Convolutional Neural Network Transfer Learning for Robust Face Recognition in NAO Humanoid Robot

Daniel Bussey¹, Alex Glandon², Lasitha Vidyaratne², Mahbubul Alam², Khan Iftekharuddin²
¹Embry-Riddle Aeronautical University, ²Old Dominion University

Background Research

• Artificial Neural Networks
  • Computing systems whose model architecture is inspired by biological neural networks [1].
  • Capable of improving task performance without task-specific programming.
  • Show excellent performance at classification based tasks such as face recognition [2], text recognition [3], and natural language processing [4].
• Deep Learning
  • A subfield of machine learning that focuses on algorithms inspired by the function and structure of the brain called artificial neural networks. The method used to allow computers to learn through a process called training.
• Transfer Learning
  • Retraining a pre-existing neural network on a new data set to perform a specific task.
• Face Recognition Software
  • Software capable of identifying a person from a digital image or video frame.

Research Goals and Objectives

• Apply transfer learning to the convolutional neural network AlexNet [5] for face recognition tasks.
• Compare the performance of the retrained AlexNet to VGG-Face [6] for face recognition using high and low resolution images at different distances.
• Develop a robust facial recognition pipeline to implement on to the humanoid robotic platform NAO.

Approach

• Retrain AlexNet on the CASIA-WebFace dataset to configure the neural network for face recognition tasks.
  • Acquire an input image using NAO’s camera or a high resolution camera to run through the convolutional neural network.
  • Extract the features of the input image using the neural networks AlexNet and VGG-Face
  • Compare the features of the input image to the features of each image in the people database.
  • Determine whether a match is detected or if the input image is a photo of a person who does not exist in the database.
  • Evaluate the overall performance of each neural network’s ability to perform face recognition tasks.

How it Works

Figure 1. Humanoid Robot Platform NAO.

Figure 2. Facial Recognition Pipeline utilizing NAO

Figure 3. Shows the different outcomes that occur when testing face recognition.

Output | Should The System Recognize You? | Does The System Recognize You?
---|---|---
True Positive (TP) | YES | YES
True Negative (TN) | NO | NO
False Positive (FP) | NO | YES
False Negative (FN) | YES | NO

Results

• VGG-Face Shows better results in every performance benchmark measured compared to AlexNet, although AlexNet is able to extract features from an image 800% faster than VGG-Face
• Resolution of the input image does not have a statistically significant impact on the performance of VGG-Face and AlexNet.
• AlexNet’s performance decreases with respect to distance from the camera where VGG-Face shows no performance loss.
• Both frameworks show excellent performance when eliminating false positives.
• Performance Charts shown below.

Table I. Face Recognition Accuracy of AlexNet and VGG-Face for High and Low Resolution Images

<table>
<thead>
<tr>
<th>Resolution</th>
<th>AlexNet TP</th>
<th>AlexNet TN</th>
<th>VGG-Face TP</th>
<th>VGG-Face TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Resolution</td>
<td>0.90</td>
<td>0.95</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>Low Resolution</td>
<td>0.80</td>
<td>0.85</td>
<td>0.82</td>
<td>0.83</td>
</tr>
</tbody>
</table>

Table II. Face Recognition Accuracy of AlexNet and VGG-Face with Respect to Distance Using High Resolution Images

<table>
<thead>
<tr>
<th>Distance</th>
<th>AlexNet TP</th>
<th>AlexNet TN</th>
<th>VGG-Face TP</th>
<th>VGG-Face TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 Feet</td>
<td>81.48%</td>
<td>100%</td>
<td>80.00%</td>
<td>90.00%</td>
</tr>
<tr>
<td>4 Feet</td>
<td>88.89%</td>
<td>100%</td>
<td>85.00%</td>
<td>95.00%</td>
</tr>
<tr>
<td>2 Feet</td>
<td>100%</td>
<td>100%</td>
<td>90.00%</td>
<td>98.00%</td>
</tr>
</tbody>
</table>

Table III. TPR and FPR using AlexNet and VGG-Face for High and Low Resolution Images

<table>
<thead>
<tr>
<th>Resolution</th>
<th>AlexNet TPR</th>
<th>AlexNet FPR</th>
<th>VGG-Face TPR</th>
<th>VGG-Face FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Resolution</td>
<td>0.90</td>
<td>0.00</td>
<td>0.85</td>
<td>0.05</td>
</tr>
<tr>
<td>Low Resolution</td>
<td>0.80</td>
<td>0.00</td>
<td>0.75</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Future Research and Applications

• Implement multi face recognition from a single image, and utilize more complex classification methods.
• Perform additional complex recognition tasks such as face expression recognition, text recognition and object recognition.
• Security Applications: espionage, defense, homeland security, surveillance.
• Medical Applications: MRI image Processing.

References