Genome Bottleneck in Deep Learning

Ivan Skorokhodov

May 5, 2020

Overview

- Animals and humans have a lot of innate abilities and inclinations:
  - A colt can walk within hours after birth
  - Turkeys can visually recognize predators shortly after hatching
  - Sharks are attracted to blood immediately after birth
  - Spiders are born ready to hunt
  - Human newborns can recognize human faces
- I.e. they have higher than random performance at initialization and good few-shot learning capabilities for critical tasks
- At birth, an infant has ≈ 86 billion neurons (this number remains almost the same during life) and 10-100 trillion synapses which grow rapidly over the first 3 years [1].
  - Even if each value is binary it is ≈ 1300 GB of information
- Human genome can store up to 800 MB of data, which is 1625 times less than amount of information in an infant brain[2]
- So, human genome cannot store all the information about the brain
- How then a brain can have good performance from initialization?

---

[1] http://www.urbanchildinstitute.org/why-0-3/baby-and-brain

[2] if all the genome stores is an information about the brain