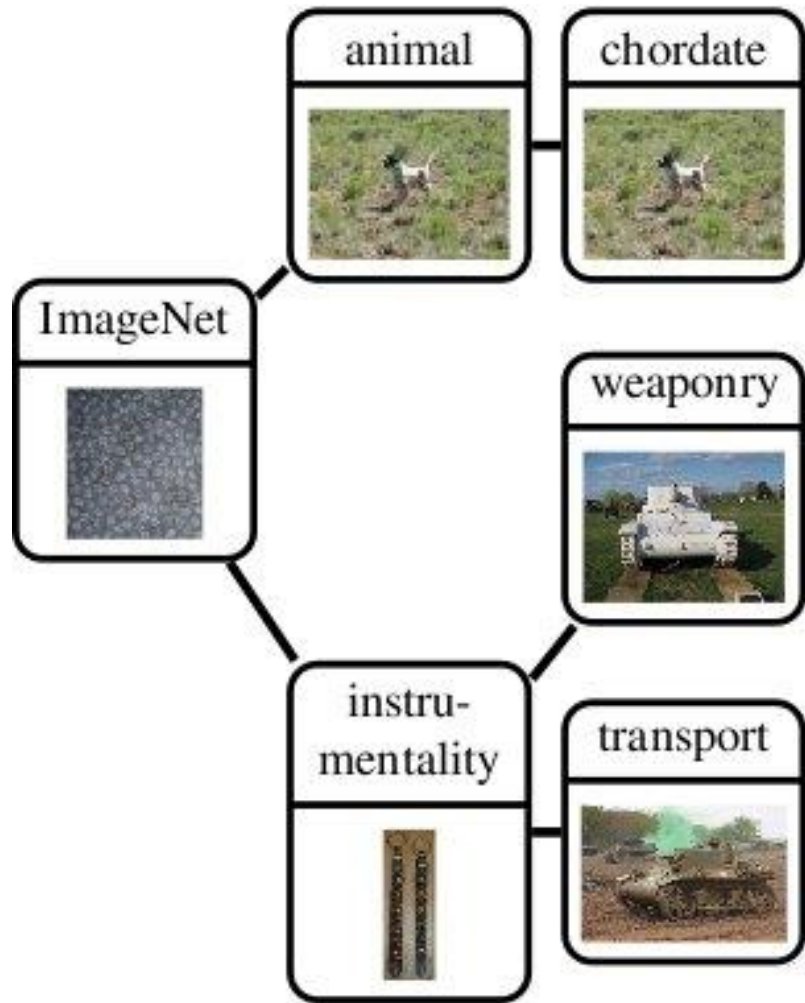


Лекция 3. Задача классификации/распознавания образов (на примере лицевой биометрии).

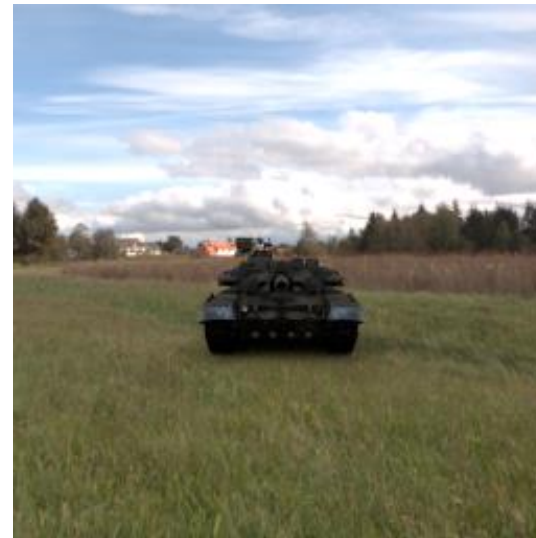
Задача классификации. Softmax with loss. Аугментация данных. Показатели качества. Классические БД. Обучение на классификацию. Задача распознавания лиц - верификация, идентификация. Показатели качества (FAR/FRR, FNIR, FPIR, Top-N). Методики подготовки данных для обучения и тестирования. DeepID(Deep features) - функции потерь и архитектуры. Триплеты/сиамские сети. Публичные БД лиц.

Горбацевич В.С

Классификация



Вход:



Выход:

Что это?

Классификация

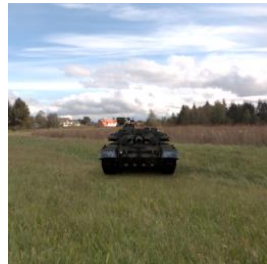
Постановка: Есть n -классов изображений необходимо каждому изображению поставить в соответствие метку класса.

Примеры:

Определение марки машины

Определение пола по фото

Определение цифры/буквы



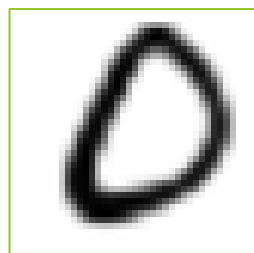
?

Классификация

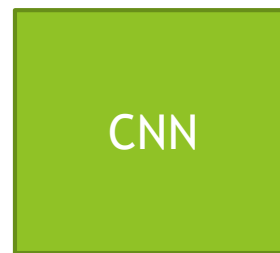
Постановка: Есть черно-белое изображение цифры от 0 до 9. Необходимо обучить ГКНС для распознавания цифры.



БД MNIST



1x18x18



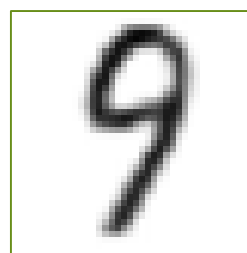
0

Классификация

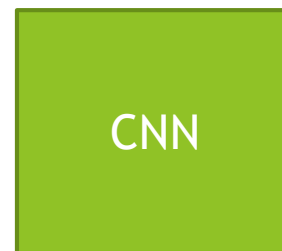
Постановка: Есть черно-белое изображение цифры от 0 до 9. Необходимо обучить ГКНС для распознавания цифры.



БД MNIST



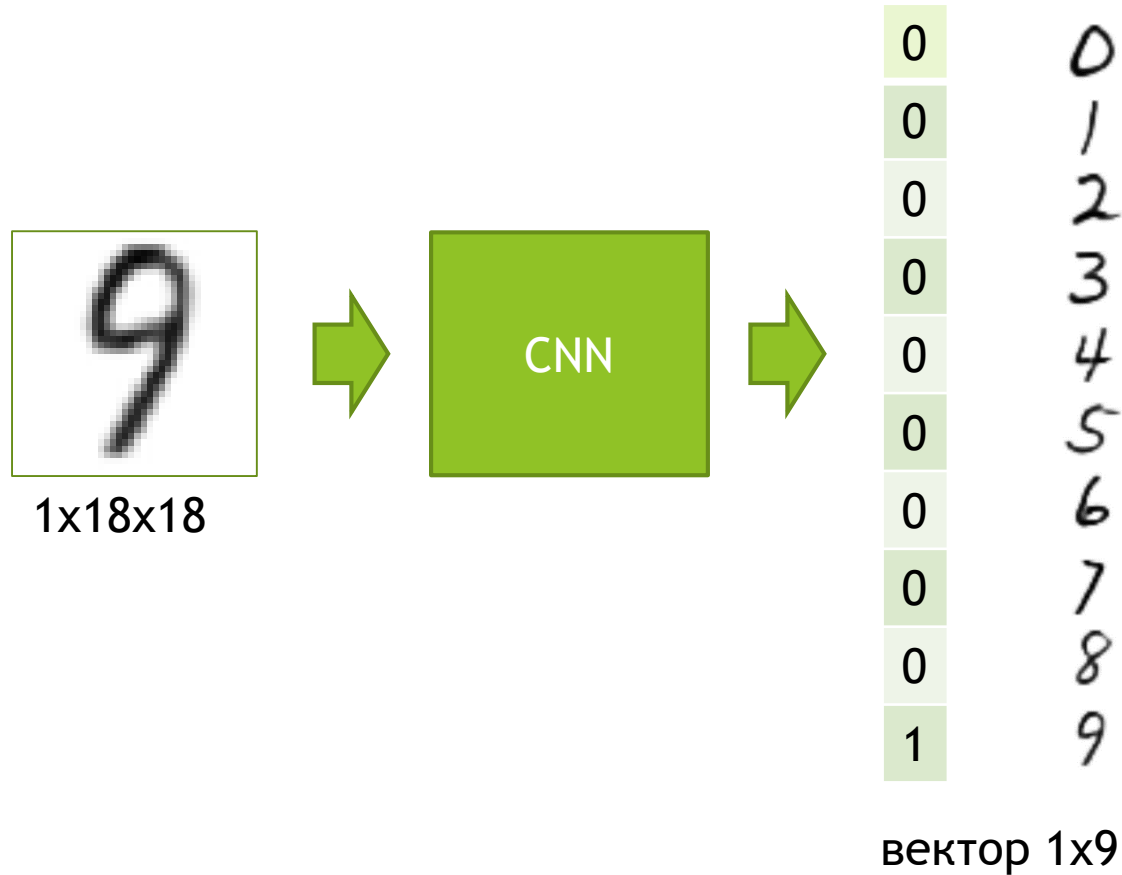
1x18x18



9

Классификация

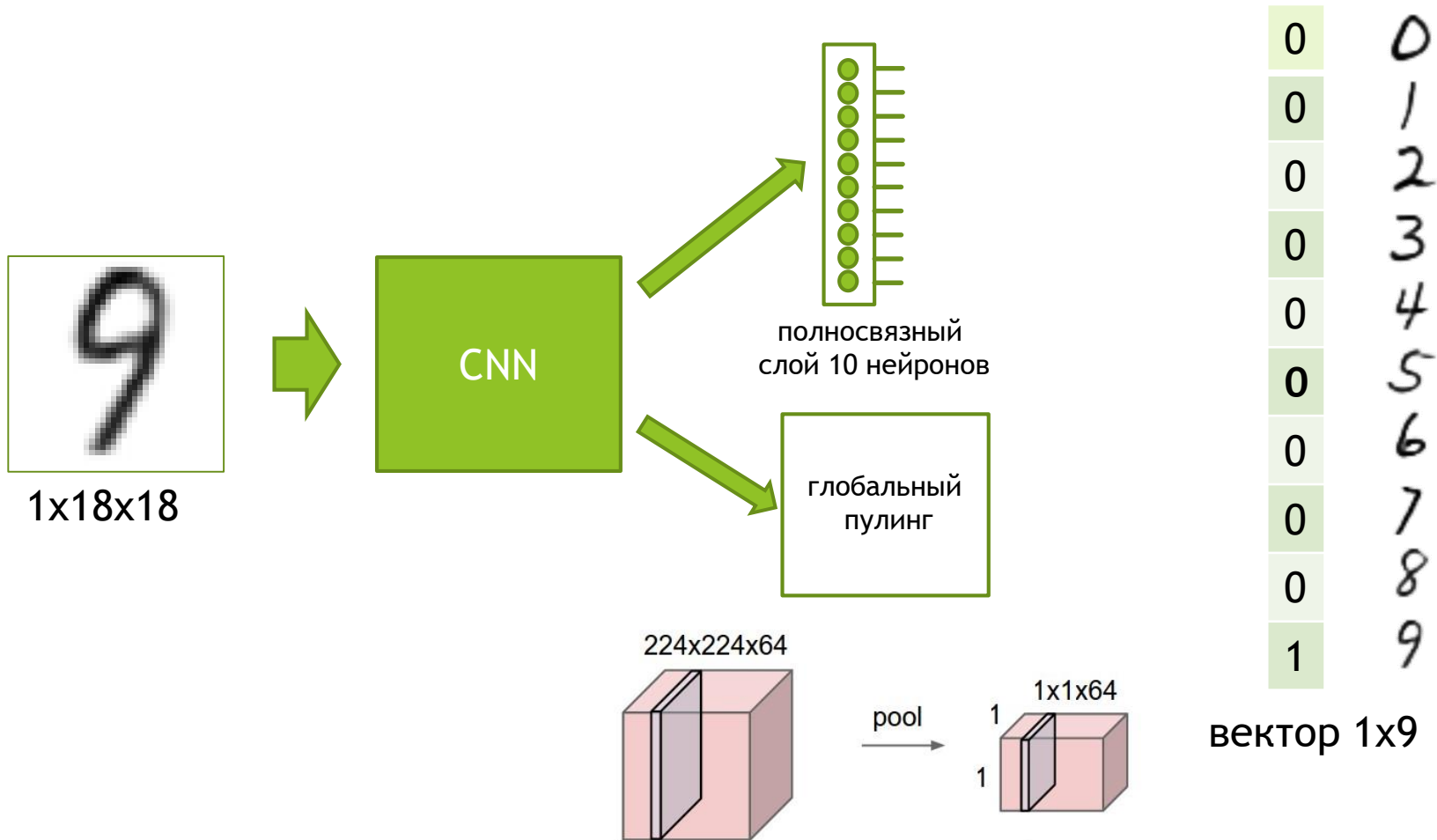
Постановка: Есть черно-белое изображение цифры от 0 до 9. Необходимо обучить ГКНС для распознавания цифры.



При решении задачи классификации в большинстве случаев выход сети - это вектор 1 x количество классов

Классификация

Постановка: Есть черно-белое изображение цифры от 0 до 9. Необходимо обучить ГКНС для распознавания цифры.




При решении задачи классификации в большинстве случаев выход сети - это вектор 1 x количество классов

Классификация

Постановка: Есть черно-белое изображение цифры от 0 до 9. Необходимо обучить ГНС для распознавания цифры.


10	0,04
-20	0,01
15	0,02
7	0,01
18	0.016
33	0.023
16	0.015
0	0.01
12	0.017
94	0.25



Sigmoid Function

$$F(X_i) = \frac{1}{1 + \text{Exp}(-X_i)}$$

dataaspirant.com




Softmax Function

$$F(X_i) = \frac{\text{Exp}(X_i)}{\sum_{j=0}^k \text{Exp}(X_j)} \quad i = 0, 1, 2, \dots, k$$

dataaspirant.com

Для нормировки результатов перед последним слоем ставят нелинейность

Классификация



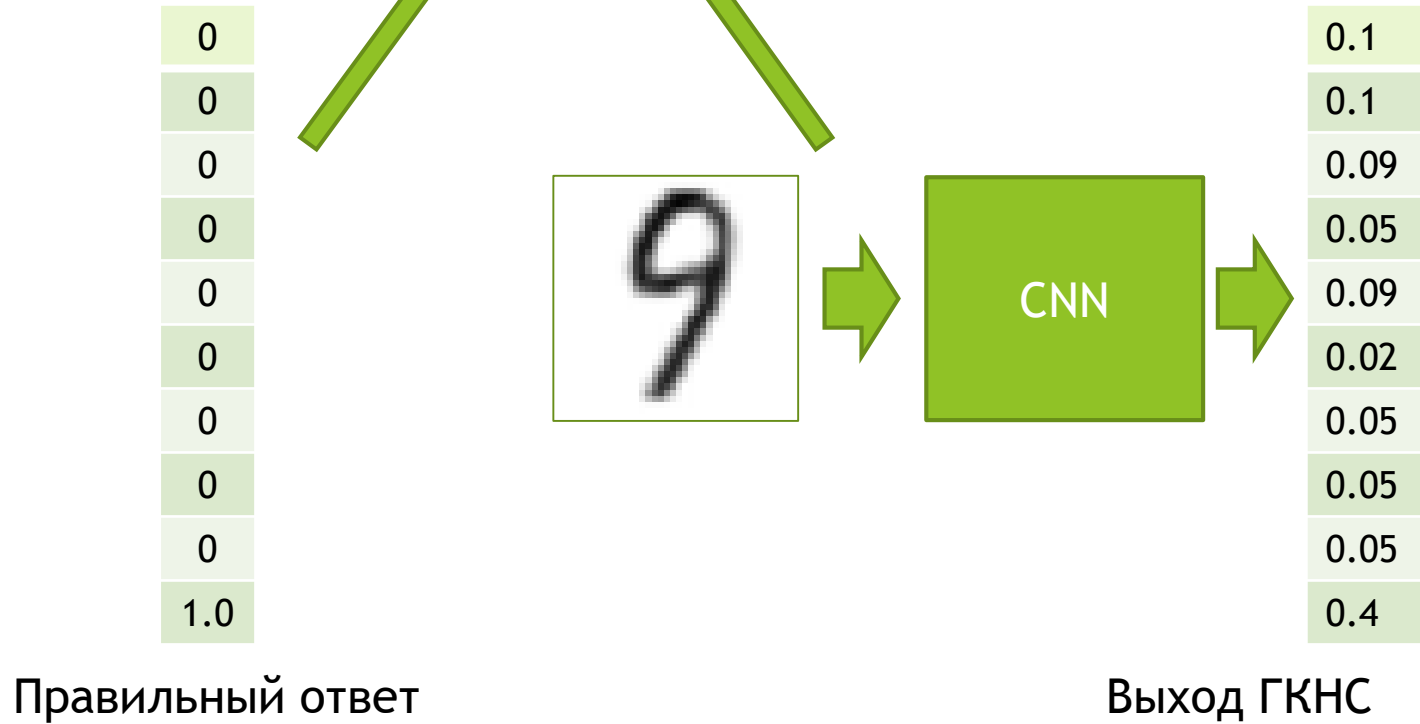
10	0,04
-20	0,01
15	0,02
7	0,01
18	0.016
33	0.023
16	0.015
0	0.01
12	0.017
94	0.25

Softmax Function	Sigmoid Function
Used for multi-classification in logistic regression model.	Used for binary classification in logistic regression model.
The probabilities sum will be 1	The probabilities sum need not be 1.
Used in the different layers of neural networks.	Used as activation function while building neural networks.
The high value will have the higher probability than other values.	The high value will have the high probability but not the higher probability.

