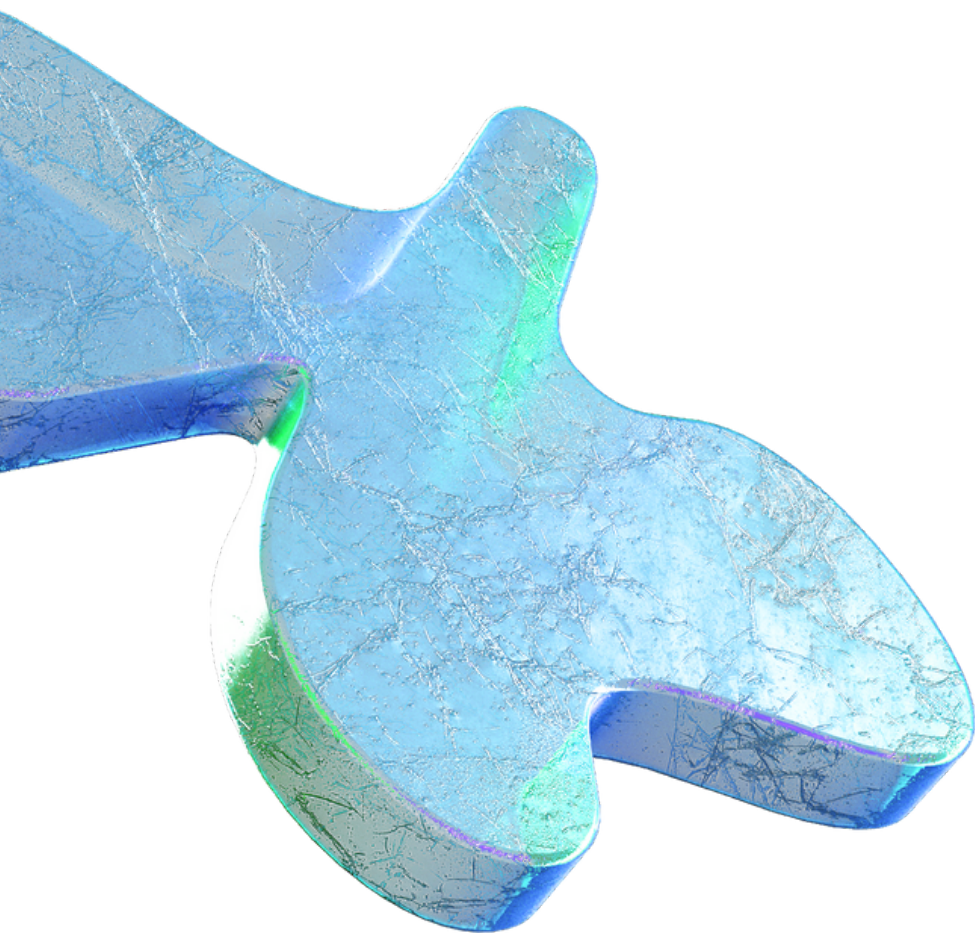


Analog of AlphaZero for chess

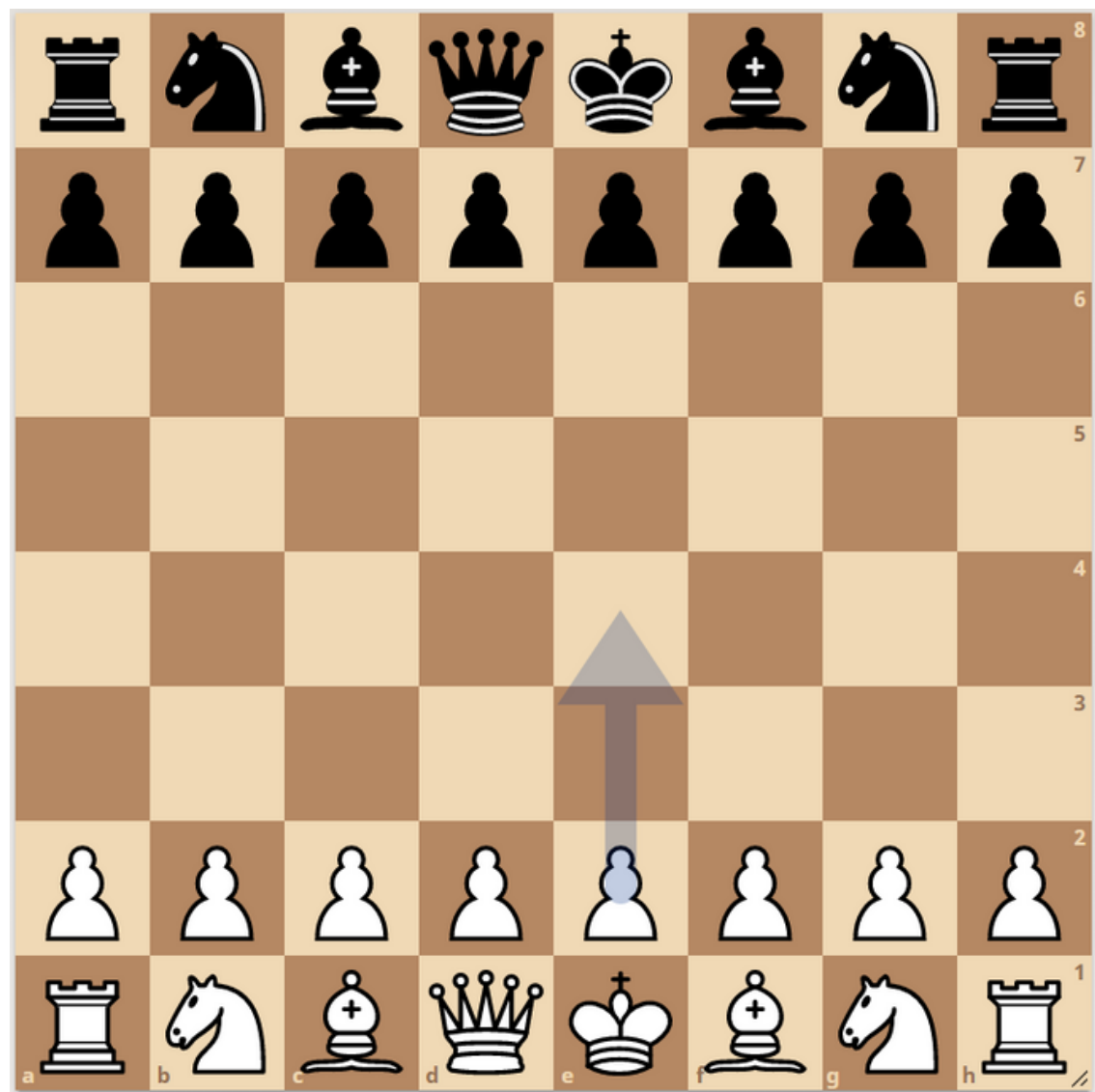


Сергеев Александр Сергеевич

Цель работы

Научить сеть играть в шахматы без датасета.

