

Машинное обучение в медицине

**Автоматическое распознавание лёгких
на флюорографических снимках.**

**Захаров Р.В.
344 гр.**

**Научные руководители:
К.Н. Невоструев,
д.ф.-м.н., проф. А.Н. Терехов**

2016 г.

Суть работы

Булгаков Андрей

Gui

Захаров Роман

Детекция легких

Шигаров Никита

Пиксельфикс

Автоэнкодер

Цель работы

**Использовать машинное обучение
для обработки снимков легких и
изучить возможности
диагностирования туберкулеза.**

Что у нас есть?

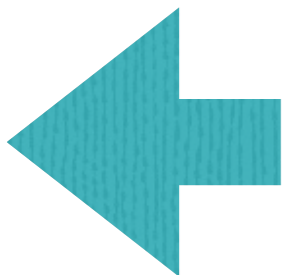
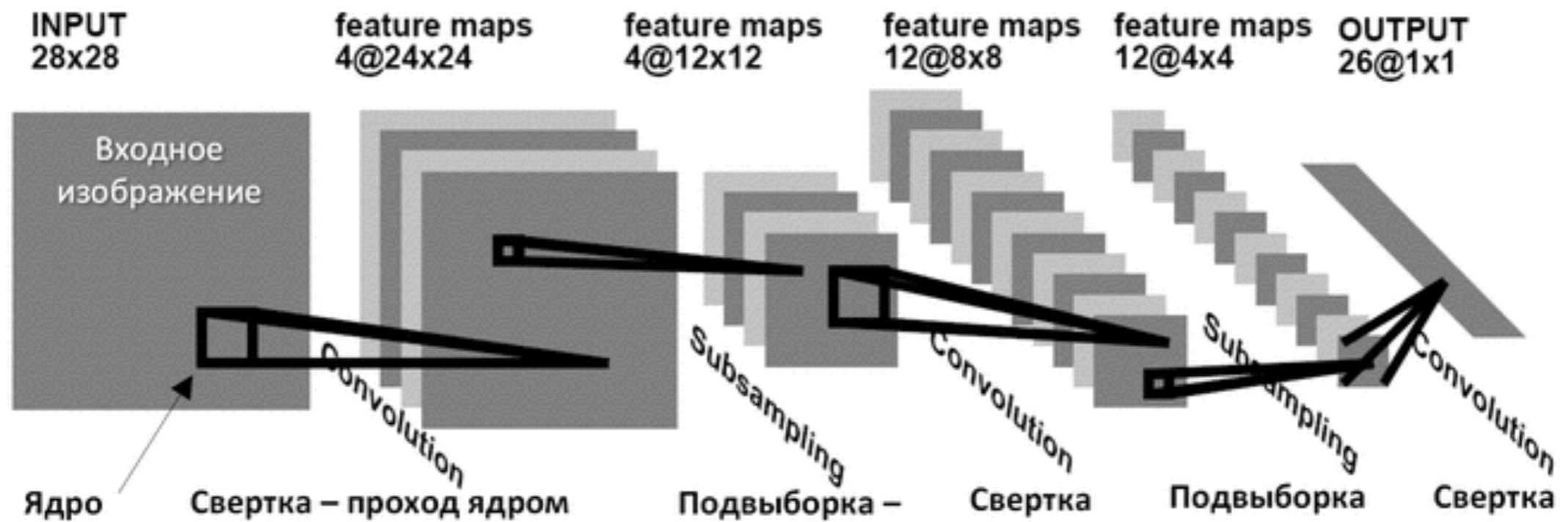
Снимки здоровых людей

Снимки нездоровых людей

Много неотсортированных снимков

Ожидается помощь специалистов

Устройство CNN



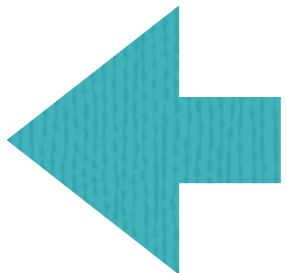
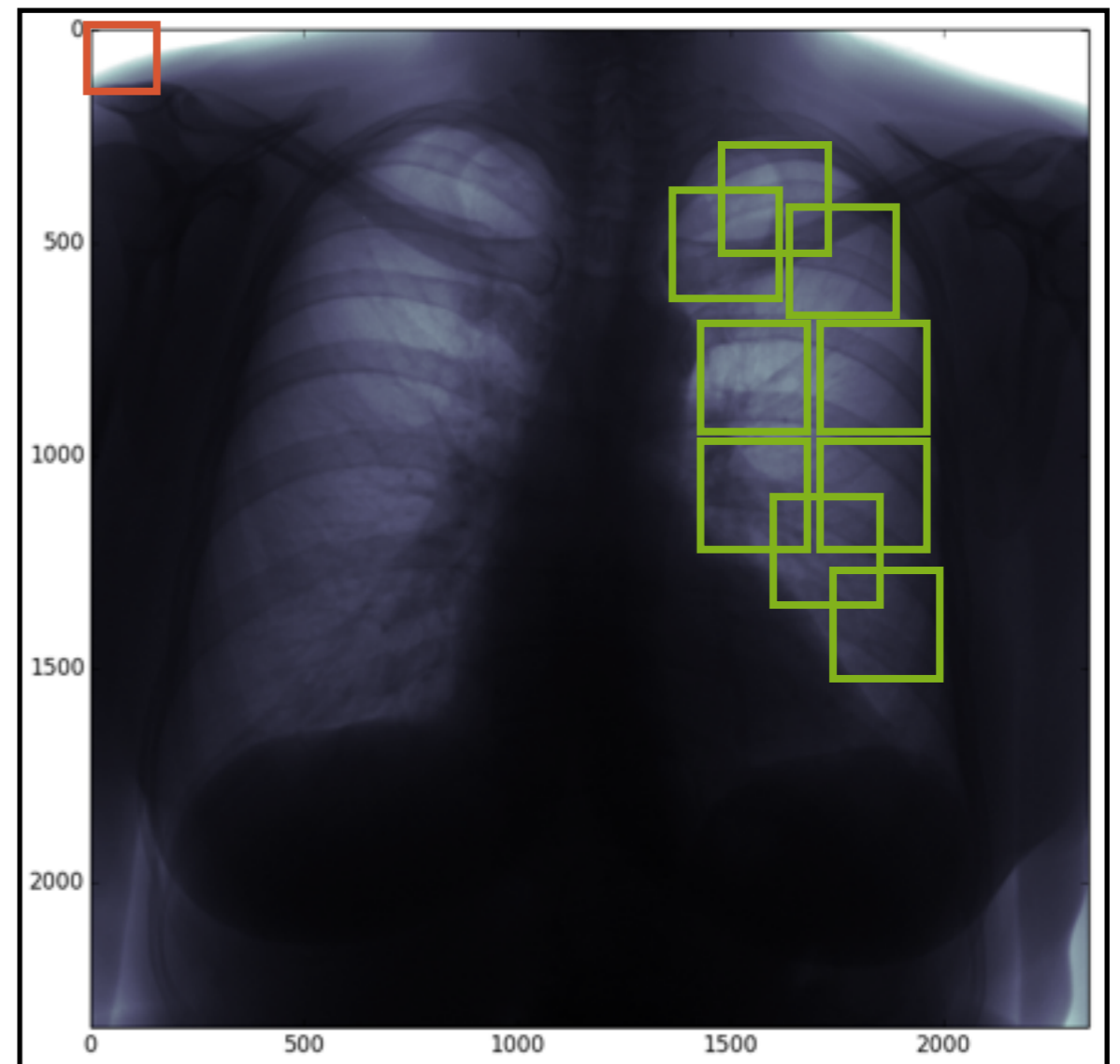
Малое количество зараженных легких

