Neurobiologically Inspired Encoding and Transfer Learning

George Sivulka (gsivulka)
Satchel Grant; Josh Melander; The Baccus Lab at Stanford University

Computational Neuroscience: the feedback loop between neuroscience discovery and machine learning breakthroughs

Present Research: a transfer learning method employing the computations and mechanisms of the human retina as an encoder for time-dependent visual stimulus—analyzing their efficacy in video classification tasks.

Recent advances in deep convolutional neural networks (CNNs) have proven to be the most successful models of the nervous system's sensory processing paradigms to date. [1]

Deep Retina (limited in cell type, quantity)

FC Deep Retina: Encoding Use

Visual Stimulus $X(0) \in \mathbb{R}^{[X \text{ X Y } \text{ #milliseconds}]}$

Visual Encodings (Features)

FC Deep Retina Output: Encodings from the filters of two learned retinal cells

Encoding Classification (RNN)

Schematic of the LSTM architecture with input $(X)$ frame bins first encoded by Deep Retina. Video classification $(y)$ is performed on a FC-layer's output of the final internal units $(h)$ of the LSTM.

Synthetic Video Data

Synthesized videos of simple artificial stimuli known to evoke retinal response—allowing for a proof of concept first pass at this transfer learning system

Synthetic Classification Results

UCF-11 Video Data

A curated dataset of action videos was used as a harder test for the system.

UCF-11 Classification Results

Future Work

Future work efforts will focus on rectifying validation and training discrepancies, further analyzing UCF-11 performance, and applying these encodings to RL and Meta-RL tasks.

Discussion

Synthetic classification yields almost perfect accuracy across the board, as expected

- Pretrained DR learns faster and both trainable and untrainable DR models have the least stochasticity

UCF11 classification is more difficult, yet DR outperforms

- Frozen DR weights have higher accuracy, quicker learning
- Trainable random init doesn’t seem to learn at all

Batch Norm Training/Validation Discrepancy

Acknowledgements

This research would not be possible without the assistance of Satchel Grant and Josh Melander of The Baccus Lab at Stanford. Additionally, this research is indebted to the teaching staff of Stanford’s CS 230: “Deep Learning” to which it was submitted as coursework.

References
