

### УСОВЕРШЕНСТВОВАНИЕ АЛГОРИТМА ИДЕНТИФИКАЦИИ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА С ИСПОЛЬЗОВАНИЕМ MFCC

В. В. Семенюк<sup>1\*</sup>, М. В. Складчиков<sup>2</sup>

<sup>1</sup> Южно-Российский государственный политехнический университет (НПИ) им. М. И. Платова,  
Новочеркасск, Россия

\* [semenuk.viktoriya@gmail.com](mailto:semenuk.viktoriya@gmail.com)

<sup>2</sup> Донецкий национальный технический университет, Донецк, Россия

**Аннотация.** Представлен подход к реализации алгоритма эмоционального состояния человека с использованием сверточных нейронных сетей. На основе общей концепции научного исследования рассмотрен вариант усложнения иерархии идентифицируемых эмоций. Проведен сравнительный анализ применения оконного преобразования Фурье и алгоритма MFCC в качестве инструмента для обработки информационных данных. Вариант усложнения предлагаемого метода рассматривается как логический переход от более простого математического аппарата, представленного в виде оконного преобразования Фурье, к мелкестральнойным коэффициентам. Благодаря данному подходу сформирован более информативный входной набор данных без усложнения архитектуры нейронной сети, скорректирована методология научного исследования и с использованием идеализированной базы данных достигнута точность идентификации, близкая к 100 %. Приведено обоснование применения Deep Network Designer как инструмента для создания архитектуры нейронной сети.

**Ключевые слова:** нейронная сеть, распознавание эмоций человека, сверточная нейронная сеть, дактилоскопия звука

**Ссылка для цитирования:** Семенюк В. В., Складчиков М. В. Усовершенствование алгоритма идентификации эмоционального состояния человека с использованием MFCC // Изв. вузов. Приборостроение. 2024. Т. 67, № 9. С. 731–740. DOI: 10.17586/0021-3454-2024-67-9-731-740.

### IMPROVEMENT OF THE HUMAN EMOTIONAL STATE IDENTIFICATION ALGORITHM USING MFCC

V. V. Semenuk<sup>1\*</sup>, M. V. Skladchikov<sup>2</sup>

<sup>1</sup> Platov South-Russian State Polytechnic University, Novocherkassk, Russia

\* [semenuk.viktoriya@gmail.com](mailto:semenuk.viktoriya@gmail.com)

<sup>2</sup> Donetsk National Technical University, Donetsk, Russia

**Abstract.** An approach to the implementation of an algorithm for the emotional state of a person using convolutional neural networks is presented. Based on the general concept of scientific research, a variant of complicating the hierarchy of identifiable emotions is considered. A comparative analysis of the application of the windowed Fourier transform and the MFCC algorithm as a tool for processing information data is carried out. The variant of complication of the proposed method is considered as a logical transition from a simpler mathematical apparatus, presented in the form of a windowed Fourier transform to the use of mel-frequency cepstral coefficients. This allowed to form a more informative input data set without complicating the neural network architecture, the methodology of scientific research was adjusted and, using an idealized database, the accuracy of identification close to 100% was achieved. The rationale for using Deep Network Designer as a tool for creating neural network architecture is given.

**Keywords:** neural network, human emotion recognition, convolutional neural network, sound fingerprinting

**For citation:** Semenuk V. V., Skladchikov M. V. Improvement of the human emotional state identification algorithm using MFCC. *Journal of Instrument Engineering*. 2024. Vol. 67, N 9. P. 731–740 (in Russian). DOI: 10.17586/0021-3454-2024-67-9-731-740.

**Введение.** В современном информационном обществе при интенсивном развитии технологий создание систем, способных распознавать эмоциональные состояния человека и адаптироваться к ним, имеет немаловажное значение. Эмоции играют ключевую роль в человеческом поведении, и точное определение эмоций может значительно повысить эффективность систем компьютерного взаимодействия, автоматизации и даже управления роботами.

Предлагаемый в настоящей статье подход обладает расширенным потенциалом по значительному улучшению производительности существующих систем идентификации эмоций, предоставляя более точные и надежные результаты. На основании достижений в области глубокого обучения и компьютерного зрения была поставлена задача создания инновационного инструмента, способного эффективно взаимодействовать с человеком и определять его эмоциональное состояние.

Известные системы идентификации эмоционального состояния, включающие различные методы и техники, такие как анализ речи, распознавание лиц, обработка биометрических данных и использование сенсоров для измерения физиологических параметров, применяются в разных областях, включая компьютерное зрение, медицину, робототехнику, маркетинг и др. Однако точность распознавания эмоционального состояния может варьироваться в зависимости от различных факторов, таких как качество входных данных, используемые алгоритмы и методы обучения.

