

УДК 004.855.5

DOI: 10.18413/2518-1092-2024-9-1-0-9

Мартон Н.А.¹
Жихарев А.Г.²
Черных В.С.¹

**ПРИМЕНЕНИЕ СОВРЕМЕННЫХ ТЕХНОЛОГИЙ СБОРА ДАННЫХ
И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ
ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ**

¹Белгородский государственный национальный исследовательский университет,
ул. Победы, 85, Белгород, 308015, Россия

²Белгородский государственный технологический университет им. В.Г. Шухова,
ул. Костюкова, 46, Белгород, 308012, Россия

e-mail: zhikharev@bsu.edu.ru

Аннотация

В работе представлен комплексный подход к разработке и оптимизации модели нейронной сети для эффективного распознавания лиц в динамичной среде. В частности, рассматривается разработанная авторами нейронная сеть, целью которой является распознавание ограниченного круга лиц. Для обучения нейронной сети была сформирована обучающая выборка.

В работе описывается процесс обучения нейронной сети с использованием библиотеки Keras, включая архитектуру сети, размеры слоев, функции активации и методы оптимизации. Также обсуждаются этапы предварительной обработки и подготовки исходных данных для обучения нейронной сети.

Полученные результаты исследования показывают, что разработанная нейронная сеть обладает высокими производительностью и точностью.

Ключевые слова: нейронная сеть; обучение нейронной сети; модель; слой; эмбединг; распознавание лиц

Для цитирования: Мартон Н.А., Жихарев А.Г., Черных В.С. Применение современных технологий сбора данных и методов машинного обучения для распознавания лиц // Научный результат. Информационные технологии. – Т.9, №1, 2024. – С. 79-87. DOI: 10.18413/2518-1092-2024-9-1-0-9

Marton N.A.¹
Zhikharev A.G.²
Chernykh V.S.¹

**APPLICATION OF MODERN DATA COLLECTION
TECHNOLOGIES AND MACHINE LEARNING METHODS FOR
FACE RECOGNITION**

¹Belgorod State National Research University,
85 Pobedy Str., Belgorod, 308015, Russia

²Belgorod State Technological University named after V.G. Shukhov,
46 Kostyukova Str., Belgorod, 308012, Russia

e-mail: zhikharev@bsu.edu.ru

Abstract

The article presents a comprehensive approach to the development and optimization of a neural network model for effective face recognition in a dynamic environment. In particular, the neural network developed by the authors is considered, the purpose of which is to recognize a limited number of people. A training sample was formed to train the neural network.

The paper describes the process of training a neural network using the Keras library, including the network architecture, layer sizes, activation functions and optimization methods. The stages of preprocessing and preparation of initial data for training a neural network are also discussed.

The obtained research results show that the developed neural network has high performance and accuracy.

Keywords: neural network; neural network training; model; layer; embedding; face recognition

For citation: Marton N.A., Zhikharev A.G., Chernykh V.S. Application of modern data collection technologies and machine learning methods for face recognition // Research result. Information technologies. – Т.9, №1, 2024. – P. 79-87. DOI: 10.18413/2518-1092-2024-9-1-0-9

ВВЕДЕНИЕ

В современном информационном обществе, на фоне стремительно развивающихся технологий, распознавание лиц становится неотъемлемой частью самых различных областей, от повседневного использования до обеспечения безопасности [1]. В данной работе будет рассмотрен процесс создания и оптимизации модели нейронной сети для эффективного распознавания лиц, работу которой мы проверили на обширном наборе данных. Помимо этого, будет рассмотрен процесс формирования обучающей выборки, а также описаны шаги предварительной обработки и подготовки исходных данных.

НАЧАЛЬНЫЕ УСЛОВИЯ

Для каждого пользователя, которого будет необходимо распознавать модели нейронной сети, необходимо подготовить набор данных, основанный на различных фото, с разными выражениями и положениями лица. Это необходимо для того, чтобы показать модели как можно больше вариантов одного и того же лица, чтобы распознать человека было возможно даже во время активного движения головой или речи.