Для этого было решено предсказывать оценку ситуации на поле (stockfish), а не сам ход.



Stockfish

Stockfish занимает первые места
большинства рейтинговых списков



Stockfish анализирует ситуацию на поле,
делая наилучший ход. Этот анализ основан
на переборе будущих ходов (обычно на 20
и более ходов)

Dataset

Был найден датасет, нормировал

В нем было:

- Немного более 2 млн позиций(FEN)
- Оценки Stokfish



Кодировка

Перекодировал шахматную доску в тензор (14, 8, 8)

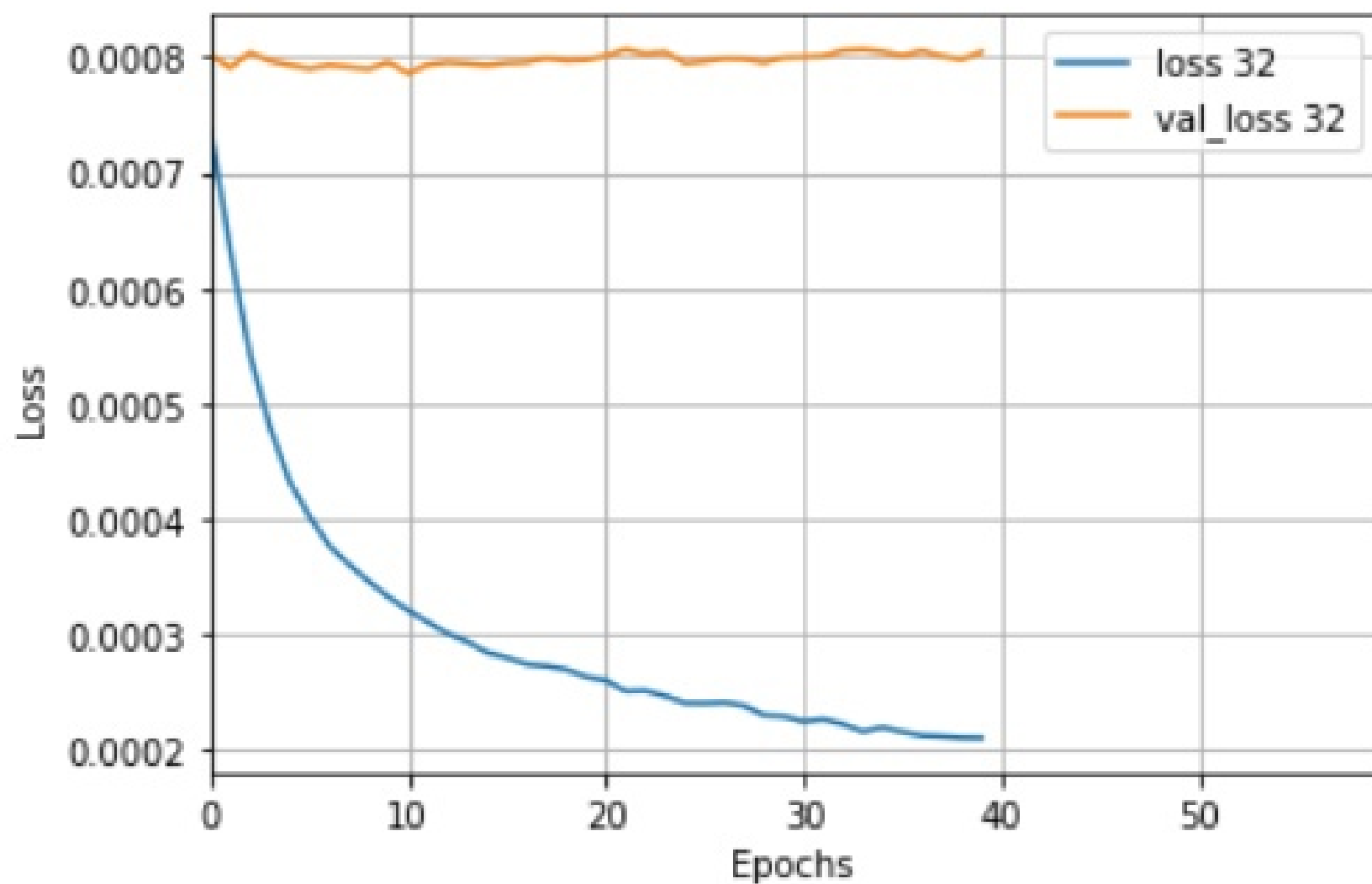
6 белых + 6 черных + возможные ходы (только 0 и 1)

