint image vary more globally than locally. Therefore, choosing a large window size would fail to capture this global variation in the fingerprint image. On the other hand, choosing a small size window would result in classifying the regions of the fingerprint image with noise as pore regions. Also, a small window size would adversely impact the processing time. Hence, the window size should be chosen suitably.

B. *Determination of centroids of candidate pores*

A single pore belongs to a specific region in the fingerprint image. The pixels belonging to this region cannot be disconnected. Therefore, typically a region of the binary map $\boldsymbol{X_b}$ that has, say, N_p pixels with the values of 1 and are connected can be considered to be a single pore region. However, it should be pointed out that this definition of pore region is not absolutely accurate, since for instance, if N_p is too small, the region may or may not considered to be a pore region, in view of the fact that a connected region of $\boldsymbol{X_b}$ with too few pixels could possibly be noise. On the other hands, if N_p is too large, it may be a situation where the region in question may be a region corresponding to two or more pores merged together. As to how such regions of the binary map are dealt with will be discussed in the next subsection. For the time being, we will consider a set of pixels of $\boldsymbol{X_b}$ that have a binary value of 1 and are connected to represent the region of a single pore. Detection of a pore means finding the coordinates of a pixel in the corresponding pore region that can be used to describe the physical location of the pore. In the previous works, pore centroids have been used to describe the locations of the pores. The centroid of a pore has been computed using only the coordinates of the pixels that belong to the pore region. However, the human visual system while determining the location of a pore focuses not only on the geometric center of the pore but it is also influenced by the intensity variation of the pixels in the pore region. Hence, the pore centroid is not a very accurate representation of the pore location. Consequently, in our scheme, the pore location is represented by its weighted centroid. We compute the weighted centroid by using the physical geometry of the pore region defined as a set of pixels that have a value of 1 and are connected in the binary map $\boldsymbol{X_b}$ as well as the pore intensity information of the underlying pixels from the fingerprint image itself. Accordingly, the weighted centroid (X, Y) of a candidate pore is computed as

$$(X, Y) = \left(\frac{\sum_{v=1}^{v=N_p}(X_v P_v)}{\sum_{v=1}^{v=N_p} P_v} \quad , \quad \frac{\sum_{v=1}^{v=N_p}(Y_v P_v)}{\sum_{v=1}^{v=N_p} P_v} \right) \tag{5}$$

where(X_v, Y_v) represents the coordinates of the v^{th} pixel in the pore region of the candidate pore from the binary map $\boldsymbol{X_b}$, P_v is the pixel gray value of the corresponding pixel in the input image $\boldsymbol{X_i}$, and N_p is the number of pixels within the pore region of the candidate pore. This process of determining the weighted centroids of the candidate pores results in the map $\boldsymbol{X_c}$ which consists of the locations of the detected candidate pores.

3.2.2 Step 2: Pore Refinement Process

In principle, one would expect that a fingerprint image has as many pores as the number of pore centroids in the map $\boldsymbol{X_c}$. However, in practice each centroid in $\boldsymbol{X_c}$ does not necessarily represent a true pore. In order to distinguish between true and false pores, we first make certain observations on the true pores in ground truth fingerprint images that will help in distinguishing a true pore from a false one.

The first observation is that a true pore has a distance from its nearest neighbor which varies from pore to pore. Fig. 5 shows an example of a ground truth fingerprint image. In this figure, the nearest neighbor of pore 1 is pore 2 and that for pore 3 is pore 4. It is seen from this example that the distance between 1 and 2 is different from that between 3 and 4. Fig. 6 shows the histograms of the frequency (numbers) of pores whose Euclidian distance from its nearest neighbor falls in a given range in all the 30 ground truth images of the Poly High-Resolution Fingerprint database [35]. For example, there are 1,287 pores in all the fingerprint images in this database that have Euclidian distances that lie in the range [11,12) from their nearest neighbors. It is seen from Fig. 6 that in none of the fingerprint images in this database there is a pore with a Euclidian distance less than 4.12 from its nearest neighbor. We denote this minimum distance by d_{min}. Hence, in a weighted centroid map $\boldsymbol{X_c}$, if a pair of neighboring pores is detected with Euclidian distance less than d_{min}, both the pores in the pair cannot be simultaneously regarded to be true pores. Note that the conclusion of d_{min}= 4.12 is reached based on 12,767 pores contained in the images of Poly High-Resolution Fingerprint database. We believe that the same conclusion would be valid for the pores in the fingerprint images from any other dataset with the value of d_{min} multiplied by a factor equal to the ratio between the resolution level of the images in the dataset considered in our study and that of the images in the other dataset.

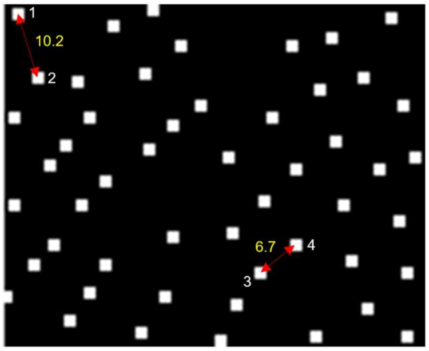


Fig. 5. Two different pairs of true pores and their corresponding Euclidian distance value in the ground truth fingerprint image.

The second observation is that, in different windows of a fingerprint image, there could be significant differences in the average intensities of their pores as well as in the variation of the intensities of the pores from the average intensity of the pores within the window. Fig. 7 shows two windows each of size 40×40 of a fingerprint image. It is seen from the fingerprint window in Fig. 7 (a) that the mean intensity of its pores is 77.3 and the variation of the pore intensities from this mean, represented by the standard deviation, is large. On the other hand, it is seen from the fingerprint window of Fig.7 (b) that the mean intensity of its pores is smaller, and the variation of the pore intensities is also larger.