В качестве обучающих данных были выбраны фотографии 9 человек. Для каждого человека была создана некоторая базовая коллекция изображений, которые затем подверглись аугментации, включая клонирование и деформацию. Этот процесс искусственно увеличил объем данных и их разнообразие, что поспособствовало более эффективному обучению модели. Для эксперимента для разных людей было взято разное количество фото. Это было сделано с целью проверить, насколько отличается точность определения при различном количестве данных. Так же, для увеличения точности распознавания, на некоторых фото были использованы различные аксессуары/помехи, вроде очков или причёски.

В ходе экспериментов было выявлено, что помимо данных девяти человек, так же было необходимо определить класс “неизвестный”, который будет отвечать за неизвестных людей. Данный класс должен включать как можно больше разнообразных видов персон, разных полов, рас, возрастов и выражений лиц. Всё это необходимо для того, чтобы составить максимально расплывчатый класс, который может подойти любому человеку, не входящему в число девяти распознаваемых классов. В рамках данной работы, для класса “неизвестный” было выбрано 47 изображений людей различных полов, рас, возрастов и т.д.

Каждый набор изображений был обработан, в результате чего для каждого класса были получены различные наборы эмбеддингов, записанных в формате “.пру”, каждый из которых хранится в соответствующих папках.

ОБРАБОТКА ДАННЫХ

Получение данных для обучения нейронной сети и распознавания происходит следующим образом:

1. Получение изображения, содержащее лицо – на данном этапе над обрабатываемым фото ведётся работа по выявлению на нём лиц. Обработка статичных фото и видео реализуется с помощью функций библиотеки OpenCV, а также каскадов Хаара, импортируемых с помощью класса CascadeClassifier [3]. Функция данного класса detectMultiScale используется для поиска лица на фото, после чего найденное изображения лица записывается в специальный файл, с которым мы будем работать в следующих пунктах.

2. Получение эмбеддинга лица – на данном этапе из фото, полученного в пункте 1, получают специальные метрики лица, которые в дальнейшем будут использоваться для

распознавания и обучения модели. Получение эмбедингов происходят благодаря функции `face_encodings`, библиотеки `face_recognition`. Сам эмбединг представляет из себя массив, размером 128 элементов, в котором каждый элемент является значением от -1 до 1 [4].

3. Сохранение полученных данных: на последнем этапе обработки данные, с помощью функции `pr.save` сохраняются в файл в формате “.пру”. Далее, при необходимости, их можно будет извлечь из файлом данного формата с помощью функции `pr.load`.