FEN: q7/6P1/6K1/8/8/8/4Q3/1k6 w - - 11 99

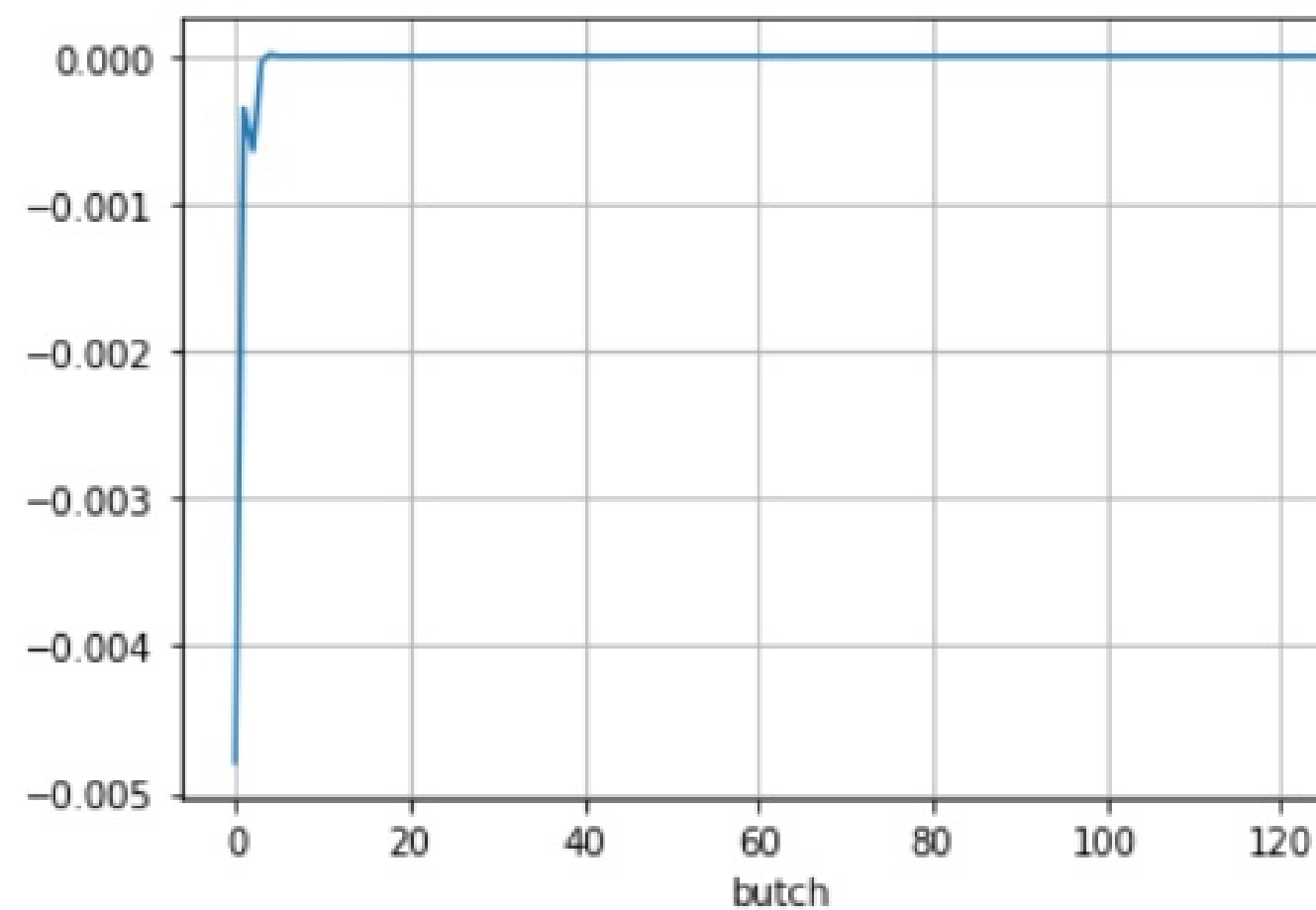
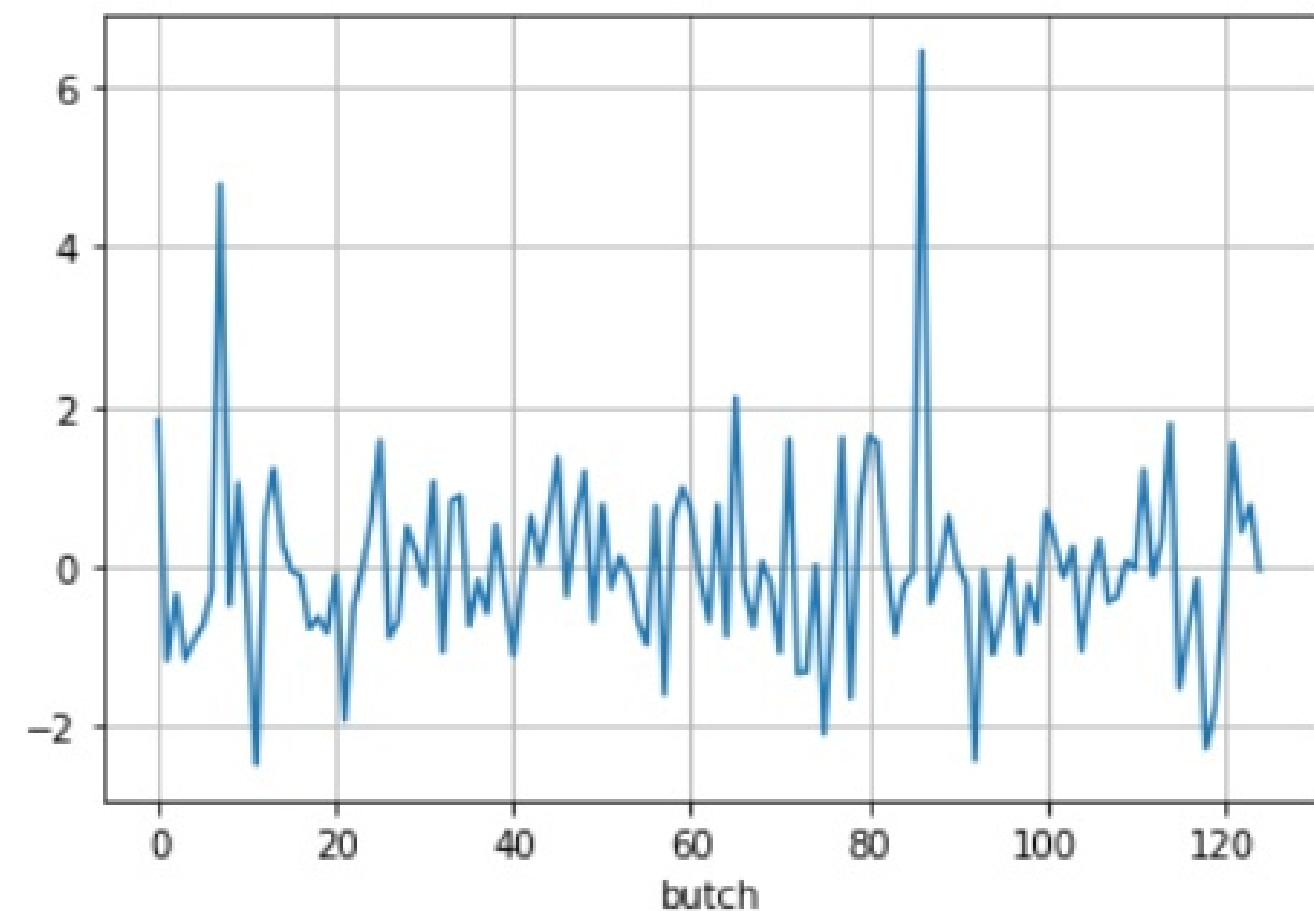
Модель