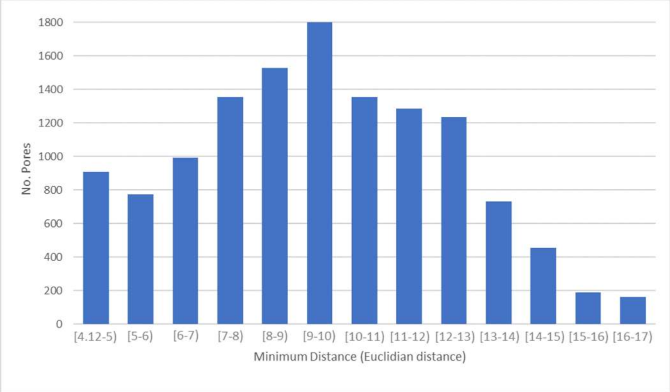
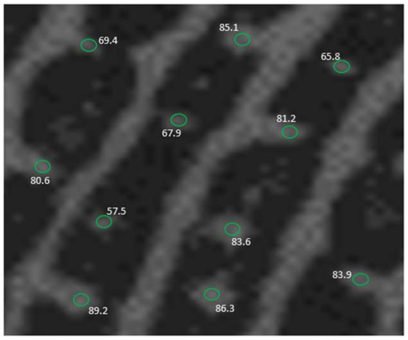
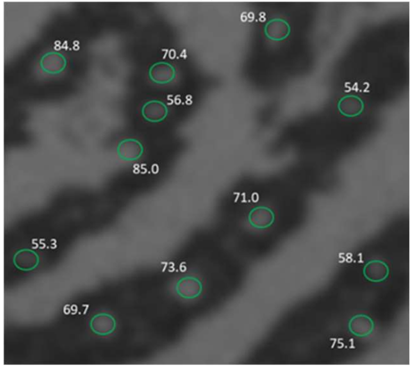


Fig. 6. A histogram of the minimum distance between two nearest pores in all fingerprint images of Poly High-Resolution Fingerprint database.



Mean = 77.3, STD = 9.85

(a)



Mean = 65.9, STD = 10.1

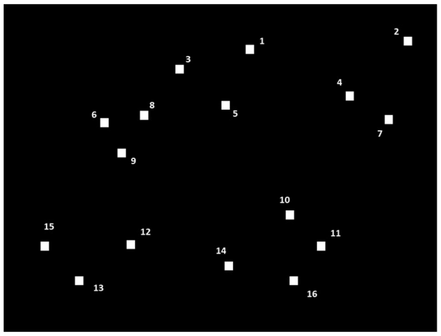
(b)

Fig. 7. Input fingerprint images with the intensity values of the true pores from the ground truth.

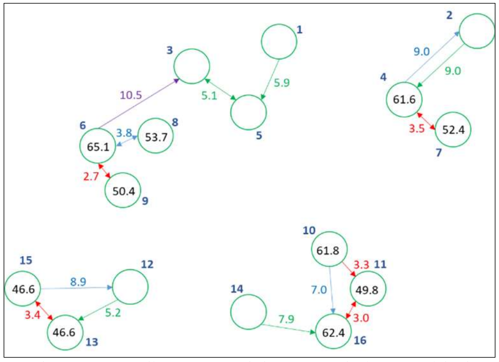
Both of the above observations could be used to distinguish true and false pores in a fingerprint image, and therefore, the weighted centroid map $\boldsymbol{X_c}$ could be refined by eliminating from it some of the centroids that correspond to false pores. The first observation that gives rise to the parameter d_{min} can be used to require the Euclidian distance between a pair of centroids in $\boldsymbol{X_c}$ to have a minimum value d_{min} for both the pores in the pair to be true pores. Based on the second observation, the mean and standard deviation values of the pore intensities in a given window in $\boldsymbol{X_c}$ have values that vary from window to window. Therefore, if the intensity of a pore in any given window has a value that is much smaller than the mean intensity of the pores in the window, then such a pore cannot be regarded to be a true pore, and hence, its centroid must be removed from $\boldsymbol{X_c}$. For the purpose of deciding whether the intensity of the pore is much smaller than the mean intensity of the pores within the window, its local standard deviation can be used. We now develop a criterion for each of the two observations to use it for refining $\boldsymbol{X_c}$.

A. *Procedure for identifying and removing false pores from X_c using the d_{min} criterion*

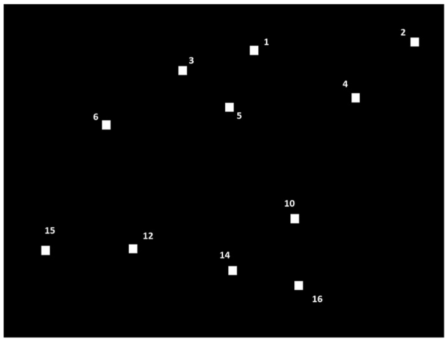
In this section, we develop a systematic procedure for identifying as many false pores as possible in the pore map $\mathbf{X_c}$ based on the minimum distance criterion between a pair of pores. We describe our procedure through an illustrative example. For this purpose, we make use of a manually constructed pore map $\mathbf{X_c}$ rather than a pore map corresponding to a natural fingerprint image so that all the different possibilities of pores can be illustrated through a single example. Fig. 8 (a) shows one such pore map $\mathbf{X_c}$ with 16 centroids numbered 1 to 16.



(a) An example of a pore map X_c with 16 candidate pores.



(b) Pore graph corresponding to X_c of the example in (a) with the distances of the nearest neighbors specified.



(c) Pore map $X_{d_{min}}$ corresponding to the pore map X_c in (a) after applying the d_{min} criterion.

Fig. 8. Illustrations of the proposed refinement of the first criterion.