**Краткий обзор существующих систем.** Ниже представлен обзор некоторых известных систем с указанием приблизительной точности распознавания ими объектов [1].

1. Microsoft Emotion API — использует анализ изображений лиц для распознавания эмоций. Система может обнаруживать следующие эмоции: счастье, грусть, удивление, гнев, отвращение и страх. Точность распознавания варьируется в зависимости от условий съемки, но обычно оценивается в 60–70%.

2. Affectiva — разработаны различные системы для распознавания эмоций на основе изображений лиц и анализа речи. Используемая технология демонстрирует точность распознавания от 70 до 90 % в зависимости от конкретного сценария и качества данных.

3. Emotient (приобретена Apple) — также специализируется на анализе эмоций на основе изображений лиц. Точность распознавания оценивается в диапазоне от 65 до 90 % в зависимости от контекста использования и качества данных.

4. IBM Watson Tone Analyzer — использует анализ текста для выявления тональности и эмоций в текстовых данных. Точность распознавания может быть высокой, особенно при анализе структурированных текстов.

5. Распознавание речи и голосовые ассистенты — голосовые ассистенты, такие как Siri, Alexa и Google Assistant, также могут распознавать эмоции в речи пользователей. Точность распознавания может варьироваться, но, как правило, достаточно хорошо распознаются основные эмоции: радость, грусть, гнев и т. д.

Общепризнанной единой метрики для оценки точности систем не существует. Кроме того, точность систем может существенно изменяться в зависимости от контекста применения и специфики данных, на которых они обучались.

**Точность идентификации эмоций.** Точность идентификации эмоций человека зависит от множества факторов, которые могут варьироваться в зависимости от используемого метода и технологии [2–12]. Перечислим некоторые из ключевых факторов, влияющих на точность идентификации эмоций.

*Качество входных данных:* качество изображений лиц, записей речи или других источников данных напрямую влияет на точность идентификации эмоций. Недостаточное освещение, низкое разрешение или шум в данных могут снизить точность распознавания.

*Разнообразие эмоций:* некоторые системы могут быть ориентированы на распознавание ограниченного набора базовых эмоций, таких как радость, грусть, гнев и т. д. Более сложные эмоции или их комбинации труднее распознавать и классифицировать.

*Культурные и индивидуальные различия:* эмоциональные выражения могут варьироваться в зависимости от культурных норм и индивидуальных особенностей. Некоторые системы могут иметь ограничения по распознаванию эмоций у людей разных культурных фонов или с разными эмоциональными выражениями.

*Обучение модели и алгоритмы:* качество обучающих данных и выбранные алгоритмы машинного обучения имеют решающее значение для точности системы. Недостаточное количество данных или неподходящие алгоритмы могут привести к низкой точности распознавания.

*Контекст и ситуация:* эмоции человека могут изменяться в зависимости от контекста и ситуации. Некоторые системы могут быть более или менее эффективными в распознавании эмоций в различных ситуациях, таких как разговор по телефону, встреча лицом к лицу или просмотр видеоконтента.

*Внешние факторы:* шум, артефакты данных, изменения освещения и другие внешние факторы могут оказывать влияние на точность идентификации эмоций.

Идентификация эмоций человека по голосу изучается в области эмоциональной речи и звука, которая анализирует физиологические изменения, происходящие в голосовых свойствах при различных эмоциональных состояниях. Рассмотрим несколько ключевых аспектов физиологии, влияющих на идентификацию эмоций по голосу.

1. Интонация и тембр голоса — эмоциональные состояния могут изменять интонацию и тембр голоса. Например, голос может стать более высоким или низким, более напряженным или расслабленным в зависимости от эмоции.

2. Скорость и ритм речи — эмоции также могут влиять на скорость и ритм речи. Например, человек может говорить быстрее и более ритмично при возбуждении или более медленно и плавно в спокойном состоянии.

3. Частотные характеристики — эмоциональное состояние может изменить частотные характеристики голоса, такие как основная и высшие частоты. Например, голос может стать более высоким или более низким в зависимости от эмоций.

4. Интонационные и темпоральные параметры — параметры, такие как интонационные контуры и длительность фраз, также могут отражать эмоциональное состояние. Например, увеличение или уменьшение интонационного разнообразия или изменение длительности пауз могут свидетельствовать о различных эмоциях.