В модель входило
4 Con2d + 6 Dense слоев

Обучение



Градиенты



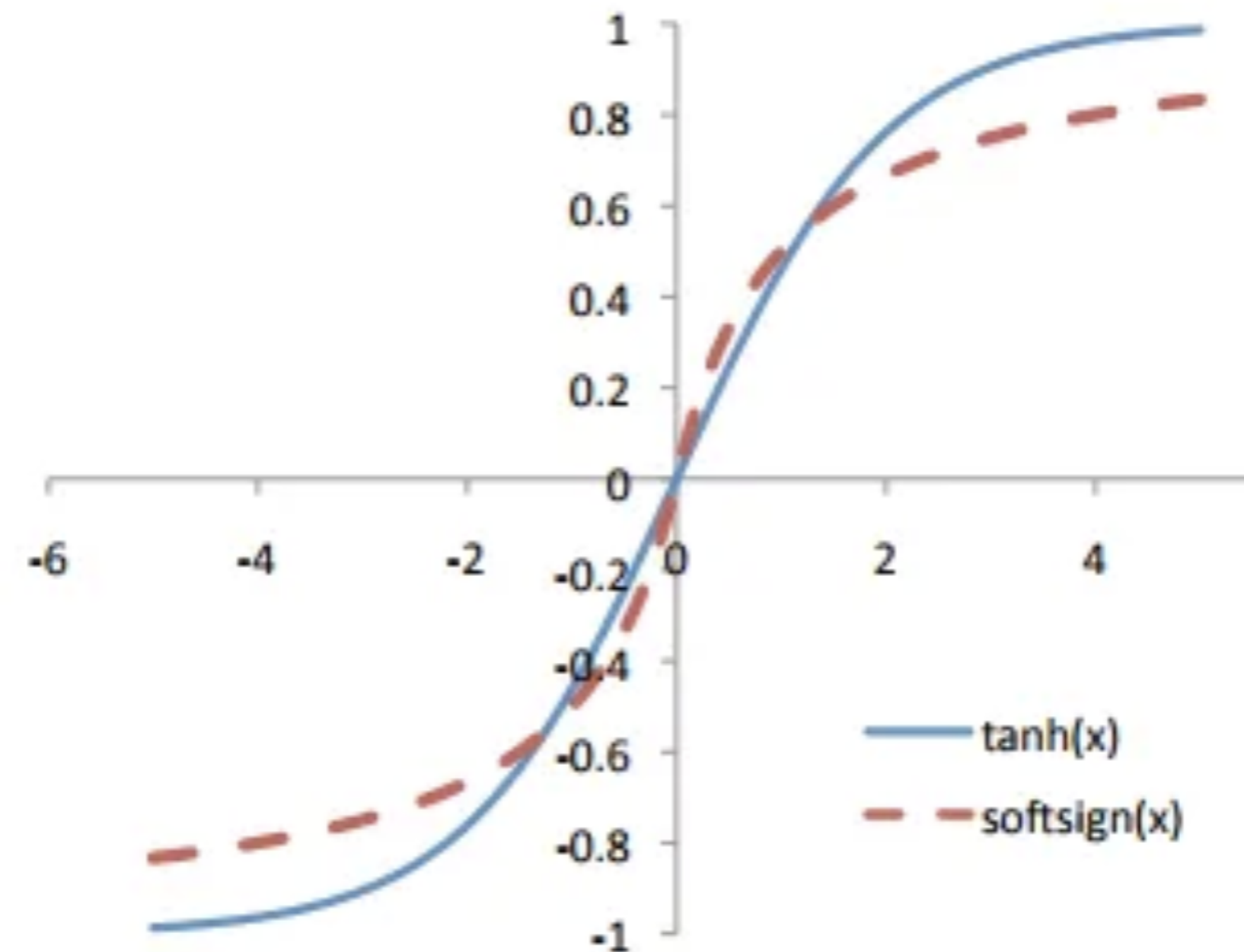
Как все исправить?

Было протестировано более 200 различных моделей, изменял:

- конфигурацию и количество слоев (residual, only dense)
- активацию слоев
- размер модели

Стало лучше когда

- оставил один сварочный слой
- два скрытых Dense слоя
- активация softsign



Однако ничего не получилось

Позже была обнаружена научная статья из Stanford University. В ней использовали такую же конфигурацию сети. Авторы утверждали, что в такой конфигурации невозможно обучить сеть

Predicting Moves in Chess using Convolutional Neural Networks

Barak Oshri
Stanford University
boshri@stanford.edu

Nishith Khandwala
[Stanford University](mailto:nishith@stanford.edu)
nishith@stanford.edu

Abstract

We used a three layer Convolutional Neural Network (CNN) to make move predictions in chess. The task was defined as a two-part classification problem: a piece-selector CNN is trained to score which white pieces should be made to move, and move-selector CNNs for each piece produce scores for where it should be moved. This approach reduced the intractable class space in chess by a square root.

reported a 44.4% accuracy in predicting professional moves in Go, a game known for its abstract logical reasoning that experts often describe as being motivated by faithful intuition [2]. This is an exciting result, indicating that CNNs trained with appropriate architectures and a valid dataset can catch up with much of the experience-based human reasoning in complex logical tasks.

The success of CNN-Go can be attributed to smooth arrangements of positions that are approximately continuous

Изменения

Из статьи была заимствована **модель кодирования** шахматной доски. (8, 8, 6)

Также в тензор, которым кодируется доска, добавлен **рандомный шум**, чтобы избежать падения градиента на свертках.

Найден **новый датасет** на kaggle



Почему тензор (8, 8, 6)

- Шахматная доска 8 на 8 плюс 6 фигур
- Где белые 1 и черные -1
- Шум (значения от 0 до 0.1)

The diagram illustrates the construction of an 8x8x6 tensor as the sum of three 8x8 matrices. Each matrix is displayed as a grid of values, with large plus signs indicating their summation.

Matrix 1 (Left):

0.	0.	0.	0.	0.	0.	0.	0.
-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
1.	1.	1.	1.	1.	1.	1.	1.
0.	0.	0.	0.	0.	0.	0.	0.
-1.	0.	0.	0.	0.	0.	0.	-1.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
1.	0.	0.	0.	0.	0.	0.	1.

Matrix 2 (Middle):

0.	-1.	0.	0.	0.	0.	-1.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	1.	0.	0.	0.	0.	1.	0.
0.	0.	-1.	0.	0.	-1.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	1.	0.	0.	1.	0.	0.