Our objective is to construct from this pore map $\mathbf{X_c}$ a refined pore map $\mathbf{X}_{\boldsymbol{d_{min}}}$(Fig. 8(c)) from which all the candidate pores in $\mathbf{X_c}$ that do not satisfy the minimum distance requirement are removed. Initially $\mathbf{X}_{\boldsymbol{d_{min}}}$is set to be $\mathbf{X_c}$. Then, starting from pore 1, we determine its Euclidian distance E_D *(1, j)* to all the other pores j (j=2,…,16) and choose the pore that has the minimum distance from pore 1 as its nearest neighbor using the centroid values Pc_k (k=1,..,16). In the example considered, the nearest neighbor of pore 1 is found to be pore 5. This is shown in Fig. 8(b) by connecting pore 1 to pore 5 with a green arrow directing from the former pore to the latter with the minimum distance of 5.9 indicated on the arrow. Since E_D *(1,5)* between these two pores is larger than $d_{min} = 4.12$, both these pores are retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. The same procedure is followed for pores 2 and 3 to find their nearest neighbors, pores 4 and 5, with E_D *(2,4) = 9* and E_D *(3,5) =5.1*. It is noted that these pairs of pores also do not violate the minimum distance criterion, and hence, they are also retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. For the next pore, i.e., pore 4, the nearest neighbor is found to be pore 7 with E_D *(4,7) = 3.5,* which is less than d_{min}. Therefore, one of these pores in this pair must be removed from $\mathbf{X}_{\boldsymbol{d_{min}}}$. In order to decide as to which pore in the pair has to be removed, we determine their pore intensities γ_4 and γ_7. The intensity of a pore j is found to be the average of the intensities of all the pixels in a 3×3 window centered at the centroid of pore j in the original fingerprint image $\boldsymbol{X_i}$ corresponding to $\mathbf{X_c}$. The intensity values γ_4 and γ_7 for pores 4 and 7 are, respectively, found to be 61.6 and 52.4, as indicated in Fig. 8(b). Since γ_7 is less than γ_4, pore 7 is removed from $\mathbf{X}_{\boldsymbol{d_{min}}}$ of Fig. 8(c). As pore 7 was the nearest neighbor of pore 4 and it has been removed, we need to find a new nearest neighbor of this latter pore. The new nearest neighbor of pore 4 is found to be pore 2 with E_D *(4,2) =9* indicated by a blue arrow in Fig. 8(b). Since E_D *(4,2)* $> d_{min}$, pore 4 is still retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. Next, since pore 5 does not violate the minimum distance criterion, it is retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. As for pore 6, its nearest neighbor is pore 9 with E_D *(6,9) = 2.7* $< d_{min}$. Since γ_6 =65.1 and γ_9 =50.4, pore 9 is removed from $\mathbf{X}_{\boldsymbol{d_{min}}}$ and the new nearest neighbor of pore 6 is found to be pore 8. However, since E_D *(6,8) = 3.8* is also less than d_{min}, the pore in the pair with the lower intensity, i.e., pore 8, is removed from $\mathbf{X}_{\boldsymbol{d_{min}}}$. Thus, we find the third nearest neighbor of pore 6, which is pore 3 with E_D *(6,3) =10.5* $> d_{min}$, indicated by a purple arrow, and consequently, both pores 6 and 3 are retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. Since pores 7, 8 and 9 are already removed from $\mathbf{X}_{\boldsymbol{d_{min}}}$, next, we consider pore 10. For this pore, the nearest neighbor is pore 11 with E_D *(10,11) =3.3* $< d_{min}$, and since $\gamma_{11} < \gamma_{10}$, pore 11 is removed from $\mathbf{X}_{\boldsymbol{d_{min}}}$. Thus, we find the next nearest neighbor of pore 10, which is pore 16 with E_D *(10,16) =7* $> d_{min}$, and consequently, both pores 10 and 16 are retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. For the next pore, i.e., pore 12, the nearest neighbor is pore 13 with E_D *(12,13) =5.2* $> d_{min}$, and thus, it is retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. Next, the nearest neighbor of pore 13 is found to be pore 15 with E_D *(13,15) =3.4* $< d_{min}$, thus, the pair of pores in question violating the minimum distance criterion. However, since in this case $\gamma_{13} = \gamma_{15} =$ 46.6, we remove the first pore in the pair, i.e., pore 13, from $\mathbf{X}_{\boldsymbol{d_{min}}}$. The last three pores, i.e., pores 14, 15 and 16, as seen from Fig. 8(b), do not violate the minimum distance criterion, and hence, retained in $\mathbf{X}_{\boldsymbol{d_{min}}}$. After taking into consideration all the pores in Fig. 8(b) and removing all those that violate the minimum distance criterion, $\mathbf{X}_{\boldsymbol{d_{min}}}$ shown in Fig. 8(c) is the final composition of the refined pore map.

The procedure for identifying the false pores by applying the minimum distance criterion and removing the corresponding centroids from $\boldsymbol{X_c}$ is given in Algorithm 1. This algorithm presents our procedure for identifying and removing the false pores from $\boldsymbol{X_c}$ by applying the d_{min} criterion. The input to this algorithm is $\boldsymbol{X_i}$, $\boldsymbol{X_c}$ and d_{min}, and the output is the refined pore map $\mathbf{X}_{\boldsymbol{d_{min}}}$ in which the centroids in $\boldsymbol{X_c}$ corresponding to the pores not satisfying the d_{min} criterion has been removed.

Algorithm 1 Algorithm for identifying and removing the false pores from X_c using d_{min}

```
Input: Fingerprint image $X_i$, pores centroids map $X_c$, $d_{min} = 4.12$
Output: The pore map $X_{d_{min}}$ obtained by refining $X_c$ using $d_{min}$ creterion.
    .The pores with centroids $P_i$ in $X_c$ are listed from 1 to $V$
    .The intensity of each pore $\gamma_i$ is calculating by taking the average of the intensities of the pixel values from $X_i$ in the 3x3 window at the center of each pore from $X_c$
    .$X_{d_{min}}$ is an initial copy of $X_c$
 1: for $i = 1$ to $V$ from $X_c$ do
 2:   if ($i$ in $X_{d_{min}}$) then
 3:     Find the nearest neighbor pore $j$ in $X_{d_{min}}$, where $j \neq i$, with the smallest Euclidean distance $E_D(i, j)$ using $P_i$ and $P_j$
 4:     if ($E_D(i, j) < d_{min}$) then
 5:       if ($\gamma_i > \gamma_j$) then
 6:         Remove $j$ from $X_{d_{min}}$
 7:         Return to step 3
 8:       else if ($\gamma_i \leq \gamma_j$) then
 9:         Remove $i$ from $X_{d_{min}}$
10:       end if
11:     end if
12:   end if
13: end for
14: return $X_{d_{min}}$
```

B. *Procedure for identifying and removing the false pores from $X_{d_{min}}$ based on the criterion of local mean and standard deviation*

In this section, we develop a procedure for identifying as many false pores as possible in the pore map $\mathbf{X}_{\boldsymbol{d_{min}}}$ based on the local means and standard deviations of the pores corresponding to the centroids still remaining in the map $\mathbf{X}_{\boldsymbol{d_{min}}}$ after carrying out the procedure of Section 3.2.2.A. Generally, the intensity of a false pore is lower than that of a true pore in a given neighborhood of the fingerprint image. The standard deviation of the pore intensities is an indication of the degree by which the intensity of a true pore could be lower than the mean intensity in the neighborhood. In our method for removing a false pore, we require that its intensity must be less than a threshold value given by $t =$ ŋ $- c$б, where ŋ and б are the mean and standard deviation of the pore intensities within a given neighborhood of the fingerprint image and c is a parameter that needs to be determined suitably. We determine the value of c empirically for the entire dataset. The images of the dataset are divided into $W \times W$ windows. First, for a given image in the dataset, we determine a value of c so that the accuracy that all the pores removed from that image are indeed false pores is more than 96%. Then, a value c_{avg}=1.12 that is the average of the values of c so determined for the individual images is used to remove pores from each of the windows of a test image. Removing all the false pores from $\mathbf{X}_{\boldsymbol{d_{min}}}$ using this procedure results in the final pore map $\mathbf{X}_{\boldsymbol{t}}$.