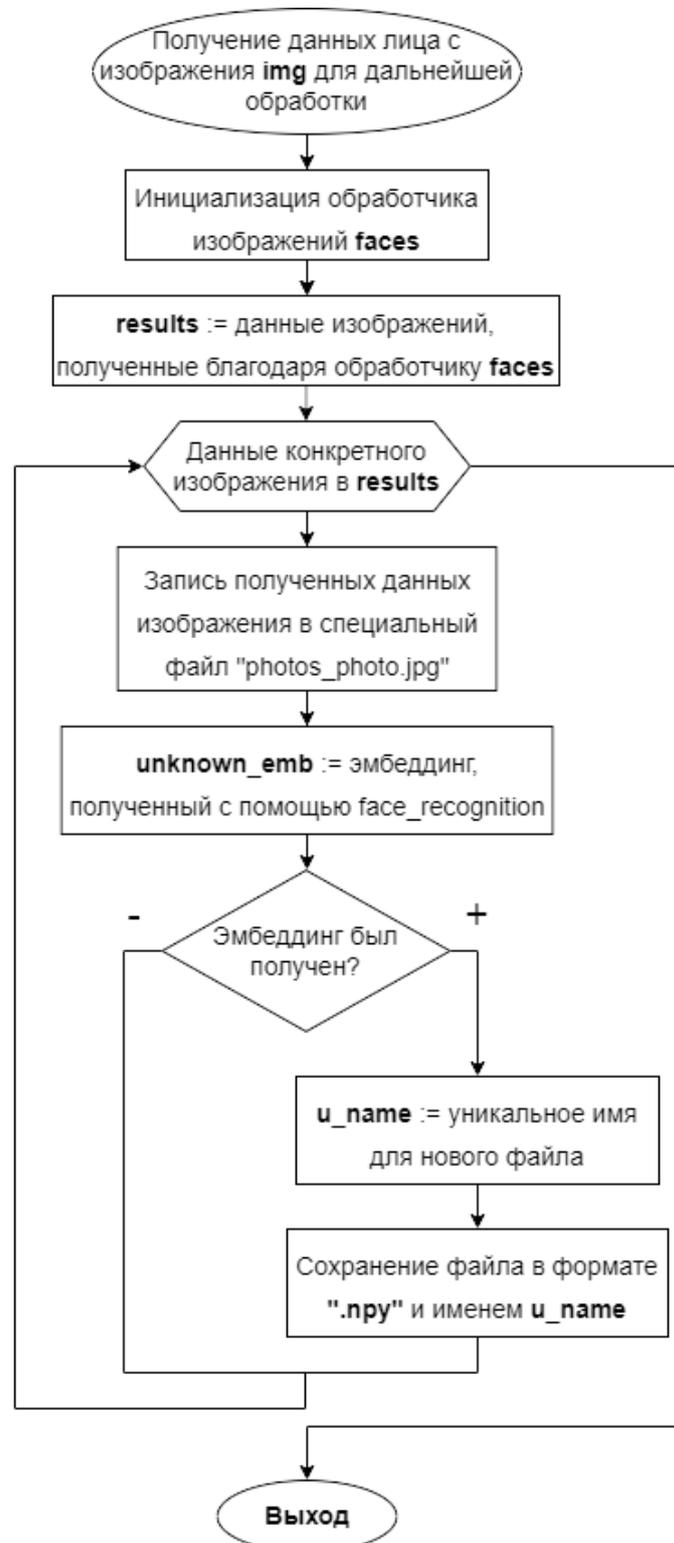


Рис. 1. Получение данных лица с изображения img
Fig. 1. Getting face data from img image

МОДЕЛЬ НЕЙРОННОЙ СЕТИ

В данной модели нейронной сети были использованы следующие слои:

BatchNormalization – данный слой нормализует входные данные путём стандартизации их по среднему значению и дисперсии. Это помогает улучшить стабильность и скорость обучения модели.

Dropout – данный слой используется для предотвращения переобучения модели путем случайного обнуления некоторых элементов выходных данных во время обучения. Это помогает улучшить обобщающую способность модели и предотвратить переобучение. На вход данному слою передаётся доля элементов, которые будет необходимо обнулить.

Dense – данные полносвязные слои применяются для классификации. Они содержат определённое количество нейронов, которые соединены со всеми нейронами предыдущего слоя. [2] На вход в данном слое подаётся количество нейронов данного слоя, функция активации нейронов, а также регулятор веса слоя, который отвечает за добавление штрафа к функции потерь, что так же способствует предотвращению переобучения модели.