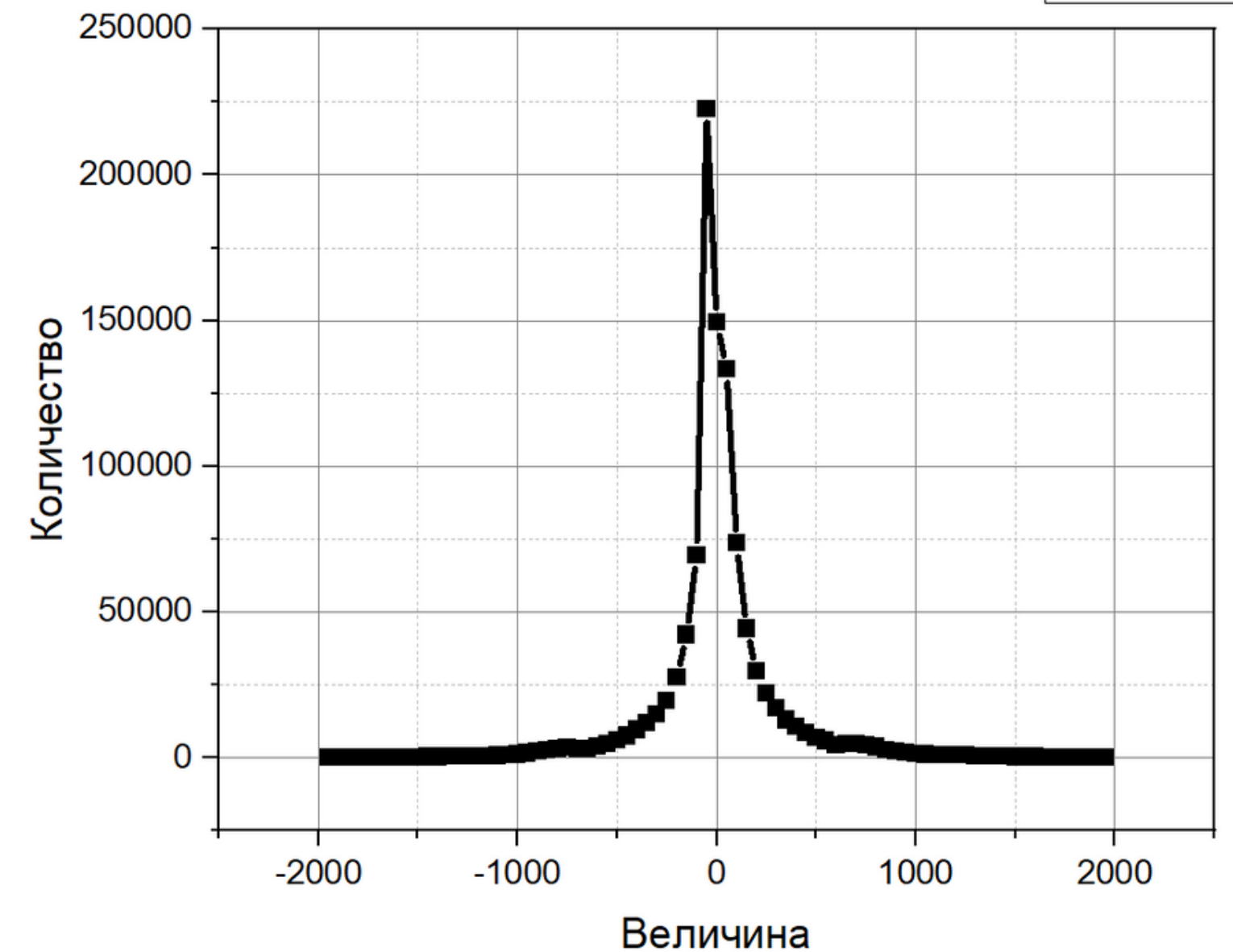
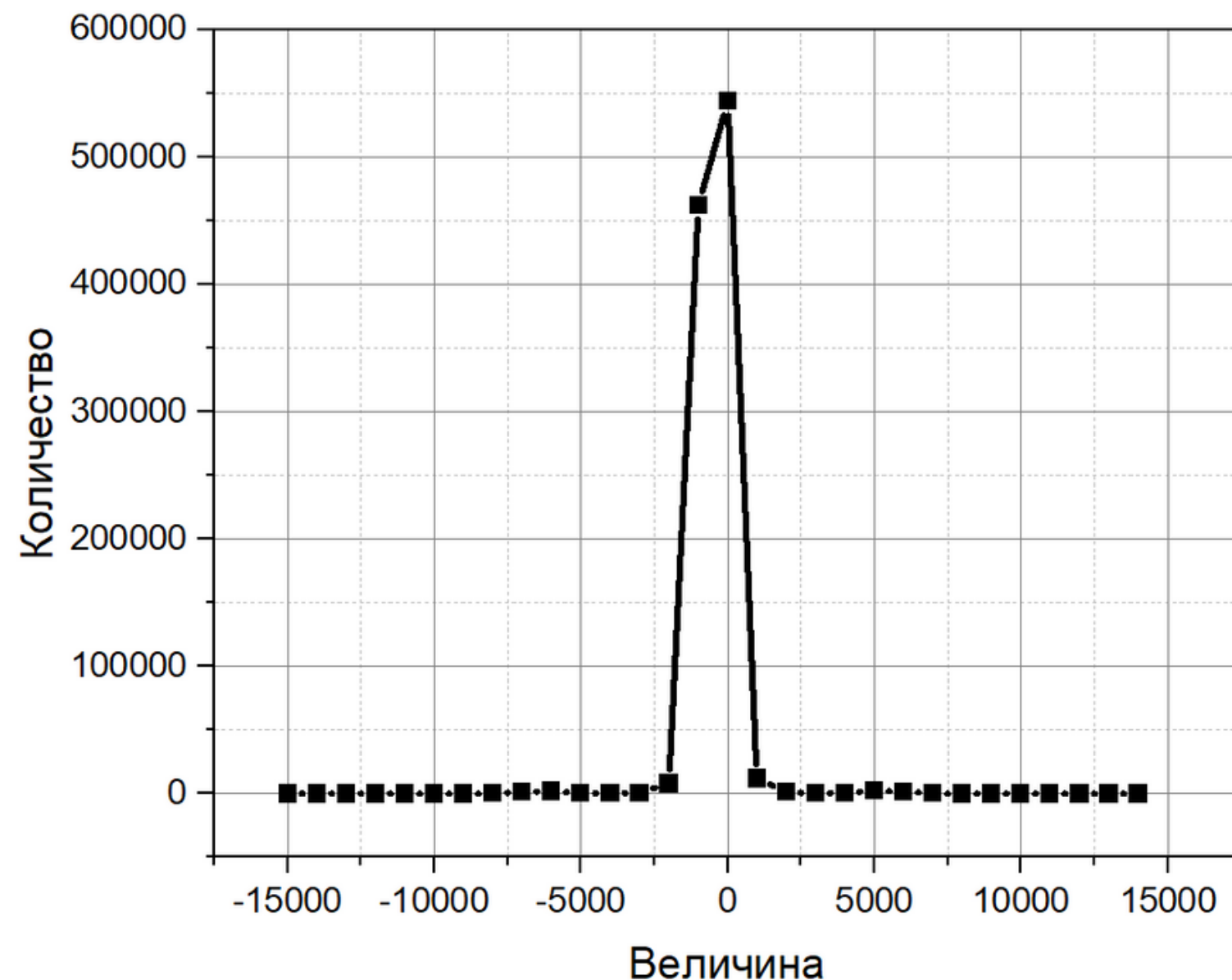
Matrix 3 (Right):

0.	0.	0.	-1.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	1.	0.	0.	0.	0.
0.	0.	0.	0.	-1.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	1.	0.	0.	0.