4 EXPERIMENTAL RESULTS

In this section, first, the dataset used and the procedure for the training of the proposed network along with the software and hardware platforms and the figures of merit for the evaluation procedure of the proposed and other schemes considered are described. Then, the results of the various experiments performed using the proposed pore detection scheme are presented and analyzed, and its performance is compared with that of the existing state-of-the-art schemes for pore detection.

4.1 Training Details and Evaluation Procedure

In our experiments, images from the PolyU High-Resolution-Fingerprint (HRF) database [35], are used for the training and testing of the proposed network and for the comparison of its performance with that of the other networks. The images in this database are acquired using an optical touch-based sensor device. The database has 30 fingerprint images of size 240×320 pixels with a resolution of 1200 dpi. The database also has a ground truth set containing the coordinates of the central position of each pore of the fingerprint images, as perceived by human experts. This ground truth set contains a total of 12767 coordinates of the pores in all the fingerprint images in the database.

The fingerprint images are divided into training and testing sets using k-fold cross-validation [36] with k = 5. Therefore, in our experiments, we have five different training and testing sets. As in [22], [23] and [24], we make sure that none of the 6 fingerprint images in each of the 5 testing sets are repeated and they are different from those in other testing sets. Thus, we have five different training models of the proposed network each of which is evaluated using a unique testing set. The results of the five evaluations are then averaged to obtain the performance of the network.

In the proposed network, we generate a labeled pore intensity map corresponding to a ground truth fingerprint image by setting all the pixel values in the 3×3 window centered at the coordinates of each true pore of the ground truth image to 1, and setting all the remaining pixels in the ground truth image to a value of 0. The fingerprint images in the training set as well as in their corresponding labels are partitioned into $p \times p$ patches using a random patch extraction with overlapping method [37], which ensures that each patch stays inside the image and all the patches together in an image cover the entire image. Therefore, in this method, regardless of the value of p, there does not exist a boundary problem. A set of 142,800 patches is obtained by randomly extracting 5,950 patches in each of the training fingerprint images. Some of the patches of a fingerprint image may overlap, i.e., the patches may contain the same part of a fingerprint image. During the testing phase, however, each test image is partitioned into non-overlapping patches of size $p \times p$. In this case, depending on the value p, there may exist a situation in which a patch may not lie totally inside the image in order for all the pixels of the image to be covered by one or more patches. If in the formation of a patch on the boundary, there are not sufficient number of pixels within the image, then the boundary pixels are simply repeated to complete the formation of that patch. It should be pointed out that the same value of p is used both in the training and testing phases of the network. In the testing phase, the network outputs a pore map corresponding to a single patch of the given input test image. When the pore maps of all the patches of a test image have been obtained, they are placed in a non-overlapping manner to compose a pore intensity map corresponding to that test input image. Data augmentations [38] of the patches through their rotations by 90° and 180°, flipping and Gaussian blurring are utilized to enrich the training dataset and boost the network performance. The convolution kernels with spatial support of $s \times s$ are randomly initialized with a Gaussian distribution having a zero mean and a variance of $\sqrt{\frac{2}{s^2 n}}$, where n is the layer width. A binary cross-entropy loss function is optimized to train the parameters of the network. The binary cross-entropy loss function given by [39]

$$L = -\frac{1}{N_B}\sum_{i=1}^{N_B}(y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)) \quad (6)$$

where y_i represents the i$^{\text{th}}$ labeled patch in a batch of size N_B and $\hat{y}_i$ represents the corresponding predicted patch, is used to train the network. We employ the stochastic gradient descent (SGD) algorithm with the Nesterov acceleration scheme [40] to update the parameters and minimize the loss. In our training, the momentum parameter and initial learning rate are set as 0.9 and 0.1, respectively, and the learning rate is decreased by 0.01 after every 10 epochs until the rate becomes 0.01, beyond which it is not decreased farther. The network is trained with a batch size of 34 and a maximum number of 100 epochs. The facts that the proposed network is not very deep, it contains a modest number of parameters, the high-resolution training images help the training process [41], [42], [43], and that each training image is undergone through different kinds of augmentations, a training set constructed using the 24 fingerprint images is sufficient to adequately train the proposed pore detection network. The experiments are performed in a Python environment on a supercomputer with 2.2 GHz Intel E5-2650 v4 Broadwell CPU, 125 GB RAM and NVIDIA P100 Pascal (12G HBM2 memory) GPU.

The figures of merit that are used for the performance evaluation are as follows:

1. True detection rate (R_T), which represents the ratio of the number of the true pores detected to the number of actual true pores present in the ground truth of a fingerprint image [16].
2. False detection rate (R_F) indicates the ratio of the number of false pores detected to the total number of detected pores [16].
3. The number of the parameters used in the CNN and the number of FLOPs, which indicate the computational cost of the CNN network.

The optimal values for R_T and R_F are one and zero, respectively. High true detection rate and low false detection rate are indicative of a superior performance of the pore detection system, and small standard deviations of R_T and R_F represent a good robustness of the system.

4.2 Performance Evaluation

The objective of this section is to study the performance of the proposed pore detection scheme and compare it with that of other schemes in the literature. We conduct this study in three parts. In the first part, we study the impact of the patch size of the images and the depth of the network on the performance of the proposed scheme. In the second part, we study the influence of replacing either the proposed network or the proposed post-processing method with the one used in other neural network-based schemes. In the third part, we compare both the quantitative and qualitative performance of the proposed scheme with those of the other schemes that are neural network-based or otherwise. The computational times of the network part and the postprocessing part of the proposed scheme along with its network complexity are also compared with those of the other network-based schemes.

