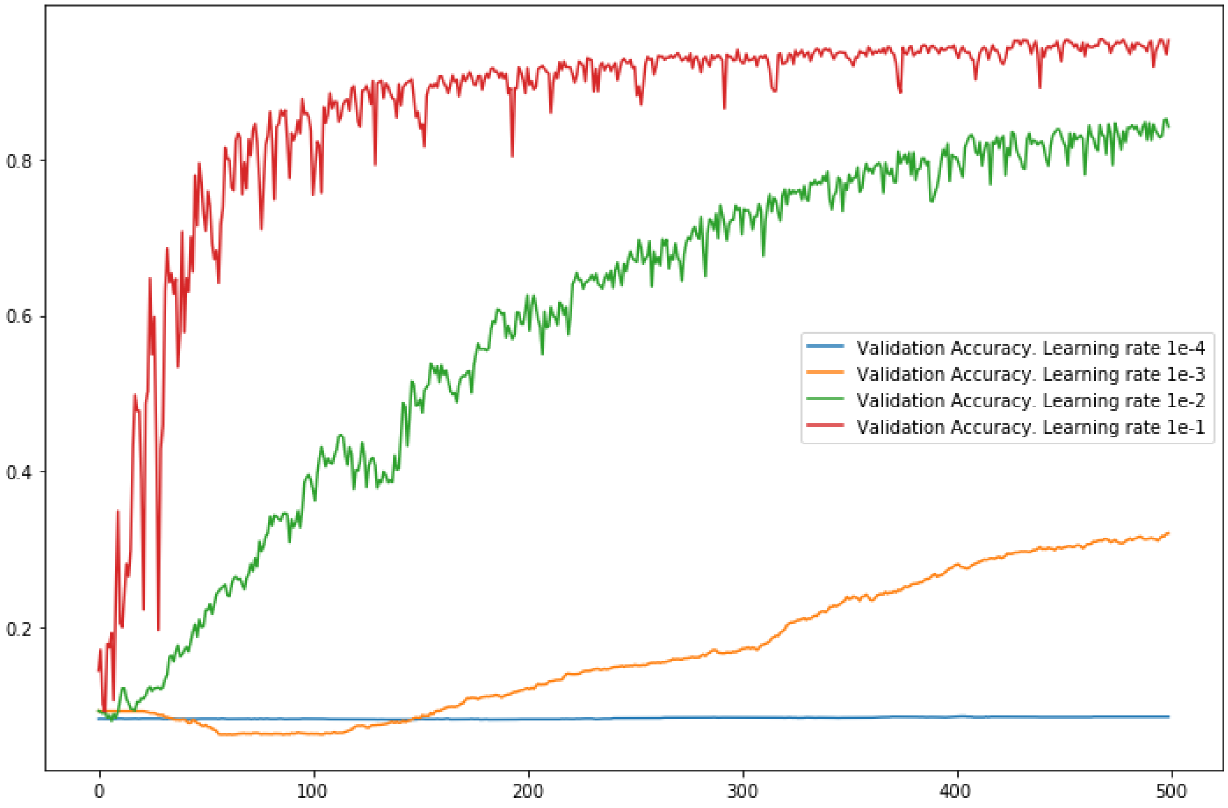
**Deep Learning Lab Course 2017**

**Assignment 2.**

**Olesya Tsapenko**

I have used the tutorials from the website https://www.tensorflow.org and implemented a CNN as it was described at the first part of this assignment. However, I used a learning rate 0,0001. Understanding that that was a mistake came to me on the second part of the assignment. Below, you can see the graphics with different learning rates.



Conclusions from this graphics:

- learning rate 0.0001 is too small that we do not have any convergence at all;
- learning rate 0.001 is still too small however we have convergence, but it may take very long;
- learning rate 0.01 is pretty good but even 500 epochs not enough for accuracy as minimum 80%;
- learning rate 0.1 is the best.