Для нормировки результатов перед последним слоем ставят нелинейность
В 99% случаев используется SoftMax

Классификация

$$H(p, q) = - \sum_i p_i \log(q_i)$$

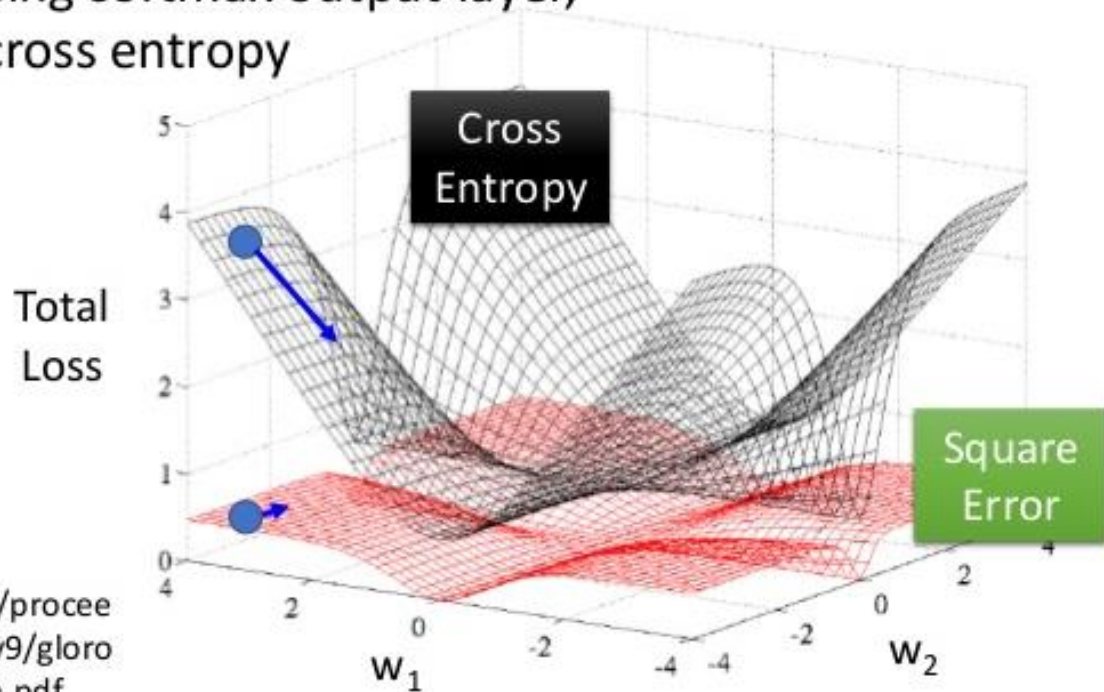


Классификация

Choosing Proper Loss

$$H(p, q) = - \sum_i p_i \log(q_i)$$

When using softmax output layer,
choose cross entropy

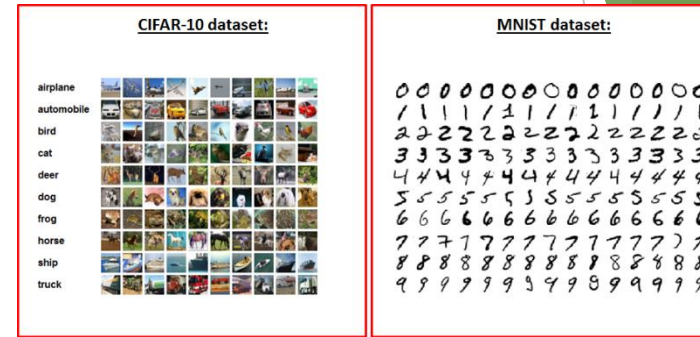


<http://jmlr.org/proceedings/papers/v9/glorot10a/glorot10a.pdf>

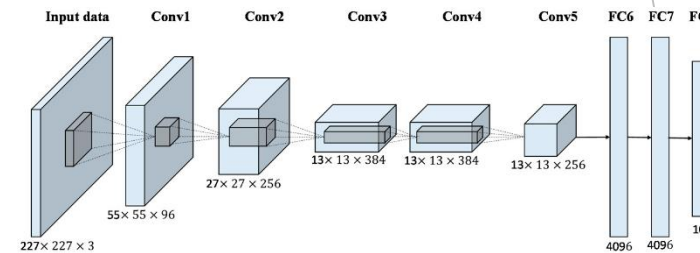
Для задачи классификации используется кросс-энтропийная функция потерь

Обучение ГКНС

1. Выбрать БД



2. Выбрать архитектуру ГКНС



3. Провести обучение ГКНС

4. Провести тестирование ГКНС

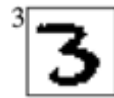


Обучение ГКНС

3 индекс класса



изображение



⁵ 5	⁰ 0	⁴ 4	¹ 1	⁹ 9	² 2	¹ 1	³ 3	¹ 1	⁴ 4
³ 3	⁵ 5	³ 3	⁶ 6	¹ 1	⁷ 7	² 2	⁸ 8	⁶ 6	⁹ 9
⁴ 4	⁰ 0	⁹ 9	¹ 1	¹ 1	² 2	⁴ 4	³ 3	² 2	⁷ 7
³ 3	⁸ 8	⁶ 6	⁹ 9	⁰ 0	⁵ 5	⁶ 6	⁰ 0	⁷ 7	⁶ 6
¹ 1	⁸ 8	⁷ 7	⁹ 9	³ 3	⁹ 9	⁸ 8	⁵ 5	⁹ 9	³ 3
³ 3	⁰ 0	⁷ 7	⁴ 4	⁹ 9	⁸ 8	⁰ 0	⁹ 9	⁴ 4	¹ 1
⁴ 4	⁴ 4	⁶ 6	⁰ 0	⁴ 4	⁵ 5	⁶ 6	¹ 1	⁰ 0	⁰ 0
¹ 1	⁷ 7	¹ 1	⁶ 6	³ 3	⁰ 0	² 2	¹ 1	¹ 1	⁷ 7
⁹ 9	⁰ 0	² 2	⁶ 6	⁷ 7	⁸ 8	³ 3	⁹ 9	⁰ 0	⁴ 4
⁶ 6	⁷ 7	⁴ 4	⁶ 6	⁸ 8	⁰ 0	⁷ 7	⁸ 8	³ 3	¹ 1

БД изображений и ответов

Обучение ГКНС

1

индекс класса

ship



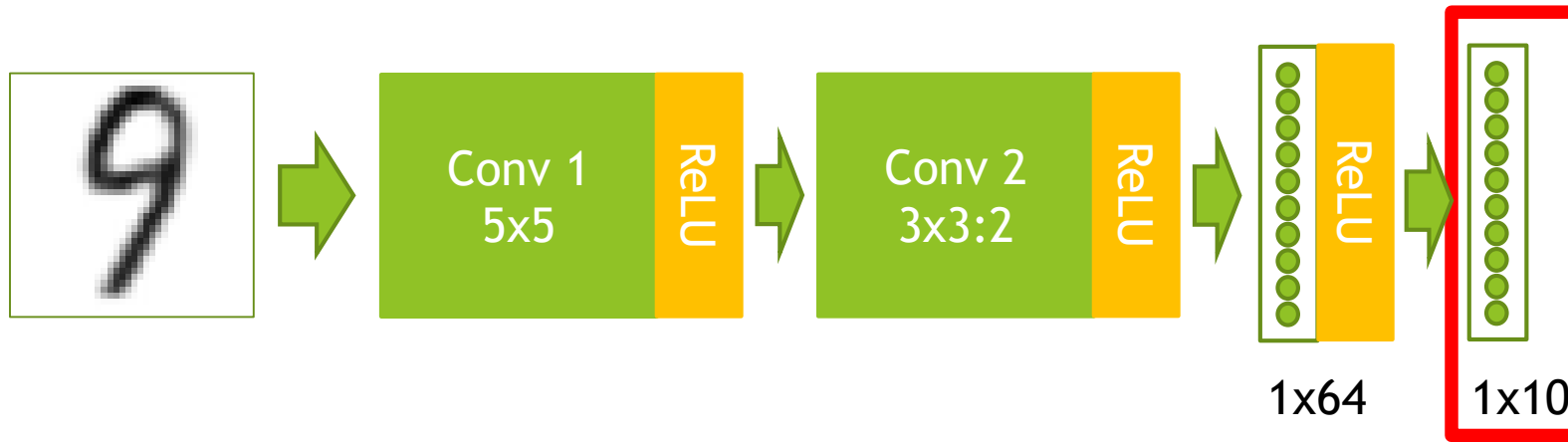
изображение



ship	dog	deer	bird	ship	cat	dog	dog
horse	horse	ship	frog	bird	ship	bird	cat
automobile	ship	deer	truck	dog	deer	automobile	horse

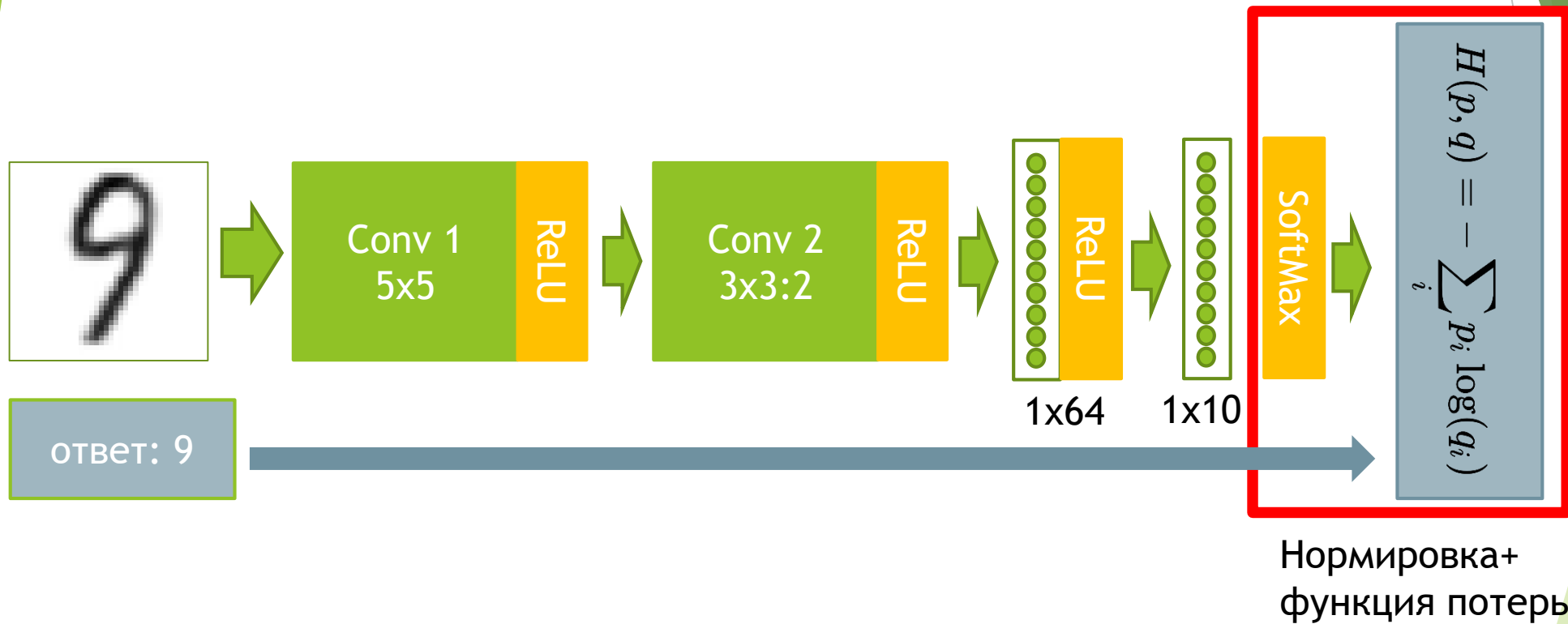
БД изображений и ответов

Обучение ГКНС



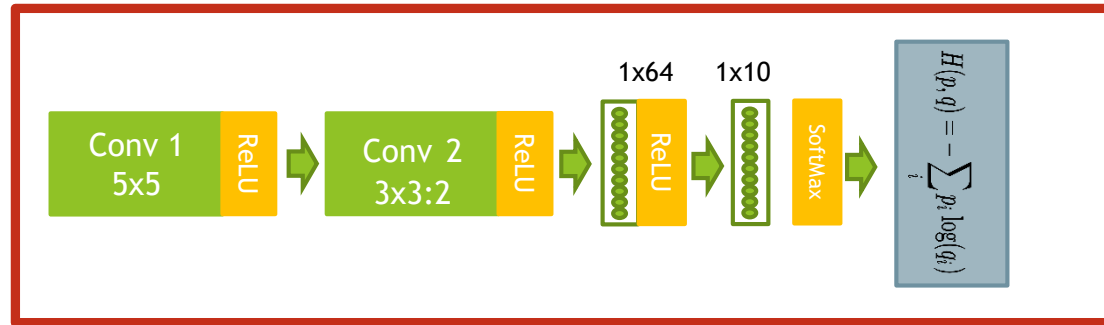
Архитектура может быть любой но она должна заканчиваться слоем, выходом которого является вектор **1 x число классов**

Обучение ГКНС



Для обучения добавляются слои нормировки(softmax) и добавляется функция потерь. Слой нормировки может быть частью ГКНС для удобства использования.

Обучение ГКНС



ГКНС

Батч = 5

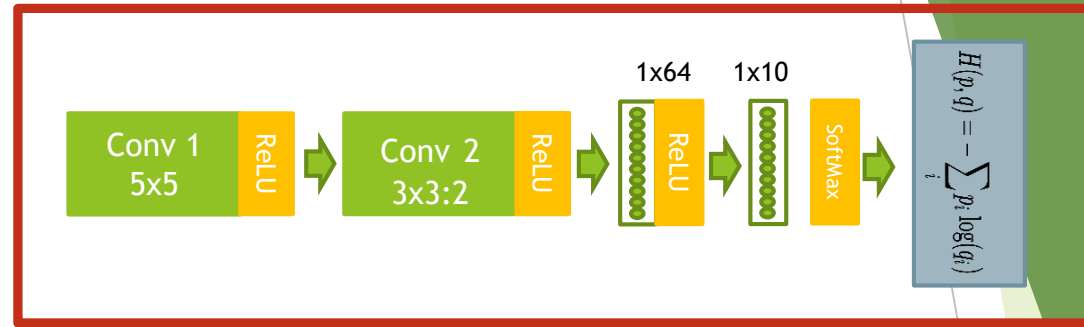


Этап 1: загрузка данных из БД: загружается батч -набор изображений и ответов



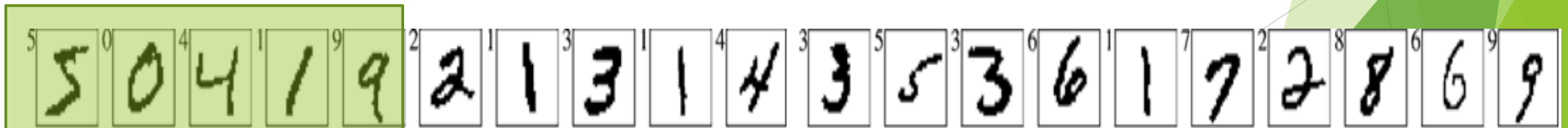
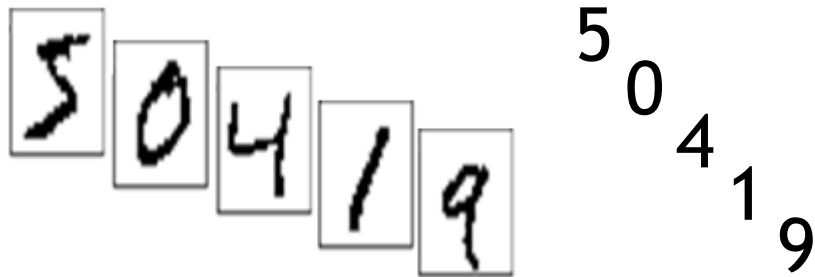
БД изображений и ответов - 20 примеров