In the first part of the performance evaluation, we first perform experiments on the proposed scheme by using different patch sizes of the input images in the dataset as input to the proposed network. Once the pore intensity maps for all the test images have been obtained, we move on to the second part of the proposed scheme that uses the hyperparameters N (the window size used in the binarization step) and W (the window size used in the refinement step). In this part of the scheme, for a test image, the pore intensity map obtained from the network patches trained by using the patches of a given size p is then used to obtain $\mathbf{X}_t$ using a given value of the parameter N and a given value of the parameter W. Using this final map of pore centroids, the values of R_T and R_F are obtained for a given test image. These values of the metrics are averaged over 30 test images. We have performed the experiments on the proposed scheme using different values for the patch size p (10, 20, 30, 40, 50, 60, 70, 80, 90, and 100) in Part 1 of the scheme and for each value of p with various values of the hyperparameters N (20, 30, 40, 50, and 60) and W (40, 50, 60, 70, and 80) in Part 2. Table 1 lists the average values of R_T and R_F and their standard deviations for each of the patch sizes used for the network training and testing, but only for those particular values of N and W that provide the best performance in terms of R_T and R_F for that patch size. From this table, it is seen that the proposed scheme of pore detection provides the best performance in terms of R_T and R_F for p =40, N =40 and W =60. Table 1 also

TABLE 1
PERFORMANCE OF THE PROPOSED SCHEME WITH DIFFERENT PATCH SIZES OF THE IMAGES INPUT TO THE NETWORK

Patch size	10 × 10	20 × 20	30 × 30	40 × 40	50 × 50	60 × 60	70 × 70	80 × 80	90 × 90	100 × 100
Hyperparameters	$N = 50$	$N = 40$	$N = 50$	$N = 40$	$N = 60$	$N = 60$	$N = 40$	$N = 40$	$N = 40$	$N = 60$
	$W = 60$	$W = 60$	$W = 70$	$W = 60$	$W = 80$	$W = 80$	$W = 60$	$W = 80$	$W = 60$	$W = 70$
R_T	91.84	93.16	95.97	**96.69**	96.51	96.43	95.93	95.51	94.22	93.96
(σ)	(1.32)	(1.48)	(1.49)	**(1.52)**	(1.83)	(1.99)	(2.57)	(2.74)	(3.2)	(3.51)
R_F	10.02	6.55	4.15	**4.18**	4.2	4.68	5.33	5.87	6.05	7.04
(σ)	(3.22)	(1.97)	(1.41)	**(1.44)**	(1.9)	(1.83)	(2.41)	(2.6)	(3.95)	(4.02)
Number of FLOPs	0.495M	0.99M	1.485M	***1.98M***	2.475M	2.97M	3.465M	3.96M	4.455M	4.95M

provides the network complexity in terms of the number of floating-point multiplication-addition operations (FLOPs). It is worth noting that the increase in the number of FLOPs by using a patch size beyond 40 × 40 actually results in deteriorating the performance of the scheme. Therefore, 40 × 40 is the optimum patch size for the best performance of the proposed scheme.

Next, we perform an experiment on the proposed network in which its depth is increased by adding additional feature extraction stages. The results are shown in Table 2. It is seen from this table that as the number of feature extraction stages in the network is increased beyond three stages, the improvement in its performance is only minimal. However, its complexity in terms of the number of FLOPs increases very significantly. Specifically, the number of FLOPs increases by more than 4 and 11 times when the number of feature extraction stages is increased to 4 and 5, respectively. On the other hand, if the number of feature extraction stages is reduced to two, the performance of the network is severely affected. Thus, the network provides optimum performance with three feature extraction stages.

TABLE 2

PERFORMANCE OF THE PROPOSED SCHEME WITH VARIOUS NETWORK DEPTHS

Number of Feature extraction stages	R_T (σ)	R_F (σ)	Number of FLOPs
2	91.68 (4.93)	6.62 (5.10)	0.86M
3	**96.69 (1.52)**	**4.18 (1.44)**	**1.98M**
4	96.89 (1.5)	4.03 (1.41)	9.49M
5	96.95 (1.49)	3.95 (1.40)	22.15M

In the second part of the performance evaluation, we examine the impact of the proposed network and the proposed postprocessing method individually on the performance of the proposed pore detection scheme, as well as on the performances of the neural network-based schemes of [22], [23] and [24]. Table 3 provides the results of this study in terms of the performance metrics R_T and R_F. The results in this table are divided into four blocks corresponding to the proposed and the schemes of [22], [23] and [24]. In each of the blocks 2, 3 and 4, the performance results of the corresponding scheme along with that of the same scheme in which either its network or its postprocessing method is replaced by the one proposed in our scheme. By comparing the first lines in each of the blocks of this table, it is observed that the proposed scheme gives the best performance among all the neural network-based schemes. By comparing the results of the first and second lines of the blocks 2, 3 and 4 of this table, it is seen that each of the schemes of [22], [23] and [24] can significantly benefit if the networks of these schemes are replaced by the proposed network. Similarly, by comparing the results of the first and third lines of the blocks 2, 3 and 4, it is seen that each of the schemes of [22], [23] and [24] can improve the performance if its postprocessing method is replaced by the proposed postprocessing method. In summary, the study of this experimental part clearly shows that both the proposed network and the postprocessing method have a significant impact on the performance of the proposed pore detection scheme.

TABLE 3

COMPARISON OF PERFORMANCES OF THE PROPOSED AND OTHER NEURAL NETWORK-BASED SCHEMES

Scheme	Detail of the scheme	R_T (σ)	R_F (σ)
Proposed scheme	Proposed network + proposed postprocessing	**96.69 (1.52)**	**4.18 (1.44)**
Scheme [22]	Network of [22] + postprocessing of [22]	84.69 (7.81)	15.31 (6.2)
Modified scheme [22] – 1	Proposed network + postprocessing of [22]	93.19 (3.04)	7.88 (3.55)
Modified scheme [22] - 2	Network of [22] + proposed postprocessing	88.21 (6.03)	5.46 (2.74)
Scheme [23]	Network of [23] + postprocessing of [23]	93.09 (4.63)	8.64 (4.15)
Modified scheme [23] – 1	Proposed network + postprocessing of [23]	95.47 (2.58)	6.52 (3.56)
Modified scheme [23] – 2	Network of [23] + proposed postprocessing	94.81 (3.05)	4.53 (2.07)
Scheme [24]	Network of [24] + postprocessing of [24]	94.49 (5.41)	8.5 (4.4)
Modified scheme [24] – 1	Proposed network + postprocessing of [24]	95.89 (2.42)	6.42 (3.41)
Modified scheme [24] - 2	Network of [24] + proposed postprocessing	95.22 (3.15)	4.38 (2.10)

