

## СИАМСКИЕ НЕЙРОСЕТИ ДЛЯ ЗАДАЧИ РАСПОЗНАВАНИЯ ЛИЦ

SIAMESE NEURAL NETWORKS  
FOR THE TASK OF FACE RECOGNITION

**T. Krutov**  
**G. Afanasyev**  
**A. Afanasyev**

*Summary.* The paper considers the architecture of the Siamese neural network in relation to the problem of facial recognition. The theoretical foundations of comparing photos of people using the neural network approach are described. The Siamese neural network was written and trained using transfer learning technology. The dataset used was a collected with a balanced number of images in each class. The results of the Siamese neural network are shown in the graphs of accuracy and loss.

*Keywords:* siamese neural network, face recognition, vector representation of images, transfer learning,.

**Крутов Тимофей Юрьевич**

Аспирант, кафедра, МГТУ им. Н.Э. Баумана  
(национальный исследовательский университет)  
timofeykrutov@gmail.com

**Афанасьев Геннадий Иванович**

К.т.н., доцент, МГТУ им. Н.Э. Баумана  
(национальный исследовательский университет)  
gaipcs@bmstu.ru

**Афанасьев Арсений Геннадьевич**

Ассистент, МГТУ им. Н.Э. Баумана (национальный  
исследовательский университет)  
wolfram\_zn@mail.ru

*Аннотация.* В работе рассматривается архитектура сиамской нейронной сети применительно к задаче распознавания лиц. Описаны теоретические основы сравнения фотографий людей с помощью нейросетевого подхода. Программно реализована и обучена сиамская нейронная сеть с применением технологии transfer learning — использования предобученной сети. В качестве данных использовался собранный набор со сбалансированным количеством изображений в каждом классе. Результаты работы сиамской нейронной сети приводятся на графиках точности и потерь.

*Ключевые слова:* сиамская нейронная сеть, распознавание лиц, векторное представление изображений, transfer learning,.

## Введение

**В** современном мире системы компьютерного зрения на основе машинного обучения находят широкое применение [2,4], в частности, системы распознавания лиц давно занимают широкую нишу в задачах обеспечения безопасности охраняемых объектов и контроля за соблюдением общественного порядка. Системы распознавания лиц позволяют существенно сократить время прохода сотрудников через контрольно-пропускные пункты, обеспечить соблюдение режима рабочего времени, снизить расходы организации на содержание большого штата охраны или отказать от традиционных способов контроля доступа, таких как бумажные пропуска и электронные карты.

## Основная часть

Одним из способов реализации системы распознавания лиц является искусственная свёрточная нейронная сеть, для которой могут быть использованы разные методы оптимизации [1, 3].

При создании системы распознавания лиц, основанных на свёрточных нейронных сетях, возникает

необходимость выбора архитектуры свёрточной нейронной сети. В качестве архитектуры нейронной сети была использована распространённая для решения задач распознавания лиц сиамская свёрточная сеть [8, 5] рисунок 1.

Основное отличие сиамской нейронной сети от простой свёрточной нейронной сети состоит в том, что сиамская сеть имеет два входа и только один выход. Такая идея предполагает использование двух одинаковых параллельных свёрточных ветвей сети, использующих одинаковые веса. Обучение сети производится на парах изображений. На этапе предобработки данных необходимо составить из обучающего набора пары изображений и присвоить бинарные метки каждой паре. Метки присваиваются в соответствии с правилом: два изображения одного человека — положительная пара, изображения разных людей — отрицательная пара.

В процессе обучения одновременно на один вход подсети подается первое изображение пары, на другой вход — второе. Сиамская нейронная сеть учится выделять схожие изображения за счёт минимизации расстояния между фотографиями одного человека. Для вы-