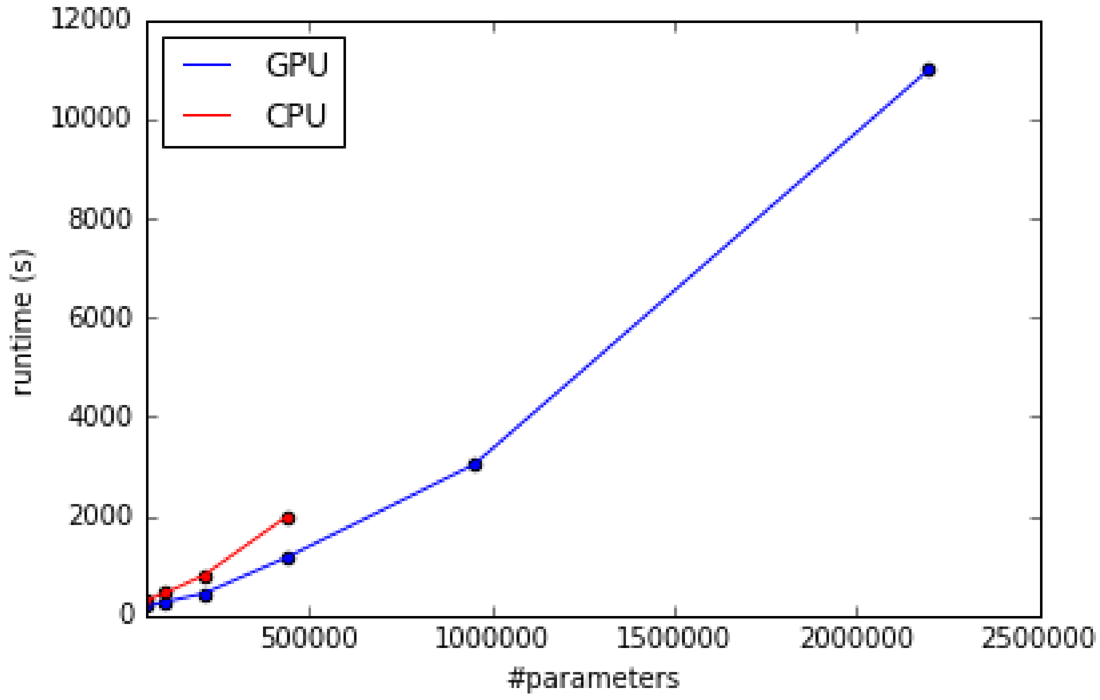
Based on this graph, I had decided to do all subsequent calculations for the third job with learning rate 0.1.

For calculations on CPU and GPU, I had made a two different virtual environments. The results (seconds for training) for the different number of filters you can see below (all numbers are calculated for 500 epochs).

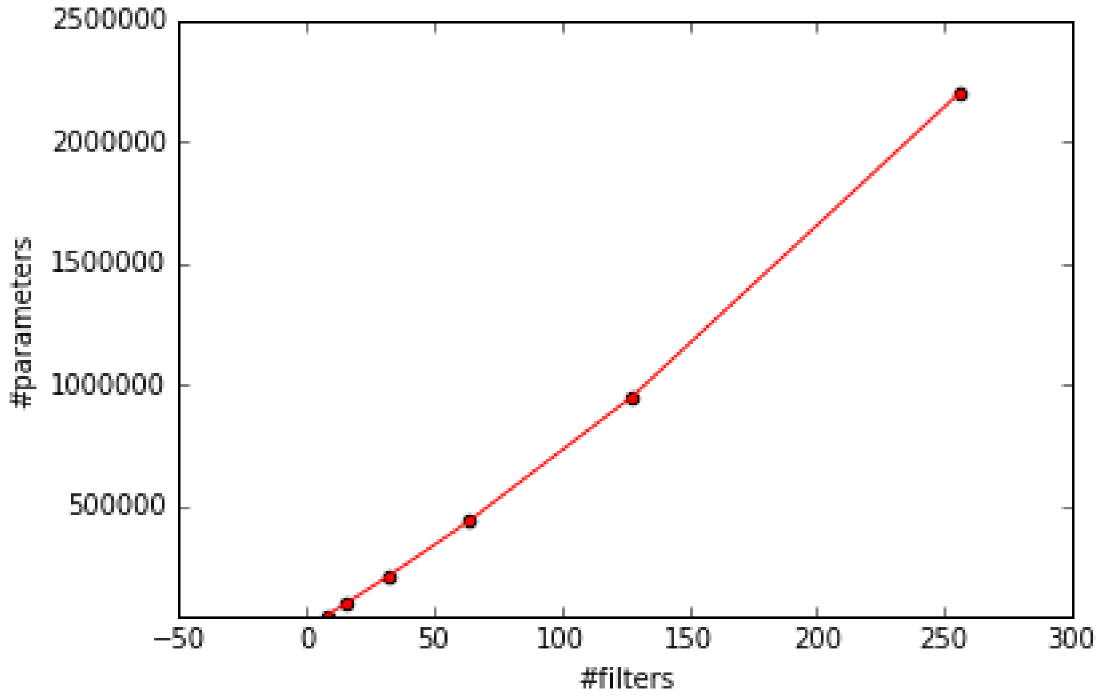
| **#filters (#all_parameters)** | **CPU**<br>**(runtime s)** | **GPU**<br>**(runtime s)** |
|---|---|---|
| **8 (52376)** | 305 | 178 |

| | | |
|---|---|---|
| **16 (104368)** | 447 | 271 |
| **32 (211808)** | 809 | 441 |
| **64 (440512)** | 2001 | 1174 |
| **128 (953216)** | tests were not conducted | 3040 |
| **256 (2199808)** | tests were not conducted | 10995 |

Below you can find two graphics with a relationship between the number of parameters and runtime for GPU and CPU respectively.



And I would like to provide one more plot with relationship between the number of filters and the total number of parameters of CNN.



Conclusions from the plots and the table above :

- the growth rate of runtime at CPU  is much higher than the one at GPU;
- using CPU for a CNN looks unreasonable even for 64 kernels;
- despite the fact that computing on GPU is faster, whole process is still very time-consuming (time and the number of parameters grow sharply);
- there is no central heating in the computer pool at the weekends.