Convolutional Neural Networks in Application to Segmentation of Fingerprint Images

Maxim PASYNKOV\textsuperscript{1} and Michael KHACHAY\textsuperscript{1,2,*}

\textsuperscript{1}Krasovsky Institute of Mathematics and Mechanics (Ekaterinburg Russia)
\textsuperscript{2}Ural Federal University (Ekaterinburg Russia)
\textsuperscript{*}Corresponding author

Keywords: Segmentation, fingerprint, Convolution neural networks.

Abstract. Segmentation of fingerprint images is one of the most important problems concerned with automatic biometric identification system. Segmentation is used to separate the area of the fingerprint (foreground) from the background and areas that cannot be recovered. We propose a new algorithm for fingerprint segmentation based on convolutional neural networks, binary region labeling technique, and morphological image processing. This approach was tested on public fingerprint dataset provided by Fingerprint Verification Competition (FVC: 2002, 2004). The experimental results showed a high quality of segmentation.

Introduction

Identification by fingerprints is the most commonly used biometric technology [1]. Under the fingerprint in such systems (Figure 1, left) is meant a specialized grayscale image.

Figure 1. Fingerprint image [2] and its manual segmentation.

In the process of identification, the image of the fingerprint is subjected to several stages of preprocessing [3]. The first is segmentation. The segmentation step is designed to separate the area of the fingerprint (foreground) from the background area and non-recoverable regions. Accurate segmentation can significantly reduce the computational costs of the entire identification, and also discard many false features. For example, the use of segmentation allows more accurately determine locations and types of singular points of the analyzed images (Figure 2). Also, segmentation is necessary at the stage of fingerprint image enhancement [4], and at the stage of final identification [5].

Segmentation of fingerprint images is studied in numerous papers. Thus, in [6], a two-stage segmentation algorithm is proposed. At the first stage a linear classifier is constructed in the feature space of 3 dimensions: coherence, average intensity value and its variance. At the second stage, preliminary segmentation results are improved by standard morphological processing procedure [7]. Training was carried to on DB1B fingerprint dataset offered to participants of Fingerprint Verification Competition (FVC 2000) [8]. Although, the algorithm performs well at this dataset (avg. error is 6.8%), its overfitting is high, since test error obtained on dataset DB2B is 14.3%.

The article [10] proposed combined method of segmentation of fingerprint that uses fingerprint orientation field and statistical information of the image. Unfortunately, this paper and some other papers (see, e.g. [11]) provide no numerical results to compare at all.
Algorithm

Input of our algorithm is a grayscale fingerprint image. At the initial stage we process the input image with a Convolutional Neural Network (CNN). The image is partitioned into small areas, with an offset corresponding to the block size (Figure 4). Then, each area is classified by previously trained neural network, and the binary mask of segmentation is collected from the obtained estimates.

At the second stage the results are refined by the known [12] Region Labeling (RL) algorithm. Its main idea employs the fact that the vast majority of the fingerprint images contain single connected foreground area. Assuming that true foreground area is the biggest one, other areas can be eliminated by the classic non-recursive algorithm for sequential region labeling.

At the final stage the analyzed image is subjected to Morphological Processing (MP). The combination of the sequential application of dilatation and erosion is called the closing operation. It allows us to get rid of breaks of a certain size. As a result of applying the dilatation to the binary image, we expand the foreground area, thus getting rid of discontinuities of a certain size. And the subsequent application of erosion to the binary image, we restore the original shape of the foreground area already without gaps. The structural scheme of the proposed algorithm is shown in Figure 3.

![Figure 3. Structural scheme of the algorithm.](image)
Convolutional Neural Network

The main classifying component of the proposed algorithm is a Convolutional Neural Network [13]. It is generally accepted [14] that this type of neural network is most suitable for effective image recognition, in addition, it is most resistant to the problem of overfitting. The most important property of neural networks is their ability to learn on incoming data in them and increase their performance [15]. Typically, the training of a neural network occurs through an iterative process of adjusting the synaptic weights and thresholds. In this paper, we use famous back propagation algorithm for training.

To setup the training phase, we manually label out all fingerprint images taken from FVC (2002 [2], 2004[16]). For this purpose, for each image, the foreground area was highlighted. The result of this selection is a manually crafted ground truth binary segmentation mask (a binary image of the same dimension encoding foreground and background pixels by 1 and 0, respectively). The resulting mask is divided into non-overlapping blocks of a certain size (Figure 4, right). Then, each block is classified by majority voting.