Обучение ГКНС



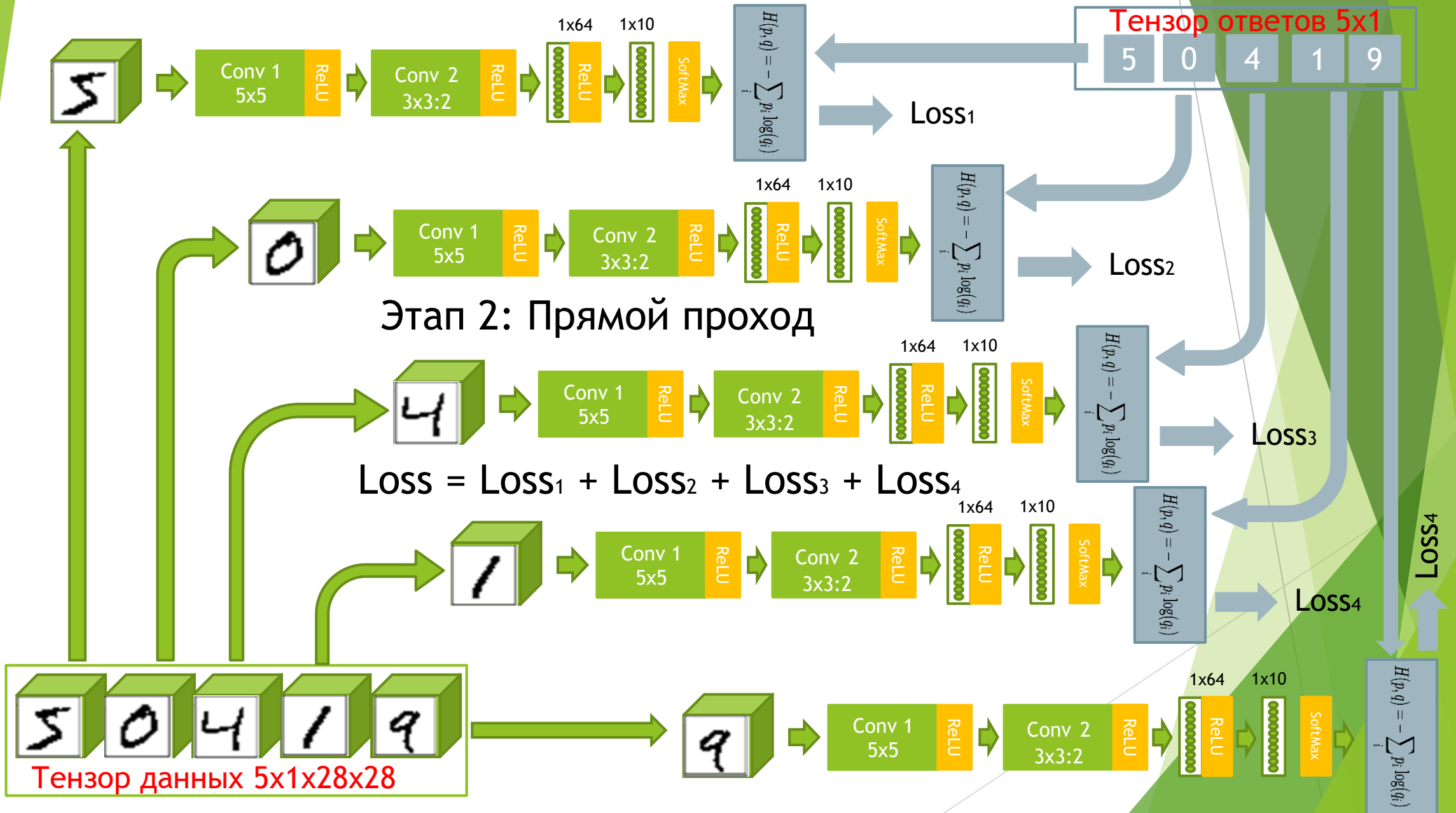
ГКНС

Этап 2: формирование батча из БД: предобработка+ трансформация в тензор /блор

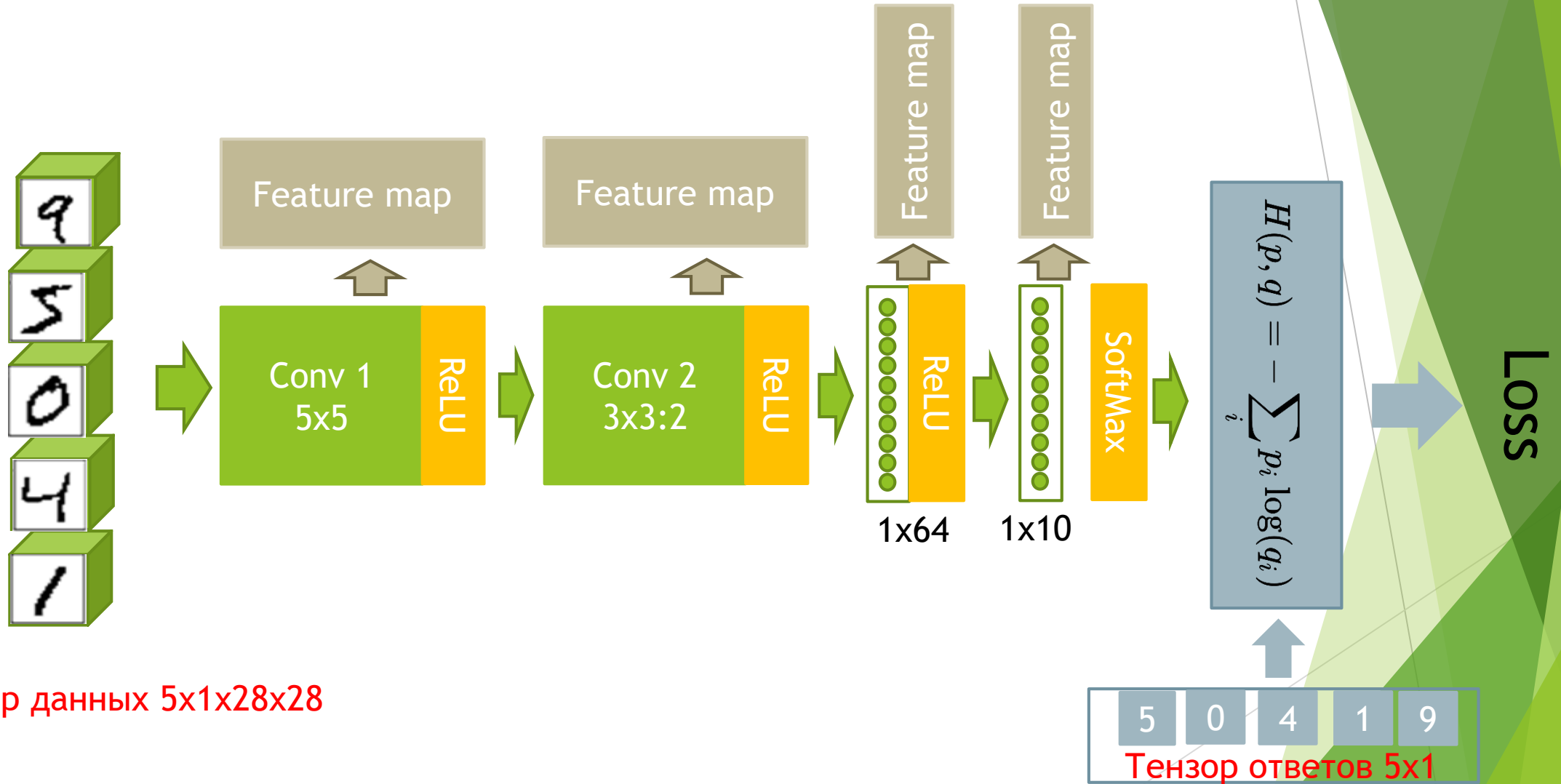


БД изображений и ответов - 20 примеров

Обучение ГКНС

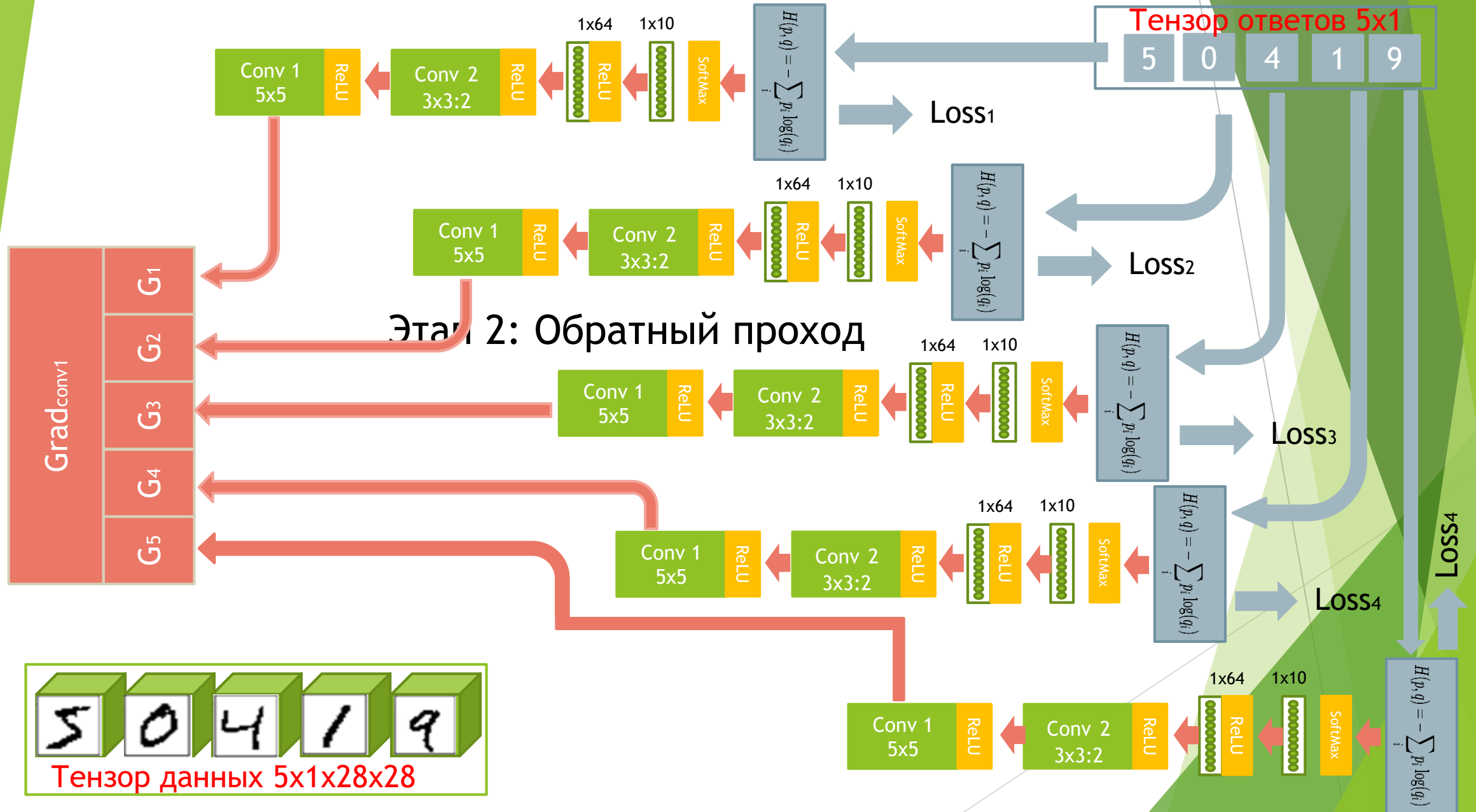


Обучение ГКНС

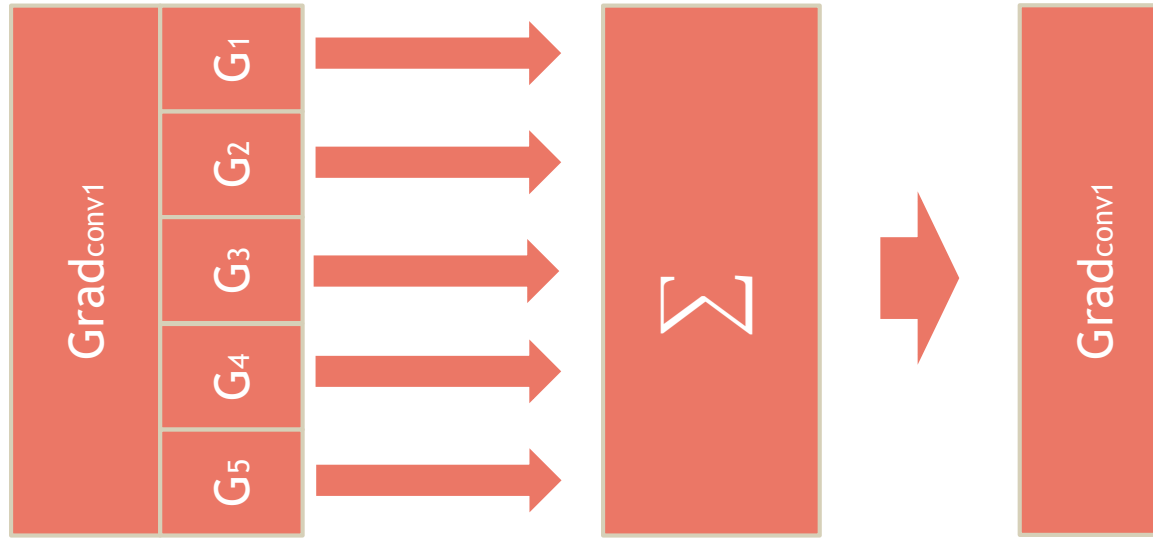


Этап 2: Прямой проход. При прямом проходе сохраняются все карты признаков

Обучение ГКНС

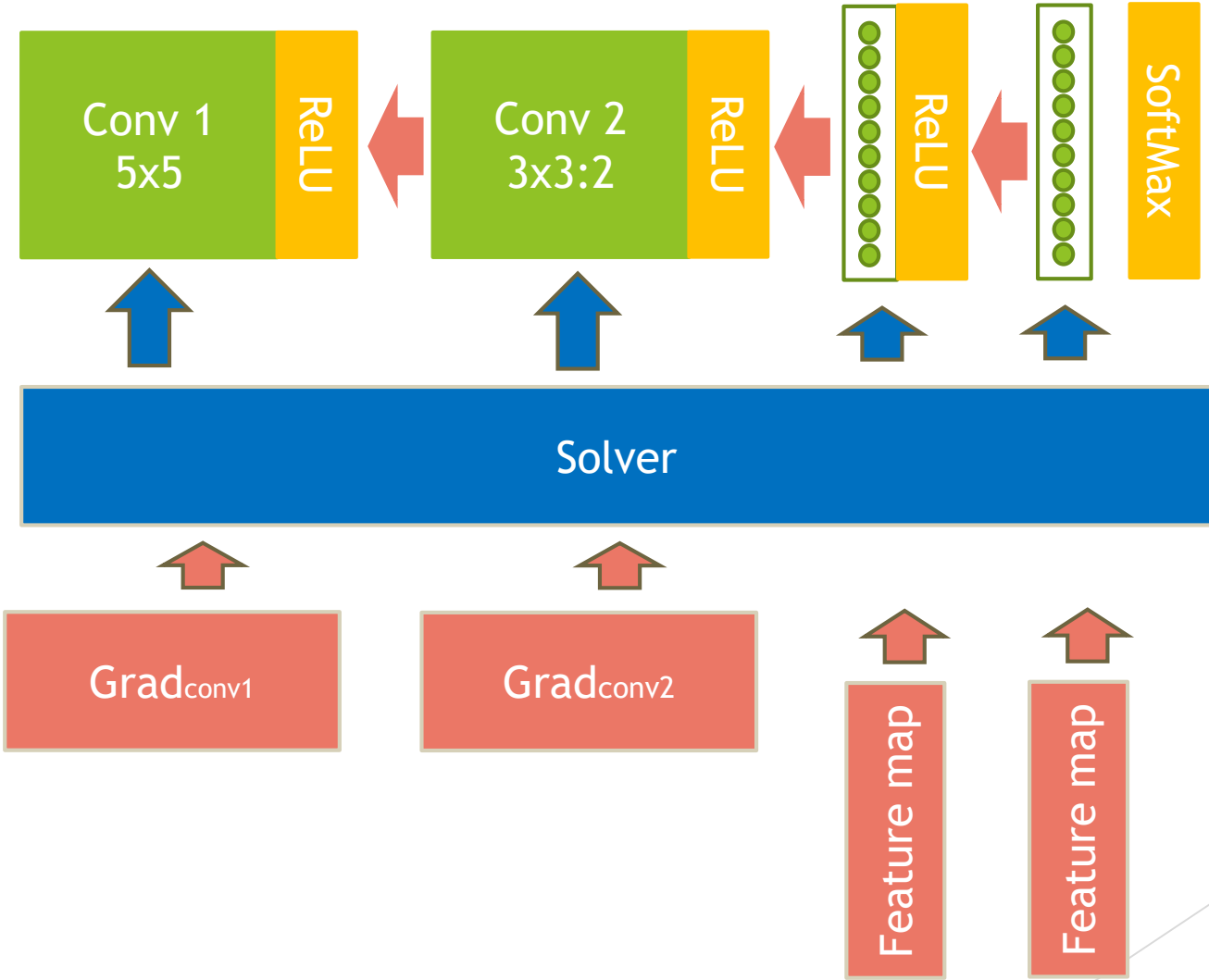


Обучение ГКНС



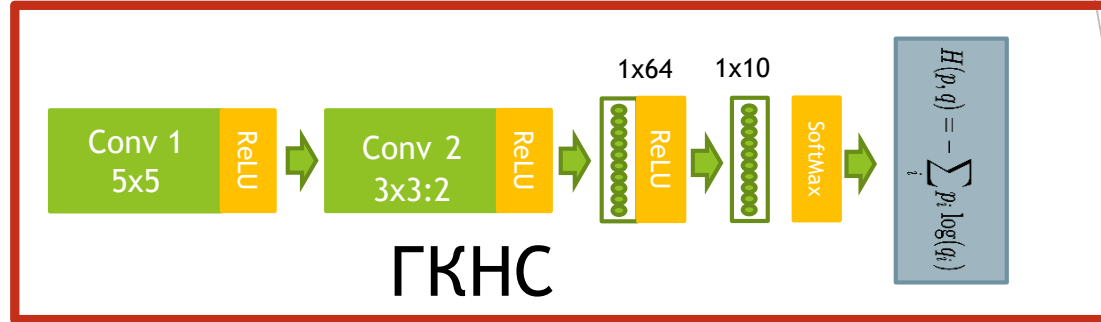
Этап 3.1: усреднение градиентов.

Обучение ГКНС



Этап 4: Обновление весов

Обучение ГКНС



Батч = 5



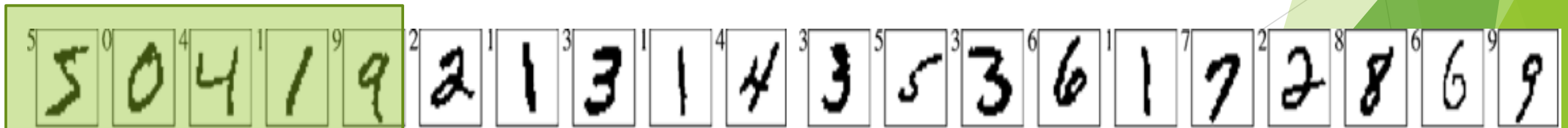
Этап 1: Загрузка данных

Этап 2: Формирование батча

Этап 3: Прямой проход

Этап 4: Обратный проход

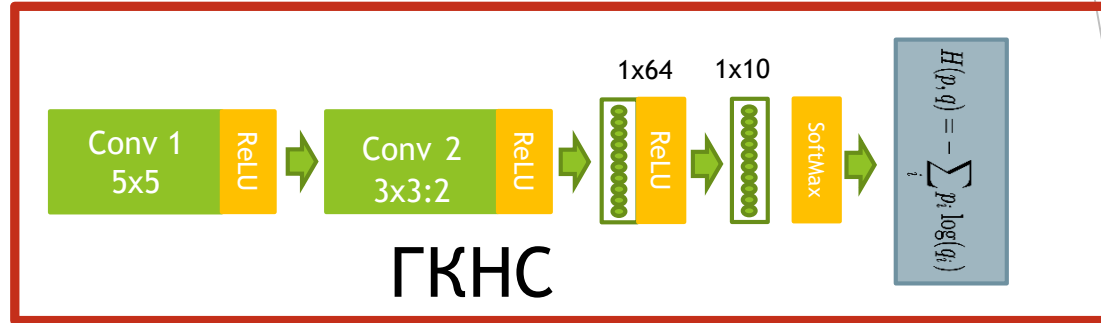
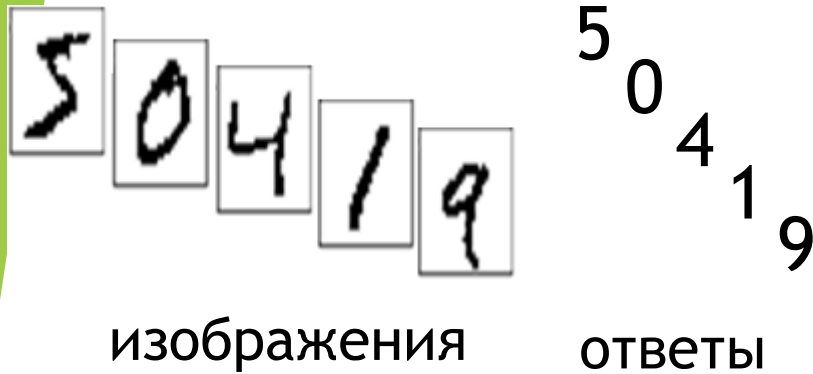
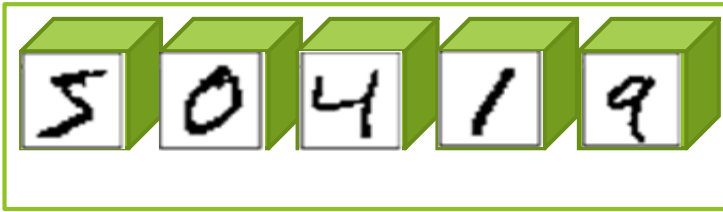
Этап 5: Обновление весов



БД изображений и ответов - 20 примеров

Обучение ГКНС

5 0 4 1 9
Тензор ответов 5x1



Этап 1: Загрузка данных

Этап 2: Формирование батча

Этап 3: Прямой проход

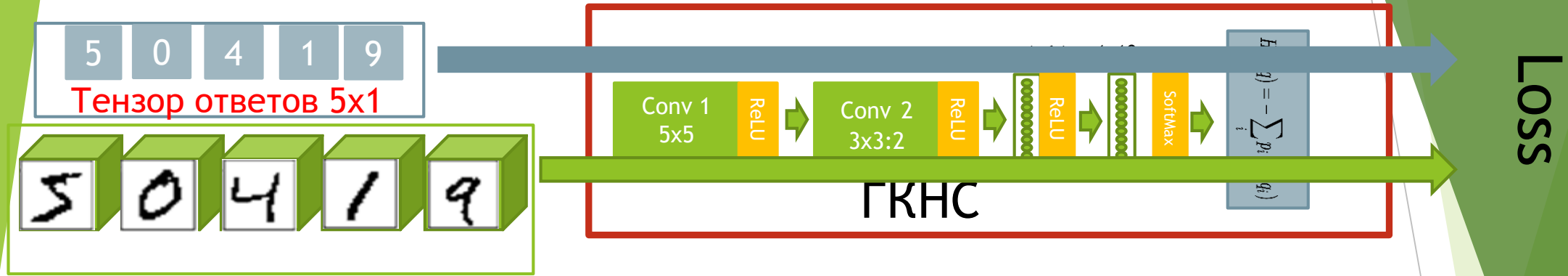
Этап 4: Обратный проход

Этап 5: Обновление весов



БД изображений и ответов - 20 примеров

Обучение ГКНС

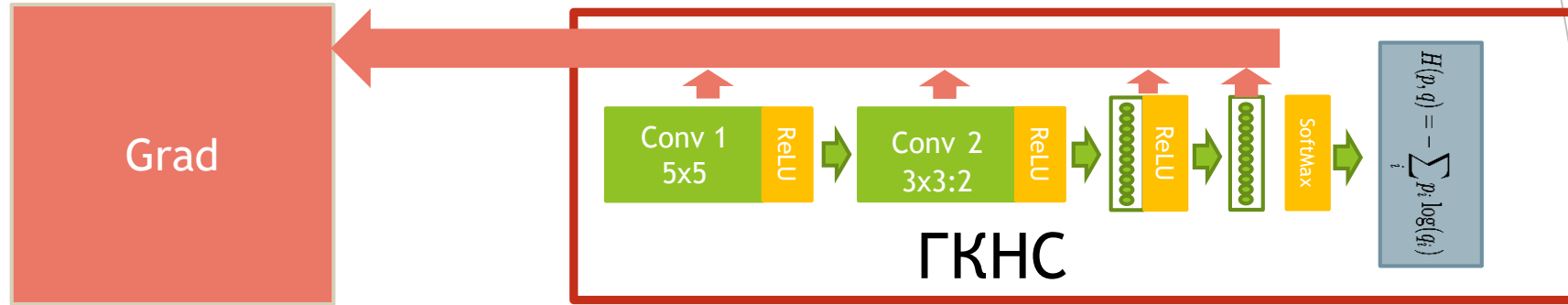


- Этап 1: Загрузка данных
- Этап 2: Формирование батча
- Этап 3: Прямой проход**
- Этап 4: Обратный проход
- Этап 5: Обновление весов