Dataset

Для обучения был выбран датасет с 13 млн позициями и их оценками. В качестве оценки использовался Stockfish

Среднее = 41,6 сп Медиана = 15 сп Дисперсия = 659,3 сп



Преобразование датасета

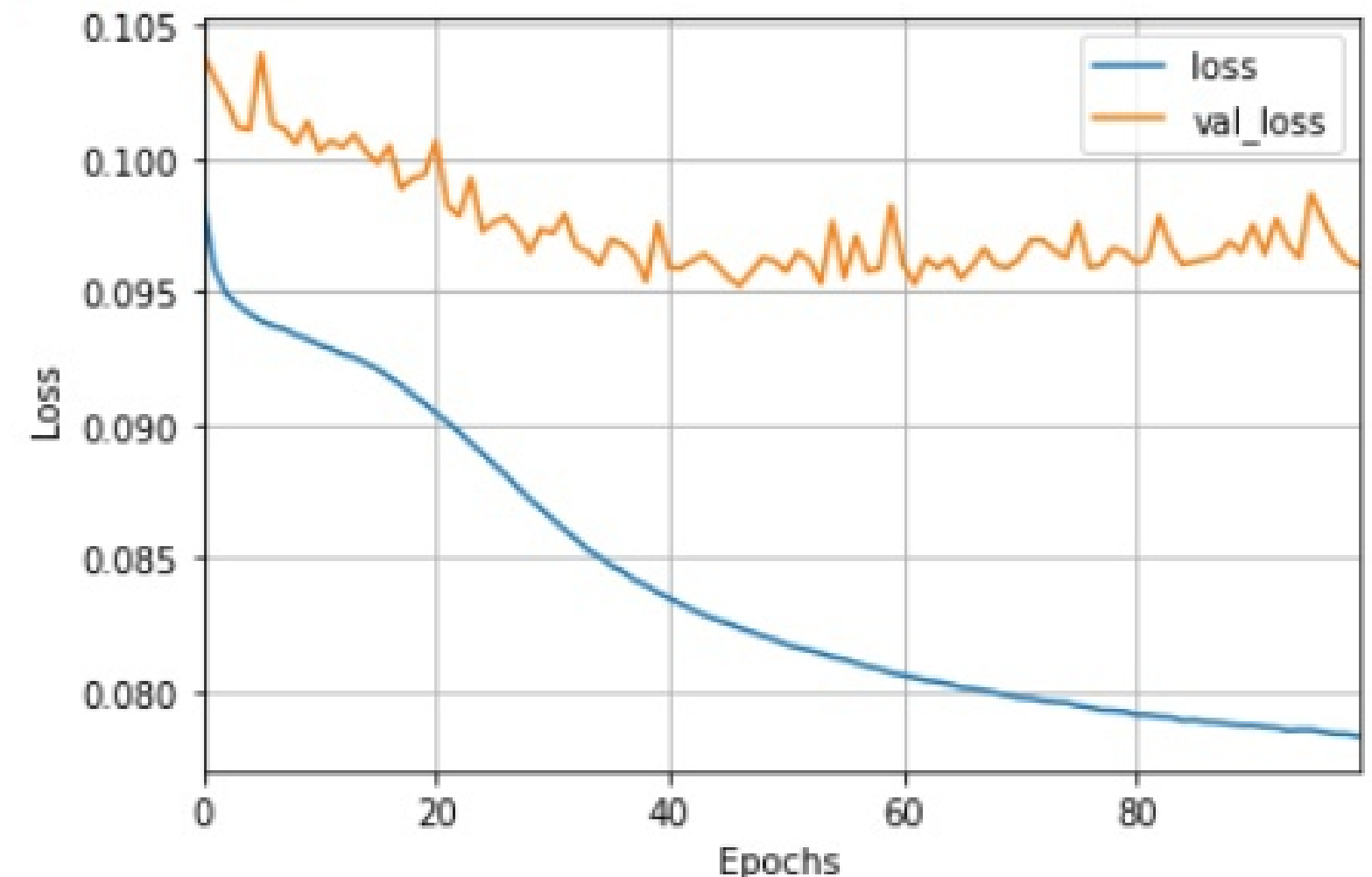
Сделал подобие z-преобразования:

