

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ COMPUTER SCIENCE

doi: 10.17586/2226-1494-2025-25-6-1107-1116

УДК 004.934.2

Многомодальный подход к определению депрессии с использованием полуавтоматической разметки данных и детерминированных методов машинного обучения Алёна Николаевна Величко¹, Алексей Анатольевич Карпов²✉

^{1,2} Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация

¹ velichko.a@iias.spb.su, <https://orcid.org/0000-0002-8503-8512>

² karpov@iias.spb.su, <https://orcid.org/0000-0003-3424-652X>

Аннотация

Введение. Исследована актуальная задача автоматического определения психоэмоциональных состояний человека. Научный интерес к исследованиям автоматического многомодального определения депрессии объясняется распространением тревожно-депрессивных расстройств и повышенной, в связи с этим, нагрузкой на первичное звено здравоохранения. Специфичность задачи обусловлена ее комплексностью, недостаточным объемом или неточностью исходных данных, дисбалансом классов. Сравнительные исследования показывают, что результаты классификации с полуавтоматической разметкой данных часто выше, чем при ее полной автоматизации. **Метод.** Предложенный многомодальный подход к определению депрессии сочетает полуавтоматическую разметку данных и детерминированные методы машинного обучения с использованием нескольких наборов признаков. Для обучения моделей применен многомодальный корпус Extended Distress Analysis Interview Corpus (E-DAIC), содержащий аудиозаписи, автоматически полученные из этих аудиозаписей тексты и вычисленные из видеозаписей видеопризнаки, а также аннотации с результатами тестов Patient Health Questionnaire-8 для каждой записи. Полуавтоматическая разметка позволяет получать точные временные метки и тексты высказываний, для снижения зашумленности данных, используемых при обучении моделей. В предложенном подходе применяется несколько наборов признаков, извлеченных из трех модальностей (акустические экспертные признаки eGeMAPS и нейросетевые акустические признаки DenseNet, визуальные экспертные признаки OpenFace и текстовые признаки Word2Vec). Их комплексная обработка минимизирует влияние дисбаланса классов в данных на результат классификации. **Основные результаты.** Экспериментальные исследования с использованием преимущественно экспертных признаков (DenseNet, OpenFace, Word2Vec) и детерминированных методов классификации (Catboost), обладающих свойством интерпретируемости результатов, на корпусе E-DAIC позволили получить значения показателей, сопоставимые с современными международными исследованиями (68,0 % и 64,3 % по показателям взвешенной средней F1-меры (Weighted F1-measure) и невзвешенной средней полноты (Unweighted Average Recall) соответственно). **Обсуждение.** Применение полуавтоматического подхода к разметке данных и объединение модальностей позволило улучшить качество разметки и распознавание депрессии по сравнению с одномодальными подходами. Достигнуты более сбалансированные результаты классификации. Применение детерминированных методов классификации на основе деревьев решений позволит в дальнейшем провести анализ результатов классификации за счет интерпретируемости их решений. Для этой цели можно также задействовать другие методы интерпретации, например, SHapley Additive exPlanations и Local Interpretable Model-agnostic Explanations.

Ключевые слова

компьютерная паралингвистика, многомодальное распознавание депрессии, деструктивные явления, разметка данных, машинное обучение

Благодарности

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 25-11-00319).

Ссылка для цитирования: Величко А.Н., Карпов А.А. Многомодальный подход к определению депрессии с использованием полуавтоматической разметки данных и детерминированных методов машинного обучения // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1107–1116.
doi: 10.17586/2226-1494-2025-25-6-1107-1116

© Величко А.Н., Карпов А.А., 2025

A multimodal approach for depression detection using semi-automatic data annotation and deterministic machine learning methods

