Lecture 13: Face

Alireza Akhavan Pour

CLASS.VISION
Face verification vs. face recognition

Verification
- Input image, name/ID
- Output whether the input image is that of the claimed person

Recognition
- Has a database of $K$ persons
- Get an input image
- Output ID if the image is any of the $K$ persons (or “not recognized”)
One-shot learning

Learning from one example to recognize the person again
Learning a “similarity” function

\[ d(\text{img}_1, \text{img}_2) = \text{degree of difference between images} \]

If \( d(\text{img}_1, \text{img}_2) \leq \tau \)

\[ \begin{align*} &> \tau \quad \text{"different"} \quad \{ \text{Veriﬁcation.} \} \\
&\quad \text{"same"} \end{align*} \]
Siamese network

\[ x^{(1)} \]

\[ x^{(2)} \]

[Taigman et. al., 2014. DeepFace closing the gap to human level performance]
Goal of learning

Parameters of NN define an encoding \( f(x^{(i)}) \)

Learn parameters so that:

- If \( x^{(i)}, x^{(j)} \) are the same person, \( \|f(x^{(i)}) - f(x^{(j)})\|^2 \) is small.
- If \( x^{(i)}, x^{(j)} \) are different persons, \( \|f(x^{(i)}) - f(x^{(j)})\|^2 \) is large.
Learning Objective

Schroff et al., 2015, FaceNet: A unified embedding for face recognition and clustering

Want: \[ \frac{\|f(A) - f(P)\|^2}{d(A,P)} + \lambda \leq \frac{\|f(A) - f(N)\|^2}{d(A,N)} \]

\[ \|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + \lambda \leq 0 \]

\[ f'(\theta) = \delta \]
Loss function

Training set: 10k pictures of 1k persons

[Schroff et al., 2015, FaceNet: A unified embedding for face recognition and clustering]
Choosing the triplets $A, P, N$

During training, if $A, P, N$ are chosen randomly, $d(A, P) + \alpha \leq d(A, N)$ is easily satisfied.

Choose triplets that’re “hard” to train on.

[Schroff et al., 2015, FaceNet: A unified embedding for face recognition and clustering]
Training set using triplet loss

<table>
<thead>
<tr>
<th>Anchor</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Anchor Image 1]</td>
<td>![Positive Image 1]</td>
<td>![Negative Image 1]</td>
</tr>
<tr>
<td>![Anchor Image 2]</td>
<td>![Positive Image 2]</td>
<td>![Negative Image 2]</td>
</tr>
<tr>
<td>![Anchor Image 3]</td>
<td>![Positive Image 3]</td>
<td>![Negative Image 3]</td>
</tr>
<tr>
<td>![Anchor Image 5]</td>
<td>![Positive Image 5]</td>
<td>![Negative Image 5]</td>
</tr>
</tbody>
</table>
Face verification and binary classification
Learning the similarity function

\[ f(x^{(i)}) \rightarrow \rightarrow \rightarrow \rightarrow f(x^{(j)}) \rightarrow \hat{y} \]

[Taigman et. al., 2014. DeepFace closing the gap to human level performance]
Face verification supervised learning

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td>1</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td>0</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td>0</td>
</tr>
<tr>
<td><img src="image4.png" alt="Image" /></td>
<td>1</td>
</tr>
</tbody>
</table>

[Taigman et. al., 2014. DeepFace closing the gap to human level performance]
منابع

• https://www.coursera.org/specializations/deep-learning