- вычел среднее
- поделил на стандартное отклонение

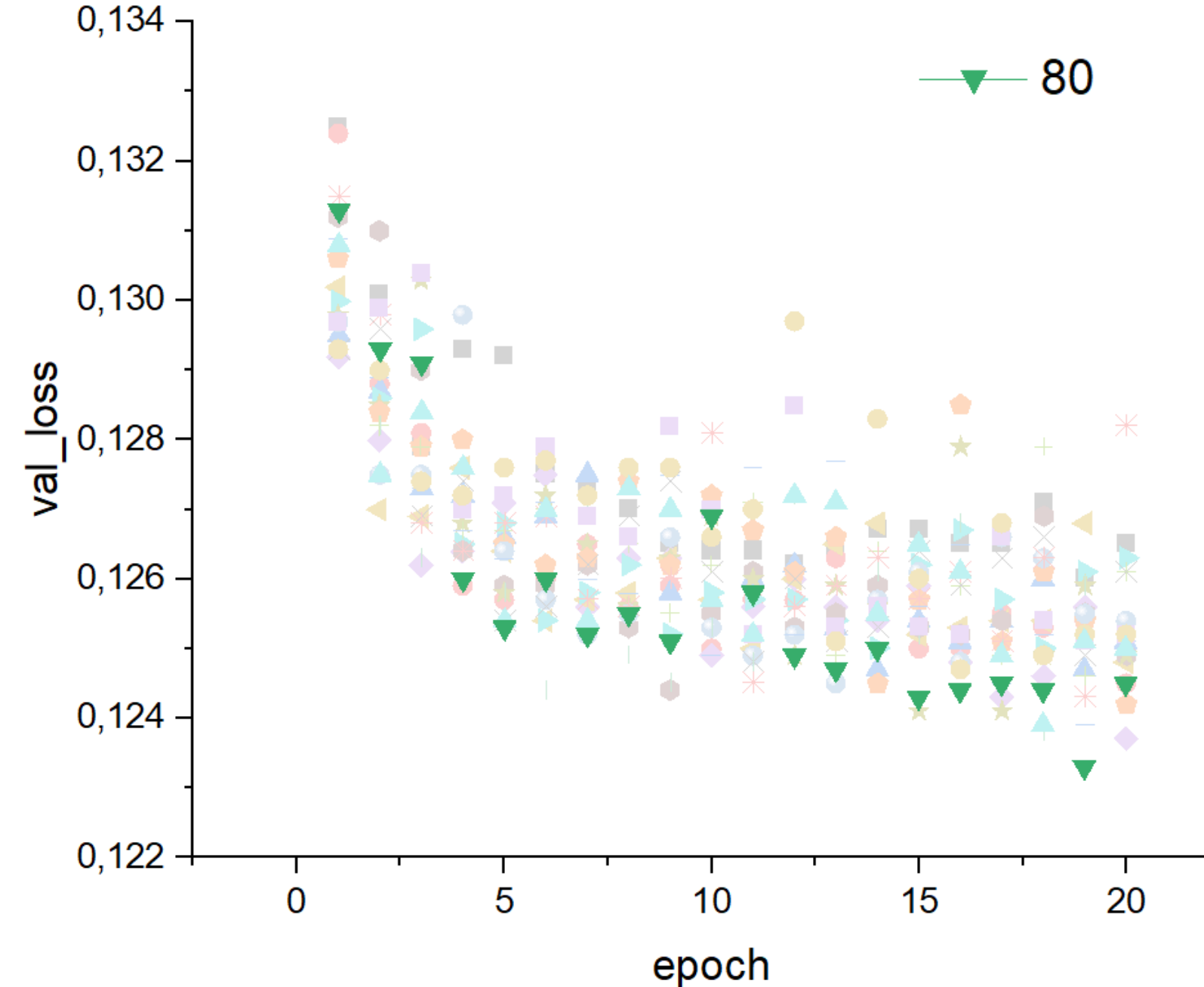
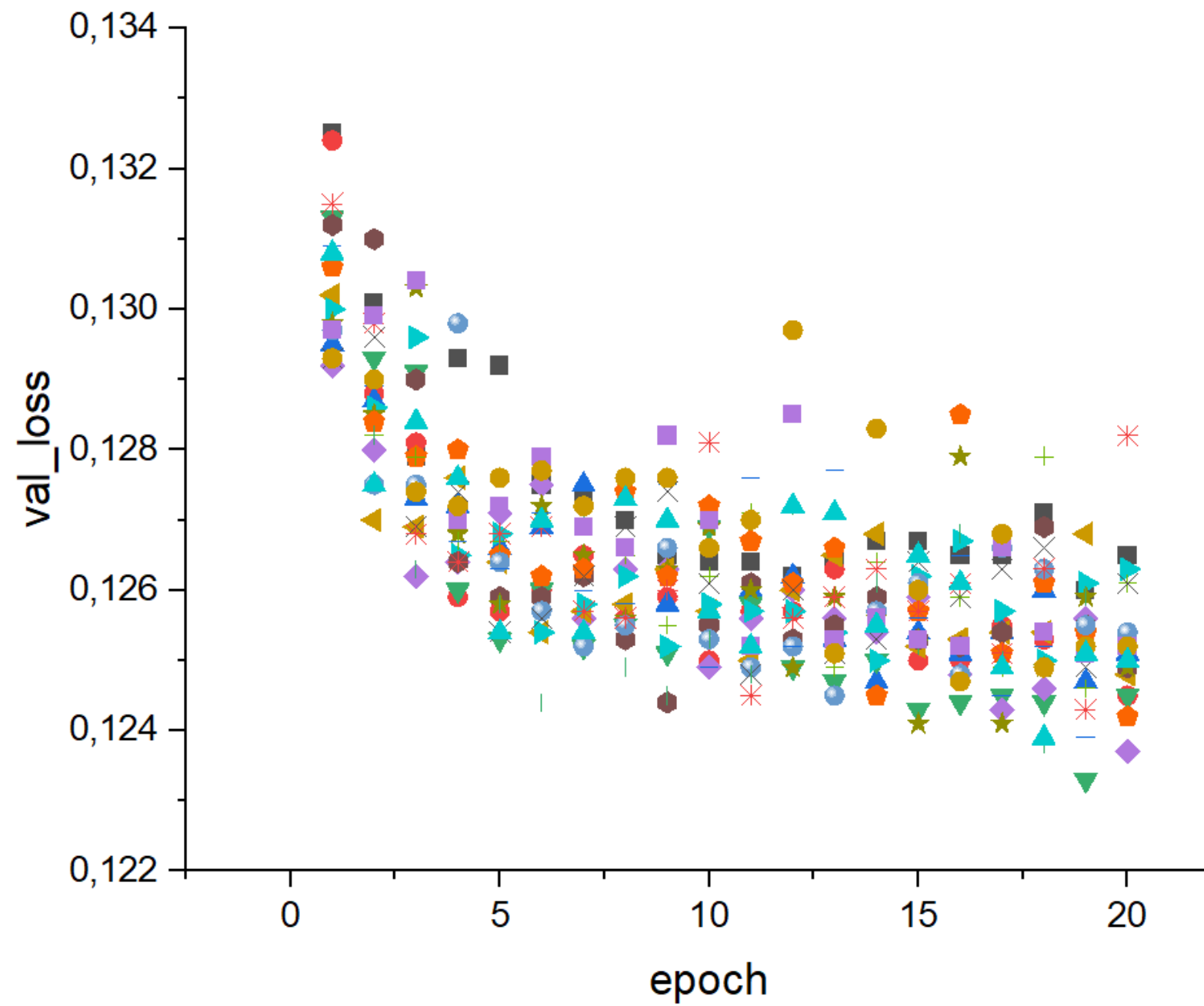
Модель

Оптимальная конфигурация :