Больное легкое заражено не целиком

Недостаточно данных для обучения

Лишние данные на изображении

Уменьшение необходимой памяти

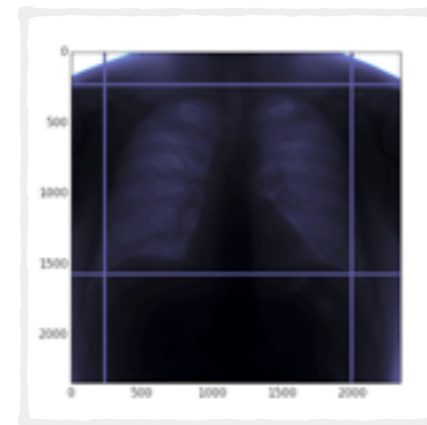
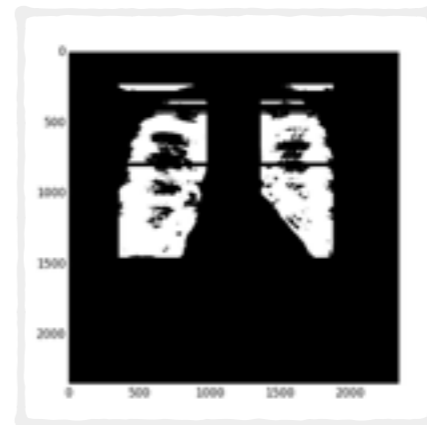
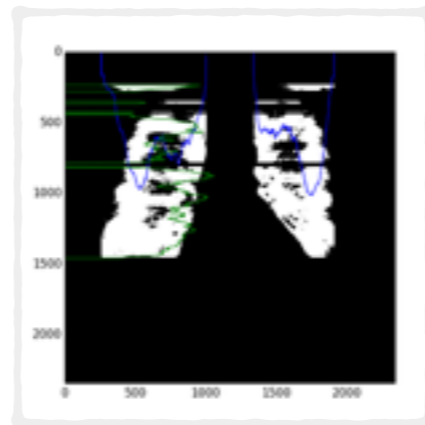
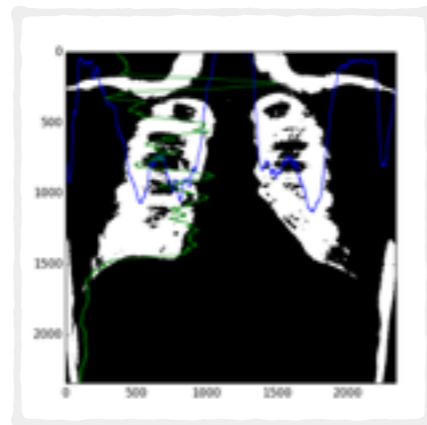
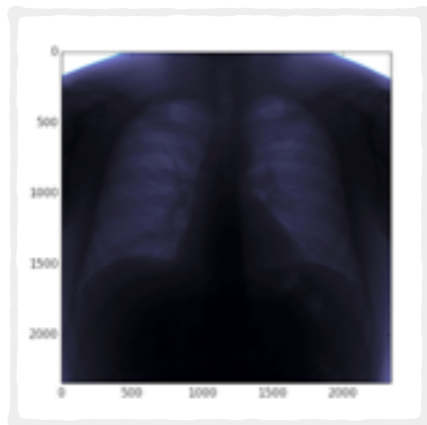


Автоматическое обнаружение легких

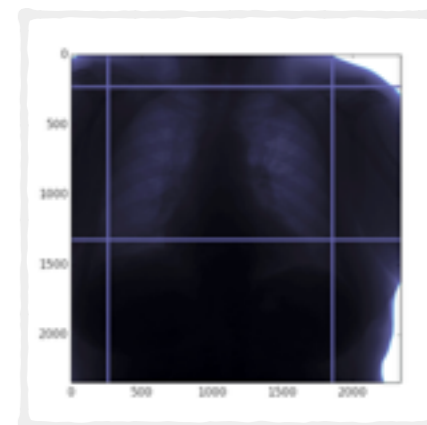
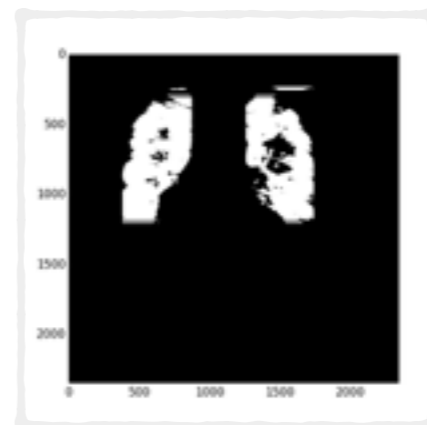
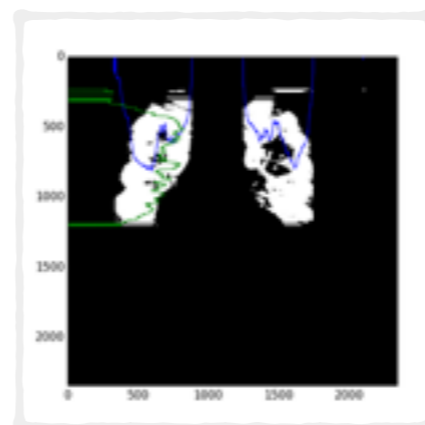
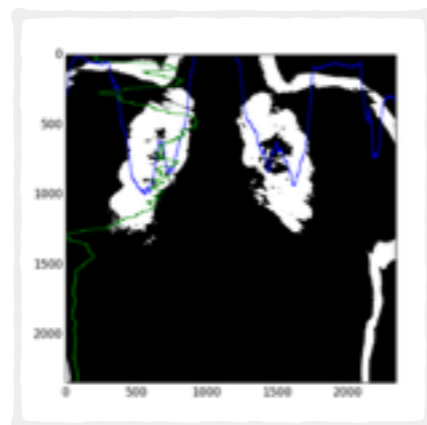
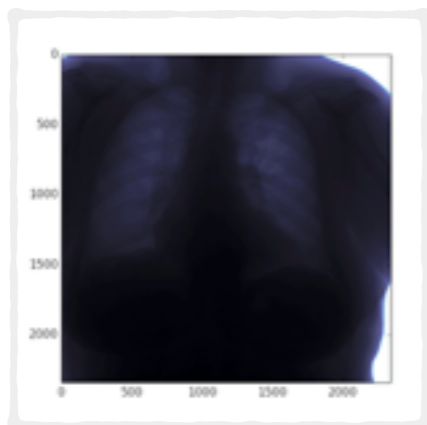
Мы не хотим вырезать легкие вручную

Первый подход
Простое обнаружение

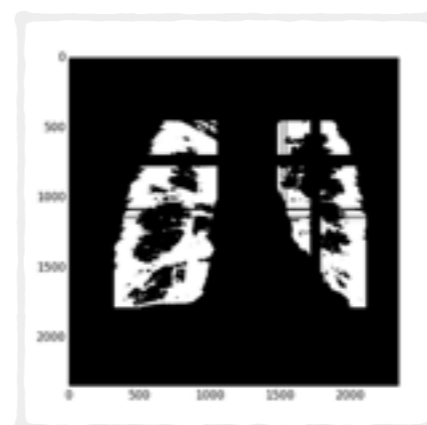
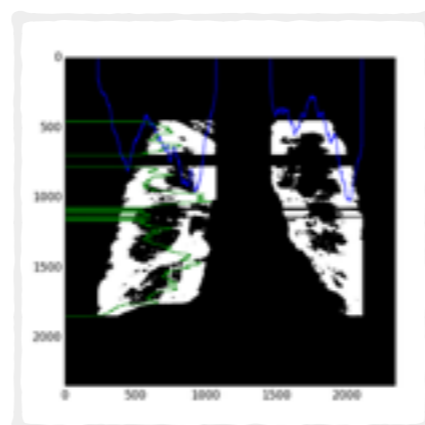
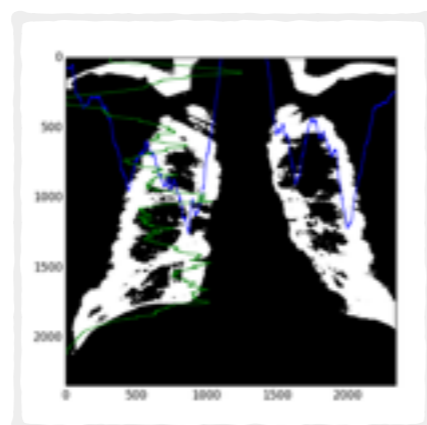
1.



2.



3.



а)

б)

в)