For classification by neural network, fingerprint image is previously partitioned onto overlapping regions of a certain bigger size with an offset step equal to the block size (Figure 4, left). Thus, we obtain a one-to-one correspondence between regions, which are fed to neural network and blocks of the manual segmentation mask.

![Figure 4. Partition of the image of the fingerprint (left) and the mask of manual segmentation (right).](image)

By construction, neural network produces a block-wise binary segmentation mask. Each block is filled with the binary output of the neural network obtained by classifying the corresponding area of the fingerprint image.

To assess the training quality, we use relative pixel-wise difference between network-generated and manual (true) segmentation masks:

\[
\text{FAR} = \frac{\text{manual} = 0, \text{algorithm} = 1}{N},
\]

\[
\text{FRR} = \frac{\text{manual} = 1, \text{algorithm} = 0}{N},
\]

\[
\text{error} = \text{FAR} + \text{FRR},
\]

where, \(N\) is the total number of pixels in the image.

To determine the optimal values of the hyper-parameters of the neural network, simple coordinate descent optimization method is performed to minimize the total error:

\[
\text{error} = \text{error}(rl, bl, cnns, cnn_i\_core\_size, cnn_i\_core\_maps, net) \rightarrow \min
\]

where, \(rl\) denotes the region size, \(bl\) is the block size, \(cnns\) is the number of convolution layers, \(cnn_i\_core\_size\) is the kernel size on the \(i\)-th convolution layer, \(cnn_i\_core\_maps\) is the number of cores on the \(i\)-th convolution layer and \(net\) is the combination of the fully connected layers.

To optimize error functional on \(rl\) and \(bl\), a simple convolutional neural network consisting of one convolution and one fully connected layer was used. The training was conducted for 10 epochs.

To train the neural network, we use fingerprint images provided by FVC (2002, 2004). In each competition, there are 4 databases of fingerprints obtained from different sensors. All bases are divided into the closed part (A) and the open part (B). In our research, we are based on part (B) ultimately. This dataset consists of 8 images for 10 different fingers per database, i.e. 80 different.
fingerprint images in total. For training, we use images of odd fingers, for testing of even. To minimize time consumption, only first 4 images per finger are used, which leads to the resulting training dataset of ≈2.5·10⁶ samples.

The absence of subsampling layers is explained by the fact that they introduce additional blurring of the pseudo-regular structure of the fingerprint, which greatly reduces the accuracy of segmentation.

As a result of the optimization, it was possible to reduce the overall error in the pixel-by-pixel comparison of the algorithm result with manually segmented images to 4.13%.

Discussion and Conclusion

The proposed algorithm showed high efficiency in the segmentation problem of fingerprint images. Thanks to the use of convolutional neural networks, the region labeling algorithm and morphological processing, a low total segmentation error of 3.151% was achieved. It is worth noting that the proposed algorithm, once tuned and trained, produces efficient and accurate segmentation on all fingerprint databases we use. This is the main difference from some other algorithms, customized for specific images.

Of the works known to us, only in [6] are numerical results of segmentation on public fingerprint databases with an estimate of accuracy at the pixel level. The authors of this work claim an average (total) error of 6.8% and 13.6%, respectively. They calculate this error based on 30 selected and manually segmented images, and for each base of prints they train their own classifier.

The results of our algorithm are given in Table 1. It is especially worth noting that these results were obtained by pixel-by-pixel comparison of algorithm segmentation with the reference (manually segmented image of the fingerprint). An illustration of the algorithm is shown in Figure 5.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>FAR</th>
<th>FRR</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>2.073%</td>
<td>2.057%</td>
<td>4.130%</td>
</tr>
<tr>
<td>CNN + RL + MP</td>
<td>2.375%</td>
<td>0.776%</td>
<td>3.151%</td>
</tr>
</tbody>
</table>

Figure 5. Segmentation with the proposed approach. To the left is a segmentation sample by CNN only; to the right by CNN+RL+MP. Additional processing allows to eliminate breaks and a spuriously segmented island.

Acknowledgments

Calculations were carried out on a cluster of the Institute of Mathematics and Mechanics of the Russian Academy of Sciences.

This research was supported by the Russian Science Foundation, grant No. 14-11-00109.
References


