Selective Refinement Network for Dark Face Detection

Track 2.2

2019.06.16
Outline

- Introduction
- Solution Pipeline
- Image Enhancement
- Face Detection
- Result
Outline

- Introduction
- Solution Pipeline
- Image Enhancement
- Face Detection
- Result
• Face Detection: Determine whether there are faces on an image, and if so, give their location.
Introduction

Face Analysis

Video Surveillance

Face Unlock
Introduction

- One of challenges: detecting faces in poor visibility environments.
Introduction

Track 2.2: (Semi-)Supervised Face Detection in Low Light Condition

• Captured during nighttime under low Light Conditions
• Training: 6,000 images with 43,849 faces
• Testing: 4,000 images with 32,571 faces
Outline

- Introduction
- **Solution Pipeline**
- Image Enhancement
- Face Detection
- Result
Solution Pipeline

1. Train SRN-Res101\textsuperscript{[1]} on WIDER FACE dataset
2. Utilize MSRCR\textsuperscript{[2]} to process DARK FACE dataset
3. Fine-tune pretrained SRN-Res101 on processed DARK FACE dataset


Outline

● Introduction
● Solution Pipeline
● Image Enhancement
● Face Detection
● Result
Image Enhancement

Image Enhancement

Outline

- Introduction
- Solution Pipeline
- Image Enhancement
- Face Detection
- Result
Selective Refinement Network (SRN): STC, STR, RFE
1. Selective Two-Step Classification (STC)

- Need to tile plenty of small anchors to detect small faces
- Cause extreme class imbalance between positives and negatives
- The number of positive samples is only a few dozen or less
- Doing two-step classification is essential to reduce the false positives
- Performing two-step classification on all pyramid levels is unnecessary

<table>
<thead>
<tr>
<th>STC</th>
<th>B</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
<td>95.1</td>
<td>95.2</td>
<td>95.2</td>
<td>95.2</td>
<td>95.0</td>
<td>95.1</td>
<td>95.0</td>
</tr>
<tr>
<td>Medium</td>
<td>93.9</td>
<td>94.2</td>
<td>94.3</td>
<td>94.1</td>
<td>93.9</td>
<td>93.7</td>
<td>93.9</td>
</tr>
<tr>
<td>Hard</td>
<td>88.0</td>
<td>88.9</td>
<td>88.7</td>
<td>88.5</td>
<td>87.8</td>
<td>88.0</td>
<td>87.7</td>
</tr>
</tbody>
</table>

- Select P2, P3, and P4 to perform two-step classification
2. Selective Two-Step Regression (STR)

- Making the location of bounding box more accurate is a challenging issue
- Current one-stage methods rely on one-step regression
- It is inaccurate the in MS COCO evaluation metric
- Blindly adding multi-step regression is often counterproductive

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy</td>
<td>95.1</td>
<td>94.8</td>
<td>94.3</td>
<td>94.8</td>
<td>95.4</td>
<td>95.7</td>
<td>95.6</td>
</tr>
<tr>
<td>Medium</td>
<td>93.9</td>
<td>93.4</td>
<td>93.7</td>
<td>93.9</td>
<td>94.2</td>
<td>94.4</td>
<td>94.6</td>
</tr>
<tr>
<td>Hard</td>
<td>88.0</td>
<td>87.5</td>
<td>87.7</td>
<td>87.0</td>
<td>88.2</td>
<td>88.2</td>
<td>88.4</td>
</tr>
</tbody>
</table>

- Select P5, P6, and P6 to perform two-step regression
3. Receptive Field Enhancement (RFE)

- Current networks possess square receptive fields
- Mismatch between receptive fields and aspect ratio of faces affect the detection performance
- Propose RFE to diversify receptive fields before predicting classes and locations
- RFE replaces the middle two convolution layers in the class and box subnet of RetinaNet
Training Detail

- Backbone: ResNet-101 with 6-level FPN
- Loss: sigmoid focal loss + smooth L1 loss
- Data augmentation: color distortions, random cropping, random flipping
- Anchor design: two specific scales $(2, 2\sqrt{2})$ and one aspect ratio (1.25)
- SGD, 0.9 momentum, 0.0001 weight decay, batch size 32
- Learning rate to 0.01, 0.001 and 0.0001 for the 100, 20 and 10 epochs
Code has been released publicly:
Outline

● Introduction
● Solution Pipeline
● Image Enhancement
● Face Detection
● Result
# Result

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Training Dataset</th>
<th>Testing Dataset</th>
<th>AP</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WIDER FACE</td>
<td>DARK FACE</td>
<td>47.96</td>
</tr>
<tr>
<td>B</td>
<td>WIDER FACE</td>
<td>Processed DARK FACE</td>
<td>69.07</td>
</tr>
<tr>
<td>C</td>
<td>WIDER FACE + Processed DARK FACE</td>
<td>Processed DARK FACE</td>
<td>83.85</td>
</tr>
</tbody>
</table>

![Bar chart showing results of experiments A, B, and C](chart.png)
Thank you!