Alena N. Velichko¹, Alexey A. Karpov²✉

^{1,2} St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation

¹ velichko.a@iias.spb.su, <https://orcid.org/0000-0002-8503-8512>

² karpov@iias.spb.su✉, <https://orcid.org/0000-0003-3424-652X>

Abstract

A trending task of automatic psycho-emotional human state detection was studied in this work. A scientific interest to researches devoted to the automatic multimodal depression detection can arise out of the widespread of anxiety-depressive disorders and difficulties of their detection in primary health care. A specificity of the task was caused by its complexity, lack of data, imbalance of classes and inaccuracies in it. Comparative researches show that classification results on semi-automatic annotated data are higher than ones on automatic-annotated data. The proposed approach for depression detection combines a semi-automatic data annotation and deterministic machine learning methods with the utilization of several feature sets. To build our models, we utilized the multimodal Extended Distress Analysis Interview Corpus (E-DAIC) which consists of audio recordings, automatically extracted from these audio recordings texts and video feature sets extracted from video recordings as well as annotation including Patient Health Questionnaire (PHQ-8) scale for each recording. A semi-automatic annotation makes it possible to get the exact time stamps and speech texts to reduce the noisiness in the training data. In the proposed approach we use several feature sets, extracted from each modality (acoustic expert feature set eGeMAPS and neural acoustic feature set DenseNet, visual expert feature set OpenFace and text feature set Word2Vec). A complex processing of these features minimizes the effect of class imbalance in the data on classification results. Experimental researches with the use of mostly expert features (DenseNet, OpenFace, Word2Vec) and deterministic machine learning classification methods (Catboost) which have the property of interpretability of classification results yielded the experimental results on the E-DAIC corpus which are comparable with the existing ones in the field (68.0 % and 64.3 % for Weighted F1-measure (WF1) and Unweighted Average Recall (UAR) accordingly). The usage of a semi-automatic annotation approach and modalities fusion improved both quality of annotation and depression detection comparing to the unimodal approaches. More balanced classification results are achieved. The usage of deterministic machine learning classification methods based on decision trees allows us to provide an interpretability analysis of the classification results in the future due to their interpretability feature. Other methods of results interpretation like SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) also can be used for this purpose.

Keywords

computational paralinguistics, multimodal depression detection, destructive behavior, data annotation, machine learning

Acknowledgements

The work is carried out within the framework of the Russian Science Foundation (project No. 25-11-00319).

Fot citation: Velichko A.N., Karpov A.A. A multimodal approach for depression detection using semi-automatic data annotation and deterministic machine learning methods. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1107–1116 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1107-1116

Введение

В последние годы в области компьютерной параллингвистики наблюдается возросший интерес к исследованиям психоэмоциональных состояний человека по его речи. Также отмечена необходимость автоматического определения речевых проявлений тревожно-депрессивных расстройств заболеваний, что является следствием как повсеместного распространения данного вида расстройств, так и повышенной нагрузки на первичное звено здравоохранения [1]. Большое депрессивное расстройство сопровождается рядом симптомов, значительно снижающих качество жизнедеятельности, а при отсутствии необходимого и своевременного лечения, может стать причиной инвалидности [2]. Кроме того, депрессия может являться как самостоятельным диагнозом, так и иметь корреляцию с другими состояниями, например болезнью Альцгеймера, деменцией [3, 4], тревожностью [5] и аффективными расстройствами (например, маниакальное состояние) [6].

На данный момент задача автоматического определения депрессии является одной из актуальных задач в мировом научном сообществе. Так, например, она

была представлена на международных соревнованиях по аудиовизуальному определению эмоционального состояния человека Audio-Visual Emotional Challenge (AVEC) AVEC'19 [7]. Среди наиболее популярных и крупных корпусов, содержащих речь людей с депрессией, можно выделить Distress Analysis Interview Corpus Wizard-of-Oz (DAIC-WOZ) и Extended Distress Analysis Interview Corpus (E-DAIC) [8], AVEC'13 и AVEC'14, а также CONVERGE [9], Chinese Multimodal Depression Corpus (CDMC) [10], и другие. Ввиду специфики задачи, многие существующие корпуса являются небольшими (относительно корпусов, использующихся для других задач) и в основном не сбалансированы по классам, однако в последние годы многие группы учебных собирают собственные корпусы (например, работы [11–13]). Среди последних работ преимущественно превалируют научные работы, анализирующие отдельные модальности (аудио (например [6, 11, 13, 14]), текст [3, 4, 15–17]), или работы, в которых рассматриваются многомодальные подходы (аудио, видео, текст) [12, 17–19]. Данная тенденция, вероятно, связана с тем, что в существующих корпусах видеомодальность отсутствует или представлена не исходными данными (виде-

озаписями), а визуальными признаками, вычисленными из видеозаписей (как, например, в корпусах DAIC-WOZ и E-DAIC), что объясняется как повышенной сложностью записи качественных видео, так и намерениями сохранить анонимность участников записи.