5. Физиологические маркеры — физиологические изменения, такие как изменение частоты сердечных сокращений или дыхательного ритма, могут также влиять на голосовые характеристики и помогать идентифицировать эмоциональные состояния по голосу.

В настоящей работе для повышения качества распознавания эмоций человека было решено перейти от преобразования Фурье к иному варианту преобразования входных данных — алгоритму MFCC (Mel-Frequency Cepstral Coefficients — мел-частотные кепстральные коэффициенты). MFCC — это метод акустического анализа звука, используемый при обработке речи и распознавании речевых команд [13]. Он основан на преобразовании звукового сигнала в частотно-временное представление, которое затем анализируется с целью извлечения характеристик, содержащих информацию о звуковых особенностях, связанных с произношением.

Метод MFCC сначала преобразует звуковой сигнал в спектрограмму, представляющую собой распределение мощности звука по частотам и времени. Затем используется мел\*-шкала, которая моделирует восприятие человеком частоты звука, для разделения спектрограммы на равномерные по мел-шкале частотные полосы. После этого применяется дискретное косинусное преобразование, чтобы извлечь наиболее значимые характеристики из полученных мел-частотных кепстральных коэффициентов.

\* Мел — единица высоты звука, основанная на восприятии этого звука органами слуха.

Метод MFCC является мощным инструментом для извлечения характеристик звуков речи, таких как форманты и другие акустические особенности, которые могут быть использованы для распознавания речи, идентификации спикеров и других задач обработки речи и аудиоаналитики. Этот метод широко применяется в сферах распознавания речи, музыкального анализа, биомедицинской диагностики и других областях, где необходимо анализировать и классифицировать звуковые данные.

В предыдущем варианте исследования [13] была реализована концепция распознавания эмоций для трех классов — положительная, отрицательная, нейтральная. Изначальная картина исследования строилась на подготовке изображений как результат применения оконного преобразования Фурье к аудиоданным: см. рис. 1, *а* — дактилоскопия звука, *б* — спектрограмма.

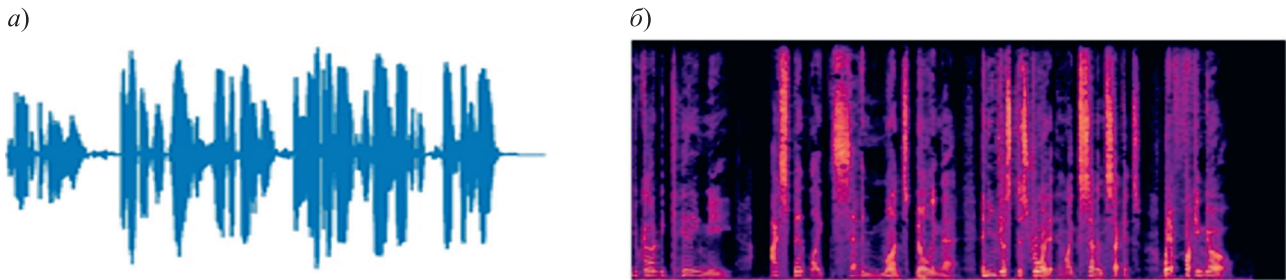


Рис. 1

Обучение было выполнено на основании RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song) — базы данных, разработанной в Университете Райерсон (Торонто, Канада) в целях предоставления исследователям доступа к аудиовизуальным данным, связанным с эмоциональными выражениями в речи. Эта база данных была создана для поддержки исследований в области распознавания эмоций, а также машинного обучения и обработки сигналов.

#### Основные характеристики RAVDESS:

— многообразие эмоций — в базе данных представлены различные эмоциональные состояния, включая грусть, радость, страх, злость, удивление и отвращение; эти эмоции записаны в речи, чтобы предоставить богатый набор данных для анализа;

— множество голосов — датасет включает записи голоса, выражающего эмоции, что позволяет исследователям анализировать аудиоданные для распознавания эмоций;

— высокое качество данных — аудиоданные имеют высокое качество и разрешение, что обеспечивает точные и надежные результаты анализа;

— сбалансированная выборка — датасет содержит сбалансированный набор данных по эмоциональным категориям, что позволяет избежать смещения в сторону определенных эмоций и обеспечивает надежные результаты исследований;

— доступность и открытость — RAVDESS является открытым ресурсом, доступным для исследователей и разработчиков по всему миру, что способствует совместной работе и обмену знаниями в области распознавания эмоций и машинного обучения.