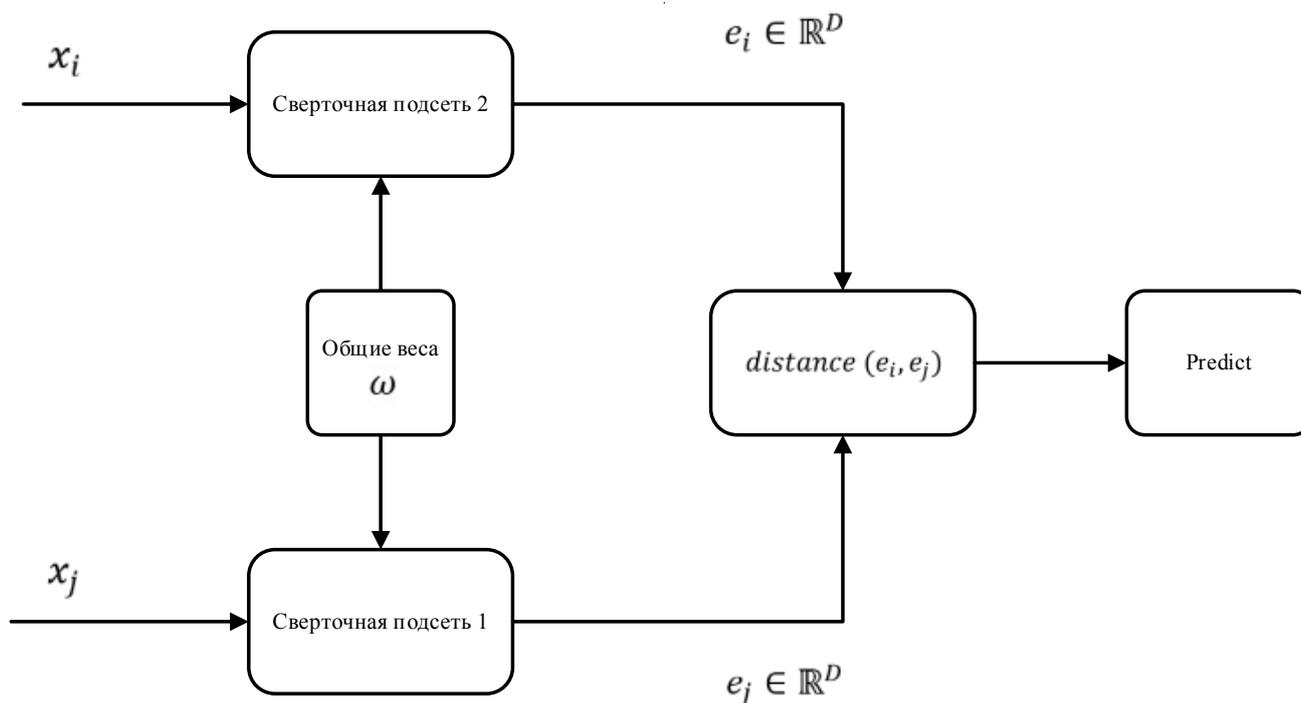


Рис. 1. Архитектура сиамской нейронной сети

деления отличий между фотографиями разных людей используется максимизации расстояния [6].

Иными словами, обучающий набор эталонных лиц людей представляется в виде точек  $(x_1^1, x_2^1, \dots, x_n^1), (x_1^2, x_2^2, \dots, x_n^2), \dots, (x_1^k, x_2^k, \dots, x_n^k)$   $n$  — мерного векторного пространства. Такое векторное представление позволяет выделить области пространства, в котором каждая новая точка будет принадлежать к определенному классу — одному конкретному человеку. Для проверки принадлежности тестовой фотографии к определенному классу используется понятие нормы. В рамках задач компьютерного зрения широко распространена  $l_2$  — норма или евклидово расстояние. Формула евклидова расстояния для двух  $n$ -мерных векторов  $x, y$ :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

После вычисления расстояния между векторами принадлежность к классу объекта будет определяться наименьшим расстоянием из вычисленных:

$$\min((d(x, y_1), d(x, y_2), \dots, d(x, y_k)))$$

Обучающий набор сиамской нейросети можно представить в виде

$$D = \{x_i, x_j, p_{i,j}\}, i, j = 1, \dots, k$$

где  $x_i$  — первое изображение в паре,  $x_j$  — второе изображение в паре,  $p_{i,j}$  — бинарная метка соответствия пар,  $k$  — размер обучающей выборки.

В качестве целевой функции используется контрастная потеря [7].

Функция контрастной потери для набора данных  $\{X_i, y_i\}, i = 1, \dots, n$  с метками  $y_i \in \{1, 2, 3, \dots, K\}$ , где  $K$  — число классов:

$$loss = 1_{y_i=y_j} \|e_i, e_j\|_2^2 + 1_{y_i \neq y_j} \cdot \max(m - \|e_i, e_j\|_2)^2$$

где  $y_i, y_j$  — метки классов первого изображения пары и второго изображения пары, соответственно;  $\|e_i, e_j\|_2$  — евклидово расстояние между выходными векторами признаков  $e_i, e_j$  изображений  $x_i, x_j$ ;  $m$  — минимально заданное расстояние между изображениями разных классов.

### Экспериментальная часть

Модель системы распознавания лиц была написана на языке Python с применением библиотек Tensorflow, Keras и библиотеки компьютерного зрения OpenCV [9]. В качестве детектора лиц используется Multi-task Cascaded Convolutional Network (MTCNN) [10].