БД изображений и ответов - 20 примеров

Обучение ГКНС



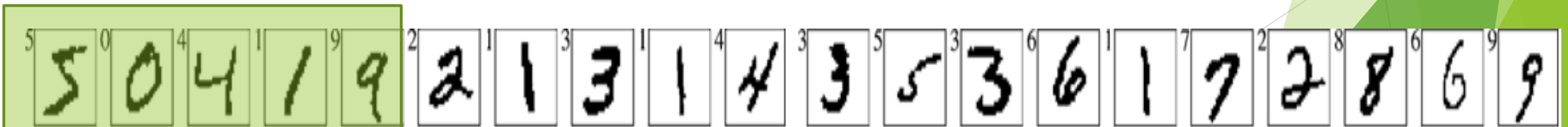
Этап 1: Загрузка данных

Этап 2: Формирование батча

Этап 3: Прямой проход

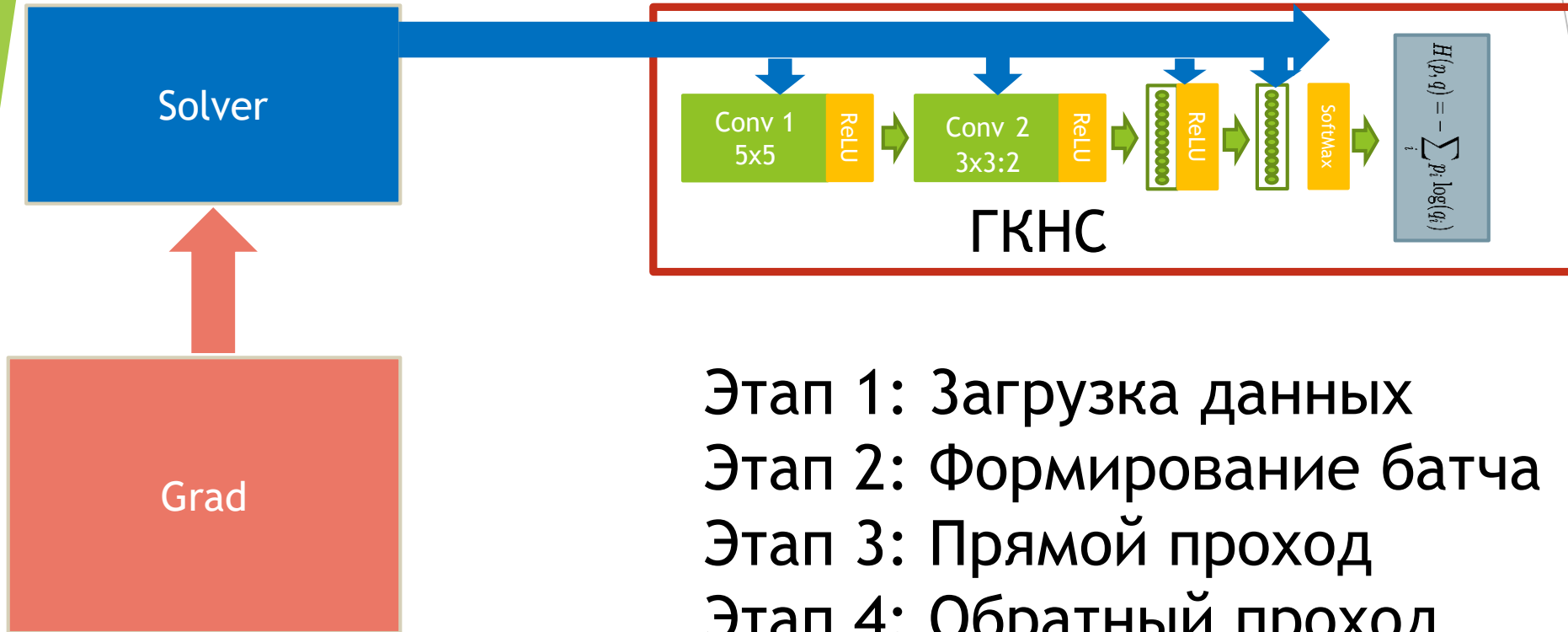
Этап 4: Обратный проход

Этап 5: Обновление весов

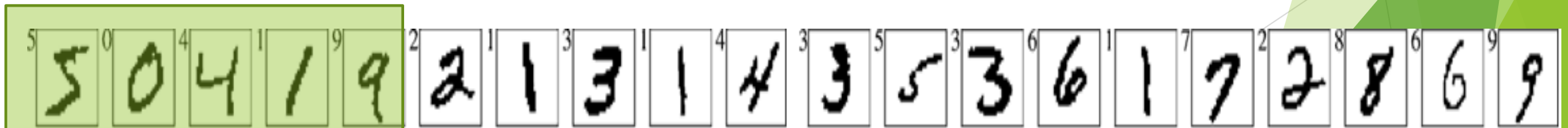


БД изображений и ответов - 20 примеров

Обучение ГКНС



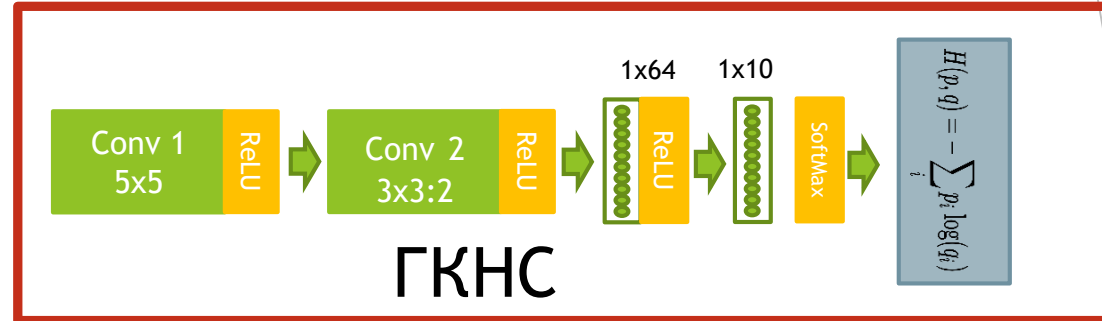
- Этап 1: Загрузка данных
- Этап 2: Формирование батча
- Этап 3: Прямой проход
- Этап 4: Обратный проход
- Этап 5: Обновление весов**



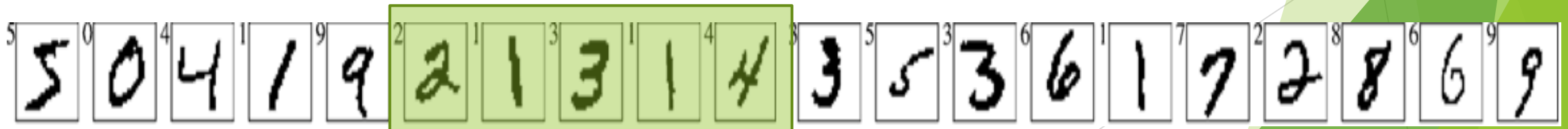
БД изображений и ответов - 20 примеров

Обучение ГКНС

Итерация -
обработка одного
батча



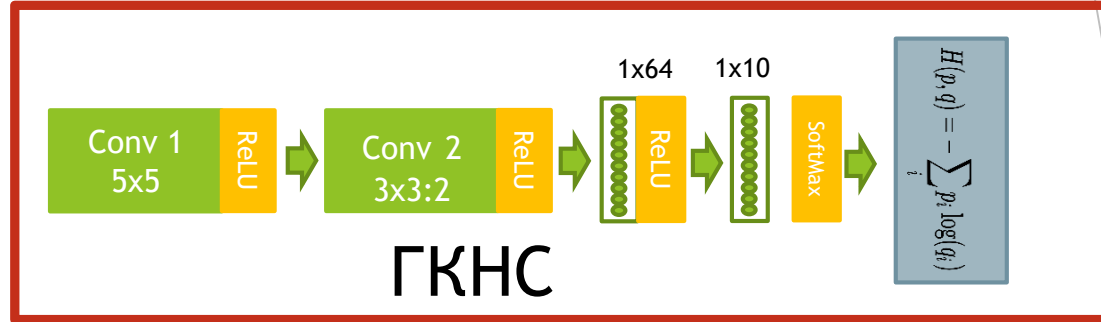
Батч = 5



БД изображений и ответов - 20 примеров

- Этап 1: Загрузка данных
- Этап 2: Формирование батча
- Этап 3: Прямой проход
- Этап 4: Обратный проход
- Этап 5: Обновление весов

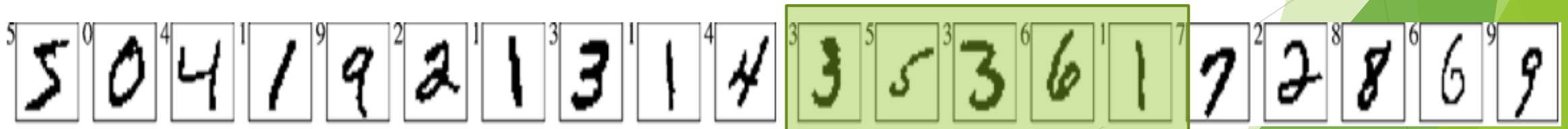
Обучение ГКНС



Батч = 5



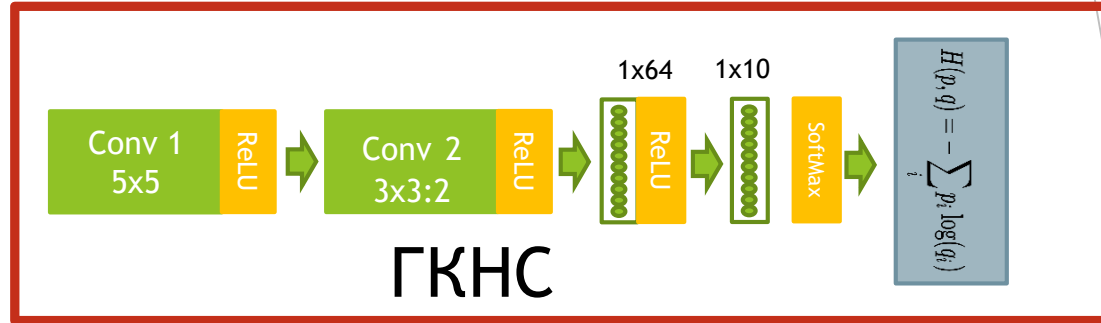
- Этап 1: Загрузка данных
- Этап 2: Формирование батча
- Этап 3: Прямой проход
- Этап 4: Обратный проход
- Этап 5: Обновление весов



БД изображений и ответов - 20 примеров

Обучение ГКНС

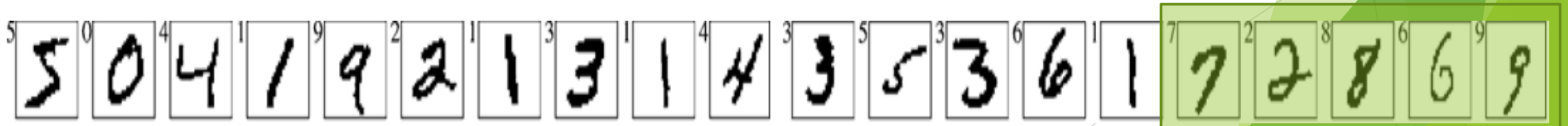
ЭПОХА!



Батч = 5



Эпоха прошла - означает, что все изображения обучающей выборки были обработаны.
По прошествии эпохи обработка начинается сначала



БД изображений и ответов - 20 примеров

Обучение ГКНС

Алгоритм обучения:

Цикл по эпохам от 0 до 20:

Цикл по итерациям от 0 до MaxIter:

Этап 1: Загрузка данных

Этап 2: Формирование батча

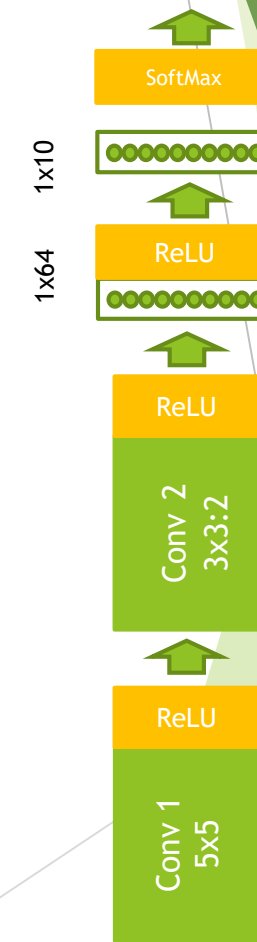
Этап 3: Прямой проход

Этап 4: Обратный проход

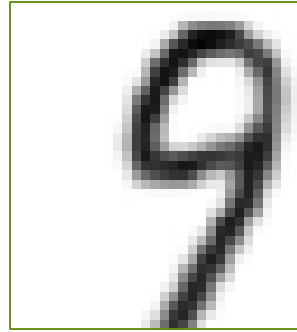
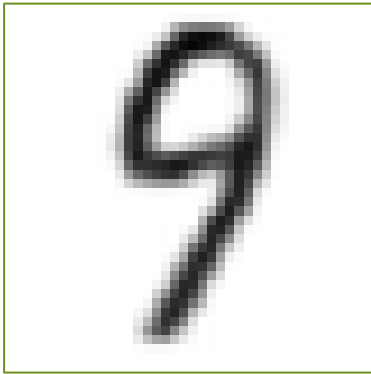
Этап 5: Обновление весов

вывод на экран ф-ции потерь

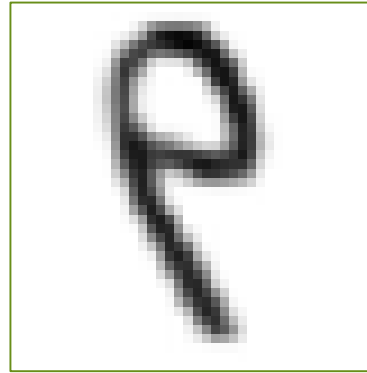
$$H(p, q) = - \sum_i p_i \log(q_i)$$



Обучение ГКНС



Crop (обрезка)



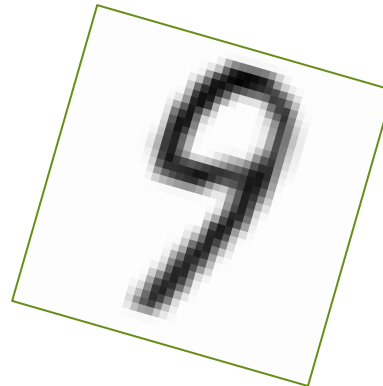
Flip (отражение)



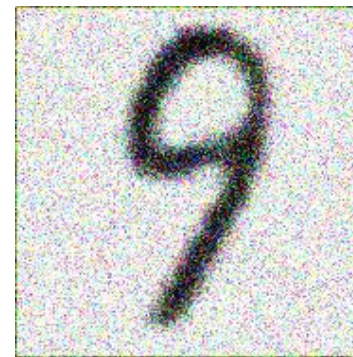
инвертировать цвета



размытие



поворот



шум

Аугментация данных - набор преобразований входных данных со случайными параметрами применяемых для повышения качества обучения

Тестирование качества классификации

		Actual class		
		Cat	Dog	Rabbit
Predicted class	Cat	5	2	0
	Dog	3	3	2
	Rabbit	0	1	11

Confusion matrix - матрица несоответствий.
Столбцы матрицы отвечают за истинные значения,
строки за предсказанные.

Тестирование качества классификации

		Actual class		
		Cat	Dog	Rabbit
Predicted class	Cat	5	2	0
	Dog	3	3	2
	Rabbit	0	1	11

$$Accuracy_{Rabbit} = \frac{11}{13} = 0.846$$

$$Accuracy_{Dog} = \frac{3}{6} = 0.5$$

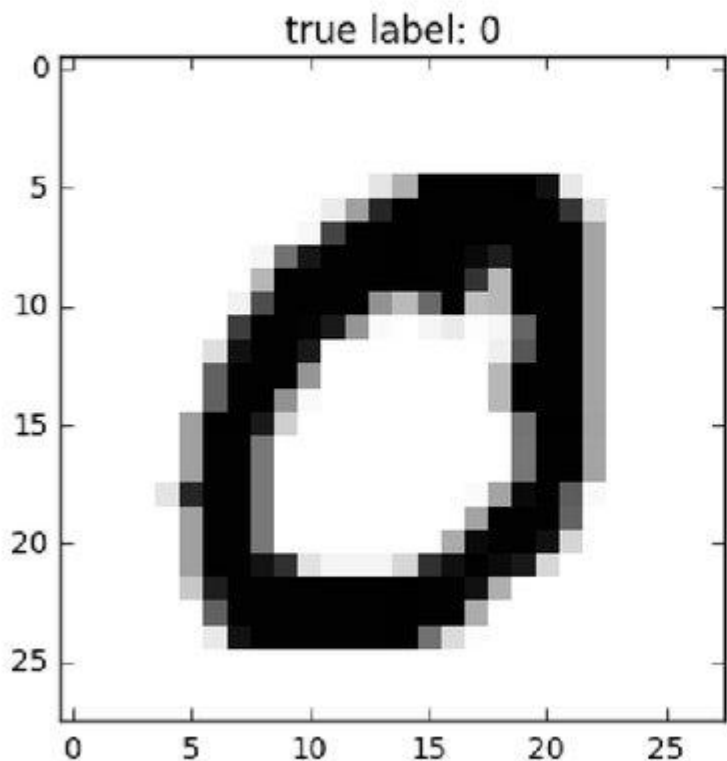
$$Accuracy_{Cat} = \frac{5}{8} = 0.625$$

$$Accuracy_{\Sigma} = \frac{\text{сумма диагональных элементов}}{\text{сумма всех элементов}} = \frac{19}{27} = 0.7$$

Accuracy - отношение числа правильно распознанных примеров к общему числу примеров

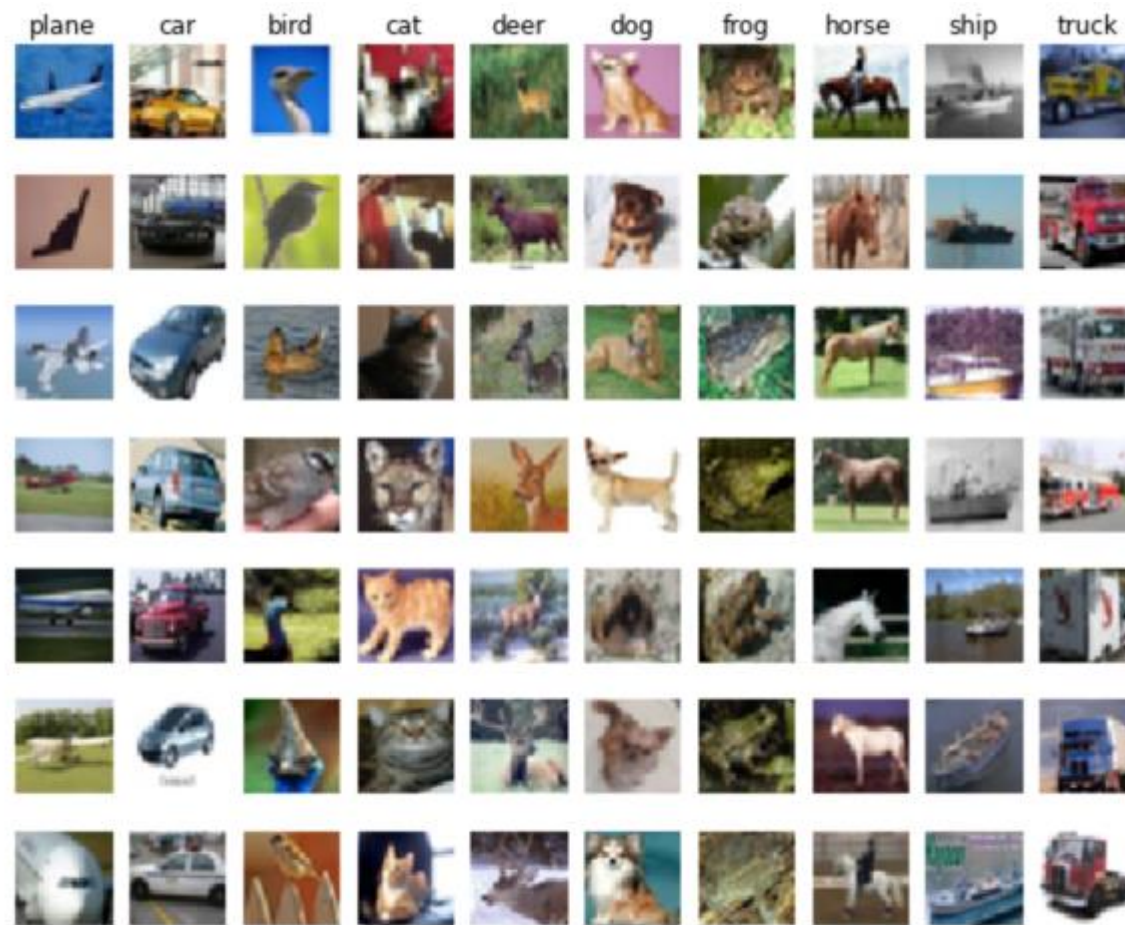
БД классификации

MNIST (сокр. от Mixed National Institute of Standards and Technology)



БД изображений рукописных цифр 28x28. Решен.
Train 60000 изображений. Test 10000

БД классификации



CIFAR: БД цветных изображений(10,100 классов)

Train 50000 изображений. Test 10000.