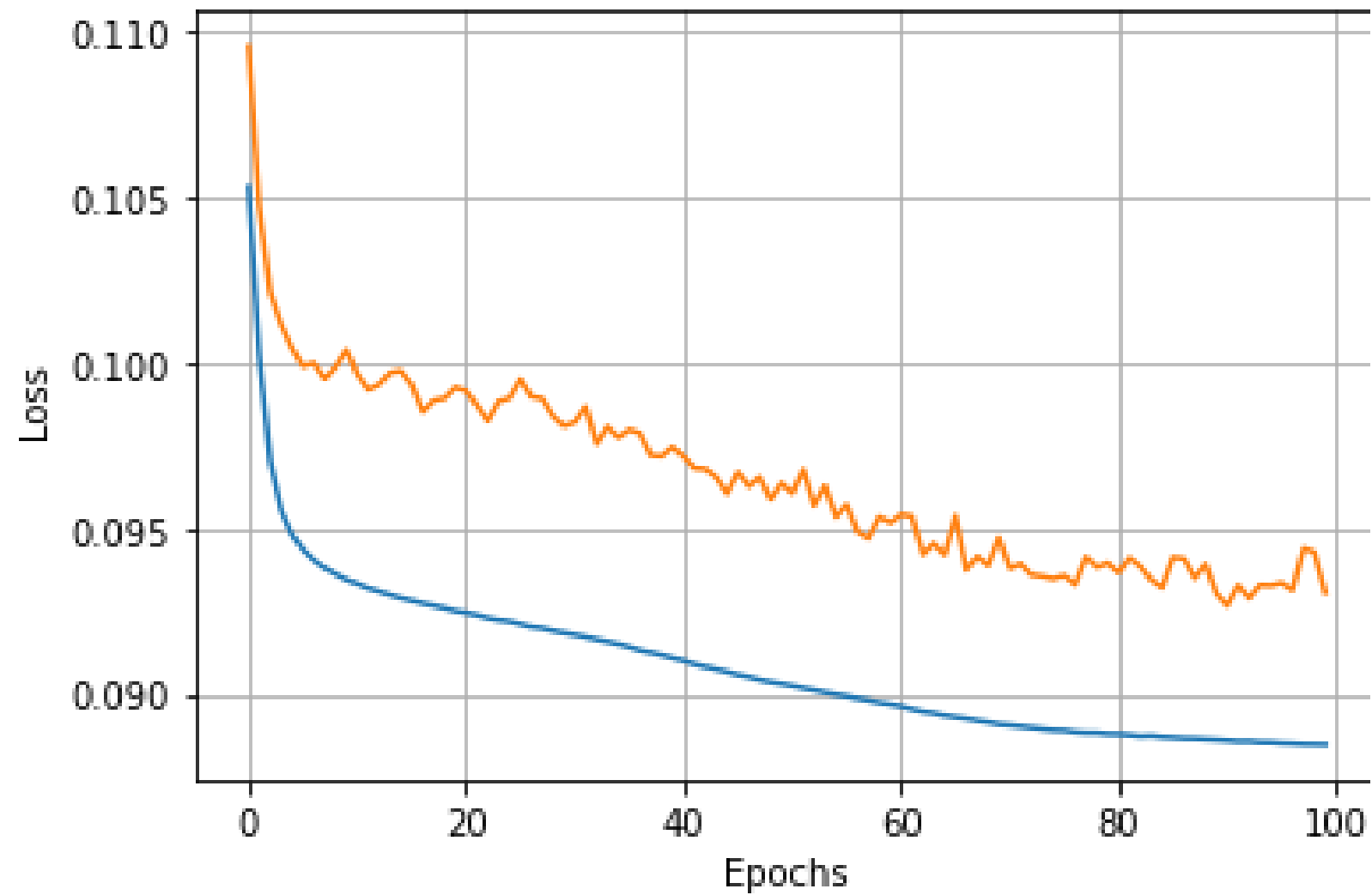
- Conv2D
- Dense*2
- Dense



Лучшая конфигурация

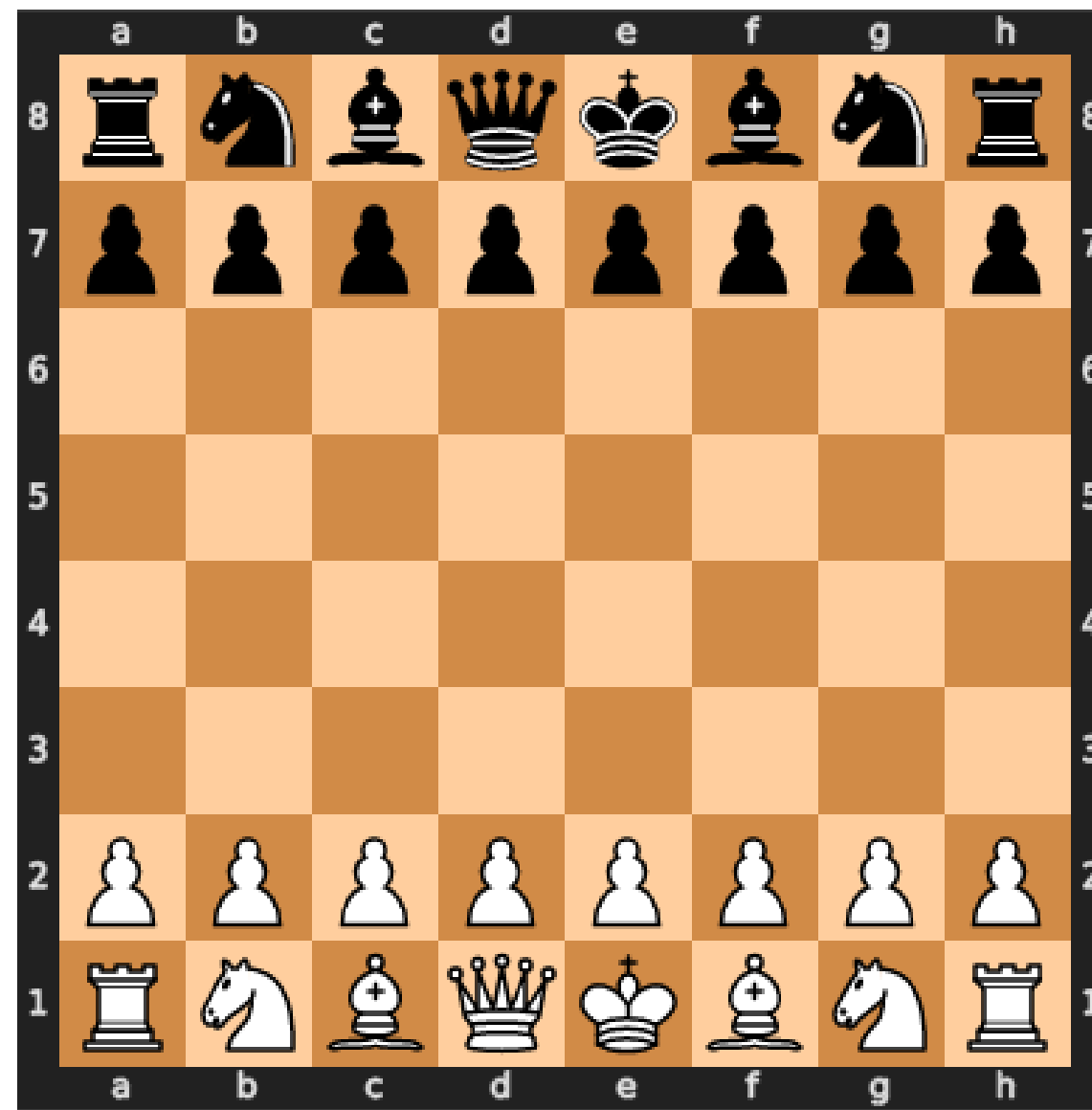


Итоги:



- **Сеть начала учиться**
- **Шум помог обучению**
- **Переиграть человека не удалось**

Демонстрация работы



FIGHT