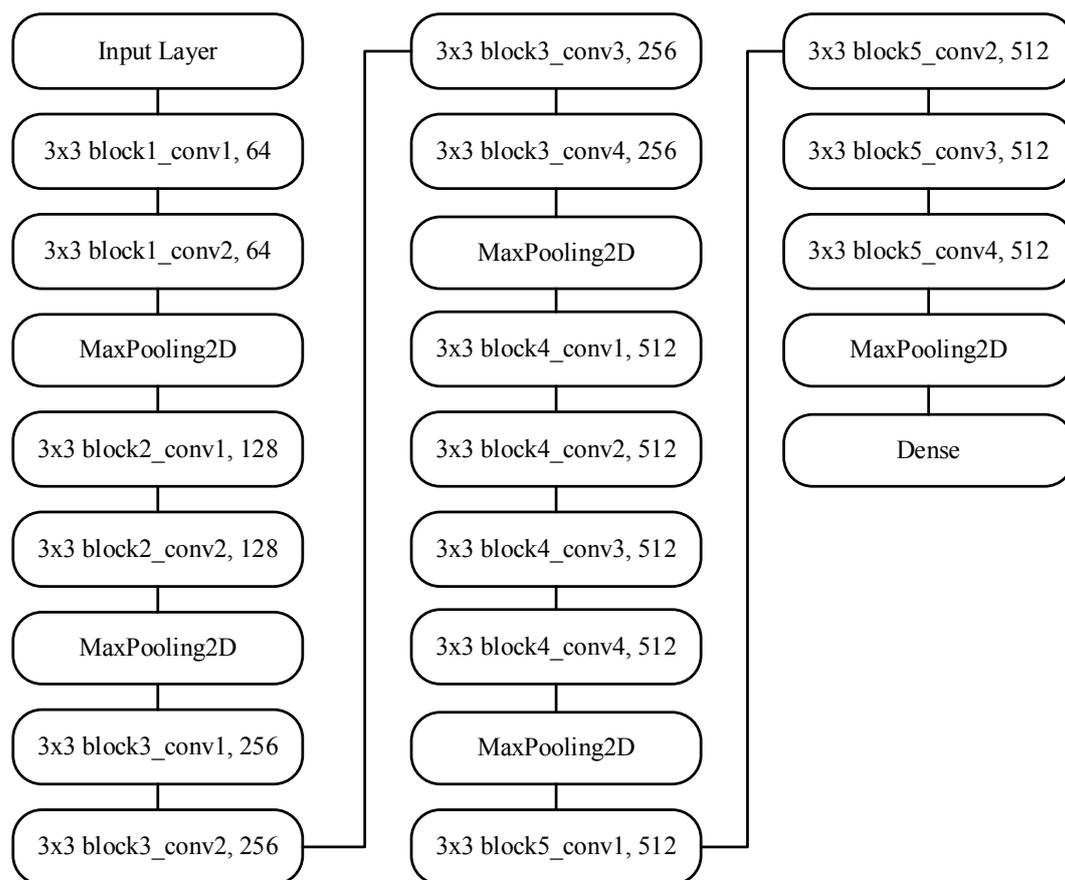


Рис. 2. Слои подсети сиамской нейронной сети

Архитектура сиамской нейронной сети была реализована с помощью предобученной на изображениях ImageNet свёрточной сети VGG19. Схема слоев подсети сиамской нейронной сети приведена на рисунке 2. Для получения векторов признаков был добавлен полносвязный слой Dense. Выходы подсетей были объединены в слое Lambda с помощью вычисления  $L2$  — нормы выходных признаков.

Обучение нейронной сети производилось на CPU без использования графического ускорителя.

Гиперпараметры нейронной сети, исследуемые в работе:

- ◆ количество нейронов выходного слоя Dense подсети
- ◆ размер мини батча

Для исследуемых гиперпараметров нейронной сети были обучены модели с использованием оптимизатора Adam, со скоростью обучения 0,001.

Для аугментированных данных количество изображений каждого класса составляет 258 файлов.

В качестве данных был использован собранный набор фотографий людей. На приведенных ниже графиках показаны кривые обучения для 12 и 2 выходных нейронов и размера мини батча — 32. На рисунке 3 демонстрируется точность обучения модели, расположенная по оси ординат, на оси абсцисс отмечены эпохи обучения модели. График потерь на этапе обучения и валидации изображен на рисунке 4. Наилучшие показатели работы модели достигаются при значении размера мини батча — 32 и количества нейронов выходного полносвязного слоя — 4. Показатели обучения приведены на рисунках 5, 6 для точности и потерь соответственно.