В целях упрощения алгоритма для распознавания был выбран более простой вариант, а именно три эмоции. По представленным на рис. 2 результатам исследования установлено, что подобный подход не позволяет добиться высокой точности идентификации ( $I$ ) эмоций.

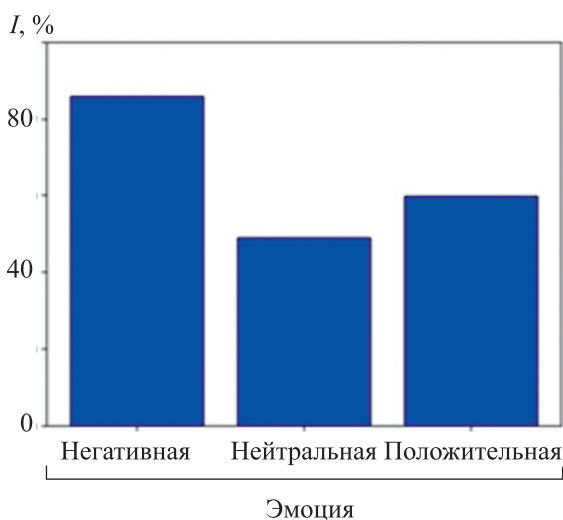


Рис. 2

Исходя из результатов, приведенных на рис. 2, было решено рассмотреть иную концепцию получения изображений для формирования базы данных. Был проведен сравнительный анализ алгоритма MFCC и оконного преобразования Фурье (STFT — Short-Time Fourier Transform) в задачах распознавания звука. Рассмотрим ключевые аспекты сравнения.

1. *Частотное разрешение:*

— MFCC — используется мел-шкала, более соответствующая восприятию звука человеческим ухом и имеющая более высокое разрешение на низких частотах и более низкое разрешение на высоких частотах;

— STFT — обеспечивается равномерное частотное разрешение по всему спектру.

2. *Информативность признаков:*

— MFCC — извлекаются характеристики, сосредоточенные на акустически важных частотах, что делает этот метод более информативным для распознавания речи и других аудиосигналов;

— STFT — предоставляется информация о спектральных характеристиках звука, но метод менее эффективен в распознавании речи вследствие меньшей специфичности для человеческого восприятия.

3. *Временное разрешение:*

— MFCC — используется временное среднее по времени для извлечения характеристик, что может снижать точность признаков на коротких временных интервалах. В данном исследовании при обучении нейронной сети основным критерием для принятия решения была мощность сигнала. Эту характеристику можно выделить из спектрограммы, которую и получают с помощью данной технологии. Однако чтобы сгладить сигнал, используется временное усреднение, а это приводит к потере деталей о быстрых изменениях в звуковом сигнале. Это снижает точность признаков на коротких временных интервалах, поскольку важная информация может быть утрачена, что усложняет анализ кратковременных звуковых событий;

— STFT — обеспечивается высокое временное разрешение и точное отражение динамических изменений в сигнале на коротких временных интервалах. Это математический метод, позволяющий анализировать частотные компоненты сигнала, изменяющегося во времени. В отличие от обычного преобразования Фурье, которое показывает только частотный состав сигнала в целом, STFT позволяет получить информацию о том, как частоты изменяются с течением времени.

4. *Размерность признаков:*

— MFCC — метод имеет, как правило, более низкую размерность признаков, что может сократить вычислительную сложность алгоритмов машинного обучения и повысить обобщающую способность модели;

— STFT — может генерировать больше признаков из-за увеличенного количества коэффициентов Фурье, что требует большего объема вычислительных ресурсов.

В целом, как было также установлено экспериментально, метод MFCC обычно предпочтительнее в задачах распознавания речи и других аудиосигналов благодаря своей специфичности для человеческого восприятия и более низкой размерности признаков. Однако STFT также может быть полезно в некоторых ситуациях, особенно в задачах, где требуется высокое временное разрешение или спектральная информация о звуке.

В результате был разработан алгоритм с использованием языка программирования MatLab, позволяющий получить мел-спектрограммы, см. рис. 3: *a* — аудиосигнал, *b* — спектрограмма MFCC.

Следующим этапом исследования была разработка архитектуры нейронной сети для обучения. Вследствие преимуществ пакета MatLab было решено разработать модель сверточной нейронной сети с использованием специализированного дополнения. Deep Network Designer (DND) — это мощный инструмент MatLab для создания, настройки и обучения нейронных сетей, позволяющий быстро прототипировать и анализировать различные архитектуры нейронных сетей без необходимости написания кода.