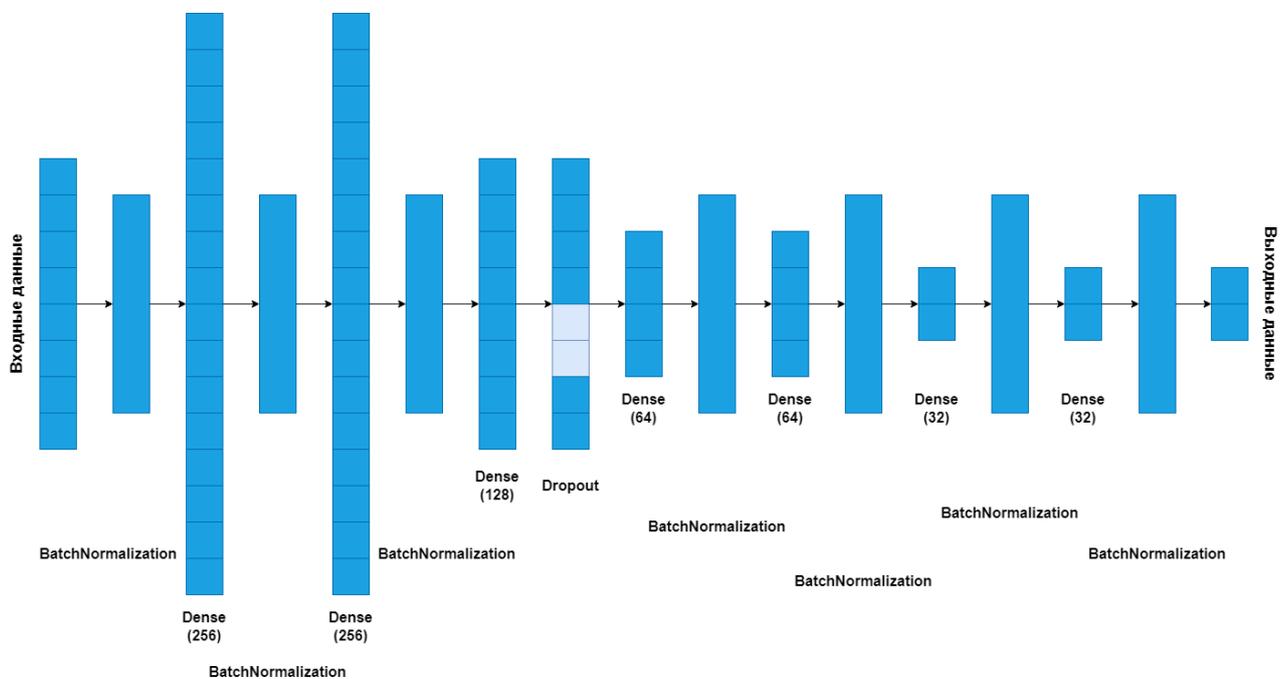


Рис. 2. Структура модели нейронной сети
Fig. 2. Structure of the neural network model

Первым слоем является BatchNormalization.

Вторым слоем является Dense с 256 нейронами, указанием формы входных данных для этого слоя в виде первого элемента в обучающем наборе данных, функцией активации tanh и L2-регулятором веса слоя со значением 0.0001. Далее во всех слоях Dense будут повторяться значения функции активации и L2-регулятора.

Третьим слоем является BatchNormalization.

Четвёртым слоем так же является Dense с 256 нейронами.

Пятым слоем является BatchNormalization.

Шестым слоем является Dense с 128 нейронами.

Седьмым слоем является Dropout со значением 0.2.

Восьмым слоем является Dense с 64 нейронами.

Девятым слоем является BatchNormalization.

Десятым слоем является Dense с 64 нейронами.

Одиннадцатым слоем является BatchNormalization.

Двенадцатым слоем является Dense с 32 нейронами.

Тринадцатым слоем является BatchNormalization.

Четырнадцатым слоем является Dense с 32 нейронами.

Пятнадцатым слоем является BatchNormalization.

Выходным слоем является Dense с количеством нейронов, равным количеству распознаваемых классов, а также функцией активации softmax.

Для разработанной модели используется функция потерь categorical_crossentropy, т.к. с помощью данной модели мы планируем решать задачу многоклассовой классификации. В то же время оптимизатор adam выбран, так как он эффективно обновляет веса модели в процессе обучения, а также обеспечивает быструю сходимость.