Не решен(~0.83). Активно используется для тестов.

БД классификации



ImageNet(ILSVRC2014): БД цветных изображений(1000 классов) 256x256
Train 1000000 изображений. Test 150000.
Не решен(~ 0.8). Активно используется для тестов.

Распознавание лиц

Задача верификации:



?



Задача верификации: попарное сравнение

Распознавание лиц

Задача идентификации:



?



Задача идентификации : поиск в базе наиболее похожего человека

Распознавание лиц

ряд кандидатов



№1



№2



№3



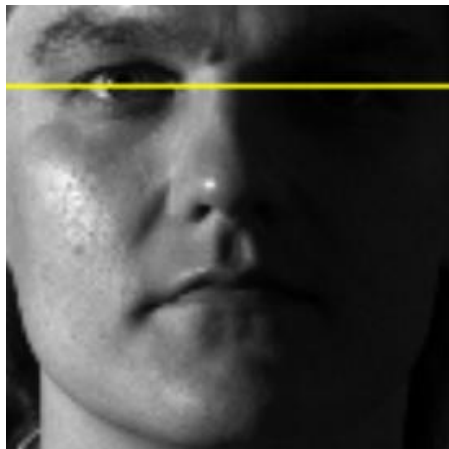
№4

Задача идентификации : формирование списка кандидатов

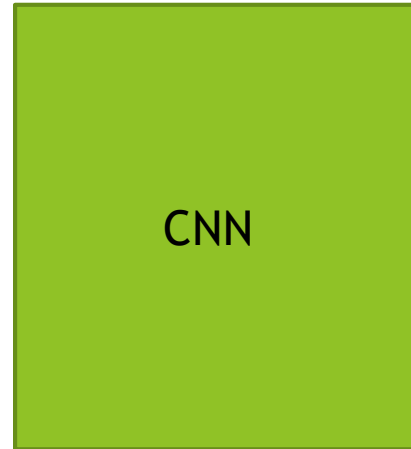
Распознавание лиц



ПОИСК ЛИЦ

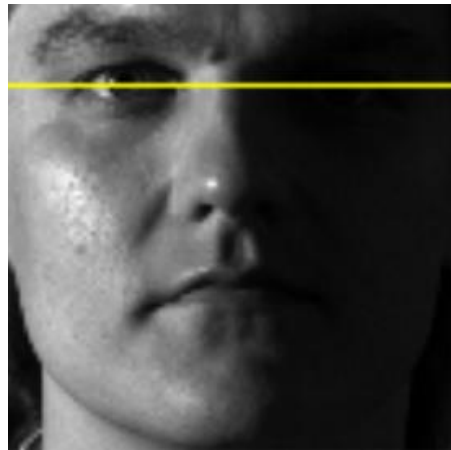


CNN



Построение
биометрического шаблона
(вектора признаков)

Биометрический шаблон



65
112
-36
13
16
64
98
41
0
16

Построение
биометрического шаблона

Биометрический шаблон - вектор чисел, являющийся внутренним представлением изображения лица.

Биометрический шаблон



65	63
112	110
-36	-15
13	8
16	15
64	68
98	110
41	40
0	2
16	14



Сравнение фотографий - сравнение биометрических шаблонов.

Биометрический шаблон



65	-32
112	70
-36	62
13	91
16	22
64	-52
98	50
41	70
0	-32
16	-14



Сравнение фотографий - сравнение биометрических шаблонов.

Биометрический шаблон

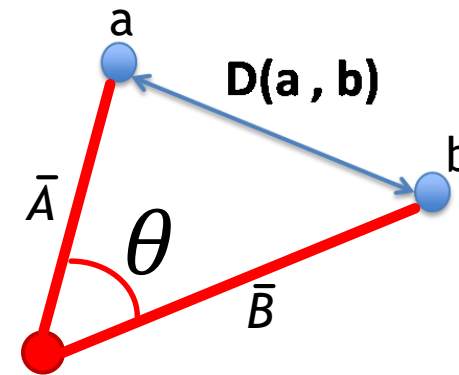
65	63	-32
112	110	70
-36	-15	62
13	8	91
16	15	22
64	68	-52
98	110	50
41	40	70
0	2	-32
16	14	-14
A	B	C

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

$$\cos(A, B) \rightarrow 1$$

$$\cos(A, C) < 0.5$$

$$D(a, b) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2}$$



Сравнение фотографий - сравнение векторов.

Косинусное расстояние либо Евклидово расстояние.

Задача верификации



65	63
112	110
-36	-15
13	8
16	15
64	68
98	110
41	40
0	2
16	14



$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \geq 0.8?$$

Задача верификации - сравнение расстояния с порогом

Задача верификации

Истинное положение дел

		Истинное положение дел	
		На фото одна персона	На фото разные персоны
Ответ алгоритма	На фото одна персона	True Positives (TP)	False Negatives (FN)
	На фото разные персоны	False Positives (FP)	True Negative (TN)

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} > 0.8?$$

Задача верификации - сравнение расстояния с порогом

Задача верификации

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} > thr?$$

FAR(false acceptance rate) - вероятность пропуска чужого

FRR(false reject rate) - вероятность не пропуска своего

$thr = 0$ $FAR = 1.0$ $FRR = 0.0$ (дверь всегда открыта)

$thr = 1.0$ $FAR = 0.0$ $FRR = 1.0$ (дверь всегда закрыта)

Как выбрать порог ? - зависит от задачи

Но всегда пара *FAR* при *FRR* !

При разных *thr* разные *FAR*/*FRR*!

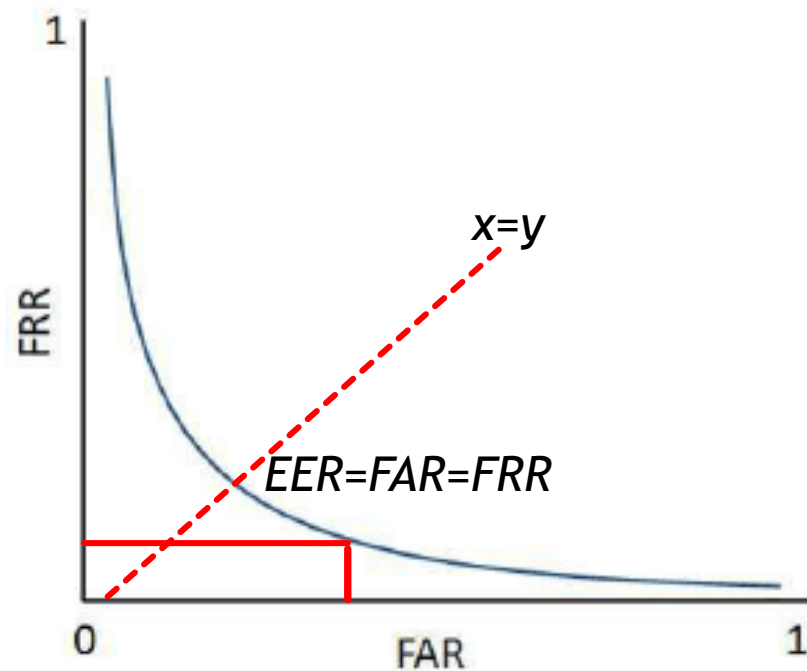
Задача верификации - сравнение расстояния с порогом

Задача верификации

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} > thr?$$

При разных thr разные FAR/FRR!

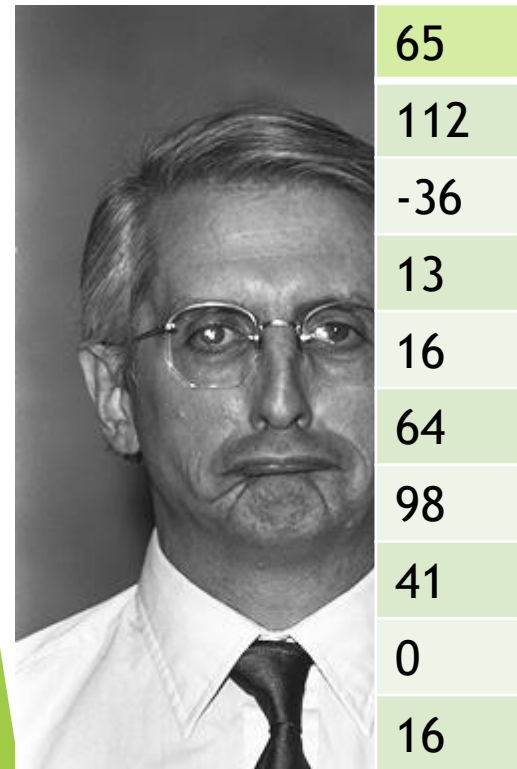
FAR - вероятность пропуска чужого
FRR - вероятность не пропуска своего



Чем ниже график(меньше площадь) FAR/FRR(Рос кривая) тем лучше алгоритм!

Задача идентификации

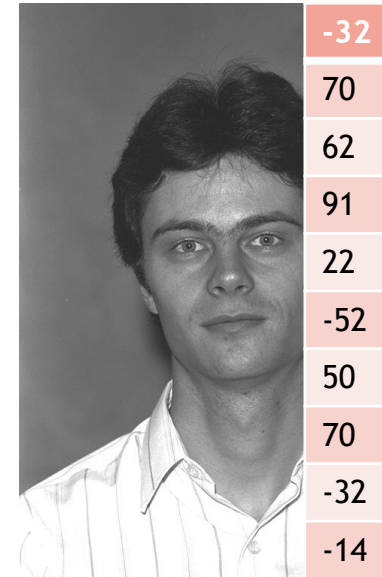
ряд кандидатов



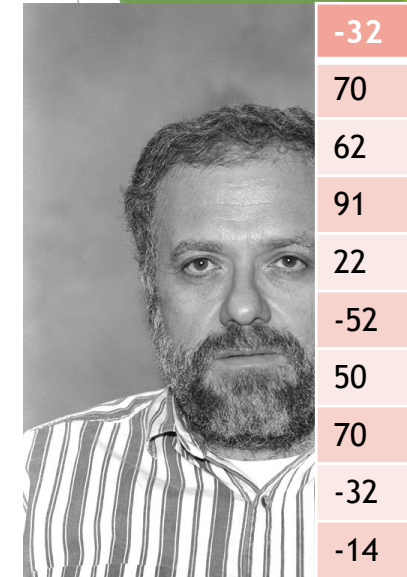
№1 S=0,99



№2 S=0,8



№3 S=0,2



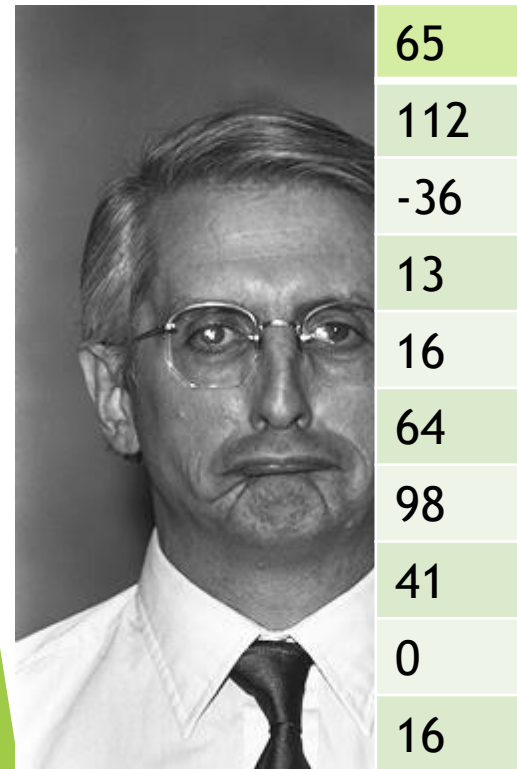
№4 S=0,15

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Задача идентификации - сравнение расстояния по базе. Поиск максимального. Формирование просмотревых списков.

Задача идентификации

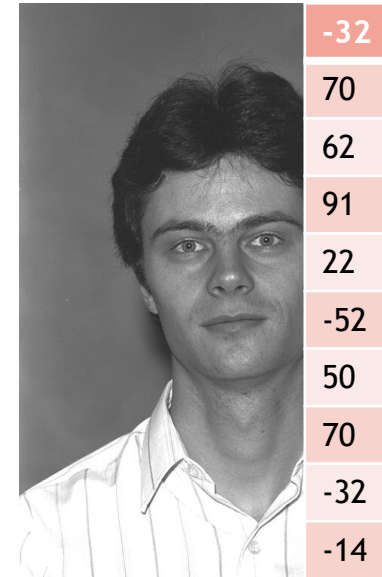
ряд кандидатов



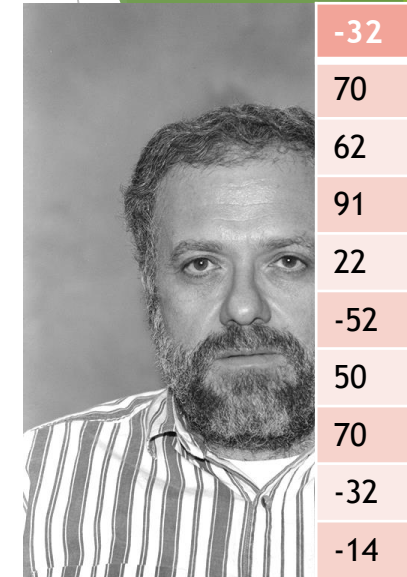
№1 S=0,99



№2 S=0,8



№3 S=0,2

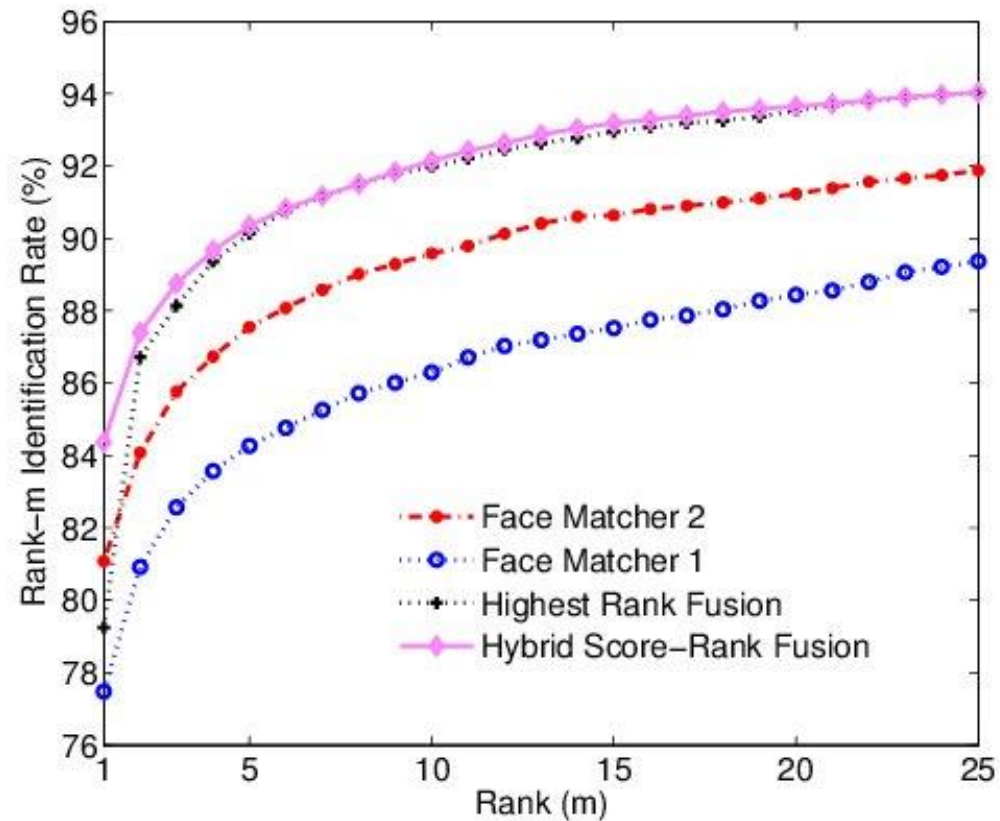
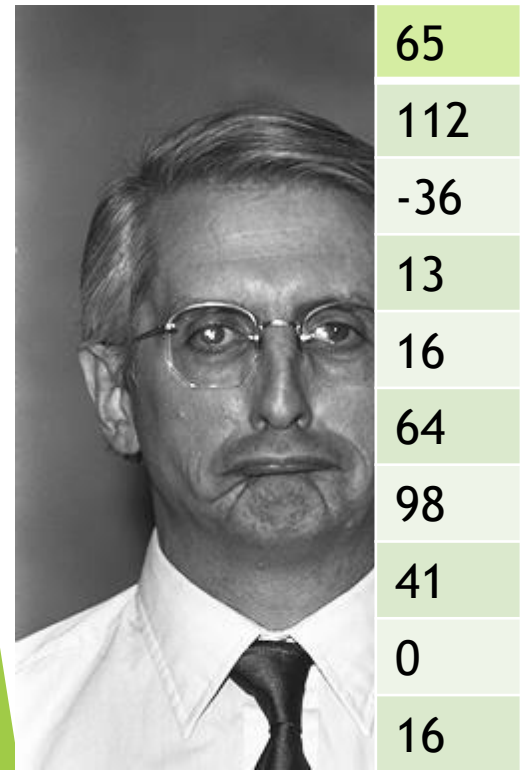


№4 S=0,15

Top N - вероятность попадания в первые N кандидатов

$$\text{Top } N = \frac{\text{число запросов с правильным фото в первых } N}{\text{общее число запросов}}$$

Задача идентификации

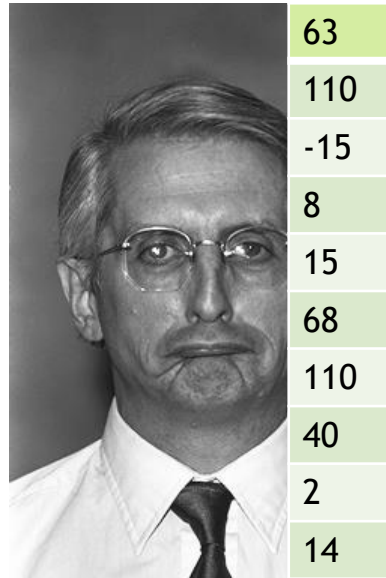
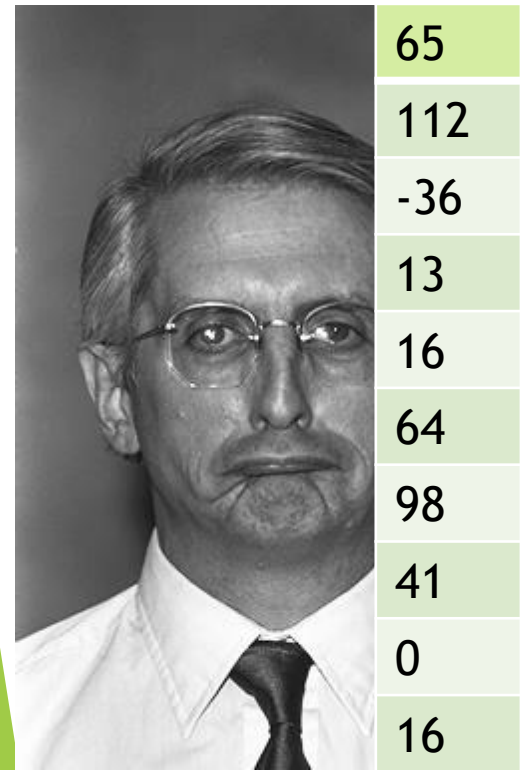


