## **Computer vision in the dark using recurrent CNNs**

## December 7 2018, by Ingrid Fadelli



Architecture and example data. a) Architecture of gruCNN. Each channels activity depends on both the current input as well as the previous state. b) Classification performance of example gruCNN and cCNN when all test sequences had an SNR of 1/4. c) Original image and image with different SNRs for a firetruck (category truck) a reindeer (category deer), and a dog, shown without jitter. d–k) Color coded predicted probabilities (output of softmax) of the correct (positive) image category for gruCNN (d–g) and cCNN (h–k).



Horizontal axes show predicted probabilities over 51 frames, vertical axes over a range of SNRs. d) & h) and e) & i) correspond to performance in the fire truck and reindeer examples, respectively. The predictive probability at low SNRs continue improving over frames for the gruCNN predictions, but are relatively constant for the cCNN. f) & j) Data for the third example (the dog), in which the gruCNN fails (which is rare) while the cCNN predicts the category correctly at most SNRs. The average predicted probability for correct (positive) image category for all 10,000 test images is displayed in g) & k). Credit: Till S. Hartmann/arXiv:1811.08537 [cs.CV].

Over the past few years, classical convolutional neural networks (cCNNs) have led to remarkable advances in computer vision. Many of these algorithms can now categorize objects in good quality images with high accuracy.

However, in real-world applications, such as autonomous driving or robotics, imaging data rarely includes pictures taken under ideal lighting conditions. Often, the images that CNNs would need to process feature occluded objects, motion distortion, or low signal to noise ratios (SNRs), either as a result of poor image quality or low light levels.

Although cCNNs have also been successfully used to de-noise images and enhance their quality, these networks cannot combine information from multiple frames or video sequences and are hence easily outperformed by humans on low quality images. Till S. Hartmann, a neuroscience researcher at Harvard Medical School, has recently carried out a study that addresses these limitations, introducing a new CNN approach for analyzing noisy images.

Hartmann, who has a background in neuroscience, has spent over a decade studying how humans perceive and process visual information. In recent years, he became increasingly fascinated by the similarities



between deep CNNs used in computer vision and the brain's visual system.

In the visual cortex, area of the brain specialized in processing visual input, the majority of neural connections are made in lateral and feedback directions. This suggests that there is a lot more to visual processing than the techniques employed by cCNNs. This motivated Hartmann to test convolutional layers that incorporate recurrent processing, which is vital for the human brain's processing of visual information.





Detailed comparison of cCNN with Bayesian inference and gruCNN performance over a large range of SNR levels. Each model architecture was tested after training with slightly higher SNRs (default training) and after training with slightly lower SNRs (low training). a) & b) Percent correct over the course of 51 frames for different SNRs (color coded) using default training for a) the cCNN (with Bayesian Inference) and b) the gruCNN. c) Dots: correct classification for the model architectures at the last frame. Jitter in SNR values was added to increase readability of plots, but was not in the data. Lines: mean performance of the five models per architecture. d) Mean performance of gruCNNs minus mean performance of cCNNs for models trained with default and lower SNRs (green and red, respectively). SNR levels used during training are indicated by dots. Credit: Till S. Hartmann/arXiv:1811.08537 [cs.CV].

Using recurrent connections within the CNN's convolutional layers, Hartmann's approach ensures that networks are better equipped to process pixel noise, such as that present in images taken under poor light conditions. When tested on simulated noisy video sequences, recurrent CNNs (gruCNNs) performed far better than classical approaches, successfully classifying objects in simulated low quality videos, such as those taken at night.

Adding recurrent connections to a convolutional layer ultimately adds spatially constrained memory, allowing the network to learn how to integrate information over time before the signal is too abstract. This feature can be particularly helpful when there is low signal quality, such as in images that are noisy or taken in poor light conditions.

In his study, Hartmann found that cCNNs performed well on images with high SNRs, gruCNNs, outperformed them on low SNR images. Even adding Bayes-optimal temporal integrations, which allow cCNNs to integrate multiple image frames, did not match gruCNN performance.



Hartmann also observed that at low SNRs, gruCNNs predictions had higher confidence levels than those produced by cCNNs.

While the human brain has evolved to see in the darkness, most existing CNN are not yet equipped to process blurry or noisy images. By providing networks with the capacity to integrate images over time, the approach devised by Hartmann could eventually enhance computer vision to the point that it matches, or even exceeds, human performance. This could be huge for applications such as self-driving cars and drones, as well as in other situations where a machine needs to 'see' under nonideal lighting conditions.

The study carried out by Hartmann could pave the way for the development of more advanced CNNs that can analyze images taken under poor light conditions. Using recurrent connections in the early stages of neural network processing could vastly improve computer vision tools, overcoming the limitations of classical CNN approaches in processing noisy images or video streams.

As a next step, Hartmann could expand the scope of his research by exploring real-life applications of gruCNNs, testing them in a wide range of real-world scenarios. Potentially, his approach could also be used to enhance the quality of amateur or shaky home videos.

**More information:** Seeing in the dark with recurrent convolutional neural networks. arXiv:1811.08537 [cs.CV]. <u>arxiv.org/abs/1811.08537</u>

© 2018 Science X Network

Citation: Computer vision in the dark using recurrent CNNs (2018, December 7) retrieved 27 April 2024 from <u>https://techxplore.com/news/2018-11-vision-dark-recurrent-cnns.html</u>



This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.