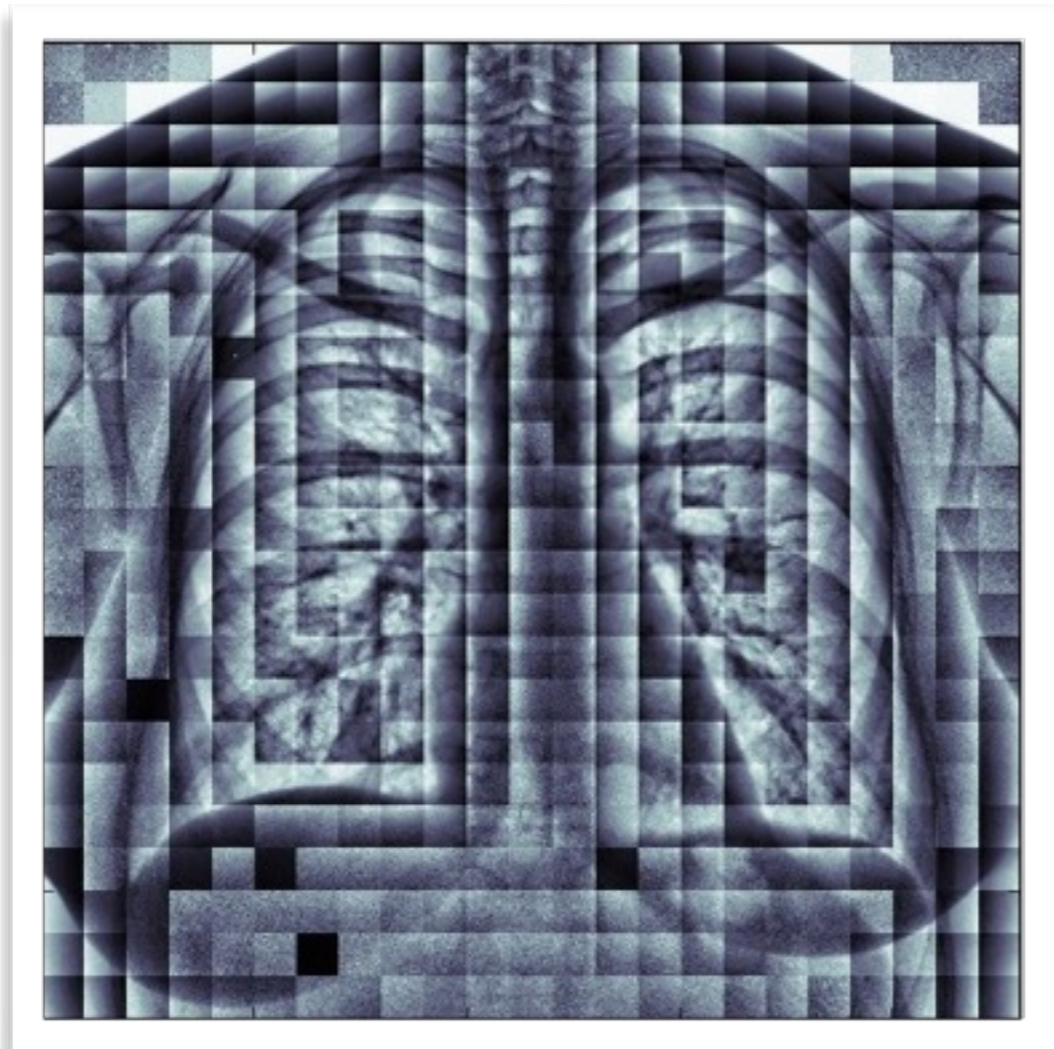
г)

д)