СМС curve - кривая TopN(N)

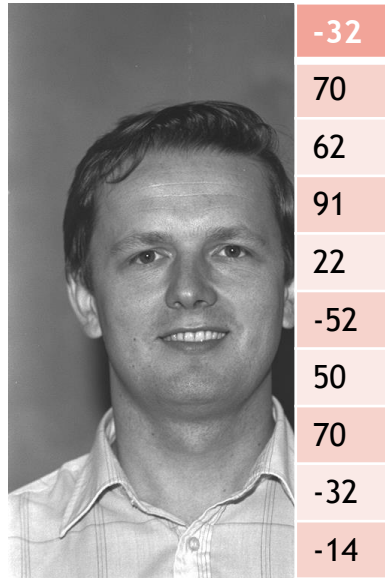
Чем больше площадь тем лучше

Задача идентификации

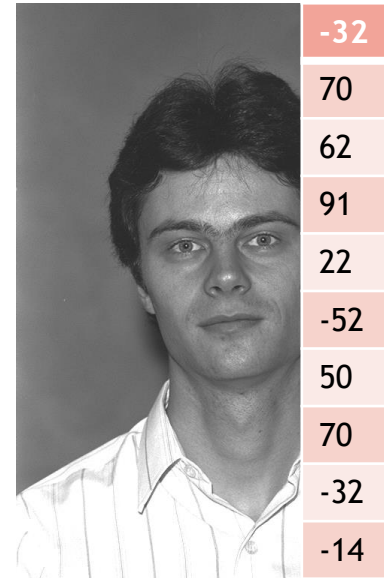
ряд кандидатов



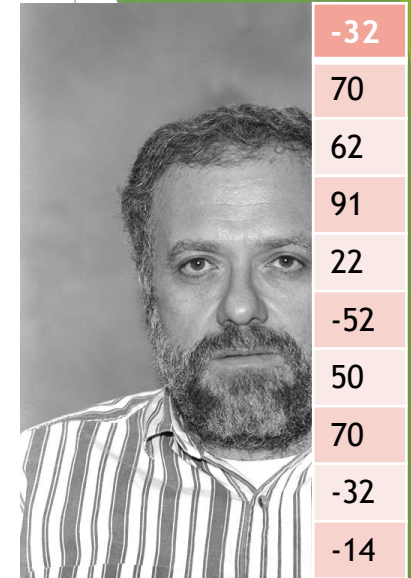
№1 $S=0,99 > thr?$



№2 $S=0,3$



№3 $S=0,2$



№4 $S=0,15$

Задача идентификации - сравнение расстояния по базе. Поиск по базе. Поиск максимального и сравнение с порогом

Задача идентификации

Истинное положение дел

		Истинное положение дел	
		Персона есть в базе	Персоны нет в базе
Ответ алгоритма	Персона есть в базе	True Positives (TP)	False Negatives (FN)
	Персоны нет в базе	False Positives (FP)	True Negative (TN)

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} > 0.8?$$

Задача идентификации - сравнение расстояния по базе. Поиск по базе. Поиск максимального и сравнение с порогом

Задача идентификации

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} > thr?$$

FNIR (false negative identification rate)

- вероятность найти персону которой нет в базе

FPIR (false positive identification rate)

- вероятность найти персону которой нет в базе

$thr = 0$ $FNIR = 1.0?$ $FPIR = 0.0$ (дверь всегда открыта)

$thr = 1.0$ $FNIR = 0.0$ $FPIR = 1.0?$ (дверь всегда закрыта)

Как выбрать порог ? - зависит от задачи

Но всегда пара *FNIR* при *FPIR* !

При разных *thr* разные *FNIR/FPIR*!

Предобработка



Обнаружения лица



Поиск особых точек



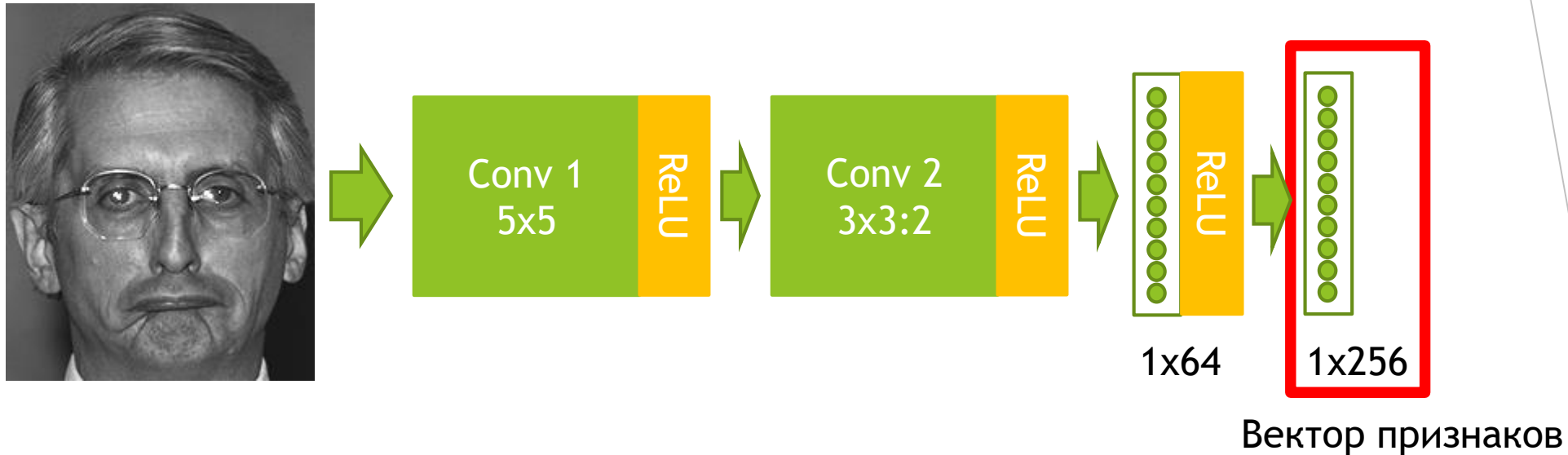
128

128

Нормализация -
поворот и масштаб

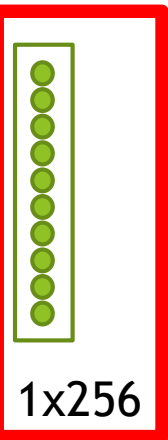
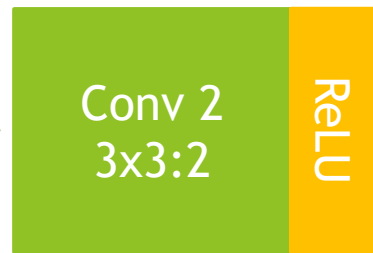
На вход алгоритму распознавания обычно подаётся нормализованное лицо. При этом на текущий момент яркостных преобразований не проводят.

Построение векторов признаков

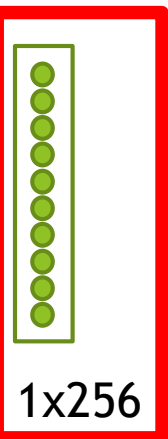
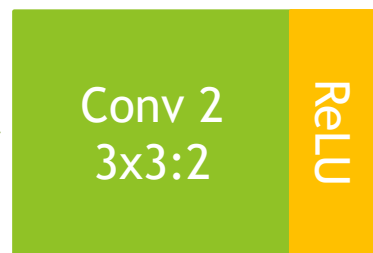
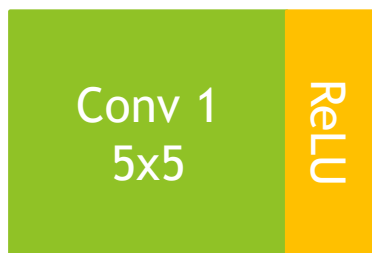


Архитектура может быть любой но она должна заканчиваться слоем, выходом которого является вектор **1 x n (n=128,256...)**

Сиамские сети



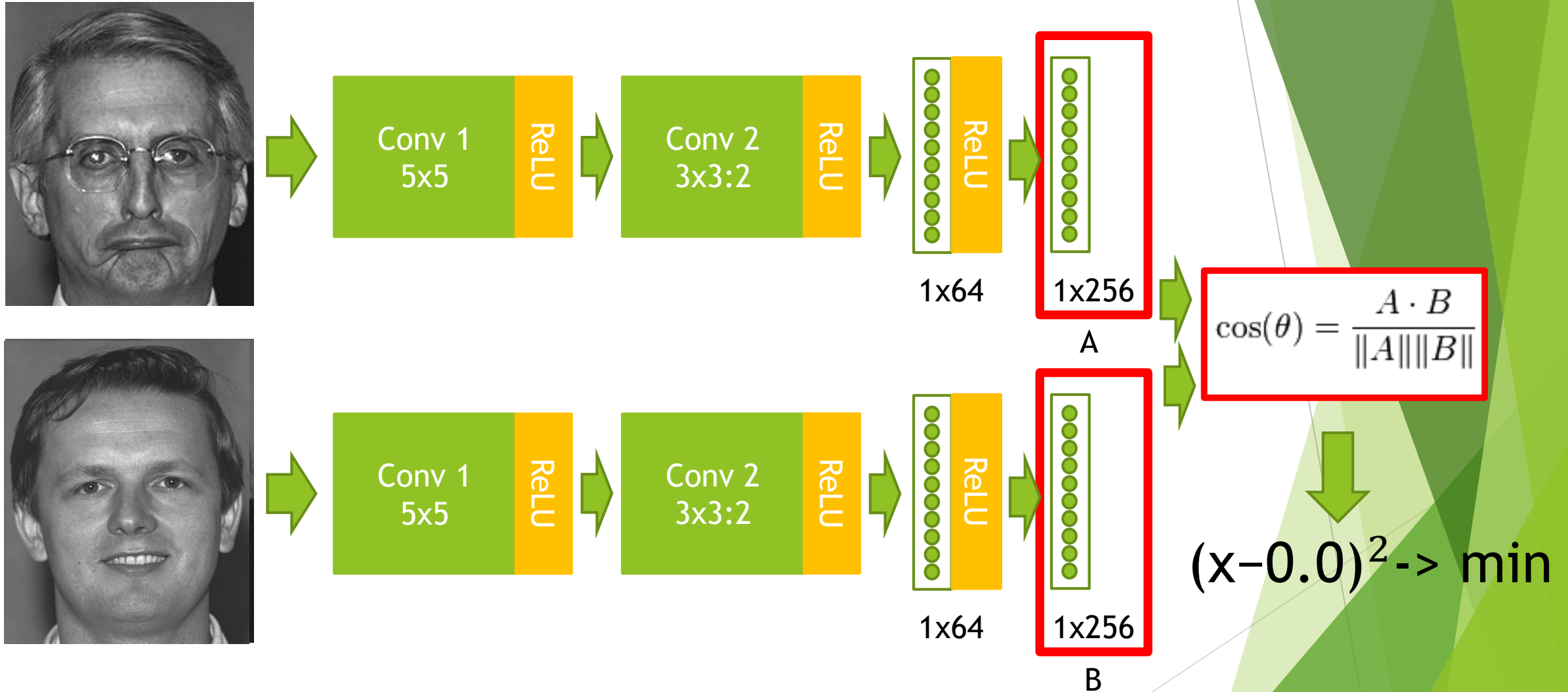
$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$



$$(x-1.0)^2 \rightarrow \min$$

Обучение парами!

Сиамские сети



Обучение парами! Если изображений N то пар N^2 !!!
Своих пар меньше чем чужих!!! Обучение на верификацию!

Триплеты



Figure 3. The **Triplet Loss** minimizes the distance between an *anchor* and a *positive*, both of which have the same identity, and maximizes the distance between the *anchor* and a *negative* of a different identity.

Триплеты

$$\sum_i^N \left[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]_+$$



α - параметр сдвига.

Если значение меньше нуля то оно считается нулевым.

Триплеты

FaceNet: A Unified Embedding for Face Recognition and Clustering

Florian Schroff

fschroff@google.com

Google Inc.

Dmitry Kalenichenko

dkalenichenko@google.com

Google Inc.

James Philbin

jphilbin@google.com

Google Inc.

Статья google на момент выхода показала лучшие результаты. 200 миллионов фото 8 млн. персон.

Сеть google-net.

До сих пор одна из лучших по разновозрастным базам.

Триплеты



N^3 триплетов - нужен алгоритм подбора сложных случаев.

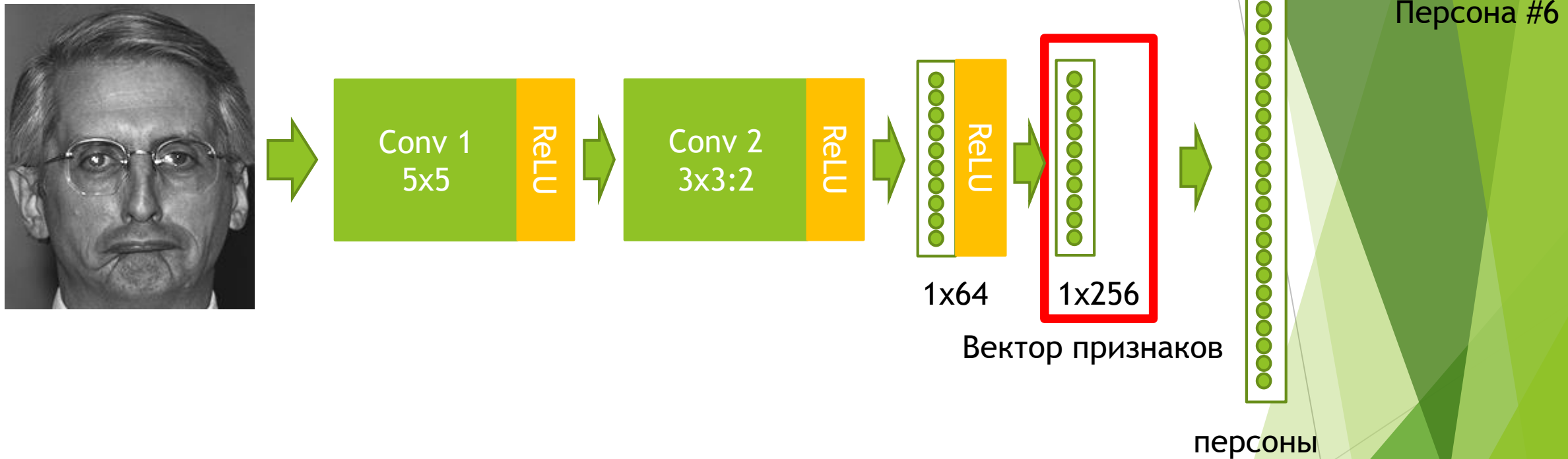
Три сети!

Очень медленно учится.

Обычно используется для дообучения.

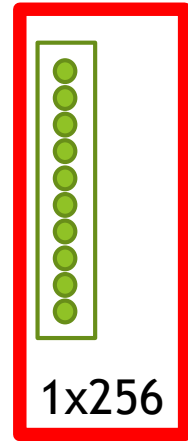
Есть варианты с cosine.

Обучение на идентификацию

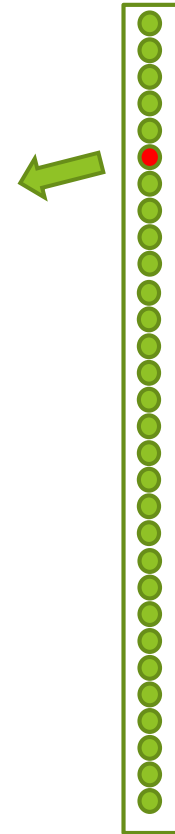
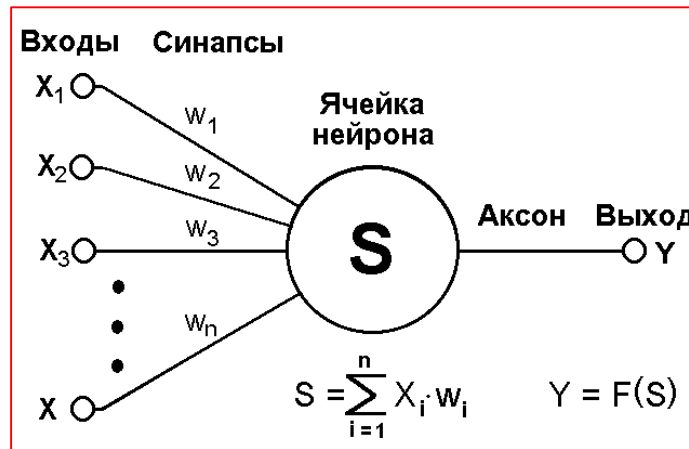


Обучение через классификацию. Каждой персона - это класс.
Вектор признаков - **предпоследний** слой сети

Обучение на идентификацию



Вектор признаков

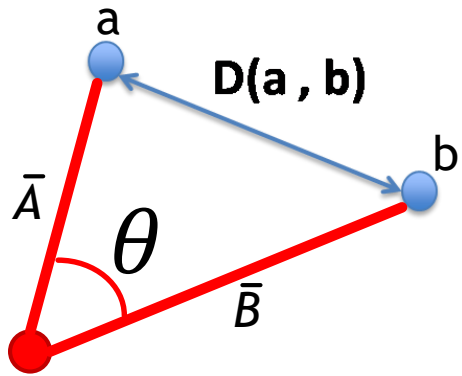


Персона #6

$$S = \sum_{i=1}^{256} x_i * w_i + B$$

Каждая персона это нейрон со своими весами.

