CONVOLUTIONAL NEURAL NETWORKS DURING OBJECT IDENTIFICATION

Katrenko M.O.
National Aviation University, Kyiv
Supervisor – Sineglasov V.M., Doctor of technical sciences, Professor

Nowadays, convolutional neural networks perform very well in identifying objects, but unfortunately, they have very important problems that are very difficult to get rid of.

Convolutional networks use multiple layers of feature detectors. Each feature detector is local, so feature detectors are repeated across the space. Pooling gives some translational invariance in much deeper layers, but only in a crude way. The psychology of shape perception suggests that the human brain achieves translational invariance in a much better way. We know that, roughly speaking, the brain has two separate pathways, a “from” and “where”.

Neurons in the “from” pathway respond to a particular type of stimulus regardless of where it is in the visual field. Neurons “where” are responsible for encoding where things are. As a side note, it is hypothesized that the “from” has a lower resolution then “where”.

The most important thing to admit about the "from" path is that it does not know where the objects are without informing the "where" path. If a person's path is "damaged," he can determine if there is an object, but cannot track where it is in sight and where it is relative to other objects. This leads to **simultanagnosia**, a rare neurological condition in which patients can see only one object at one time.

And the second problem. Pooling was introduced to reduce the redundancy of the presentation and reduce the number of parameters, given that the exact location is not important for object detection. However, deep learning engineers have not given much thought to routing. Pooling performs routing in a very rough way - for example, with maximum pooling, the neuron with the maximum activation is selected, and not the one that is most likely related to the task at hand.

So, this leads to very important problems:
- poor translational invariance and lack of orientation information
- severe loss of information on pooling stage.

Resources:

1. Geoffrey E. Hinton, A. Krizhevsky & S. D. Wang. URL: http://www.cs.toronto.edu/~fritz/absps/transauto6.pdf (Last accessed: 17.02.2021).
2. Sara Sabour, Nicholas Frosst, Geoffrey E Hinton. URL: https://arxiv.org/abs/1710.09829 (Last accessed: 14.02.2021).
3. Geoffrey E. Hinton. URL: https://u.to/Ov4rGw (Last accessed: 14.02.2021).
4. Anish Athalye, Logan Engstrom, Andrew Ilyas & Kevin Kwok. URL: https://www.labsix.org/physical-objects-that-fool-neural-nets/ (Last accessed: 14.02.2021).