Второй подход

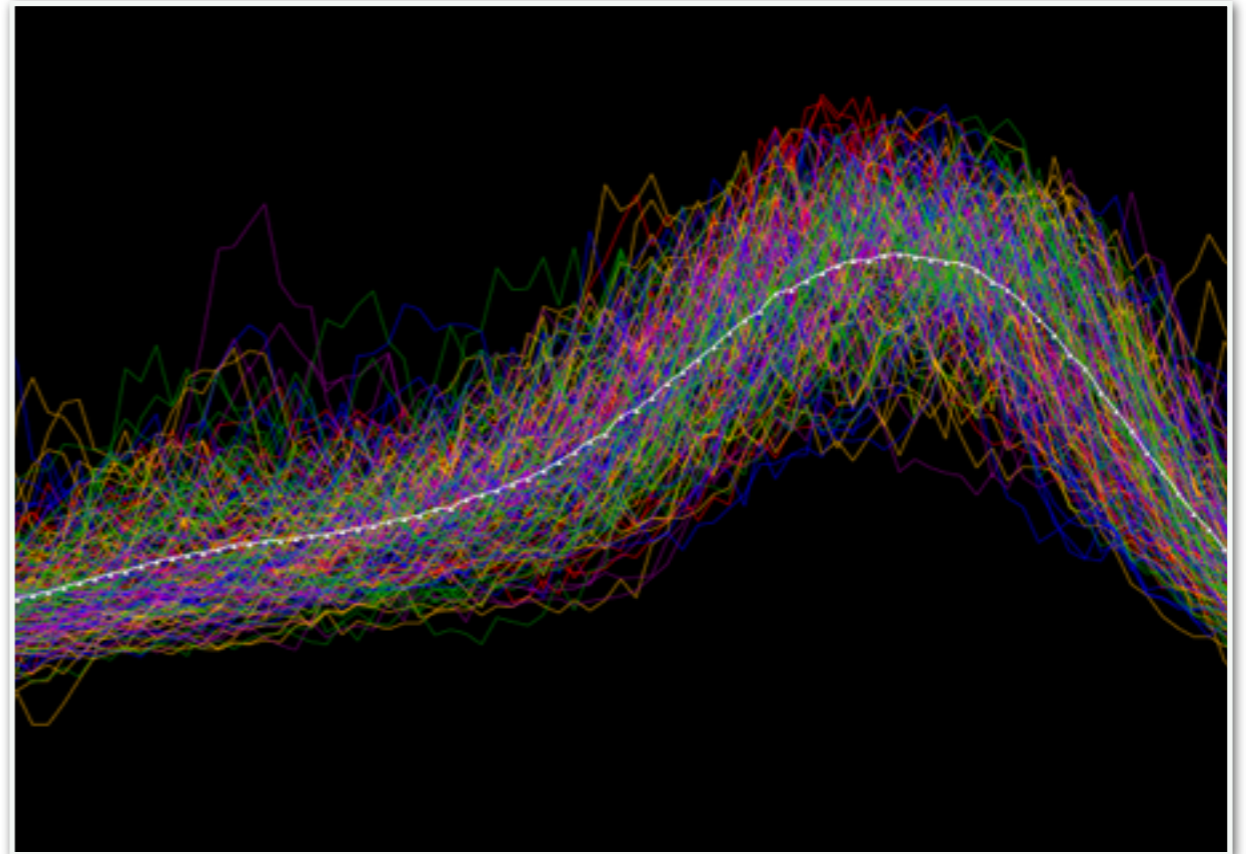
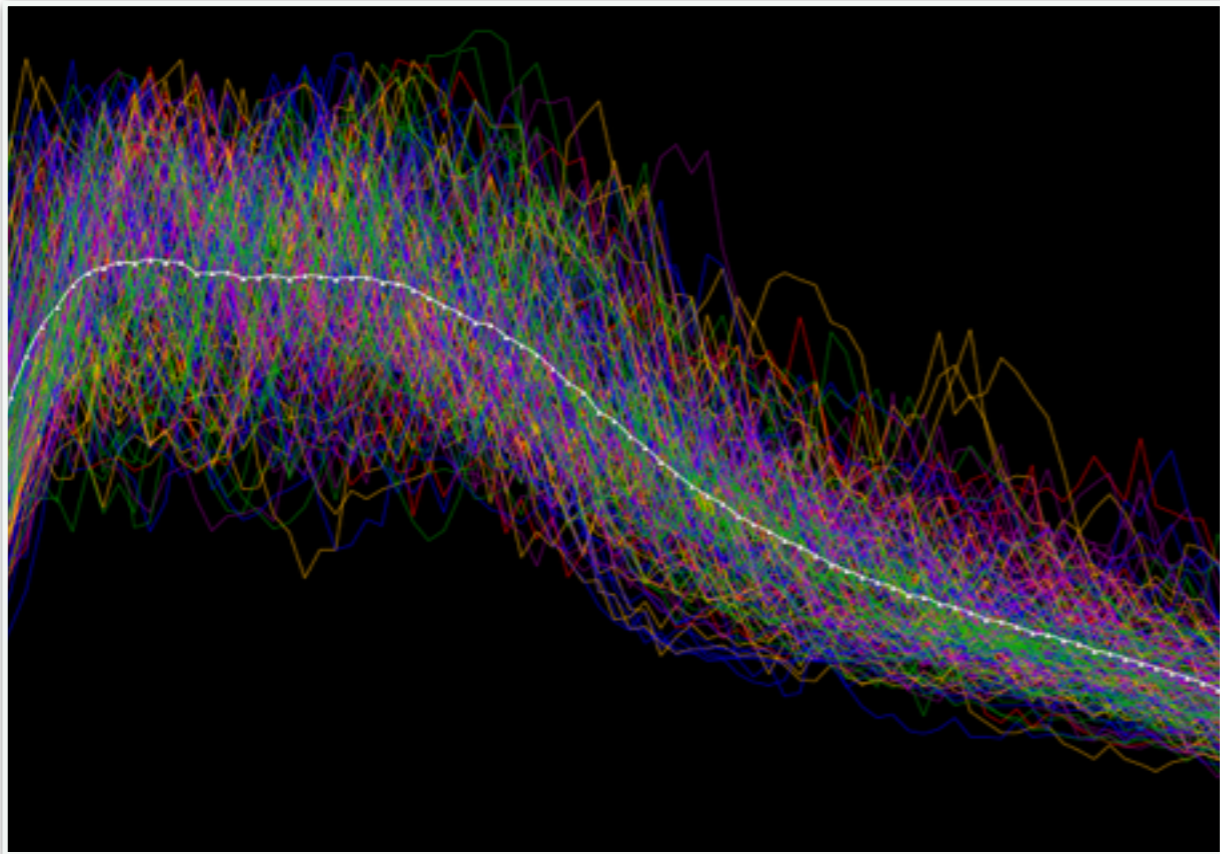
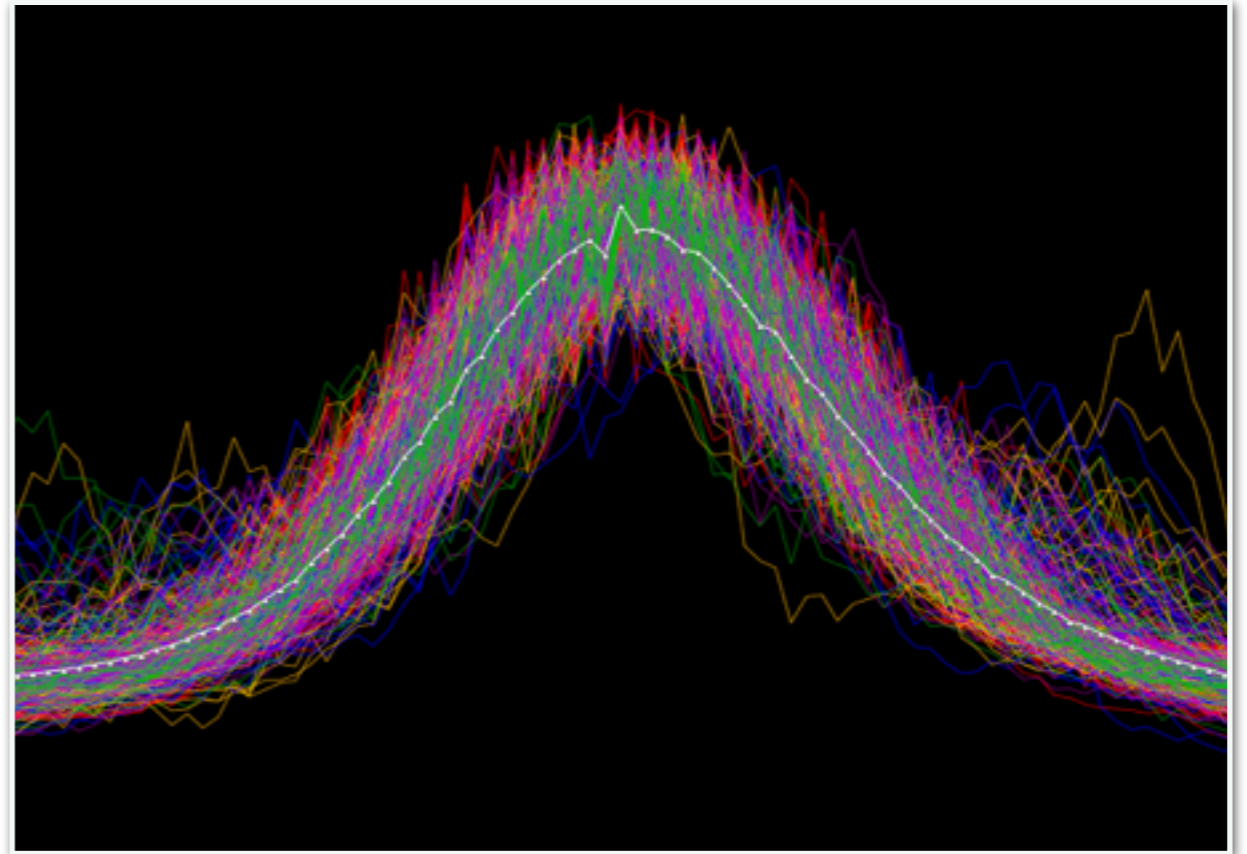
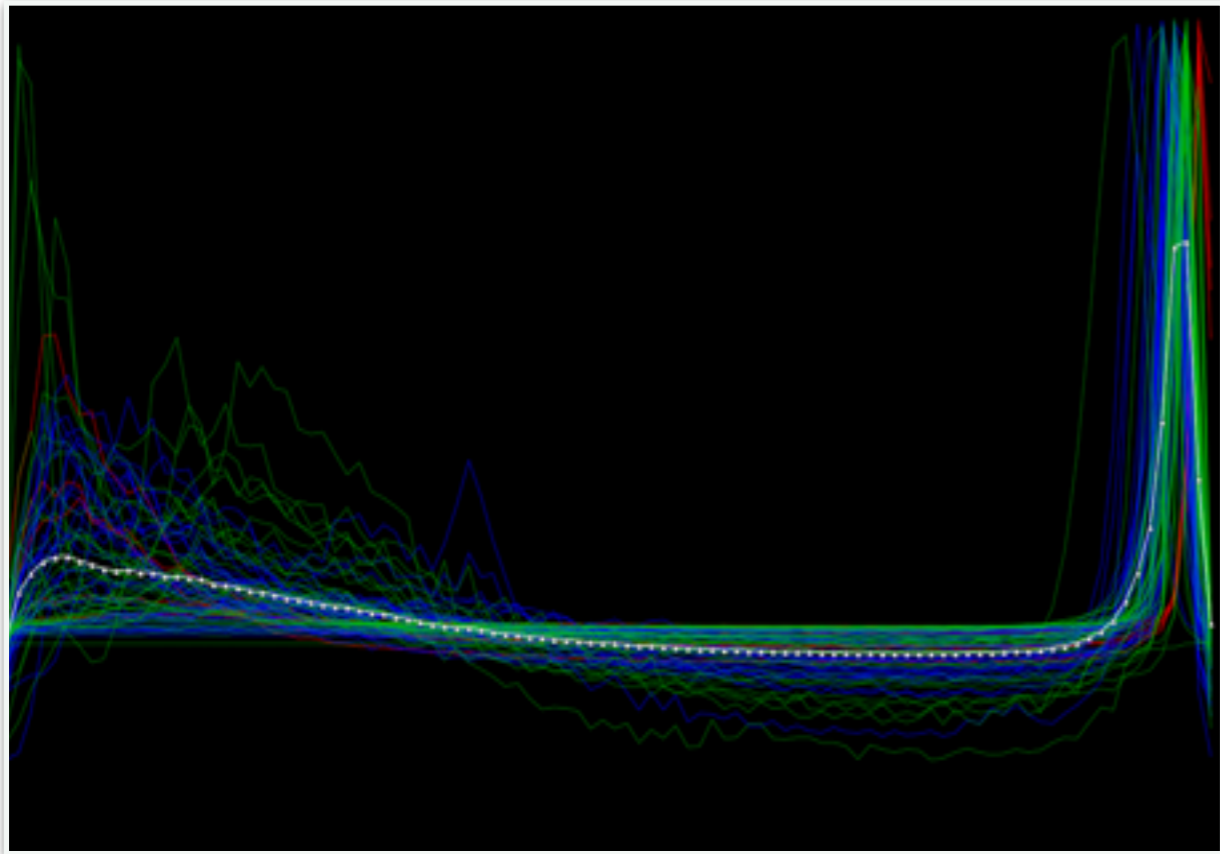
Машинное обучение: K-means + Random Forest

Рентгенокубизм

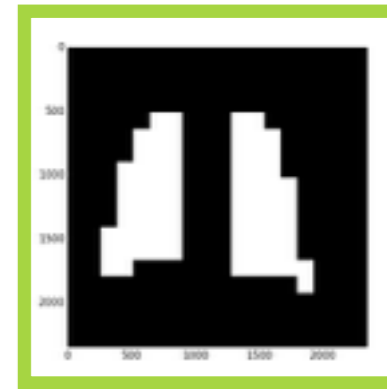
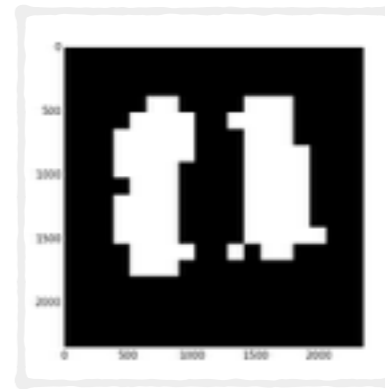
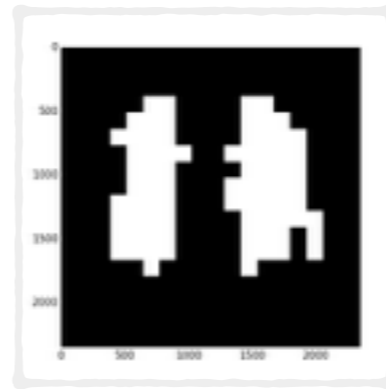
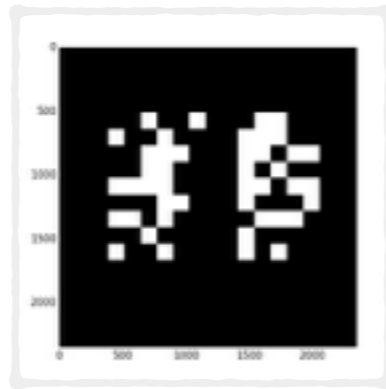
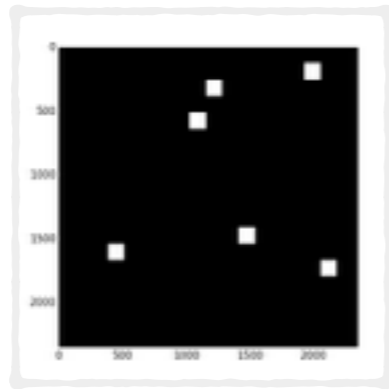
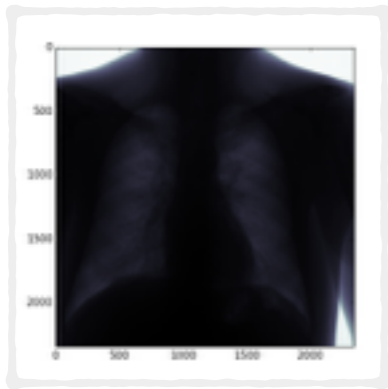


