Comparing Human and Deep Convolutional Neural Network Performance on Twin Identification

Introduction

- Deep convolutional neural networks (DCNNs) achieve human-level accuracy on face recognition tasks [1].
- High performance across variable images.
- Rarely tested on highly similar identities.
- What about identical twins? Extreme case of high similarity.
- Simulates large-scale face recognition with highly similar identities.
- Previous face-recognition algorithms: front-facing, same lighting, same day (3, 4, 5, 6).

Goals

- Human and machine comparison for a face-identification task involving highly-similar identities (i.e., identical twins).
- Face-identification performance for twins across variation in viewpoint.

Procedure

- Face-matching task over viewpoint change that includes identical twins.
- Viewpoint conditions between subject: 0°–90°, 0°, 45°, 0°, 0°.
- Imposter type (with subject): twin-imposter or general-imposter.
- Dependent variables: Area under receiver operating characteristic (AUC) curve.

Stimuli:

- NO-TWINS-090-2010
- 200 identities—120 pairs
- Neutral expression.
- Same identity (N = 80).
- Twin-imposer = twin siblings (N = 40).
- General-imposer = two unrelated individuals (N = 40).

Identification-Matching Task:

- Human: 5-point scale.
- “Sure different” to “sure same.”
- DCNN: Similarity between DCNN embeddings of images.
- Cosine between embeddings.
- Performance measured using AUC.

Correct responses: Same identity/image pairs.
False alarms: Twin-imposer or general-imposer/image pairs.

DCNN Methods

- DCNN trained for face identification [7].
- Trained with approximately 3.6 million images of 58,000 identities [8].
- ResNet-101 based architecture.
- Crystal Loss function, alpha parameter set to 50.
- Generate DCNN embedding for each image.
- Embedding: 512 dimensions.

Face-Matching Accuracy: Human Participants and DCNN

- Corelating Human and DCNN Responses by Image-Pair Type

Correlation between human and machine performance

- Perceived similarity for each image pair.
- Human participants: Average rating across all participants.
- DCNN: Cosine similarity between embeddings.
- Significant correlation for 8 of 9 comparisons.
- p’s for twin-imposer: 0.0°–90°.
- Evidence for overlap in perceived similarity for human participants and a DCNN.

Conclusions

- A DCNN trained for face identification outperforms human participants on a face-identification task that includes identical twins.
- Human participants and the DCNN exhibit a similar decrease in face-identification accuracy as the difference in viewpoint between images increases.

Summary

- DCNNs now achieve state-of-the-art performance on benchmarks used for human participants when recognizing identical twins. This increased performance generalizes across changes in viewpoint. This indicates that accurate performance can be maintained even in large-scale applications. In addition, there is overlap in the perceived similarity between image pairs for human participants and a DCNN.

References

- Funding provided by National Eye Institute Grant R01EY029692.

Acknowledgements

Funding provided by National Eye Institute Grant R01EY029692 to AOT and CDC.

For information, please contact Connor.Pardo@utdallas.edu