In the third part of our experimental study, we compare the quantitative and qualitative performance of the proposed pore detection scheme with that of both the traditional schemes and neural network-based schemes. The traditional schemes that are used for comparison are those reported in [3] and [16], whereas the neural network-based schemes are the same as those used in Table 3, i.e., the schemes of [22], [23] and [24]. Table 4 gives the pore detection accuracy in terms of R_T and R_F provided by the various schemes. It is seen from this table that, as in the case of other neural network-based schemes with the exception of that reported in [22], the proposed scheme outperforms the traditional schemes by significant margins. It is to be pointed out that, since the network in [22] relies only on the network width to extract the pore features, it is not able to learn the high-level features, and hence, this network is not able to provide much improvement over the non-neural network-based methods. As already noted from Table 3, the proposed scheme provides performance much superior to that provided by the other neural network-based schemes. Table 4 provides the performance results using both with and without the pore refinement process introduced in Section 3.2.2. Note that the pore refinement process was intended to reduce the number of false pores. It is seen from the performance results given in Table 4 that this objective is achieved by providing the value of R_F that is almost one-half of that provided without using the pore refinement process while maintaining almost the same value for R_T.

TABLE 4

PERFORMANCE COMPARISON AGAINST STATE-OF-THE-ART METHODS

Scheme	R_T (σ)	R_F (σ)
Gabor Filter [3]	75.90 (7.5)	23.00 (8.2)
DoG [16]	80.80 (6.5)	22.20 (9.0)
DAPM [16]	84.80 (4.5)	17.60 (6.3)
Labati *et al.* [22]	84.69 (7.81)	15.31 (6.2)
DeepPore [23]	93.09 (4.63)	8.64 (4.15)
DeepResPore [24]	94.49 (5.41)	8.5 (4.4)
Proposed without Refinement	**96.85 (1.53)**	**8.69 (2.12)**
Proposed with Refinement	**96.69 (1.52)**	**4.18 (1.44)**

Table 5 demonstrates the average run times per image of the various CNN-based pore detection schemes for Parts 1 and 2 of the schemes individually as well as together. All the schemes have been implemented in the same software and hardware environment of the proposed scheme as stated in Section 4.1. It is seen from this table that the time complexities of the second part of all the four schemes are about the same. On the other hand, the proposed scheme takes a significantly lower time for its first part in comparison to the times taken by the other three schemes. This table also provides information on the complexity of the various neural network-based schemes in terms of the number of parameters and the number of FLOPs. Note that the number of parameters and the number of FLOPs refers only to the numbers used by the network parts of these schemes. It is particularly important to note that the superiority in performance by the proposed scheme is achieved along with its very significantly lower complexity. Specifically, the number of parameters and the number of FLOPs of the network in the proposed scheme are several orders of magnitude lower than those of the network in the scheme of [24], the best performing scheme in the literature.

Fig. 9 shows the qualitative performance of the proposed scheme and the schemes of [23] and [24] by providing visual illustrations of the pores detected. Fig. 9 (a) shows an original fingerprint image from the HRF dataset, whereas Figs. 9 (b), (c) and (d) show the results of the pore detection by the three schemes. In this figure, the pores shown in green, red and yellow colors represent, respectively, true positive, false positive and false negative pores. It is seen from this figure that the proposed scheme is more successful in detecting the true pores by providing only one false positive pore and much fewer false negative pores than that provided by the other two schemes.

TABLE 5
COMPARISON OF AVERAGE TIME PER IMAGE AND NETWORK COMPLEXITY OF THE PROPOSED AND OTHER NEURAL NETWORK-BASED SCHEMES

Scheme parts	T_{CNN} (ms) Part 1: Neural Network	T_P (ms) Part 2: Postprocessing	T_T (ms)	Number of Network Parameters	Number of FLOPs used by Network
Labati *et al.* [22]	2.76	15.54	18.3	1,531	50.03M
DeepPore [23]	29.85	16.1	46.36	297,361	335.12 M
DeepResPore [24]	218.02	16.24	234.67	11M	70,438.1M
Proposed scheme	**1.08**	**15.92**	**17**	**1,288**	**1.98M**

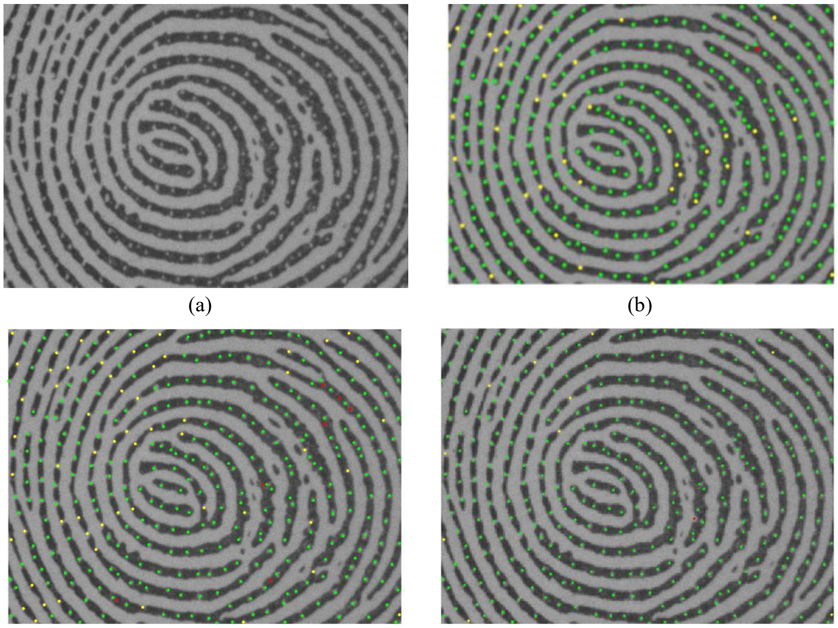


Fig. 9. (a) Original fingerprint image. The pore detection results of (b) DeepPore (c) DeepResPore (d) Proposed work. The green, yellow and red dots in these images represent the true positive, false negative and false positive, respectively.