В ходе экспериментов во время работы над проектом были испробованы различные варианты архитектуры модели нейронной сети для распознавания лиц, и именно описанный выше вариант показал наиболее перспективные результаты, как на обучающих и тестовых данных, так и во время тестирования обученной модели на неизвестных ранее фотографиях и видео.

Так как данная модель нейронной сети предназначена для распознавания лиц людей, стоит реализовать возможность добавления новых классов к уже существующей модели, а также удаление старых. Данную возможность необходимо обеспечить, так как в организациях, в которых данная модель может быть использована, часто может меняться состав людей, которых модели необходимо опознавать. К примеру, в рамках использования внутри высшего учебного заведения, студенты добавляются и уходят в больших количествах каждый год, а значит обеспечить безопасное добавление и удаление новых классов для распознавания просто необходимо.

Добавление нового класса для распознавания реализовано следующим образом:

1. Копируем веса и смещения последнего слоя – на данном этапе происходит копирование весов и смещения последнего слоя модели нейронной сети, которые мы будем использовать в дальнейшем.

2. Удаление последнего слоя – на данном этапе мы удаляем ранее скопированный последний слой из модели нейронной сети.

3. Создание нового слоя – на этом этапе мы создаём новый слой модели нейронной сети, но на один нейрон больше скопированного последнего слоя. Размещение нового нейрона зависит от условий решаемой задачи.

4. Копируем веса прошлого последнего слоя обратно – на данном этапе происходит копирование весов предыдущего последнего слоя в новый.

5. Инициализация веса для нового класса – на данном этапе инициализируем веса нового класса, для чего используем средний вес.

6. Тренировка новой модели – на этом этапе происходит тренировка изменённой модели.

Удаление класса, в целом, похоже на добавление, и происходит следующим образом:

1. Копируем веса и смещения последнего слоя – на данном этапе происходит копирование весов и смещения последнего слоя модели нейронной сети, которые мы будем использовать в дальнейшем.

2. Удаление последнего слоя – на данном этапе мы удаляем ранее скопированный последний слой из модели нейронной сети.

3. Создание нового слоя – на этом этапе мы создаём новый слой модели нейронной сети, но размером на один нейрон меньше скопированного последнего слоя.

4. Копируем веса прошлого последнего слоя обратно – на данном этапе происходит копирование весов предыдущего последнего слоя в новый, исключая веса и смещения того класса, который нам необходимо удалить.

5. Тренировка новой модели – на этом этапе происходит тренировка изменённой модели.

РЕЗУЛЬТАТЫ РАБОТЫ

Исходные данные были разделены на обучающие и тестовые. Для обучения было выделено 70% исходных данных, в то время как для тестов использовались оставшиеся 30%. Само распределение данных по классам выглядело следующим образом:



Рис. 3. Распределение данных между классами
Fig. 3. Distribution of data between classes

В результате обучения спроектированной модели нейронной сети мы получили высокие показатели точности распознавания. При анализе графиков точности и потерь на этапах обучения и проверки, можно заметить, что ещё на начальных эпохах точность модели достигает высоких показателей, а потери стремятся к нулю, что свидетельствует об успешном обучении и классификации, как на обучающих, так и новых данных. Так же было выяснено, что лучшие показатели были достигнуты на 278-й эпохе обучения. Точность модели достигла до **94.83%**, в то время как потери на данном этапе составили **0.1944**.

