Dual Dense Uncertainty Embedding for Iris Recognition

Yunlong Wang¹, Jianze Wei², Zhaofeng He³, Zhenan Sun¹, Tieniu Tan¹,⁴
¹Center for Research on Intelligent Perception and Computing (CRIPAC), Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing, P.R.China
⁴Nanjing University, Nanjing, P.R.China

Abstract

In this paper, D²UE is proposed to model data uncertainty of iris recognition in a pixel-level manner, which can attenuate the interference caused by uncertain acquisition factors. D²UE takes the intermediate feature map of any dense DL framework as input and generates the UEDR and VSM of an iris image. VSM is then leveraged to produce a binary mask through adaptive threshold masking, and thus pixelwise iris segmentation is on longer needed. Experimental results on several public iris datasets demonstrate the superiority of D²UE in improving the recognition performance of baseline methods, and it is a remarkably lightweight building block.

Motivation

Uncertain acquisition factors inevitably affect the process of iris imagery formation. Deep uncertainty embedding (DUL) is leveraged to represent the iris image using a Gaussian distribution.

Framework

- **Uncertainty Embedded Dense Representation (UEDR)**

  Model each pixel’s uncertainty by a univariate Gaussian distribution, μ-branch and σ-branch are constructed by a simple transformation with a convolutional layer and an instance normalization (IN) layer.

  - Training phase
    \[ P_x = I_{UE}(\text{Conv}(\mathbf{H})) \]
  - Inference phase
    \[ S_{ij} = P_{ij} + \epsilon_{ij}^\mu \quad S_{ij} = P_{ij}^\sigma \]

- **Adaptive Threshold Masking**

  Find an optimal α for each pixel in an iris image through predicting a variance scaling map (VSM), taking not only the intensity distribution of the iris image but also each pixel’s low-level uncertainty into consideration.

  \[ A = \text{Sigmoid}(\text{Conv}(\mathbf{H})) \]

  \[ M_{ij} = \begin{cases} 1, & \text{if} \quad I_{ij}^\mu - I_{ij}^\sigma < 3A_{ij}\epsilon^\sigma \text{otherwise} \\ 0 \end{cases} \]

- **Loss function**

  Extended Triplet Loss (ETL)

  \[ \mathcal{L}_{etl} = \frac{1}{N} \sum_{n=1}^{N} \max(D(S_n^p, S_n^q) - D(S_n^p, S_n^t) + \gamma, 0) \]

  Kullback-Leibler (KL) Regularization

  \[ \mathcal{L}_{kl} = E\left[ KL[N(S_{ij}, \mu_{ij}, \sigma_{ij}) || N(c_i(\mathbf{1}), \mathbf{1})] \right] = \frac{1}{2NHW} \sum_{h=1}^{H} \sum_{w=1}^{W} \sum_{l=1}^{L} (1 + \log(\sigma_{ij}^2) - \mu_{ij}^2 - \sigma_{ij}^2) \]

  Total loss

  \[ \mathcal{L}_{total} = \mathcal{L}_{etl} + \lambda \mathcal{L}_{kl} \]

The schematic of the proposed framework D²UE

Experimental Results

Within-database comparisons

Cross-database comparisons

Visualizations

Conclusions

In this paper, D²UE is proposed to model data uncertainty of iris recognition in a pixel-level manner, which can attenuate the interference caused by uncertain acquisition factors. D²UE takes the intermediate feature map of any dense DL framework as input and generates the UEDR and VSM of an iris image. VSM is then leveraged to produce a binary mask through adaptive threshold masking, and thus pixelwise iris segmentation is on longer needed. Experimental results on several public iris datasets demonstrate the superiority of D²UE in improving the recognition performance of baseline methods, and it is a remarkably lightweight building block.