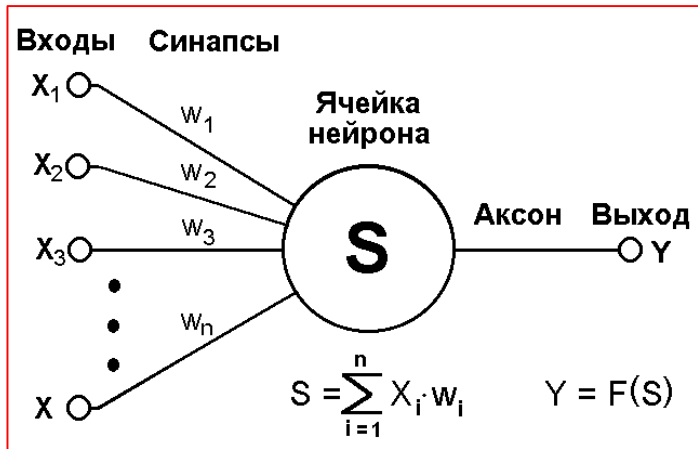
Обучение на идентификацию



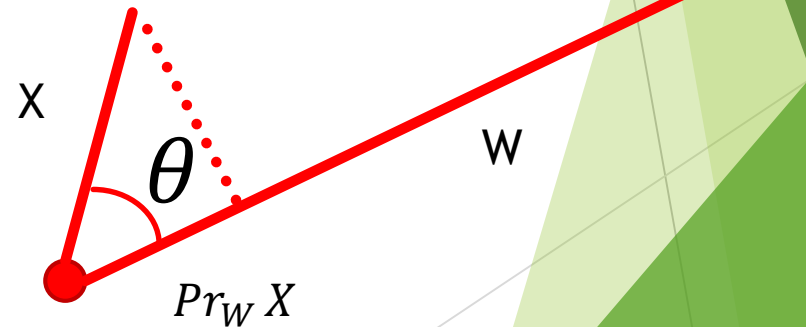
$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

$$S = \sum_{i=1}^{256} x_i * w_i + B$$

ПОРОГ

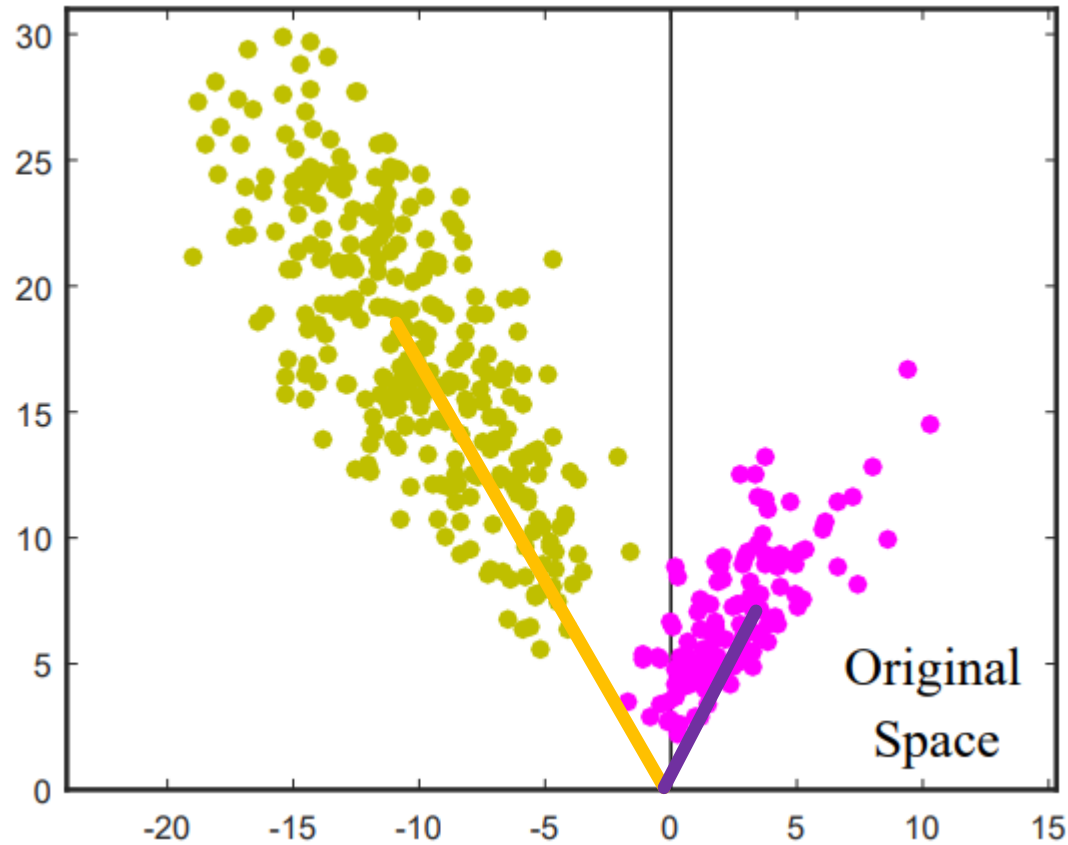


Персона #6



Каждая персона это нейрон со своими весами.

Обучение на идентификацию



Каждая нейрон - биометрический шаблон класса

Шаблоны изображений должны быть близки с шаблоном класса
Соответственно и друг с другом.

Обучение на идентификацию

A Light CNN for Deep Face Representation with Noisy Labels

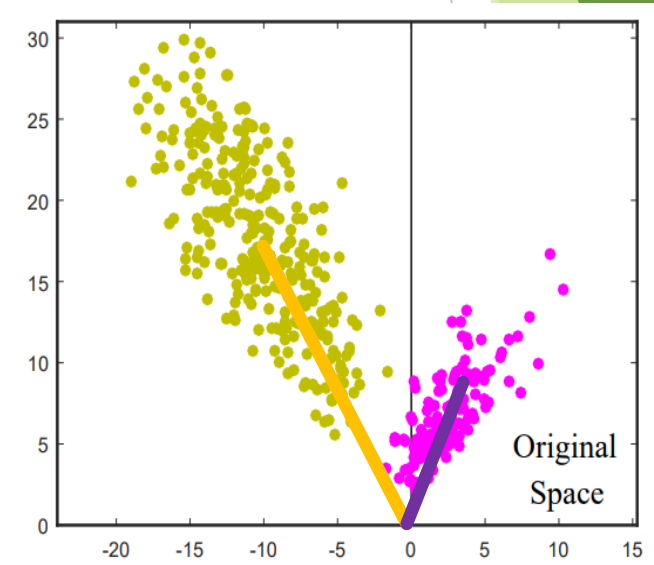
Xiang Wu, Ran He, *Senior Member, IEEE*, Zhenan Sun*, *Member, IEEE*, and Tieniu Tan, *Fellow, IEEE*

Одна из наиболее быстрых сетей

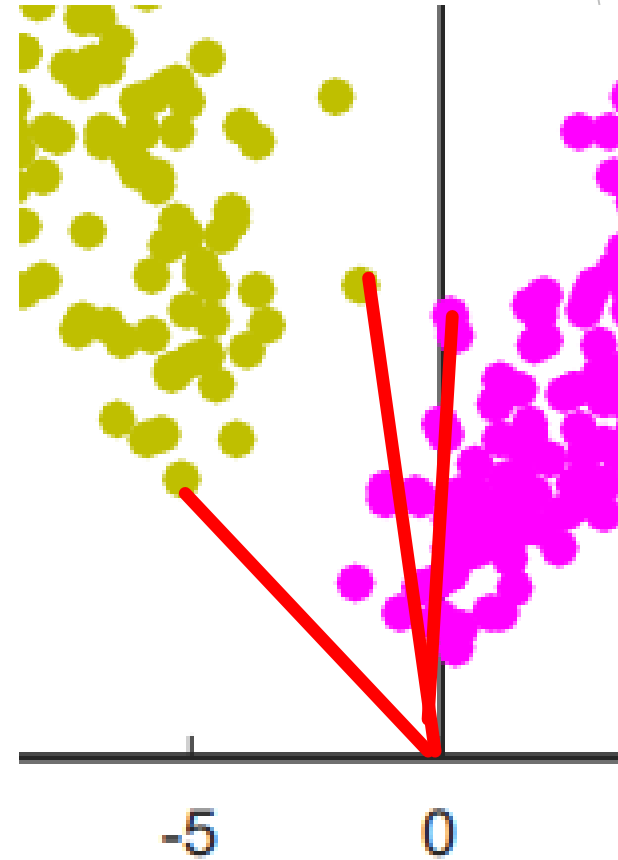
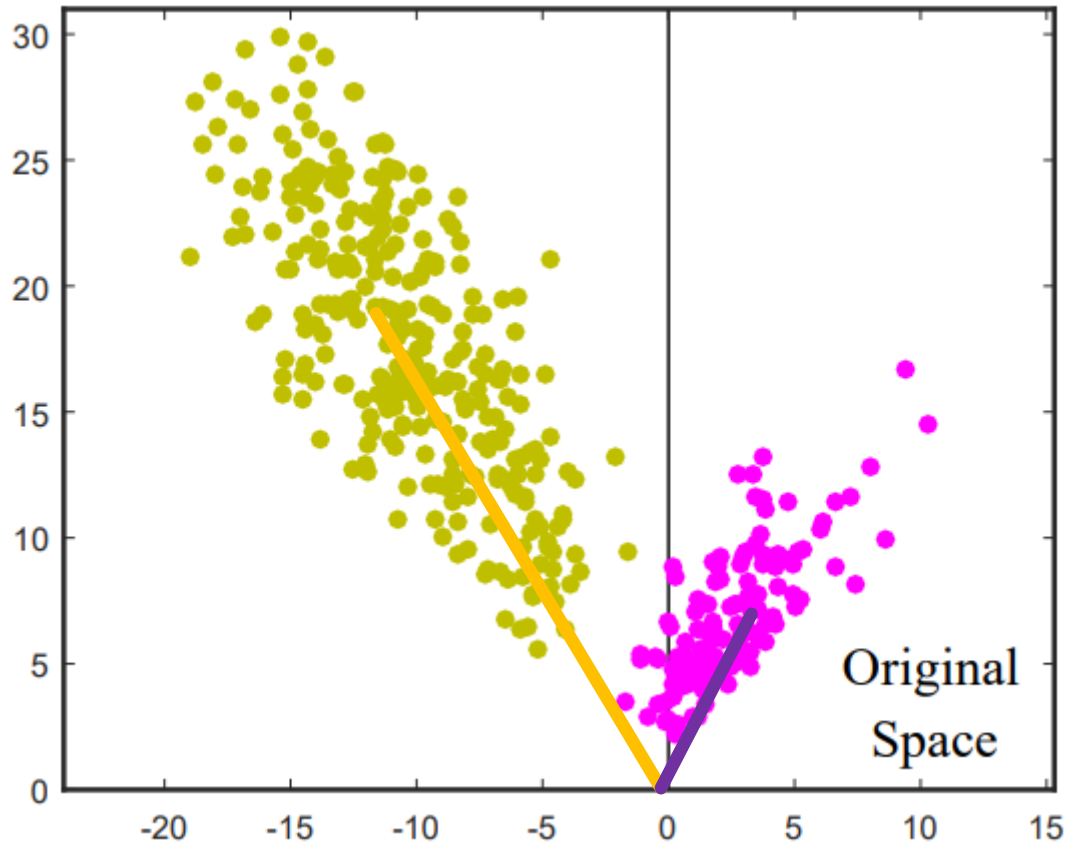
Одна из лучших по скорость/качество

До сих пор актуальна

github!



Обучение на идентификацию

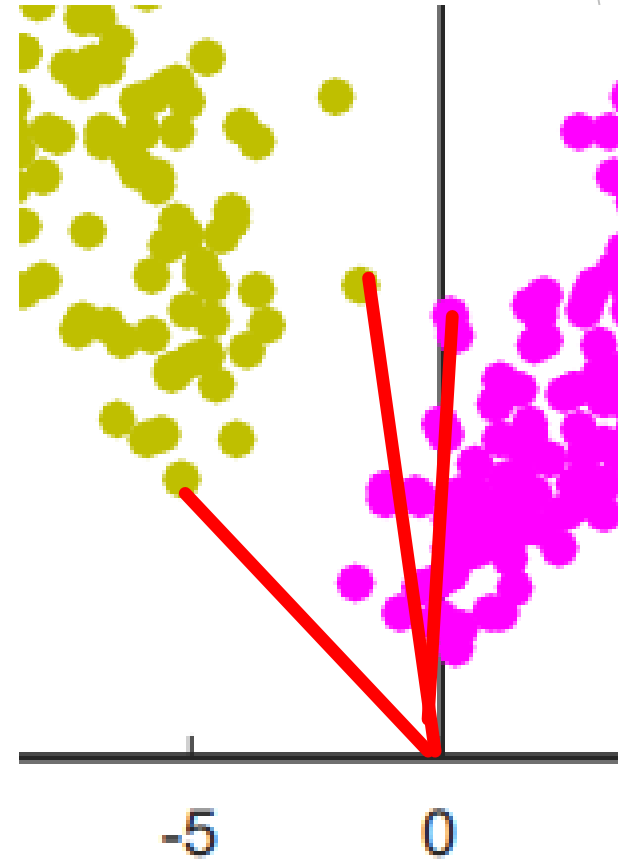
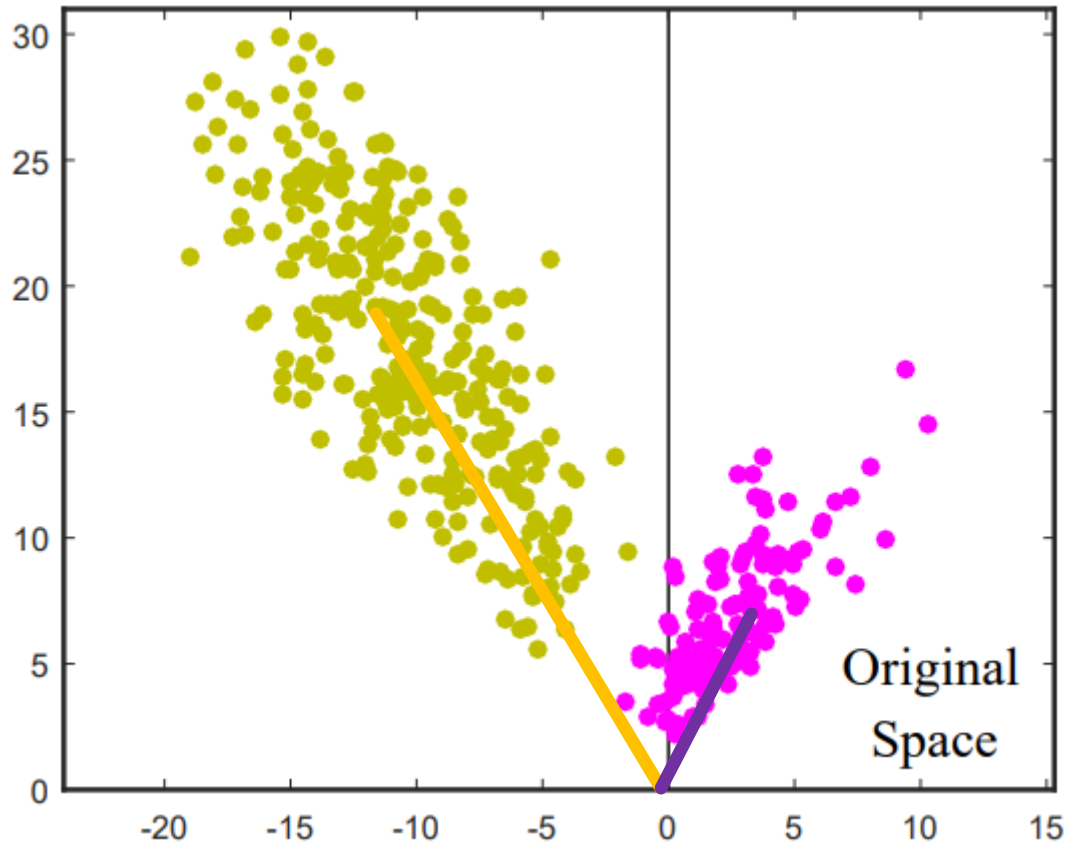


Проблема мы решаем не ту задачу

Шаблоны изображений должны быть близки с друг с другом
а не с вектором класса

Отсутствует нормировка

Обучение на идентификацию

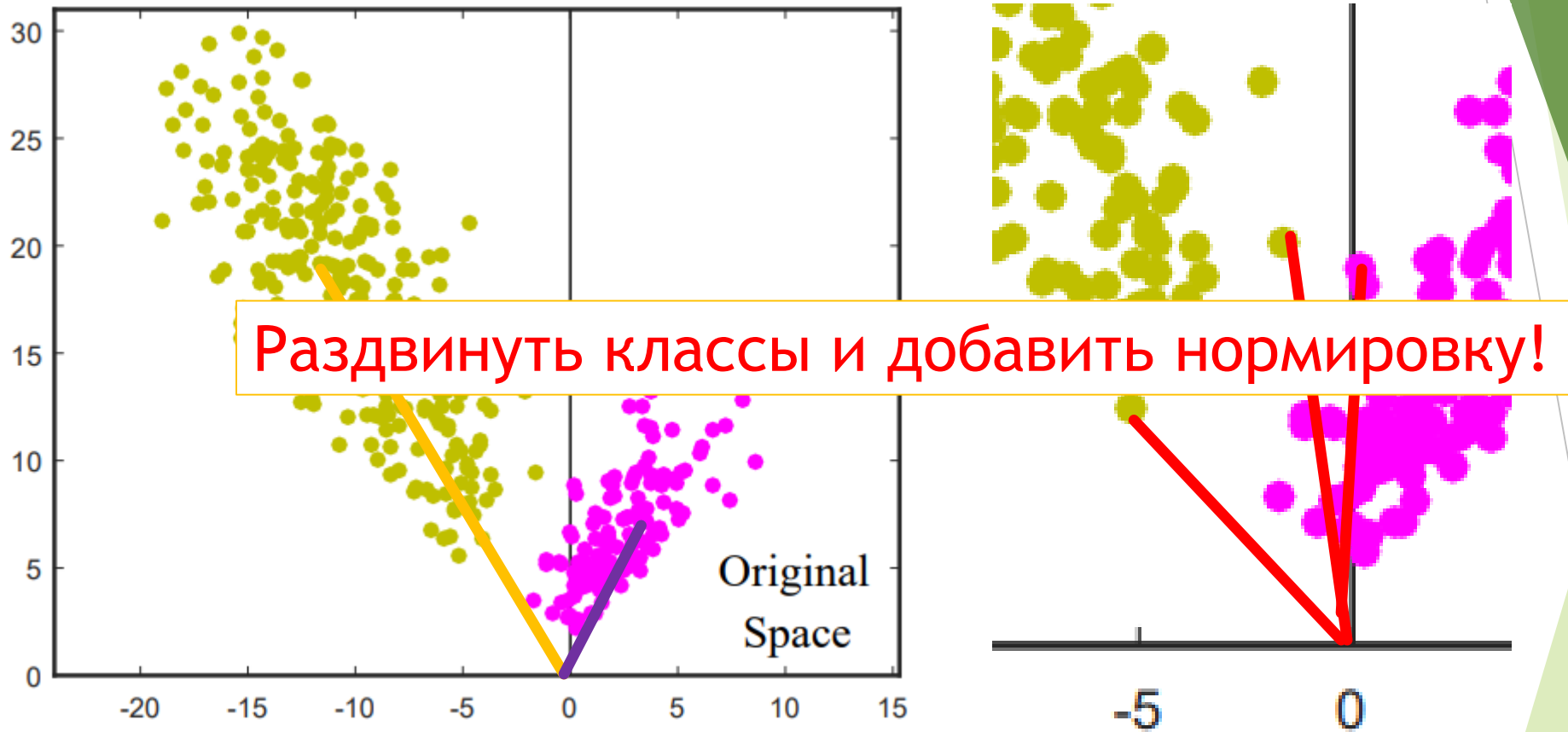


Проблема мы решаем не ту задачу

Шаблоны изображений должны быть близки с друг с другом
а не с вектором класса

Отсутствует нормировка

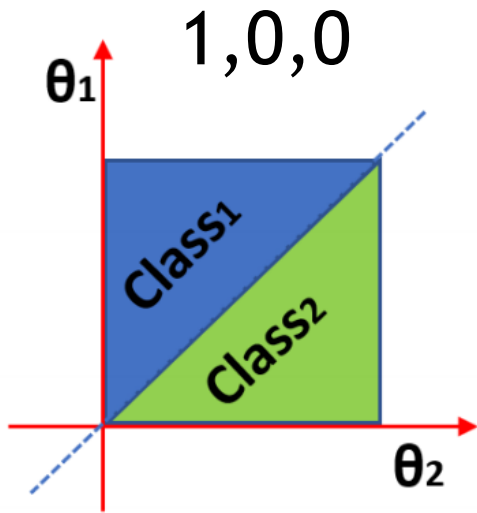
Обучение на идентификацию



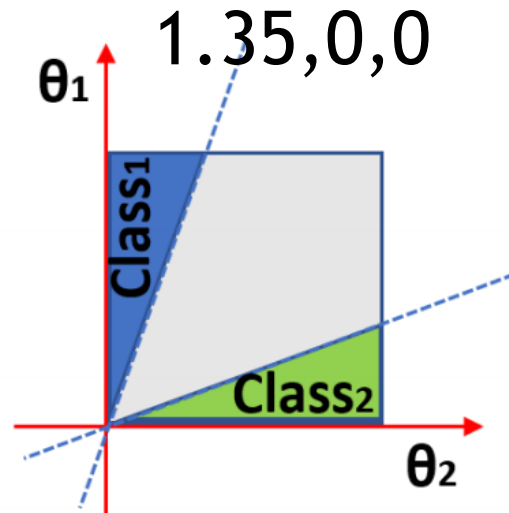
Проблема мы решаем не ту задачу
Шаблоны изображений должны быть близки с друг с другом
а не с вектором класса
Отсутствует нормировка