Битые пиксели

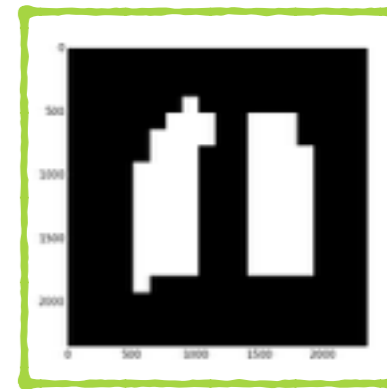
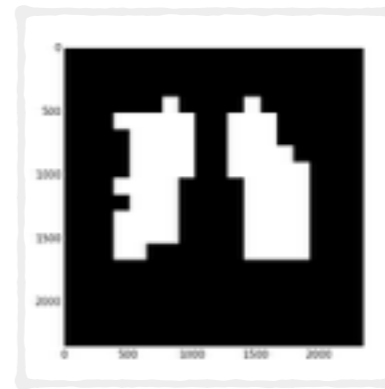
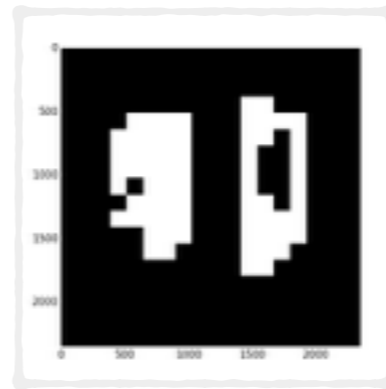
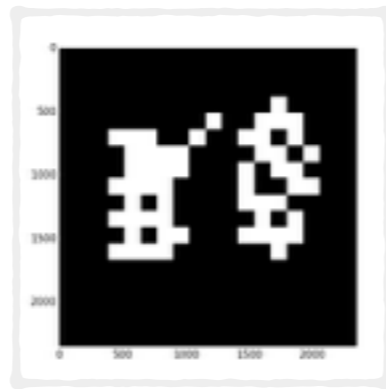
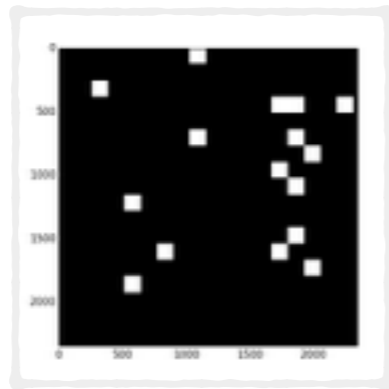
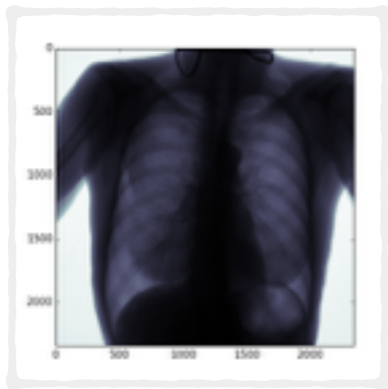
Гистограммы



1.



2.



а)

б)

в)

г)

д)

е)

Точность:

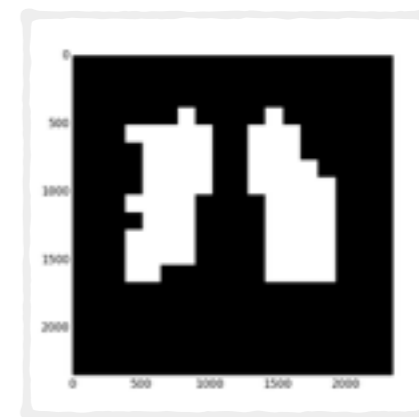
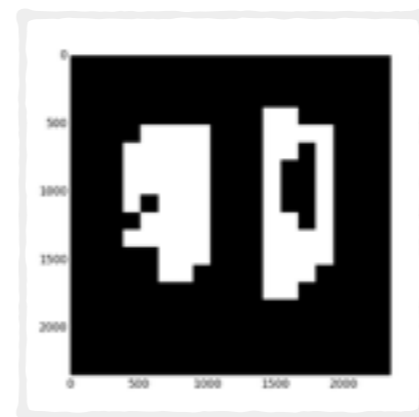
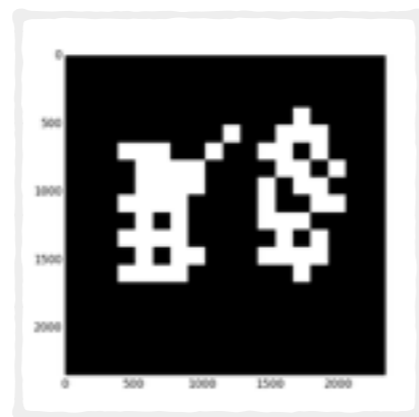
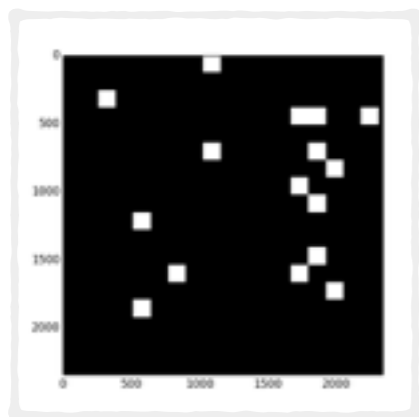
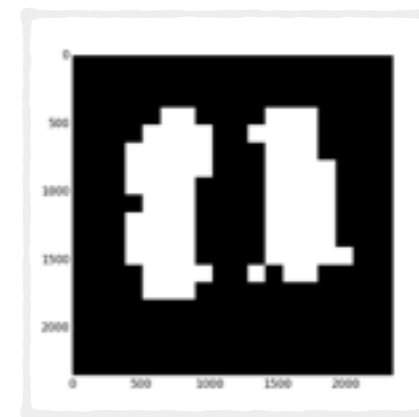
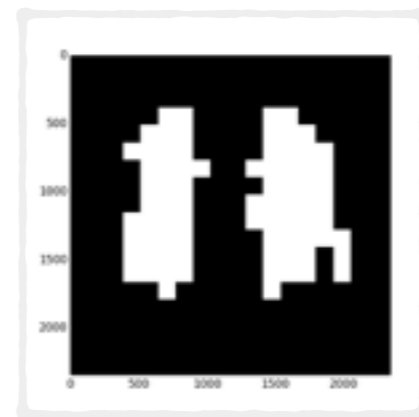
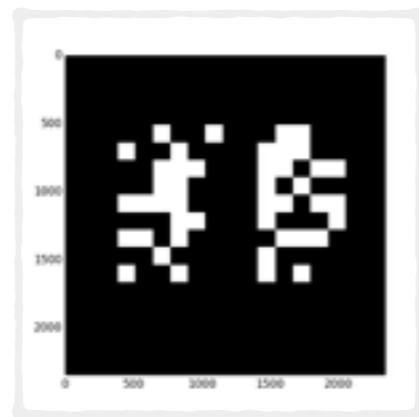
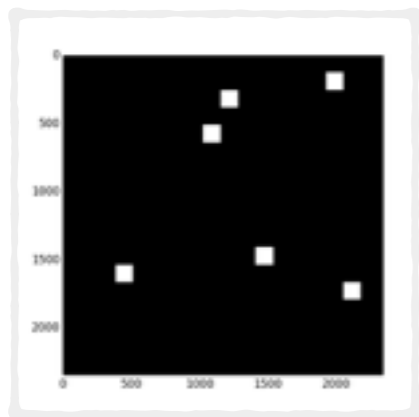
~65%

