Iris Segmentation using Machine Learning

Aashutosh Khandelwal
IIT2016514
Sourin Chakrabarti
IIT2016513
Harshit Agarwal
IIT2016108
Nistala Venkata Kameshwar Sharma
ISM2016005

Introduction

One crucial step in iris recognition systems is the **segmentation** of the iris region from the input image. This step has traditionally been solved using manually designed segmentation techniques and considerable performance has already been achieved on numerous datasets of variable quality. However, with the success of **deep-learning** models for other vision problems, researchers are increasingly looking into convolutional neural networks (CNNs) to further improve on the performance of existing iris segmentation techniques.

In this project, we have used and studied the utility of **U-Net**, a deep convolutional neural network (CNN) commonly used for image translation tasks, for the problem of iris segmentation and compared it with the several existing methods of iris segmentation such as **Hough** and **Daugman Transforms** and **thresholding** which fail to give competitive results.

Datasets used

Our main resource for training and testing our model will be the **CASIA-IrisV4** and **CASIA-IrisV1** dataset. CASIA-IrisV4 contains a total of 54,601 iris images from more than 1,800 genuine subjects and 1,000 virtual subjects. All iris images are 8 bit gray-level JPEG files, collected under near infrared illumination or synthesized. CASIA Iris Image Database Version 1.0 (CASIA-IrisV1) includes 756 iris images from 108 eyes. For each eye, 7 images are captured in two sessions with our self-developed device CASIA close-up iris camera. All images are stored as BMP format with resolution 320*280.

Dataset Preprocessing

In the pre-processing step, all the images need to be **normalized** and converted to **320x320** single channel **grayscale** image. Then, the ground truth values of all the images we fetched from the **CC dataset (IRISSEG-CC Dataset)**. For the CC dataset the parameters define circles which give the iris boundaries and eyelid masks. Eyelashes occlusion is not included in the segmentation data.

Related Work

Initially we tried to **localize** the iris boundary using **thresholding and canny edge detection**. If the intensity of a pixel I(x, y) is greater than a threshold, it is marked as white, otherwise it is marked black. Thresholding produces a **binary** image. The image obtained was further processed using **morphology** techniques with openCV. On this preprocessed image, a **Canny edge operator** is applied to obtain a set of edge points. A **voting** procedure is employed to ascertain whether that edge point lies on a circular region or not. Finally, all the regions were classified using a **binary classifier**.

Next method used for segmentation was **Hough transform**. Specular reflections in the eye lead to incorrect detection of the pupil/iris boundary. Hence, all kinds of **illumination** was removed from the image.

First, the image **complement** was analyzed. Areas of dark pixels surrounded by lighter pixels were replaced by the lighter pixels. The general Hough transform can be used to detect shapes that can be written in **parametric** form such as lines, circles, parabolas, and hyperbolas . The circular Hough transform can be used to detect the circles of a known radius in an image. As we don’t know the definite radius of the pupil or iris, the transform must be computed for a **range of radii**. For each radius tested, the **location** and **value** of the **maximum** is stored. The radius with the **highest peak** indicates the most likely radius and center coordinate for the boundary. Another method used for our task was segmentation using **Daugman’s Integro Differential Operator (IDO) algorithm**.

Deep Learning Architecture

The **U-Net** model represents a popular CNN architecture for solving biomedical problems and other image translation tasks. The main advantage of this model is it's ability to learn relatively accurately from (very) small datasets, which is a common problem for data-scarce computer-vision tasks, including iris segmentation. U-Net uses an **encoder-decoder** architecture. The architecture is divided into corresponding encoder and decoder convolutional layers. The left side of the model is the encoder path and the right side is the decoder path. The encoder follows a typical CNN architecture, as popularized by the VGG network consisting of two 3×3 **convolutional** layers followed by a **ReLu activation** layer with max pooling.

Each **down-sampling** step of the encoder **doubles** the amount of feature channels and decreases the image resolution by **half**. The decoder part then **up-samples** the feature maps of the lower layer while also concatenating and cropping the output of the encoder part of the same depth. This process ensure that information is propagated from the encoder to the decoder at all scales and no information is lost during the down-sampling operations in the encoder. The final layer of the network is a 1×1 convolutional layer that mixes the output channels of the preceding layer and produces the **segmentation maps** (one per class - iris vs. non-iris) that represent the output of the U-Net model

Results

In our series of experiments, we found that the U-Net architecture gave significantly **better** results than the other methods on images without specular illumination. The model achieved an overall accuracy of **94.80** percent which is significantly higher than the traditional methods. For images with illumination, we found that the accuracy dipped significantly to **85.70** percent which is below par in case of iris images which have high amounts of details. Traditional methods of illumination removal seemed to work better than deep learning techniques in terms of average precision. Hence, the next step would be to take the best of both worlds and work on an architecture that removes illumination first and then goes for iris segmentation.

Future Scope

Hence, we see that the U-Net based segmentation model proved to be very successful at segmenting the iris, while also outperforming all considered baseline methods(for certain cases). The model didn’t require a great amount of training data and worked well without the use of data augmentation during training. In conclusion the use of deep learning methods in iris recognition may provide an increase in performance in an area of biometrics compared to conventional methods