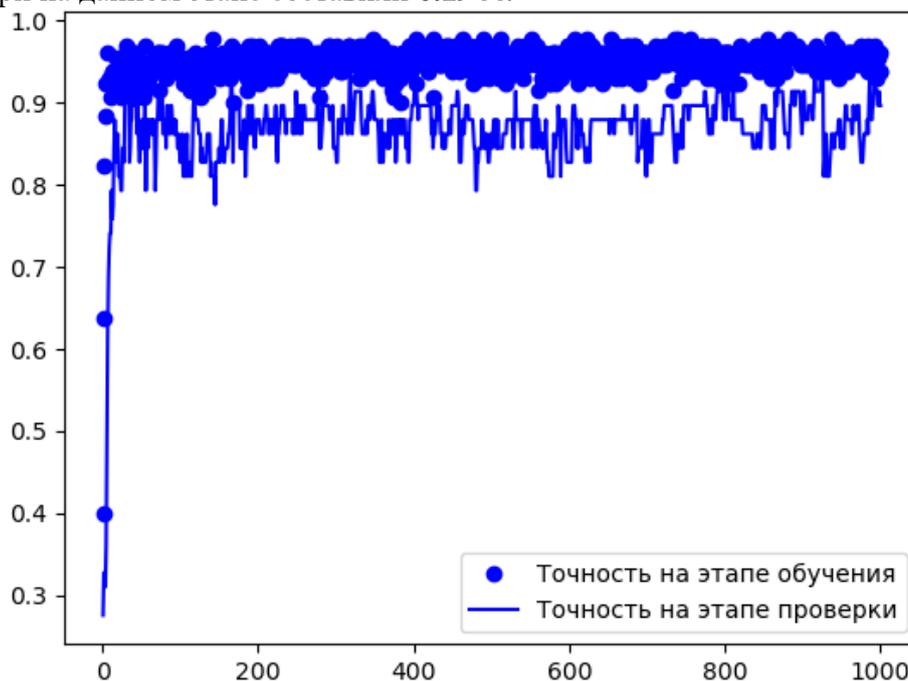


Рис. 4. Точность на этапах обучения и проверки
Fig. 4. Accuracy at the stages of training and verification

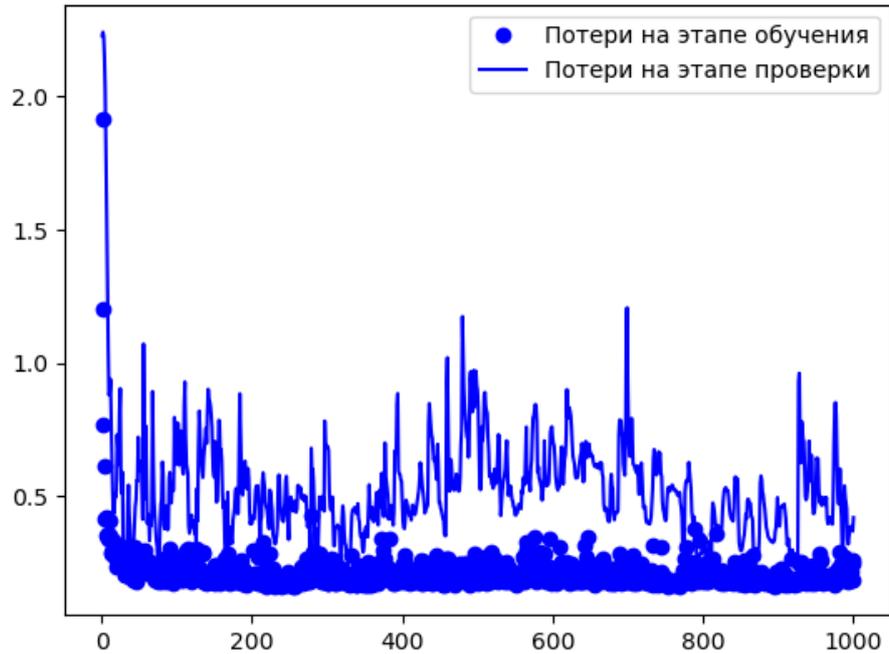


Рис. 5. Потери на этапах обучения и проверки
Fig. 5. Losses at the stages of training and verification

После обучения модель была протестирована с фотографиями, не входящими ни в набор обучающих, ни в набор тестовых данных. Так же провелось несколько тестов с использованием видекамеры. В тестировании принимали участие, как персоны, распознавать которых модель была обучена, так и ранее неизвестные ей лица.