~75%

~80%

~87%

100%

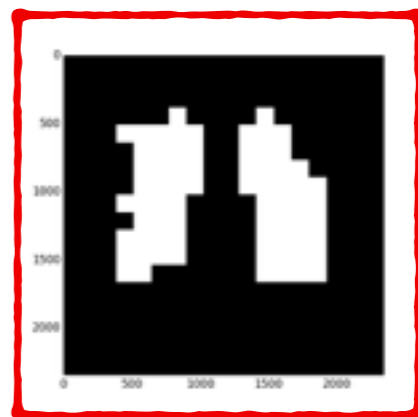
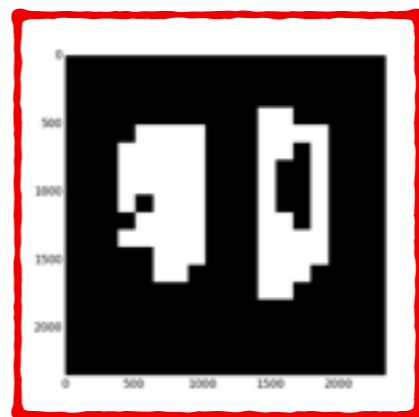
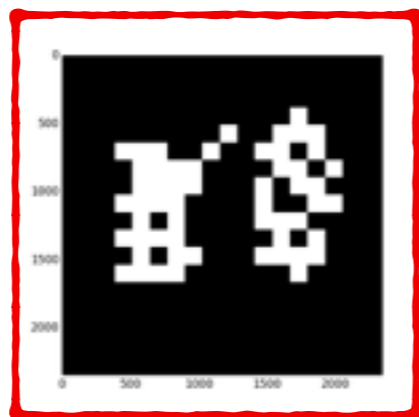
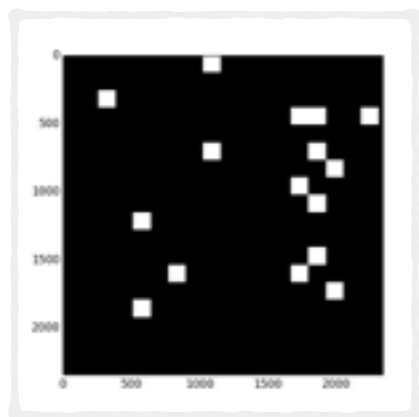
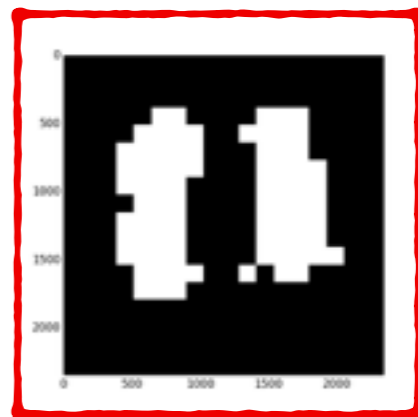
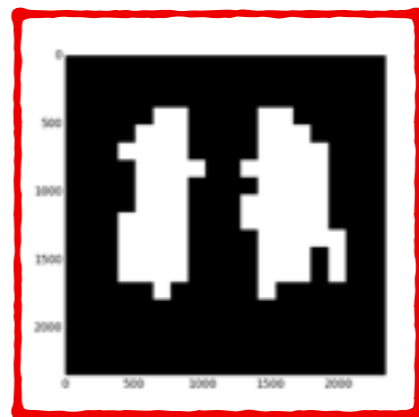
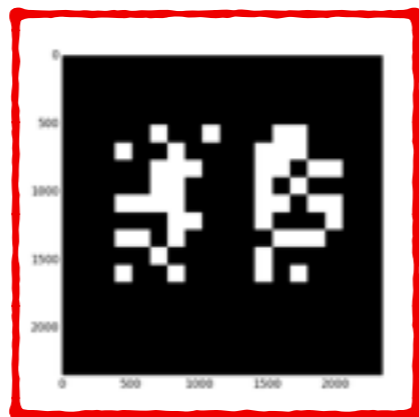
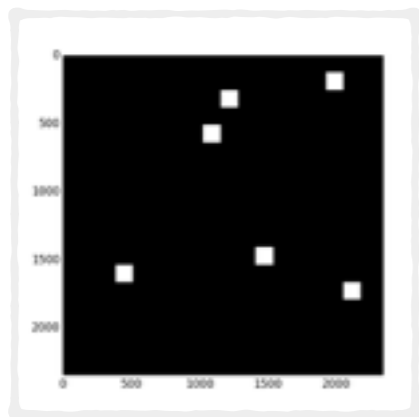