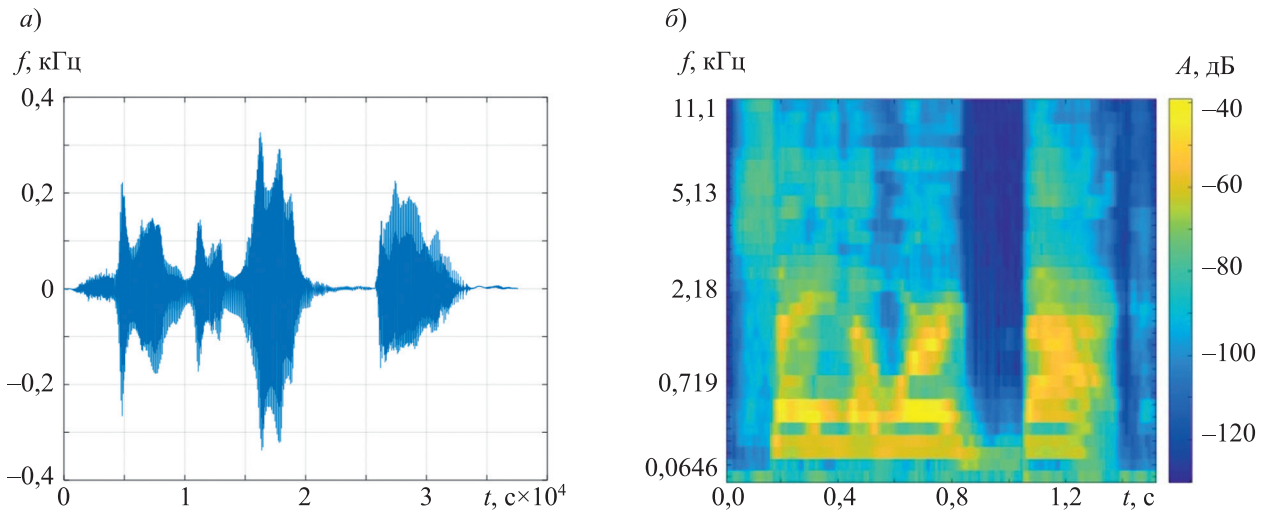


Рис. 3

**Алгоритм идентификации.** Благодаря описанным выше параметрам выбора, было решено усовершенствовать алгоритм идентификации эмоций, для чего следовало перестроить архитектуру нейронной сети, а также применить MFCC-преобразование для создания входного набора данных. Под усложнением алгоритма также понимается использование более расширенного варианта входного набора данных, позволяющего идентифицировать уже 7 эмоций. Новая архитектура нейронной сети приведена на рис. 4 (скриншот).

На рис. 5 представлена матрица ошибок и несколько других метрик, характеризующих результаты работы этой нейронной сети, обученной для классификации эмоций.