## ВЫВОДЫ

По результатам экспериментальной части можно сказать, что архитектура сиамской нейронной сети может успешно применяться для построения систем распознавания лиц. В ходе эксперимента было получена точность модели 83%. При увеличении размеров выборки, количества классов изображений и дополнительной оптимизации гиперпараметров модели следует ожидать улучшения точности модели.

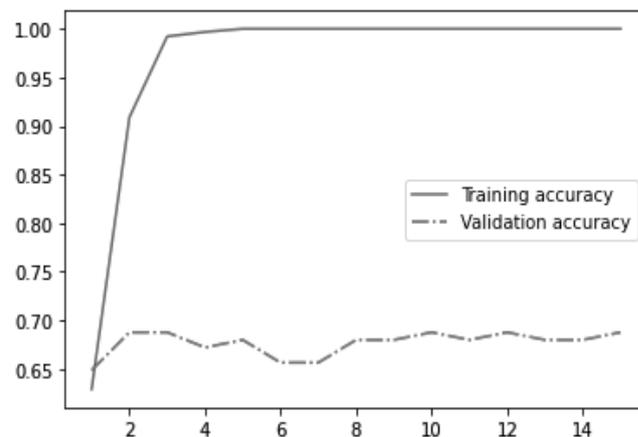
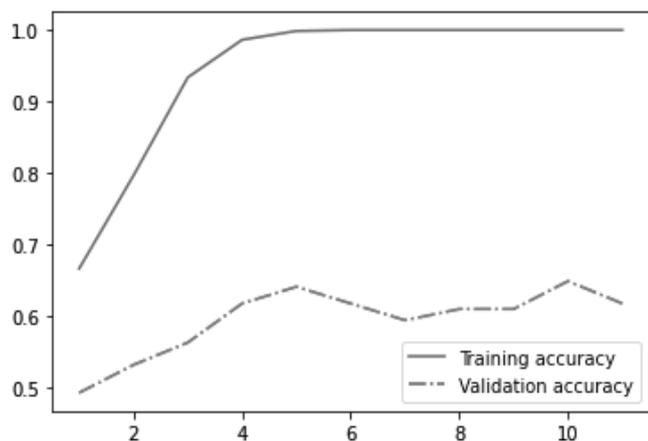


Рис. 3. Точность модели на обучающих и валидационных данных для сети с 12 нейронами полносвязного слоя — слева, для сети с 2 нейронами — справа.

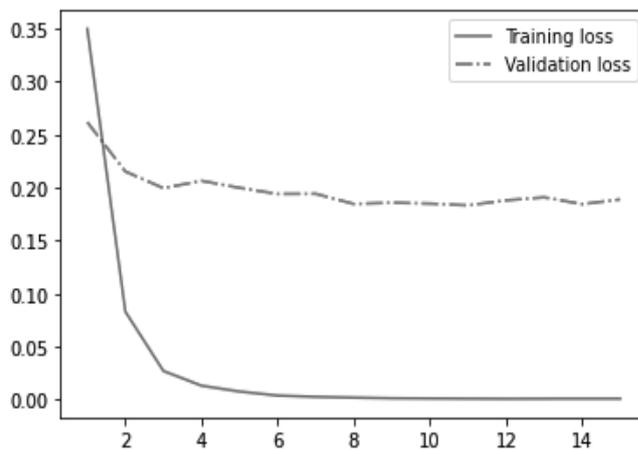
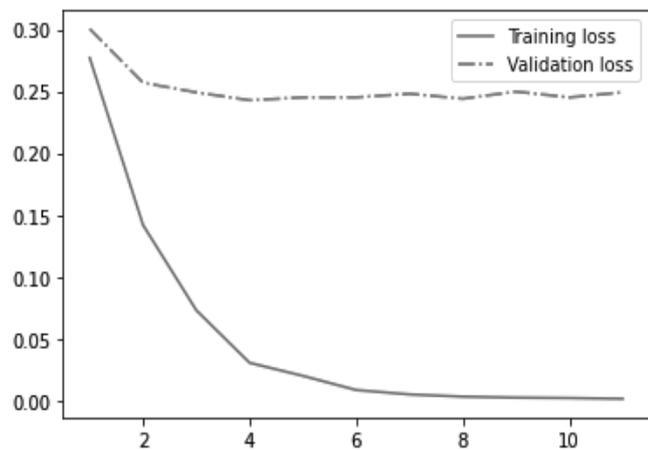


Рис. 4. Потери на обучающем и валидационном наборе для сети с 12 нейронами полносвязного слоя — слева, для сети с 2 нейронами — справа.

