**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]

**Evaluation of the gaze map**

- Observed that participants frequently concentrated their gaze on the head region

**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

- **Evaluation of the gaze map**

- **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

Correction function $C( )$ that emphasizes values when gaze locations are somewhat gathered

- Training sample $f(x, y)$

- Gaze map $g(x, y)$

- Weight $w(x, y) = C(g(x, y))$

- Test sample $\tilde{f}(x, y)$

- Weighted pedestrian image $\tilde{f}(x, y) = w(x, y) f(x, y)$

- 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 x 80 x 3 dimensions)

**Comparison 1**

Combine the gaze map with existing classifiers

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

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Gaze map</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>with</td>
<td>75.2 ± 1.4</td>
</tr>
<tr>
<td></td>
<td>without</td>
<td>69.7 ± 1.1</td>
</tr>
<tr>
<td>LMNN</td>
<td>with</td>
<td>72.1 ± 1.0</td>
</tr>
<tr>
<td></td>
<td>Without</td>
<td>68.0 ± 1.2</td>
</tr>
</tbody>
</table>

- 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 inverted weights F3.
- The gaze map contained meaningful cues to classify the gender of the pedestrian images because F1, F2 is superior to AI.