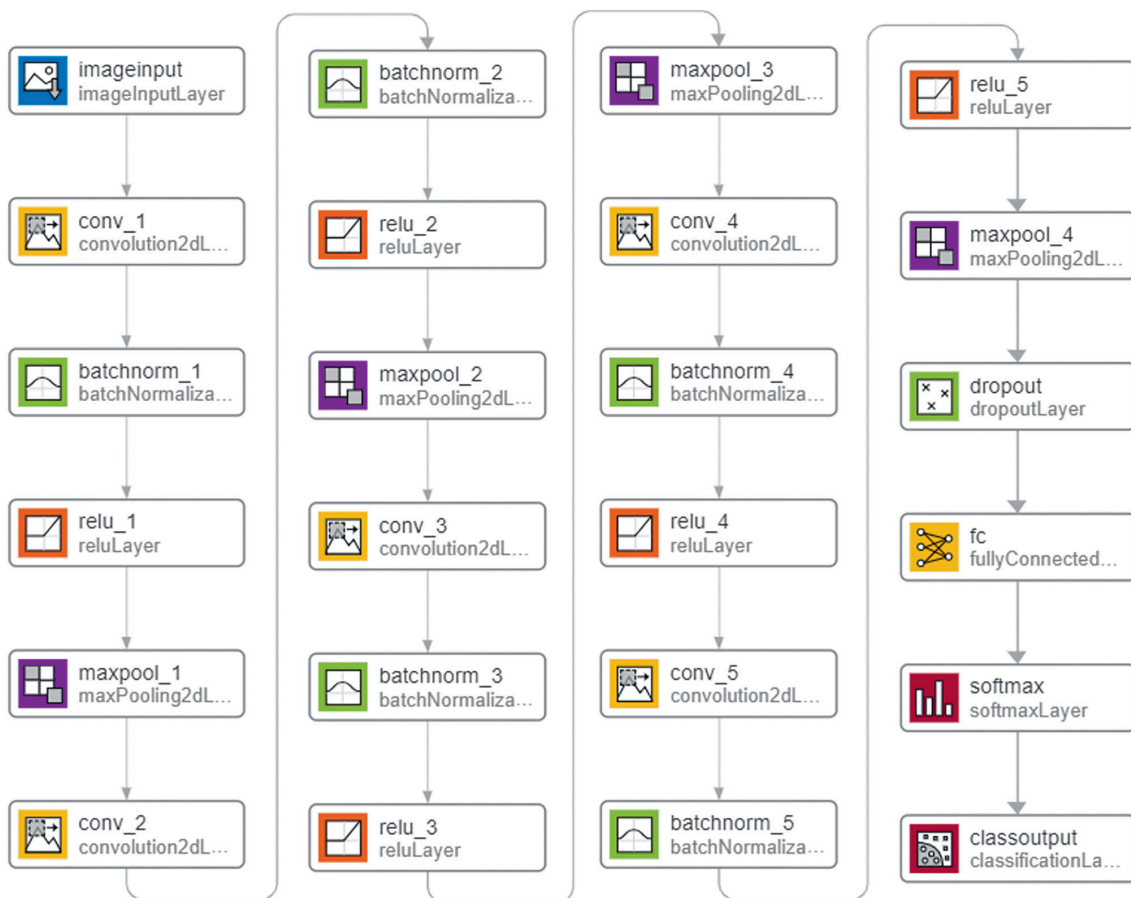


Рис. 4

Точность классификации на наборе данных, не участвовавших в обучении нейронной сети

96,3 %	3,7 %
96,3 %	3,7 %
90,0 %	10,0 %
95,0 %	5,0 %
93,8 %	6,3 %
98,8 %	1,2 %
88,8 %	11,3 %

Матрица ошибок

Агрессия	77	2		1			
Отвращение		77		1	1		1
Страх	1	1	72	1			5
Счастье		2		76			2
Нейтральное состояние			2		75		3
Печаль		1				79	
Удивление	2	7					71

Точность классификации на наборе данных, участвовавших в обучении нейронной сети

96,3 %	85,6 %	97,3 %	96,2 %	98,7 %	95,2 %	91,0 %
3,7 %	14,4 %	2,7 %	3,8 %	1,3 %	4,8 %	9,0 %

Агрессия Отвращение Страх Счастье Нейтральное состояние Печаль Удивление

Рис. 5

Матрица ошибок позволяет оценить, насколько успешно модель распознает каждую из эмоций и где она допускает ошибки.

Общее понимание матрицы ошибок: на диагонали матрицы приведены значения, показывающие, сколько раз моделью правильно предсказана каждая эмоция: например, эмоция „агрессия“ была правильно классифицирована 77 раз, „отвращение“ также 77 раз и т. д. Остальные значения вне диагонали представляют собой количество ошибочных классификаций: например, эмоция „удивление“ неправильно классифицирована как „отвращение“ 7 раз, а „нейтральное состояние“ — как „печаль“ 3 раза.

Точность классификации на наборе данных, участвовавшем в обучении нейронной сети:

— агрессия: этот класс моделью предсказан с точностью 96,3 %;

— отвращение: точность 85,6 %;

— страх: высокая точность — 97,3 %;

— счастье: точность 96,2 %;

— нейтральное состояние: эмоция предсказана с точностью 98,7 %;

— печаль: точность 95,2 %;

— удивление: точность 91,0 %.

Точность классификации на наборе данных, не участвовавшем в обучении нейронной сети:

— агрессия: моделью правильно классифицировано 96,3 % случаев, но 3,7 % случаев были ошибочно отнесены к другим классам;

— отвращение: правильная классификация составила 96,3 % при 3,7 % ошибок;

— страх: точность 90,0 %, но в 10,0 % случаев ошибки;

— счастье: правильное распознавание в 95,0 % случаев, 5,0 % — ошибки;

— нейтральное состояние: 93,8 % случаев распознаны верно, 6,3 % ошибок;

— печаль: очень высокая точность — 98,8 %, ошибки составили лишь 1,2 %;

— удивление: сложности в распознавании, так как точность составила 88,8 %, а ошибок 11,3 %.