Обучение на идентификацию

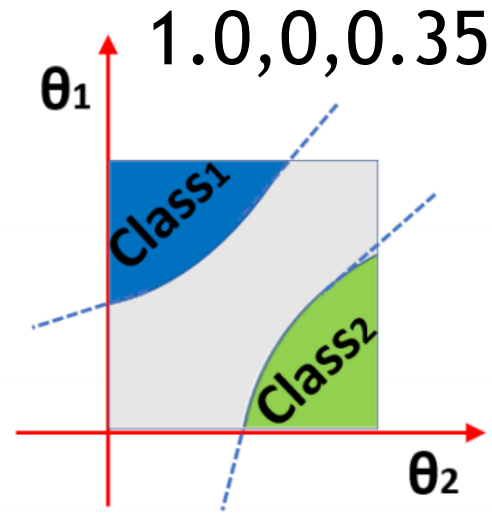
m_1, m_2, m_3



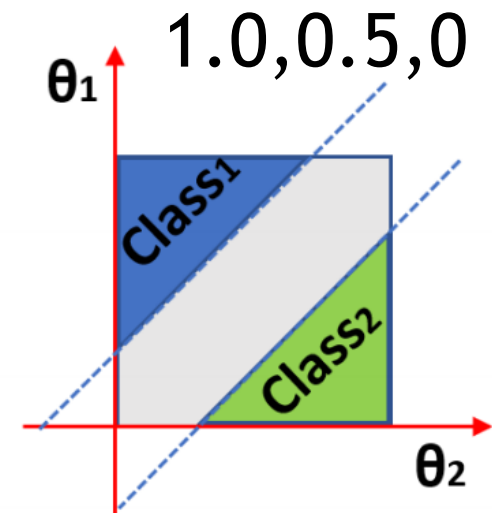
Softmax



SphereFace



CosFace



ArcFace

Раздвинуть классы и добавить нормировку!

$$L_4 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(m_1 \theta_{y_i} + m_2) - m_3)}}{e^{s(\cos(m_1 \theta_{y_i} + m_2) - m_3)} + \sum_{j=1, j \neq y_i}^n e^{s \cos \theta_j}} \quad (4)$$

Обучение на идентификацию

ArcFace: Additive Angular Margin Loss for Deep Face Recognition

Jiankang Deng *
Imperial College London
j.deng16@imperial.ac.uk

Jia Guo *
InsightFace
guojia@gmail.com

Niannan Xue
Imperial College London
n.xue15@imperial.ac.uk

Stefanos Zafeiriou
Imperial College London
s.zafeiriou@imperial.ac.uk

State of art!

Сети ResNet, LResNet50E

github!

Обучение на идентификацию

ArcFace: Additive Angular Margin Loss for Deep Face Recognition

Method	#Image	LFW	YTF
DeepID [30]	0.2M	99.47	93.20
Deep Face [31]	4.4M	97.35	91.4
VGG Face [22]	2.6M	98.95	97.30
FaceNet [27]	200M	99.63	95.10
Baidu [15]	1.3M	99.13	-
Center Loss [36]	0.7M	99.28	94.9
Range Loss [43]	5M	99.52	93.70
Marginal Loss [8]	3.8M	99.48	95.98
SphereFace [17]	0.5M	99.42	95.0
SphereFace+ [16]	0.5M	99.47	-
CosFace [35]	5M	99.73	97.6
MS1MV2, R100, ArcFace	5.8M	99.83	98.02

Verification EER

Обучение на идентификацию

ArcFace: Additive Angular Margin Loss for Deep Face Recognition

Top1

@FAR=10⁻⁶

Methods	Id (%)	Ver (%)
Softmax [17]	54.85	65.92
Contrastive Loss[17, 30]	65.21	78.86
Triplet [17, 27]	64.79	78.32
Center Loss[36]	65.49	80.14
SphereFace [17]	72.729	85.561
CosFace [35]	77.11	89.88
AM-Softmax [33]	72.47	84.44
SphereFace+ [16]	73.03	-
CASIA, R50, ArcFace	77.50	92.34
CASIA, R50, ArcFace, R	91.75	93.69
FaceNet [27]	70.49	86.47
CosFace [35]	82.72	96.65
MS1MV2, R100, ArcFace	81.03	96.98
MS1MV2, R100, CosFace	80.56	96.56
MS1MV2, R100, ArcFace, R	98.35	98.48
MS1MV2, R100, CosFace, R	97.91	97.91

Megaface challenge 1
База 1000000

БД для обучения и теста

Faces in the Wild



Carlos Ortega (3)



Carlos Quintanilla
Schmidt (2)



Carlos Ruiz (3)



Carlos Vives (4)



Carly Fiorina (3)



Carmen Electra
(6)



Carol Burnett (2)



Carol Moseley
Braun (2)



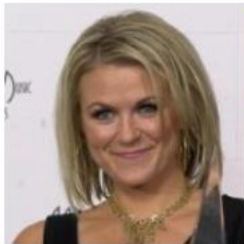
Carolina Klüft (3)



Carolina Moraes
(2)



Caroline Kennedy
(3)



Carolyn Dawn
Johnson (3)



Carrie-Anne Moss
(5)



Carson Daly (2)



Carson Palmer
(4)

13000 изображений 1600
персон.
Не актуальна.
Первая сложная база лиц

БД для обучения и теста

Casia Webface



500000 изображений 10575 персон.

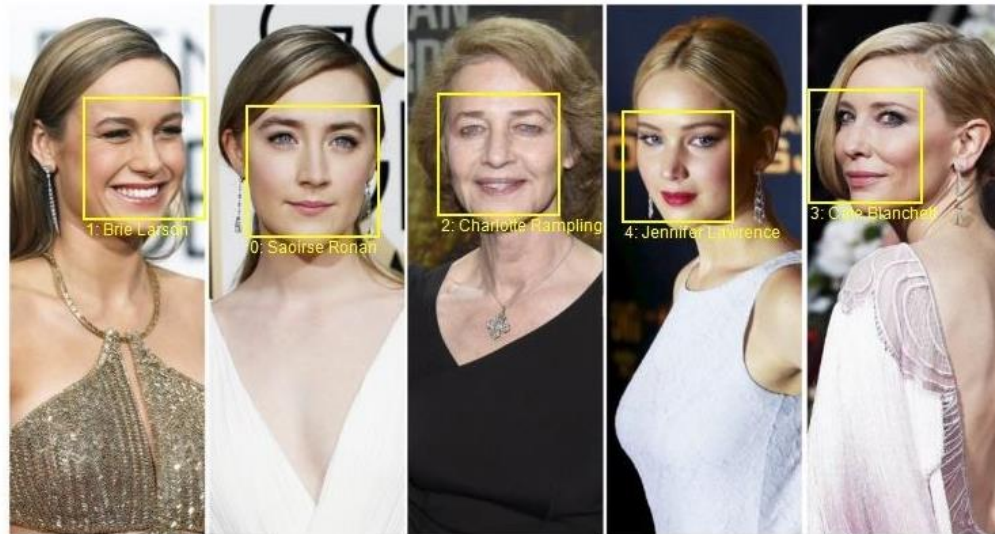
Первая база лиц для глубокого обучения

Используется для обучения по малым базам

БД для обучения и теста

Celeb1M

MSR Image Recognition Challenge (IRC)
@ACM Multimedia 2016



10 млн. изображений 100 тыс. персон.

Самая большая база.

Очень “грязная” (есть фото Аристотеля и Платона)

После чистки 5 млн. фото 80 тыс. персон.

БД для обучения и теста

MegaFace

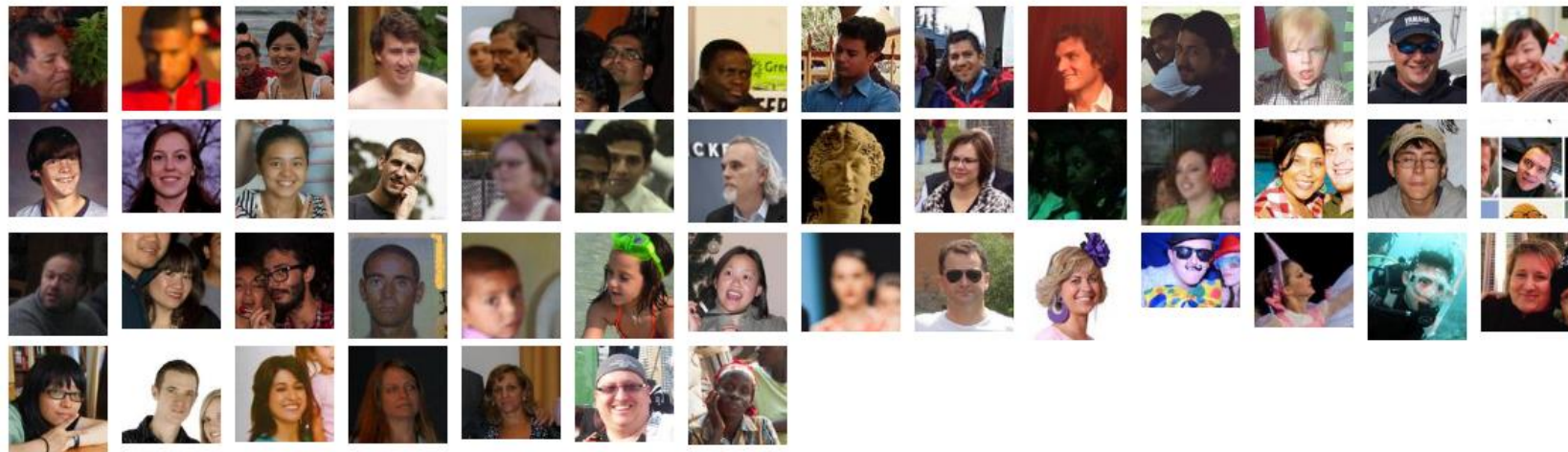
Challenge 1: Train on any dataset, test your method with 1 million distractors

Participate and download Challenge 1



Challenge 2: Training on 672K identities (4.7 Million photos), test at Million scale

Participate and download Challenge 2



Обучение 4,7 млн фото 672 тыс. персон(challenge 2)

Самая большая база для теста.

Два протокола

Face scurb и FGNet

БД для обучения и теста

MegaFace

Два протокола
FaceScurb и FGNet

Test Sets

FaceScrub Celebrities

FGNet Age-invariant non-celebrities

FaceScurb содержит 530 знаменитостей с несколькими фото и 1 миллион фото произвольных людей.
FGNet не знаменитости разновозрастные
1 миллион фото произвольных людей.

БД для обучения и теста

MegaFace FaceScurb

Identification Rate vs. Distractors Size

Algorithm	Date Submitted	Set 1	Set 2	Set 3	Data Set Size
Sogou AIGROUP - SFace	9/5/2018	99.939%	99.939%	99.939%	Large
SRC-Beijing-FR(Samsung Research Institute China-Beijing)	8/15/2018	99.888%	99.888%	99.888%	Large
SenseTime PureFace(clean)	6/13/2018	99.801%	99.801%	99.801%	Large
EI Networks	8/10/2018	99.414%	99.414%	99.414%	Large
ICARE_FACE_V1	9/13/2018	99.3198%	99.3198%	99.3198%	Large
Sogou	6/12/2018	99.2%	99.2%	99.2%	Large
Uniview Technology	8/12/2018	99.196%			Large
QINIU ATLAB - FaceX V1 (iBUG cleaned data)	7/23/2018	99.132%	99.132%	99.132%	Large
Visual Computing-Alibaba-V1(clean)	8/6/2018	99.126%			Large

БД для обучения и теста

MegaFace FaceScurb

Verification

Algorithm	Date Submitted	Set 1	Set 2	Set 3	Data Set Size
Sogou AIGROUP - SFace	9/5/2018	99.939%	99.939%	99.939%	Large
SRC-Beijing-FR(Samsung Research Institute China-Beijing)	8/15/2018	99.8%	99.8%	99.8%	Large
SenseTime PureFace(clean)	6/13/2018	99.626%	99.626%	99.626%	Large
ICARE_FACE_V1	9/13/2018	99.49%	99.49%	99.49%	Large
BingMMLab V1(iBUG cleaned data)	4/10/2018	99.487%	99.487%	99.487%	Large
Visual Computing-Alibaba-V1(clean)	8/6/2018	99.465%			Large
EI Networks	8/10/2018	99.286%	99.286%	99.286%	Large
Uniview Technology	8/12/2018	99.196%			Large
QINIU ATLAB - FaceX V1 (iBUG cleaned data)	7/23/2018	99.155%	99.155%	99.155%	Large
iBUG_DeepInsight	2/8/2018	98.948%	98.948%	99.008%	Large
ATLAB-FACEX (QINIU CLOUD)	6/23/2018	98.69%	98.69%	98.69%	Large
cyberlink_resnet-v2	9/5/2018	98.406%	98.406%	98.406%	Large

БД для обучения и теста

MegaFace FGNet

Identification Rate vs. Distractors Size

Algorithm	Date Submitted	Set 1	Set 2	Set 3	Data Set Size
THU CV-AI Lab	12/12/2017	77.977%	77.995%	77.968%	Large
EI Networks	8/10/2018	75.974%	75.974%	75.974%	Large
Google - FaceNet v8	10/23/2015	74.594%	74.585%	74.558%	Large
QINIU ATLAB - FaceX V1 (iBUG cleaned data)	7/23/2018	72.438%	72.438%	72.438%	Large
SIATMMLAB TencentVision	12/1/2016	71.247%	71.283%	71.256%	Large
SRC-Beijing-FR(Samsung Research Institute China-Beijing)	8/15/2018	70.0%	70.0%	70.0%	Large
BingMMLab V1(iBUG cleaned data)	4/10/2018	69.109%	69.109%	69.109%	Large
BingMMLab-v1 (non-cleaned data)	4/10/2018	69.108%	69.108%	69.108%	Large
ATLAB-FACEX (QINIU CLOUD)	6/23/2018	68.59%	68.59%	68.59%	Large
TencentAILab_FaceCNN_v1	9/21/2017	67.873%	67.855%	67.918%	Large
iBUG_DeepInsight	2/8/2018	67.476%	67.512%	67.575%	Large

БД для обучения и теста

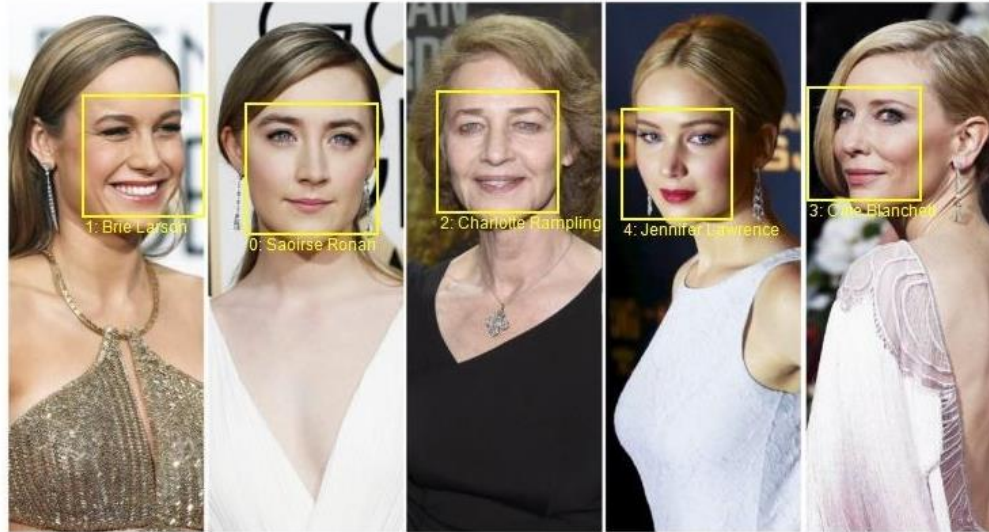
MegaFace FGNet

Verification

Algorithm	Date Submitted	Set 1	Set 2	Set 3	Data Set Size
Google - FaceNet v8	10/23/2015	75.55%	75.55%	75.55%	Large
EI Networks	8/10/2018	70.119%	70.119%	70.119%	Large
QINIU ATLAB - FaceX V1 (iBUG cleaned data)	7/23/2018	69.957%	69.957%	69.957%	Large
SRC-Beijing-FR(Samsung Research Institute China-Beijing)	8/15/2018	69.0%	69.0%	69.0%	Large
THU CV-AI Lab	12/12/2017	68.261%	68.513%	68.188%	Large
SIATMMLAB TencentVision	12/1/2016	67.954%	67.954%	67.954%	Large
FaceTag V1	12/18/2017	65.789%	65.789%	65.789%	Large
TencentAILab_FaceCNN_v1	9/21/2017	64.886%	64.724%	64.724%	Large
ULUFace	5/7/2018	64.6878%	64.6878%	64.6878%	Large
iBUG_DeepInsight	2/8/2018	64.291%	64.291%	64.291%	Large
Yang Sun	06/05/2017	63.623%	64.688%	65.247%	Large

БД для обучения и теста

Trillion Faces



Train:

MS-Celeb-1M-v1c with 86,876 ids/3,923,399

Asian-Celeb 93,979 ids/2,830,146

Test:

ELFW: Face images of celebrities in LFW name list.

There are 274k images from 5.7k ids.

DELFW: Distractors for ELFW. There are in total 1.58 million face images from Flickr.

БД для обучения и теста

Identification	Verification	1-FNIR@FPIR=10 ⁻³			
user	description	metric	status	TPR@FPR=1e-3	date
nttstar	arcface-r100-alpha	cos	Finished	0.85673	2018-11-29
CIGIT_IRSEC	p1_a31.bin	cos	Finished	0.84917	2018-11-26
CIGIT_RCIST	Public(M) R100 Cos Single-1	cos	Finished	0.84043	2018-10-20
IrvingShu	r100-newloss-m2.0	cos	Finished	0.83990	2018-12-01
faceface	merge_190	cos	Finished	0.83790	2018-10-10
tuoyuxiang	r100-cm-tpv1ds-97	cos	Finished	0.83744	2018-10-21
wanghai	r34_r100_merge_51 2	cos	Finished	0.83390	2018-12-02
leonleon	trillion_merge_600	cos	Finished	0.83390	2018-12-03
FaceFace	test4	cos	Finished	0.83241	2018-07-26
tingelam	100-msv2-off	cos	Finished	0.83241	2018-09-28

БД для обучения и теста

Identification

Verification

1-FRR@FAR=10⁻⁹

user	description	metric	status	TPR@FPR=1e-9	date
CIGIT_IRSEC	m5_glint_48	cos	Finished	0.81887	2018-11-16
nttstar	arcface-100-gf	cos	Finished	0.81591	2018-12-05
IrvingShu	faceemore2-newloss	cos	Finished	0.81448	2018-11-06
wanghai	r34_r100_merge_51 2	cos	Finished	0.81208	2018-12-02
leonleon	trillion_merge_600	cos	Finished	0.81207	2018-12-03
CIGIT_RCIST	Public(M) R100 Cos Single-1	cos	Finished	0.81117	2018-10-20
tuoyuxiang	r100-cm-tpv1b-01	cos	Finished	0.80753	2018-10-05
tingelam	r100-2-187	cos	Finished	0.80315	2018-10-03