~65%

~75%

~80%

~87%



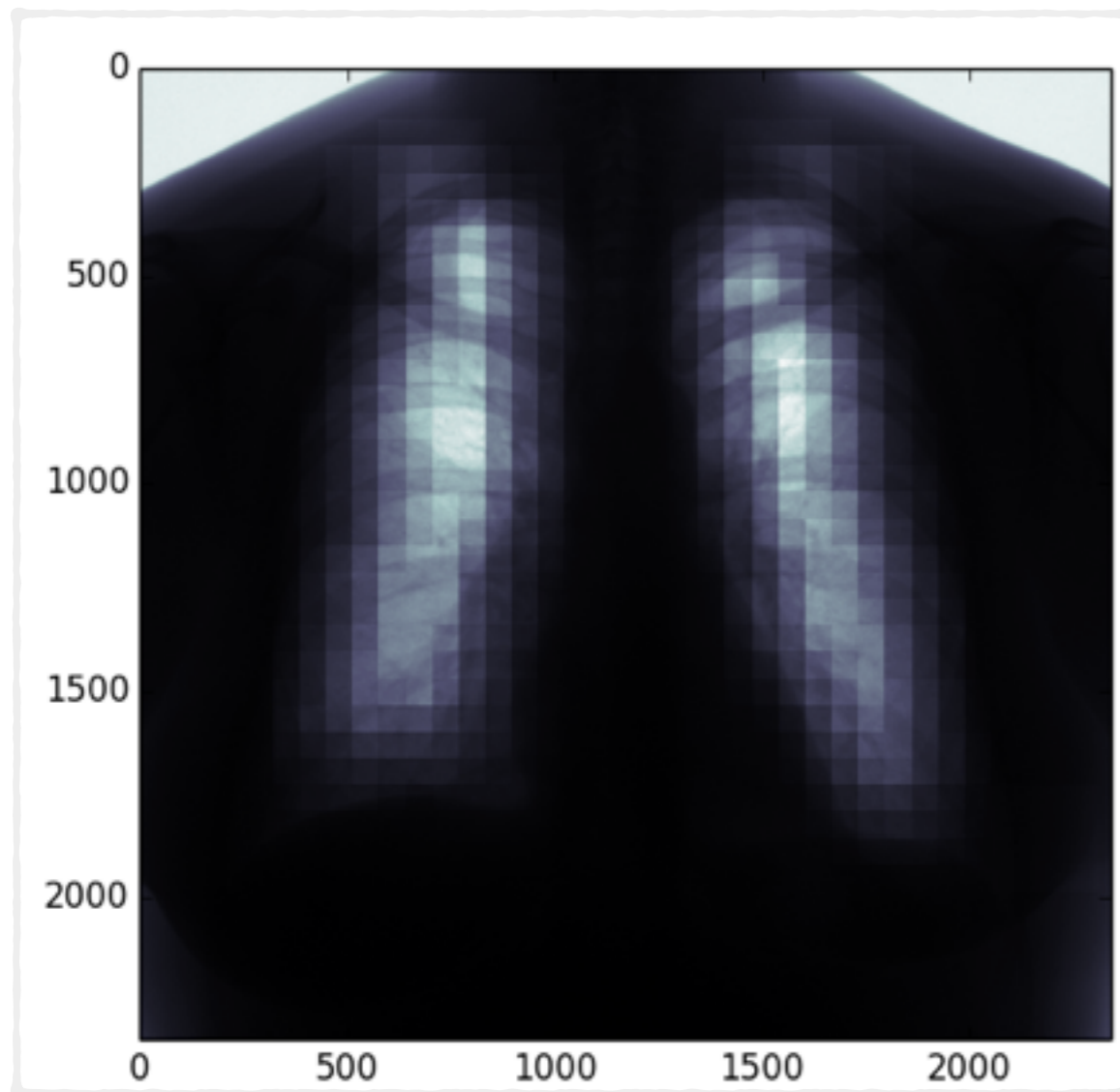
~65%

~75%

~80%

~87%

Без координат



Больше выборка

Больше кластеров

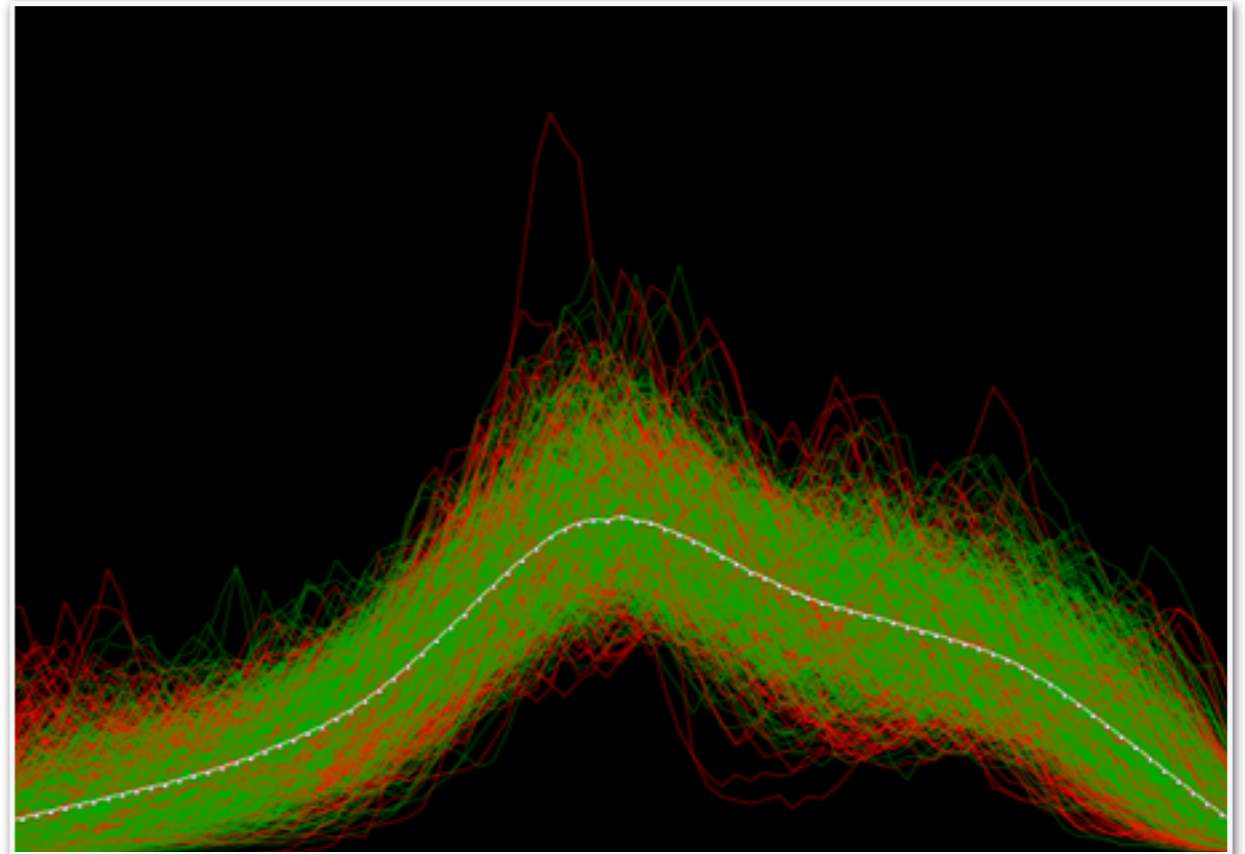
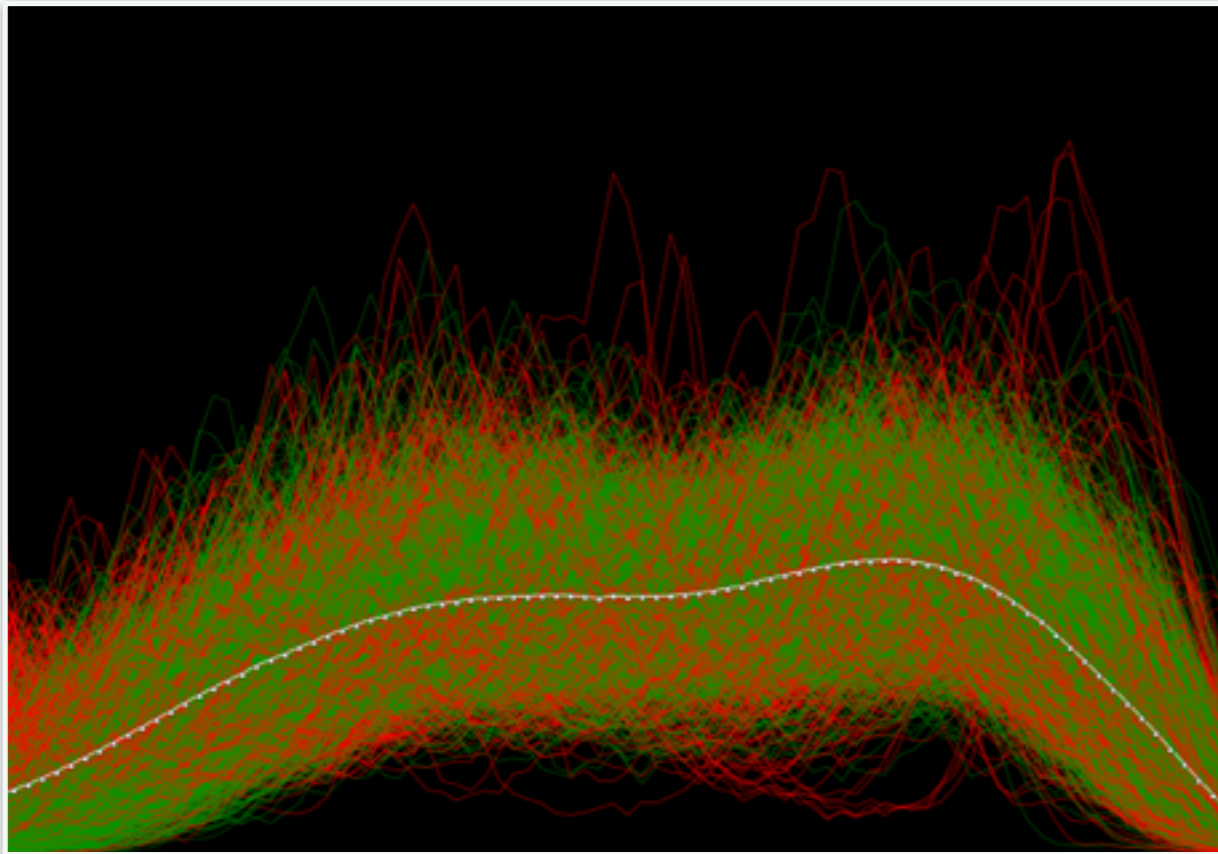
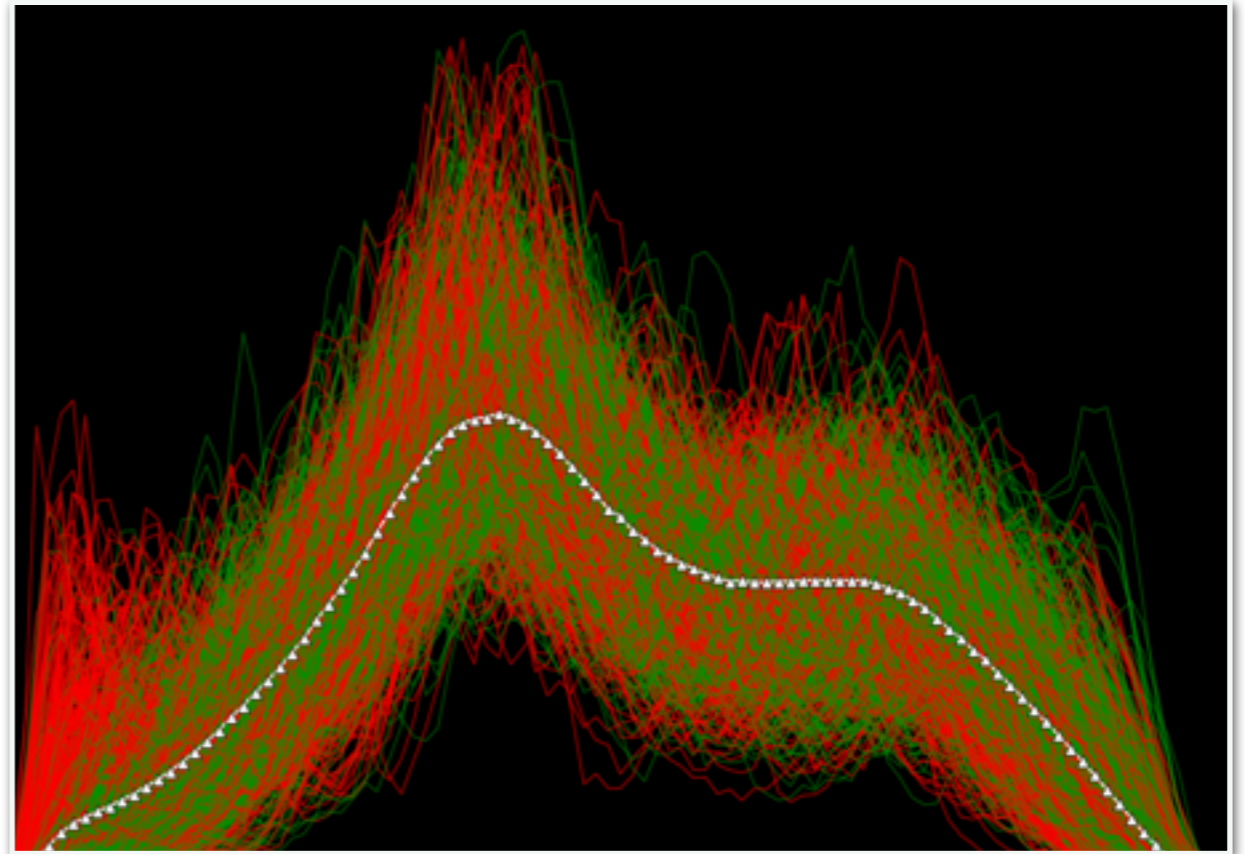
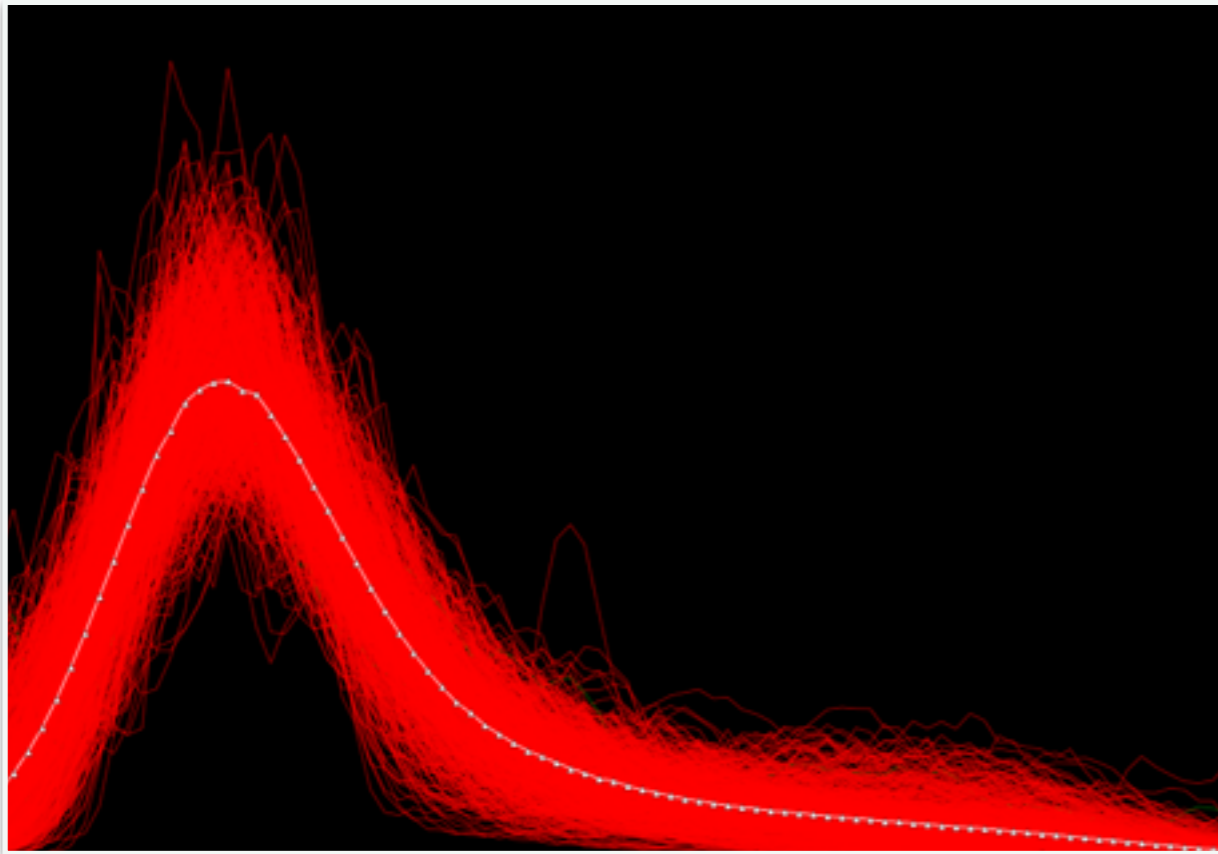
Увеличение размера модели