Нейронная сеть продемонстрировала хорошие результаты в распознавании эмоций, особенно для таких классов, как „печаль“ и „нейтральное состояние“, где точность была наиболее высокой. Однако сеть испытывает сложности в распознавании эмоций „удивление“ и „отвращение“, что проявляется в большем числе ошибочных классификаций и сравнительно низкой точности для этих классов. В целом сеть обладает высокой точностью и может быть эффективной в задачах классификации эмоций, но дальнейшие улучшения могут потребоваться для снижения ошибок в специфических случаях.

**Выводы.** По результатам разработки усовершенствованного алгоритма идентификации эмоций человека можно сделать следующие выводы.

1. Разработанный алгоритм на основе сверточной нейронной сети демонстрирует высокую эффективность в задаче идентификации эмоций человека. Полученные результаты подтверждают способность модели точно распознавать и классифицировать эмоциональные состояния на основе входных данных.

2. Проведенный сравнительный анализ алгоритма с другими методами и алгоритмами показывает преимущество разработанного подхода по эффективности и точности распознавания эмоций. Сверточные нейронные сети оказываются более адаптированными к анализу изображений лиц и способны извлекать более информативные признаки для идентификации эмоций.

3. Алгоритм представляет собой перспективное направление в области распознавания эмоций человека и может быть успешно применен в различных сферах, таких как компьютерное зрение, робототехника, медицина, психология и др.

4. Полученные результаты имеют практическую значимость и могут быть использованы для создания автоматизированных систем, способных анализировать эмоциональное состояние человека на основе визуальных данных. Это может быть полезно, например, при разработке систем управления эмоциональными интерфейсами, систем безопасности или в медицинской

диагностике. Практические результаты представленного исследования были внедрены в производство ООО „Русбирсофт“ (Биробиджан, Россия).

В ходе исследования были выявлены определенные ограничения для улучшения алгоритма. Дальнейшие исследования могут быть направлены на усовершенствование архитектуры нейронной сети, оптимизацию процесса обучения и расширение набора данных для обеспечения более высокой точности и обобщающей способности модели.

Таким образом, разработанный алгоритм представляет собой значимый вклад в область распознавания эмоций человека, обеспечивая высокую точность и перспективы для будущих исследований и практического применения.

#### СПИСОК ЛИТЕРАТУРЫ

1. Топ-5 сфер применения систем распознавания объектов [Электронный ресурс]: <https://habr.com/ru/company/toshibarus/blog/433544/>, 10.03.2022.
2. Li X., Chen Y., Hu J., Zhang J., Zhang Z. Recognizing emotion in human-computer interaction-a survey / IEEE Transactions on Affective Computing. 2016. N 7 (2). P. 149–166.
3. Шиллер А. В. Выражение модернизируемых эффектов эмоций у искусственных агентов как визуальный язык // Праксема. Проблемы визуальной семиотики. 2019. № 4 (22). С. 223–243.
4. Рюмина Е. В., Карнов А. А. Аналитический обзор методов распознавания эмоций по выражениям лица человека // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20, № 2. С. 163–176.
5. Xu H., Liang X., Sun M., Chen S., Feng J. Fusion of Visual and Affective Features for Emotion Recognition in the Wild // IEEE Transactions on Affective Computing. 2018. N 9 (2). P. 261–274.
6. Zhao M., Adelhardt J., Kummert F. Real-time emotion recognition from facial images using deep neural networks // Proc. of the ACM on Intern. Conf. on Multimodal Interaction. 2018. P. 440–447.
7. Shan H., Kan M., Wang S., Yan S. Deep learning for emotion recognition: A survey // IEEE Transactions on Affective Computing. 2017. N 8 (4). P. 512–527.
8. Хрусталёв В. О., Зубков А. В. Разработка технологии распознавания эмоций человека на основе изображений, поступающих с веб-камеры // XXIV региональная конференция молодых ученых и исследователей Волгоградской области. 2020. С. 223–224 [Электронный ресурс]: <https://elibrary.ru/item.asp?ysclid=m0425zwq1q320826398&id=42400458>.
9. Курицкий В. Ю., Садов С. В. Нейросетевой алгоритм распознавания эмоций по изображению лица // Компьютерные технологии и анализ данных (СТДА'2020). 2020. С. 245–248.
10. Хнюнин М. В., Ганелина Н. Д. Перспективы применения личных мобильных устройств для распознавания эмоций человека по фотографиям лица // Интеллектуальный потенциал Сибири. 2021. С. 179–182.
11. Артемьева Ж. Г., Крушная Н. А. Распознавание эмоций у других людей детьми с нарушениями опорно-двигательного аппарата // Человеческий фактор: социальный психолог. 2020. № 1 (39). С. 288–294.
12. Abadi M. K., Zeng Z. Emotion recognition from speech using deep learning // Proc. of the Interspeech. 2017. P. 2362–2366.
13. Семенюк В. В. Повышение качества компьютерного распознавания эмоционального состояния человека по голосу // Науч.-практ. конф. „Научные революции: Сущность и роль в развитии науки и техники“: Сб. докл. Челябинск: НИЦ „АЭТЕРНА“, 2021.