Таблица 1

Тестирование обученной модели

Table 1

Testing the trained model

	Тест фото 1	Тест фото 2	Тест фото 3	Тест камера 1	Тест камеры 2
Известный 1	99.94%	99.92%	98.37%	99.97%	99.54%
Известный 2	99.99%	99.98%	94.03%	99.77%	99.98%
Неизвестный 1	74.59%	96.66%	89.24%	73.77%	83.87%
Неизвестный 2	46.38%	92.4%	93.01%	81.95%	61.17%



Рис. 6. Примеры фото обучающего набора Мужчины 8
Fig. 6. Examples of photos of the training set Men 8



Рис. 7. Успешно распознанное фото Мужчины 8, не входящее в обучающий набор
Fig. 7. Successfully recognized photo of Man 8, not included in the training set

В результате тестирования можно прийти к выводам, что получившаяся модель отлично распознаёт тех людей, которых была обучена распознавать, а также в большей степени хорошо распознаёт, если человек ей неизвестен.

ЗАКЛЮЧЕНИЕ

Благодаря библиотекам Keras и face-recognition нами была разработана модель нейронной сети, показавшая высокую точность распознавания и довольно низкие потери. Графики точности и потерь подтверждают успешность обучения модели, которая достигла наилучших показателей на 278-й эпохе обучения с точностью в 94.83% и потерями в 0.1944.

Дальнейшее улучшение получившейся модели можно реализовать благодаря увеличению размера обучающего набора данных, особенно отмечая увеличение класса “неизвестный”, что улучшит большую обобщающую способность модели.

Список литературы

1. Шолле Ф. Глубокое обучение на Python. 2-е межд. издание / Ф. Шолле; – СПб.: Питер, 2023. – 576 с. – ISBN 978-5-4461-1909-7
2. Жихарев А.Г., Черных В.С. Классификация речевых данных по эмоциональному фону // Научный результат. Информационные технологии. – Т.8, №3, 2023. – С. 34-44. DOI: 10.18413/2518-1092-2022-8-3-0-5
3. Визуализация каскадов Хаара. [Электронный ресурс] – Электрон, дан., 2020. – URL: <https://habr.com/ru/articles/504288/>
4. Техническая документация библиотеки face-recognition. [Электронный ресурс] — Электрон, дан., 2020. – URL: <https://libraries.io/pypi/face-recognition>

References

1. Chollet F. Deep Learning with Python. 2nd interd. edition / F. Chollet; – Spb.: St. Petersburg, —2023. – 576 p. – ISBN 978-5-4461-1909-7
2. Zhikharev A.G., Chernykh V.S. Classification of speech data by emotional background // Research result. Information technologies. – Т.8, №3, 2023. – P. 34-44 DOI: 10.18413/2518-1092-2022-8-3-0-5
3. Visualization of Haar cascades. [Electronic resource] – Electronic data, 2020. – URL: <https://habr.com/ru/articles/504288/>
4. Technical documentation of the face-recognition library. [Electronic resource] – Electronic data, 2020. – URL: <https://libraries.io/pypi/face-recognition>

Мартон Никита Андреевич, студент 1 курса магистратуры, кафедры прикладной информатики и информационных технологий, институт инженерных и цифровых технологий

Жихарев Александр Геннадиевич, доктор технических наук, доцент, доцент кафедры программного обеспечения вычислительной техники и автоматизированных систем

Черных Владимир Сергеевич, студент 1 курса магистратуры кафедры информационных и робототехнических систем, институт инженерных и цифровых технологий

Marton Nikita Andreevich, 1st year Master's student, Department of Applied Informatics and Information Technologies, Institute of Engineering and Digital Technologies

Zhikharev Alexander Gennadievich, Doctor of Technical Sciences, Associate Professor, Associate Professor of the Department of Software for Computer Engineering and Automated Systems

Chernykh Vladimir Sergeevich, 1st year Master's student, Department of Information and Robotic Systems, Institute of Engineering and Digital Technologies