88%

Confusion Matrix

```
[[1676 107]
 [ 137 258]]
```

	precision	recall	f1-score	support
not lungs	0.92	0.94	0.93	1783
lungs	0.71	0.65	0.68	395
avg / total	0.88	0.89	0.89	2178

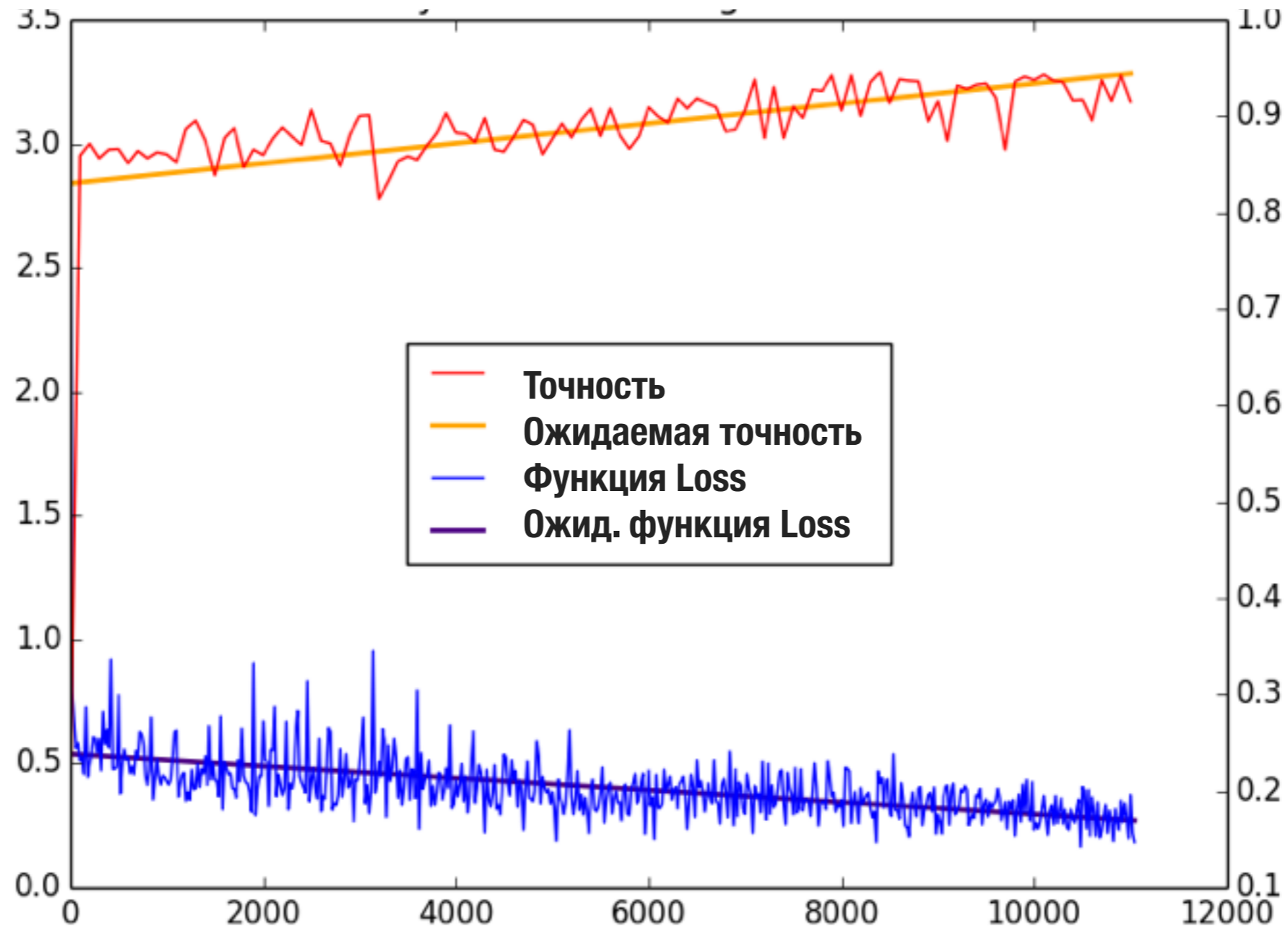


Третий подход

Машинное обучение: CNN

CNN

Зависимость точности от количества итераций



Сравнение методов

	Простой метод	K-means + RF	CNN
Выборка	Небольшая	30000 + 10000	30000 + 10000
Достигнутая точность	Низкая	88 %	95 %
Применимость	Низкая	Средняя	Высокая
Обучение	Отсутствует	Среднее	Долгое

Результаты

**Рассмотрены и применены методы машинного обучения
(K-means, Random Forest, CNN)**

Проведено сравнение точности алгоритмов

**Достигнута большая точность распознавания легких
(>95%)**