#### СВЕДЕНИЯ ОБ АВТОРАХ

**Виктория Валерьевна Семенюк** — аспирант; Южно-Российский государственный политехнический университет им. М. И. Платова, кафедра программного обеспечения вычислительной техники; E-mail: [semenuk.viktoriya@gmail.com](mailto:semenuk.viktoriya@gmail.com)

**Максим Владимирович Складчиков** — аспирант; Донецкий национальный технический университет, кафедра электропривода и автоматизации промышленных установок; E-mail: [maxsklad19981@yandex.ru](mailto:maxsklad19981@yandex.ru)

Поступила в редакцию 02.04.2021; одобрена после рецензирования 19.04.2024; принята к публикации 23.07.2024.

## REFERENCES

1. <https://habr.com/ru/company/toshibarus/blog/433544/>. (in Russ.)
2. Li X., Chen Y., Hu J., Zhang J., & Zhang Z. *IEEE Transactions on Affective Computing*, 2016, no. 2(7), pp. 149–166.
3. Shiller A.V. *Praxema. Journal of Visual Semiotics*, 2019, no. 4(22), pp. 223–243. (in Russ.)
4. Ryumina E.V., Karpov A.A. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, no. 2(20), pp. 163–176. (in Russ.)
5. Xu H., Liang X., Sun M., Chen S., & Feng J. *IEEE Transactions on Affective Computing*, 2018, no. 2(9), pp. 261–274.
6. Zhao M., Adelhardt J., & Kummert F. *Proceedings of the 2018 ACM on International Conference on Multimodal Interaction*, 2018, pp. 440–447.
7. Shan H., Kan M., Wang S., & Yan S. *IEEE Transactions on Affective Computing*, 2017, no. 4(8), pp. 512–527.
8. Khrustalev V.O., Zubkov A.V. *XXIV regional'naya konferentsiya molodykh uchonykh i issledovateley Volgogradskoy oblasti* (XXIV Regional Conference of Young Scientists and Researchers of the Volgograd Region), Volgograd, 2020, pp. 223–224. (in Russ.)
9. Kuritsky V.Yu., Sadvov S.V. *Komp'yuternyye tekhnologii i analiz dannykh (CTDA'2020)* (Computer Technologies and Data Analysis (CTDA'2020)), Proceedings of the II International Scientific and Practical Conference, Minsk, April 23–24, 2020, pp. 245–248. (in Russ.)
10. Khnyunin M.V., Ganelina N.D. *Intellektual'nyy potentsial Sibiri (RNSK-2021)* (Intellectual Potential of Siberia (RNSC-2021)), Collection of scientific papers of the 29th Regional Scientific Student Conference, Novosibirsk, 2021, vol. 3, pp. 179–182. (in Russ.)
11. Artemyeva Zh.G., Krushnaya N.A. *Human factor: social psychologist*, 2020, no. 1(39), pp. 288–294. (in Russ.)
12. Abadi M.K. & Zeng Z. *Proceedings of the INTERSPEECH*, 2017, pp. 2362–2366.
13. Semenyuk V.V. *Nauchnyye revolyutsii: Sushchnost' i rol' v razvitii nauki i tekhniki* (Scientific Revolutions: Essence and Role in the Development of Science and Technology), Scientific and practical conference, Chelyabinsk, 2021, 142 p. (in Russ.)

## DATA ON AUTHORS

- Victoria V. Semenuk** — Post-Graduate Student; Platov South-Russian State Polytechnic University, Department of Computer Software; E-mail: [semenuk.viktoriya@gmail.com](mailto:semenuk.viktoriya@gmail.com)
- Maxim V. Skladchikov** — Post-Graduate Student; Donetsk National Technical University, Department of Electric Drives and Automation of Industrial Installations; E-mail: [maxsklad19981@yandex.ru](mailto:maxsklad19981@yandex.ru)

Received 02.04.2021; approved after reviewing 19.04.2024; accepted for publication 23.07.2024.