Paper: [Can a CNN Automatically Learn the Significance of Minutiae Points for Fingerprint Matching?]

Author: Anurag Chowdhury, Simon Kirchgasser, Andreas Uhl, Arun Ross (WACV 2020)

Reviewer: Aditi Ganesh Joshi (180020010)

Problem Statement

The paper aims to determine whether a CNN can automatically learn the significance of minutiae points(ridge anomalies- endings and bifurcations), as suggested by humans from experiments on large-scale fingerprint data. Our aim certainly differs from previously developed CNNs like JudgeNet, LocateNet, FingerNet, MinutiaeNet, which have used minutiae points as features for fingerprint recognition and matching.

Approach

Firstly, fingerprint image pairs are generated using **SIFT(Scale Invariant Feature Transform)-based patch alignment** method(aligning the overlapping regions only). As this method may consider minutiae points as interest points, we also use another strategy called '**SIFT without minutiae**', which discards all interest points within a 10*10 pixel region surrounding a minutiae point. But the difference is small since only a few key points satisfy the SIFT without minutiae criterion. These overlapping regions are of size 128*128 pixels called **'image fingerprint pair' (D_t)**, which are further tessellated into 16 32*32 pixel fingerprint non- overlapping patches. Picking up patches from the same corresponding locations of image pairs gives a **genuine patch-pair(D_v)**, whereas picking up patches from different locations gives an **imposter patch-pair(D_v)**.

These patch-pairs(Dv) are supplied to the CNNs. The paper employs two CNN architectures viz., **Multi-scale Dilated Siamese CNN(MD-CNN)**, and a modified **All-CNN**. MD-CNN uses dilated 3*3 convolutions, which behave like sparse 5*5, 7*7 kernels, hence, reducing the number of parameters but decreasing the effective receptive field. Hence, the responses from these convolutions are summed up together to form a multi-scale dilated convolution. This model uses data augmentation, i.e., adding randomly rotated images to the input, to ensure rotation invariance. The modified version of All-CNN architecture has two important changes :

1. The maximum number of convolutions has been reduced to 64 from 192.

2.RELU has been replaced by **SELU** non-linearity.

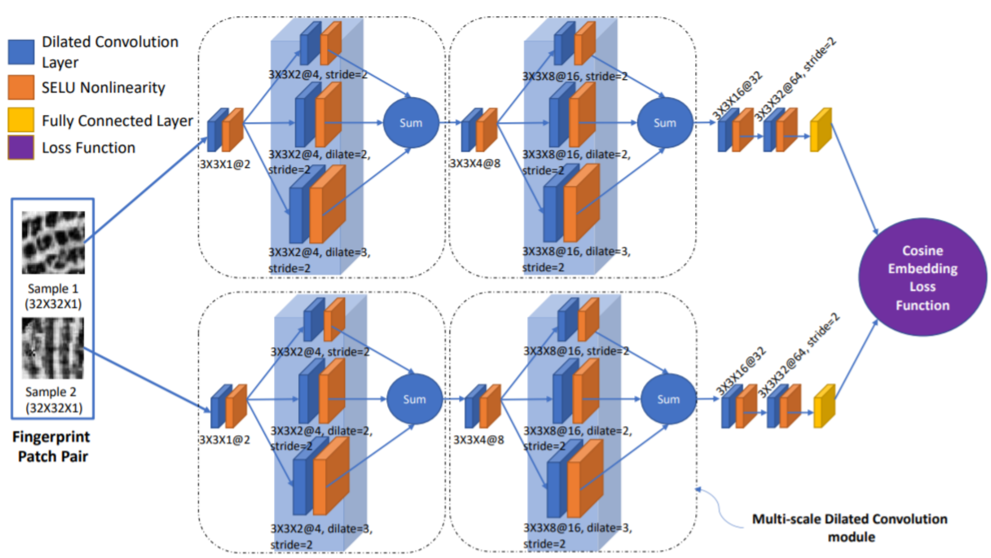

Figure 1: An illustration of the proposed Multi-scale Dilated Siamese CNN (MD-CNN) architecture during training.

Cosine Embedded Loss Function is utilized in the final layer. It ensures that the distance between genuine pairs is lesser than that of imposter pairs.

Experiments, Results, and Analysis

"CASIA fingerprint Subject Ageing Version 1.0" and the **"FVC-2000"** fingerprint datasets are used for analyzing the CNNs. The fingerprint match scores in range[-1,1](1:perfect match, -1:perfect non-match) using True Match Rate at False Match Rate(TMR @ FMR) of 0.1 and 1 are reported and it is observed that the presence of minutiae points at higher TMR gives higher match scores and average comparison scores, irrespective of the alignment method used. Also, **Gradient-weighted Activation Mapping(Grad-CAM)** is used for analyzing which local regions of the input image are important through gradient information flowing through the CNNs. These observations are also compared with a qualitative patch-based fingerprint matching. This also clearly establishes that CNNs focus on minutiae points.

Strengths :

- **Reduces the number of trainable parameters** substantially through the use of dilated convolutions(MD-CNN) and reducing the maximum number of convolutions(modified All-CNN).
- The introduction of **SELU non-linearity** instead of RELU helps in faster convergence and better accuracy without overfitting.
- As a result of **multi-scale dilated convolutions** used, in the absence of minutiae points, MD-CNN focuses along the ridges, whereas this isn't true in the case of the modified All-CNN architecture.
- In my view, this paper successfully establishes, both quantitatively and qualitatively, that CNN can learn the significance of minutiae points for fingerprint recognition and matching, irrespective of the model architecture used.

Weaknesses and Scope of improvement :

- Though the aim is to emphasize the importance of minutiae points, the accuracy of models as classifiers for genuine and imposter pairs could be analyzed and compared with existing CNNs and traditional methods.
- Doesn't provide insights about information(other than minutiae points) extracted from the given models, which can be further utilized for better fingerprint recognition, especially in the case of partial fingerprints.
- As suggested by the author, this can be further extended to be employed for determining the importance of minutiae points as a function of different types of fingerprints.