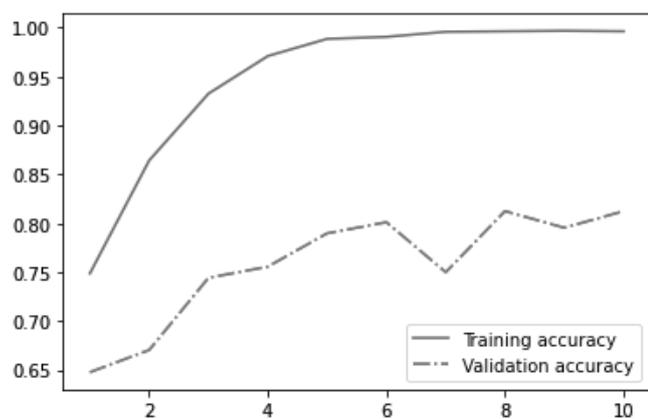


Рис. 5. Точность модели на обучающих и валидационных данных

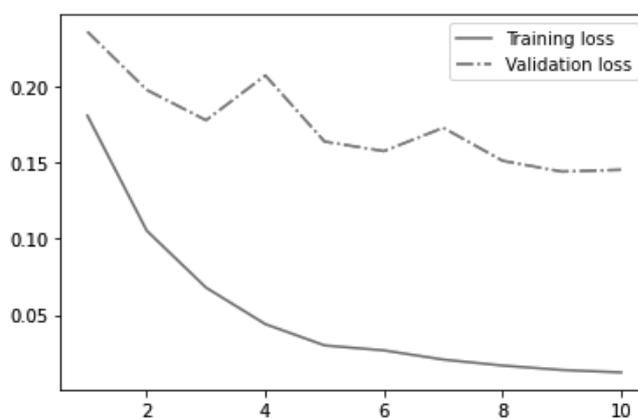


Рис. 6. Потери на обучающем и валидационном наборе

ЛИТЕРАТУРА

1. Афанасьев Г.И., Абулкасимов М.М., Сурикова О.В. Алгоритмы оптимизации, используемые в нейронных сетях, и градиентный спуск // Аспирант и соискатель. 2019. № 6 (114). С. 81–86.
2. Галичий Д.А., Афанасьев Г.И., Нестеров Ю.Г. Распознавание эмоций человека при помощи современных методов глубокого обучения // E-Scio. 2021. № 5 (56). С. 316–329.
3. Крутов Т.Ю., Афанасьев Г.И., Нестеров Ю.Г. Оптимизационные методы нейронных сетей для решения задачи бинарной классификации изображений // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2021. № 5–2. С. 68–76.
4. Подход к разработке модели цифрового двойника водителя высокоавтоматизированного транспортного средства на основе гибридной интеллектуальной информационной системы / Аладин Д.В., Горячкин Б.С., Варламов О.О., Афанасьев Г.И., Тимофеев В.Б. // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. 2022. № 2–2. С. 34–40.
5. Франсуа Шолле. Глубокое обучение на Python // СПб.: Питер. 2019. — 400 с.
6. Николенко С., Кадурич А., Архангельская Е. Глубокое обучение. // СПб.: Питер. 2020. 480 с.
7. Hadsell R., Chopra S., LeCun Y. Dimensionality Reduction by Learning an Invariant Mapping [Electronic resource] // URL: <http://yann.lecun.com/exdb/publis/pdf/hadsell-chopra-lecun-06.pdf> (дата обращения 15.11.2022).
8. Koch G. Neural Networks for One-Shot Image Recognition. [Electronic resource] // URL: <http://www.cs.toronto.edu/~gkoch/files/msc-thesis.pdf> (дата обращения 15.11.2022).
9. OpenCV (Open Source Computer Vision). [Electronic resource] // Documentation page. URL: <https://docs.opencv.org/master/index.html> (дата обращения 15.11.2021).
10. Zhang K., Zhang Z., Li Z., Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks. [Electronic resource] // URL: <https://arxiv.org/ftp/arxiv/papers/1604/1604.02878.pdf> (дата обращения 15.11.2022).

---

© Крутов Тимофей Юрьевич (timofeykrutov@gmail.com),

Афанасьев Геннадий Иванович (gaipcs@bmstu.ru), Афанасьев Арсений Геннадьевич (wolfram\_zn@mail.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»



Московский государственный технический университет имени Н.Э. Баумана