Feature Extraction using Gaze of Participants for Classifying Gender of Pedestrians in Images

鳥取大学 Tottori University

Riku Matsumoto, Hiroki Yoshimura, Masashi Nishiyama, and Yoshio Iwai, Tottori University, Japan

Media Understanding Laboratory Tottori University

Purpose

To increase the classification performance when the number of training samples is small

Generation : Gaze map from the distribution of gaze locations recorded while participants viewed some training samples

Extraction : Gaze locations play an important role in the gender classification of pedestrian images

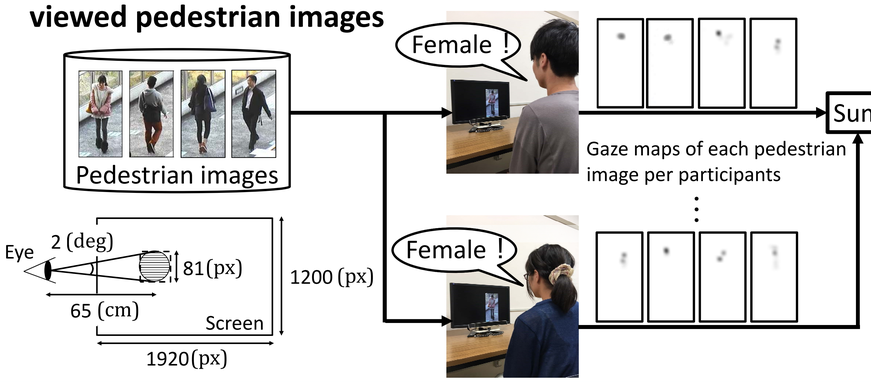
Existing methods : The performance has been found to decrease when the number of training samples is small
e.g. Deep Learning[Antipov, ACM'15], Metric Learning[Lu, TIFS'14]

Human vision : Humans have the visual capability to extract features from an individual and identify them as male or female
e.g. People correctly classify gender from facial images. [Bruce, Perception'93]

Generation of the gaze map from the distribution of gaze locations

Generation procedures

Generate the gaze map from the distribution of gaze locations measured while participants viewed pedestrian images



1. Display a pedestrian image on the screen for and the participant reported the inferred gender of the pedestrian
2. Generate each gaze maps per pedestrian images
3. Aggregate to each gaze maps for all participants and all pedestrian images

Evaluation of the gaze map

Observed that participants frequently concentrated their gaze on the head region

Gaze maps of each pedestrian image
The dark regions in the gaze maps represent the gathered gaze locations from the participants

- Dataset: CUHK included in the PETA dataset [Deng, ACM'14]
- Shown images: 30 images (15 male and 15 female images)
- Participants: 14 people (22.4±0.8, 7 males and 7 females)

(a) (b)

(a) Aggregated gaze map for gender classification
(b) Average intensities of the pedestrian images

- **The head region gathered a large number of gaze locations**
- **The chest region gathered a moderate of gaze locations**
- **The whole body region gathered few gaze locations**

Extraction of features using the gaze map

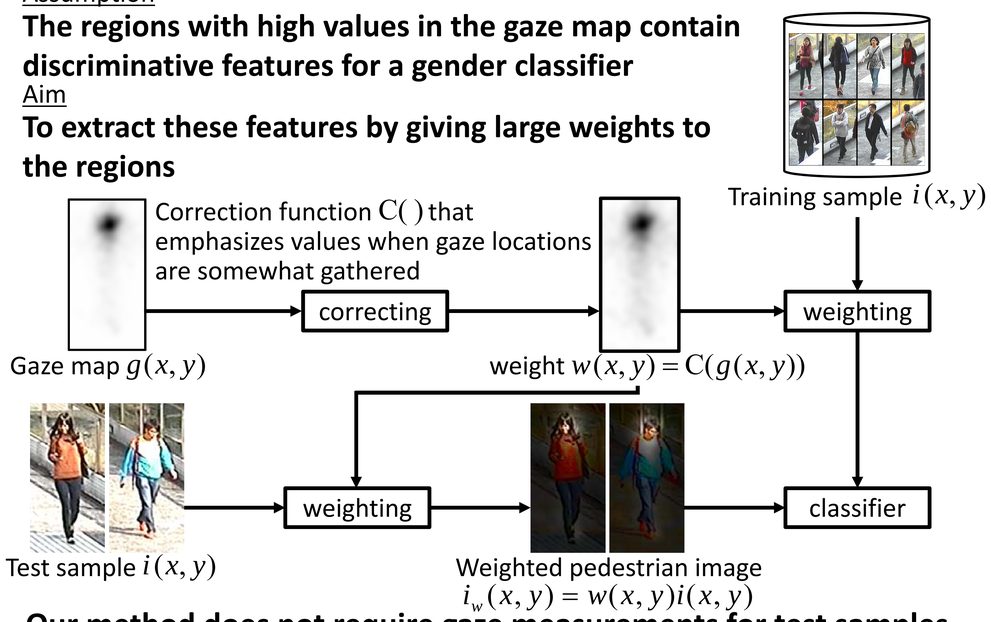
Feature Extraction algorithm

Assumption

The regions with high values in the gaze map contain discriminative features for a gender classifier

Aim

To extract these features by giving large weights to the regions



Our method does not require gaze measurements for test samples

Evaluations of gender classification

Condition

- Data set: CUHK included in the PETA dataset
- Training samples: 2000 images
- Test samples: 400 images, 5 sets
- Feature: raster scanning of RGB values (40 × 80 × 3 dimensions)

Comparison 1

Combine the gaze map with existing classifiers

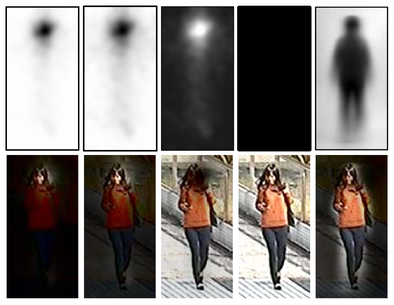
- Classifier 1: Mini-CNN [Antipov, ACM'15]
- Classifier 2: LMNN [Weinberger, MLR'09]

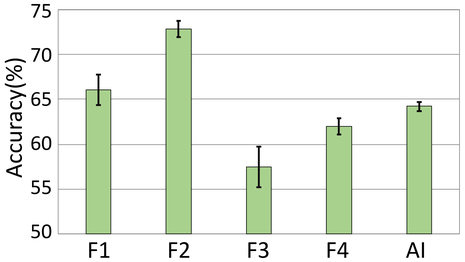
Classifier	Gaze map	Accuracy(%)
CNN	with	**75.2±1.4**
	without	69.7±1.1
LMNN	with	**72.1±1.0**
	Without	68.0±1.2

Comparison 2

Compare the correction functions

- Classifier: k-NN (k = 20)





- **The correction function improved accuracy** because F2 is superior to F1
- **The head region contained discriminative features for a gender classifier** because F2 is superior to the inversed weights F3
- **The gaze map contained meaningful cues to classify the gender of the pedestrian images** because F1, F2 is superior to AI