5 CONCLUSION

The process of pore detection can be divided into two parts. In the first part, a pore intensity map is obtained by extracting the pore features from a gray-level input fingerprint image, whereas in the second part, the centroids of all the true pores in a fingerprint image are obtained from the pore intensity map. In recent years, pore detection schemes have been developed using a hybrid approach, in which a convolutional neural network architecture is used to obtain the pore intensity map in the first part of the schemes to provide a detection accuracy much higher than that provided by the traditional schemes. However, these schemes do not perform well in the presence of incomplete and distorted pores in the fingerprint image, in view of the fact that (i) the networks, despite their large complexity, are still not able to learn all the relevant pore features, and (ii) the second part, i.e., the postprocessing part is not able to take into consideration fully and efficiently the existing knowledge on the pores and pore fingerprint images. In this paper, a high-performance ultralight hybrid pore detection scheme has been developed. In the first part, a neural network architecture has been developed to generate and extract low, middle and high level hierarchical features in a residual framework and then to fuse them to obtain a very rich set of pore features that lead to a highly representational pore intensity map. In this part, the goals of high-performance and low complexity have been achieved by designing the network based on depthwise convolutional layers. The focus in the second part of the pore detection scheme has been on efficient use of the existing knowledge on fingerprint pores, such as the pore intensity, their variations from one region to another region of the fingerprint image and the minimum distance between two neighboring pores, in order to accurately determine the pore centroids from the pore intensity map.

The proposed pore detection scheme has been extensively experimented on the fingerprint images of the PolyU High-Resolution-Fingerprint database to evaluate its objective performance in terms of the true and false pore detection rates as well as the subjective quality of the detected pores. It has been shown that both the designed network and the postprocessing method devised for the proposed scheme have a significant impact on its performance. The performance of the proposed scheme has been compared to that of the traditional schemes as well as to that of the state-of-the-art convolutional neural network-based schemes and has been shown to outperform all of them very significantly. In terms of the subjective quality, it has been shown that the proposed scheme detects almost no false positive pores and very few false-negative pores in comparison to the detections provided by the other schemes. It has also been shown that other schemes can benefit in improving their performance by replacing their network part or the postprocessing part individually with the corresponding part of the proposed scheme. Finally, it is worth mentioning that the much superior performance of the proposed scheme has been achieved with the number of the parameters and the number of arithmetic operations employed by the proposed network that are only very small fractions of those employed by the other state-of-the-art schemes.

ACKNOWLEDGMENT

This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and in part by the Regroupement Stratgique en Microlectronique du Qubec (ReSMiQ). Computations were made on the supercomputer Hélios from Laval University, managed by Calcul Québec and Compute Canada. The operation of this supercomputer is funded by the Canada Foundation for Innovation (CFI), Ministère de l'Économie, des Sciences et de l'Innovation du Québec (MESI) and le Fonds de recherche du Québec – Nature et technologies (FRQ-NT).

REFERENCES

[1] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, "Introduction," in *Handbook of Fingerprint Recognition*, London, U.K.:Springer-Verlag, 2009.

[2] R. D. Labati and F. Scotti, "Fingerprint," in *Encyclopedia of Cryptography and Security*, H. C. A. van Tilborg and S. Jajodia, Ed., Boston, MA: Springer US, 2011, pp. 460-465.

[3] A. K. Jain, Y. Chen, and M. Demirkus, "Pores and Ridges: High-Resolution Fingerprint Matching Using Level 3 Features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 15–27, 2007.

[4] D. Zhang, F. Liu, Q. Zhao, G. Lu, and N. Luo, "Selecting a Reference High Resolution for Fingerprint Recognition Using Minutiae and Pores," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 3, pp. 863–871, Mar. 2011.

[5] K. M. Kryszczuk, P. Morier, and A. Drygajlo, "Study of the distinctiveness of level 2 and level 3 features in fragmentary fingerprint comparison," presented at *Int. Workshop on Biometric Authentication*, 2004, pp. 124–133.

[6] Q. Zhao, L. Zhang, D. Zhang, and N. Luo, "Direct pore matching for fingerprint recognition," in *Int. Conf. on Biometrics*, 2009, pp. 597–606.

[7] F. Liu, Q. Zhao, and D. Zhang, "A novel hierarchical fingerprint matching approach," *Pattern Recognit.*, vol. 44, no. 8, pp. 1604–1613, 2011.

[8] Y. Xu, G. Lu, Y. Lu, F. Liu, and D. Zhang, "Fingerprint pore comparison using local features and spatial relations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 10, pp. 2927–2940, 2018.

[9] Y. Xu, G. Lu, Y. Lu, and D. Zhang, "High resolution fingerprint recognition using pore and edge descriptors," *Pattern Recognit. Lett.*, vol. 125, pp. 773–779, 2019.

[10] Y. Xu, Y. Lu, G. Lu, J. Li, and D. Zhang, "Fast Pore Comparison for High Resolution Fingerprint Images Based on Multiple Co-Occurrence Descriptors and Local Topology Similarities," *IEEE Trans. Syst. Man Cybern. Syst.*, 2019.

[11] V. Anand and V. Kanhangad, "PoreNet: CNN-Based Pore Descriptor for High-Resolution Fingerprint Recognition," *IEEE Sens. J.*, vol. 20, no. 16, pp. 9305–9313, Aug. 2020, doi: 10.1109/JSEN.2020.2987287.

[12] F. Liu, Y. Zhao, G. Liu, and L. Shen, "Fingerprint pore matching using deep features," *Pattern Recognit.*, vol. 102, p. 107208, 2020.

[13] D. R. Ashbaugh, "The Identification Process," in *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology*, Boca Raton, FL, USA, CRC press, 1999, pp. 95–156.

[14] J.D. Stosz and L.A. Alyea, "Automated System for Fingerprint Authentication Using Pores and Ridge Structure," in *Proc. SPIE*, vol. 2277, p. 2277, Oct. 1994.

[15] K. Kryszczuk, A. Drygajlo and P. Morier, "Extraction of Level 2 And Level 3 Features for Fragmentary Fingerprints," in *Proc. 2nd COST Action 275 Workshop*, 2004, pp. 83–88.

[16] Q. Zhao, D. Zhang, L. Zhang, and N. Luo, "Adaptive Fingerprint Pore Modeling and Extraction," *Pattern Recognit.*, vol. 43, no. 8, pp. 2833–2844, Aug. 2010.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," in *Commun. ACM*, vol. 60, no. 6, 2017, pp. 84–90.

[18] J. Long, E. Shelhamer and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in *Proc. of the IEEE Conf. on Comput. Vis. and Pattern Recognit.*, 2015, pp. 3431-3440.

[19] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," in *Int. J. Comput. Vis.*, vol. 115, no. 3, 2015, pp. 211–252.

[20] S. Lee et al, "Image Classification Based on The Boost Convolutional Neural Network," *IEEE Access*, vol. 6, pp. 12755-12768, 2018.

[21] B. Li and Y. He, "An Improved ResNet Based on the Adjustable Shortcut Connections," *IEEE Access*, vol. 6, pp. 18967-18974, 2018.

[22] R. Donida Labati, A. Genovese, E. Muñoz, V. Piuri, and F. Scotti, "A Novel Pore Extraction Method for Heterogeneous Fingerprint Images Using Convolutional Neural Networks," *Pattern Recognit. Lett.*, vol. 113, pp. 58–66, 2018.

[23] H.-U. Jang, D. Kim, S.-M. Mun, S. Choi, and H.-K. Lee, "DeepPore: Fingerprint Pore Extraction Using Deep Convolutional Neural Networks," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1808–1812, 2017.

[24] V. Anand and V. Kanhangad, "Pore Detection in High-Resolution Fingerprint Images Using Deep Residual Network," *J. Electron. Imaging*, vol. 28, no. 02, p. 1, 2019.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016 pp. 770–778.

[26] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Int. Conf. on Mach. Learn.*, 2013, pp. 1310–1318.

[27] J. Zhang, T. He, S. Sra, and A. Jadbabaie, "Why gradient clipping accelerates training: A theoretical justification for adaptivity," 2019. [Online]. Available: arXiv:1905.11881v2.

[28] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 1251–1258.

[29] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.

[30] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” 2015. [Online]. Available: arXiv: 1502.03167 [cs.LG].

[31] V. Nair and G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines,” in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 807–814.

[32] T. M. Mitchell, “Evaluating Hypotheses,” in *Mach. Learning*, McGraw-Hill, 1997, ch. 5, pp. 128-153.

[33] W. Niblack, *Introduction to Digital Image Processing*. London, England: Pearson Education, 1988.

[34] N. B. Rais, M. S. Hanif and I. A. Taj, “Adaptive Thresholding Technique for Document Image Analysis,” in *8th International Multitopic Conference*, *Proceedings of INMIC*, 2004, pp. 61-66.

[35] PolyU HRF Database, The Hong Kong Polytechnic University, 2009, [Online]. Available: http://www4.comp.polyu.edu.hk/~biometrics/HRF/HRF_old.htm.

[36] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Nashville, TN: John Wiley & Sons, 2012.

[37] Luca Antiga. “Retina Blood Vessel Segmentation with A Convolution Neural Network (U-Net).” GitHub repository. https://github.com/orobix/retina-unet (accessed Sep. 22, 2017).

[38] Z. Hussain, F. Gimenez, D. Yi, and D. Rubin, “Differential Data Augmentation Techniques for Medical Imaging Classification Tasks,” in *AMIA Annu. Symp. Proc.*, vol. 2017, 2017, pp. 979–984.

[39] K. P. Murphy, *The Machine Learning: A Probabilistic Perspective*. MIT Press, 2014.

[40] S. Ruder, “An Overview of Gradient Descent Optimization Algorithms,” 2016. [Online]. Available: arXiv 1609.04747v2 [cs.LG].

[41] H. T. Le et al., “Study of CNN Based Classification for Small Specific Datasets,” in *Modern Approaches for Intell. Inf. and Database Systems, Springer*, 2018, pp. 355–365.

[42] Y. Zhu, Q. Ouyang, and Y. Mao, “A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy,” *BMC Bioinformatics*, vol. 18, no. 1, pp. 1–10, 2017.

[43] M. Dvornik, “Learning with Limited Annotated Data for Visual Understanding,” *Université Grenoble Alpes*, 2019.

MOHAMMED ALI received the B.Sc. degree in electrical and electronic engineering from the University of Khartoum, Khartoum, Sudan, in 2012, the M.Sc. degree in networking and computer architecture from the University of Khartoum, Khartoum, Sudan, in 2015, and the second M.Sc. degree in electrical engineering from the KFUPM, Dhahran, Saudi Arabia, in 2017. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Concordia University, Montreal, QC, Canada, under the supervision of Dr. M. Omair Ahmad and Dr. Chunyan Wang. He has authored in the area of CNC, power amplifier, 5G transceiver, signal and image processing, and holds one patent. His research interests include the areas of machine and deep learning, signal processing, biometrics and computer vision.

CHUNYAN WANG received the B.Eng. degree in electronics from Shanghai Jiao Tong University, Shanghai, China, and the M.Eng. and Ph.D. degrees from Universite Paris-Sud, Paris, France. She is currently a Professor of electrical and computer engineering with Concordia University, Montreal, QC, Canada. Her current research interests include digital image processing and VLSI circuits for signal processing.

M. OMAIR AHMAD (S'69-M'78-SM'83-F'01) received the B.Eng. degree in electrical engineering from Sir George Williams University, Montreal, QC, Canada, and the Ph.D. degree in electrical engineering from Concordia University, Montreal. From 1978 to 1979, he was a Faculty Member with New York University College, Buffalo, NY, USA. In 1979, he joined the Faculty of Concordia University, as an Assistant Professor of computer science. Subsequently, he joined the Department of Electrical and Computer Engineering, Concordia University, where he was the Chair, from 2002 to 2005, and is currently a Professor. He was a Founding Researcher of Micronet and a Canadian Network of Centers of Excellence, from 1990 to 2004. He is also the Concordia University Research Chair (Tier I) of multimedia signal processing. He has authored in the area of signal processing. He holds four patents. His current research interests include the areas of image and speech processing, biomedical signal processing, watermarking, biometrics, video signal processing and object detection and tracking, deep learning techniques in signal processing, and fast signal transforms and algorithms. In 1988, he was a member of the Admission and Advancement Committee of the IEEE. He was a recipient of numerous honors and awards, including the Wighton Fellowship from the Sandford Fleming Foundation, an induction to Provosts Circle of Distinction for Career Achievements, and the Award of Excellence in Doctoral Supervision from the Faculty of Engineering and Computer Science, Concordia University. He was a Guest Professor with Southeast University, Nanjing, China, and the Local Arrangements Chairman of the 1984 IEEE International Symposium on Circuits and Systems. He has served as the Program Co-Chair for the 1995 IEEE International Conference on Neural Networks and Signal Processing, the 2003 IEEE International Conference on Neural Networks and Signal Processing, and the 2004 IEEE International Midwest Symposium on Circuits and Systems. He was the General Co-Chair of the 2008 IEEE International Conference on Neural Networks and Signal Processing. He is the Chair of the Montreal Chapter IEEE Circuits and Systems Society. He was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART I: FUNDAMENTAL THEORY AND APPLICATIONS, from 1999 to 2001.