За последние годы среди методов машинного обучения для задачи определения депрессии по текстовым данным набирают популярность большие языковые модели (Llama-2, GPT-3.5, GPT-4) [18–20]. Также актуальными являются различные варианты модели Bidirectional Encoder Representations from Transformers (BERT: BERT, LongBERT, RoBERTa, DepRoBERTa), например, в работах [20–22]. Среди машинных классификаторов используются как нейросетевые подходы (Convolutional Neural Networks (CNN), Gated Recurrent Unit (GRU) и Long Short-Term Memory (LSTM)), например, в работах [23–25], так и детерминированные методы на основе деревьев решений (XGBoost, Random Forest) [5, 12, 18] и векторов (X-vectors, Support Vector Machine (SVM), Support Vector Regression (SVR)) [4, 5, 11, 12–14, 18, 21].

Существует также ряд отечественных исследований по определению депрессии, в основном по текстовой модальности. Имеются различные текстовые корпусы, как, например, RusNeuroPsych [26], корпус профилей социальной сети ВКонтакте [27] и корпус эссе [28], на основе которых проводятся как исследования с использованием современных языковых моделей [29, 30], так и исследования информативных параметров поведения, а также текстовых маркеров депрессии [31].

Вопрос интерпретируемости результатов классификации поднимался в работе [16], где был применен метод LIME, и [5, 12], если учесть использование методов машинного обучения с высокой степенью интерпретируемости (детерминированные методы на основе деревьев решений). Стоит отметить, что критерий интерпретируемости является одним из важнейших для внедрения и практического использования систем автоматического определения психоэмоциональных состояний человека, в том числе депрессии. В большинстве перечисленных работ корпуса DAIC-WOZ и E-DAIC были использованы в том виде, в котором они представлены авторами, несмотря на ряд неточностей в аннотации этих корпусов, которые потенциально могут повлиять на качество обучения машинных классификаторов. Настоящая работа опирается на работу [25], в которой корпус E-DAIC был предобработан с применением полуавтоматической экспертной разметки.

Описание предлагаемого подхода

Для распознавания депрессии (двух классов: наличие/отсутствие) предложен многомодальный подход, представленный на рисунке. Он является логическим продолжением опубликованной работы [25], описывающей процесс полуавтоматической разметки и экспериментальные исследования подхода для аудиомодальности, и использует результаты, полученные в ней, для предобработки данных других модальностей. Разработанный подход основан на комбинировании признаков, извлеченных из трех модальностей (акустиче-

ские экспертные признаки eGeMAPS и нейросетевые акустические признаки DenseNet [32], визуальные экспертные признаки OpenFace [33] и текстовые признаки Word2Vec [34]), а также их комплексной обработке с целью минимизации влияния дисбаланса классов в данных на результат классификации.

В качестве экспериментальных данных в настоящей работе используется многомодальный корпус E-DAIC [8]. Данный корпус является расширенной версией корпуса DAIC-WOZ, и содержит 275 интервью, представленных в виде аудиозаписей, автоматически полученных из этих аудиозаписей текстов и видеопризнаков, вычисленных из видеозаписей, а также аннотации, где представлены результаты тестов PHQ-8 для каждой записи. Ввиду того, что в корпусе был обнаружен ряд неточностей, которые потенциально могут повлиять на качество распознавания депрессии, была проведена полуавтоматическая экспертная разметка данных корпуса, которая включала: использование программного инструментария Faster-Whisper¹ для автоматического распознавания англоязычной речи, ряда скриптов и программного инструментария Praat [35] для постобработки и проверки полученных орфографических транскрипций, выравнивание частоты дискретизации и количества каналов в аудиозаписях по всему корпусу, а также проверку бинарных значений в аннотации корпуса. Результатом полуавтоматической разметки является не только корректная разметка классов, но также временные метки и соответствующий им произнесенный текст. Кроме того, при последующих экспериментальных исследованиях по распознаванию депрессии на аудиоданных было выявлено, что, несмотря на понижение средних значений показателей качества классификации, был сокращен разрыв значений показателей для каждого класса за счет повышения качества распознавания депрессии [25, 36].

Видеоданные в обоих корпусах DAIC-WOZ и E-DAIC представлены в виде признаков, вычисленных из сырых видеозаписей: Visual Geometry Group (VGG), ResNet, OpenFace, мешок видео слов (Bag of Video Words, BOVW). Имеющиеся видеопризнаки, представленные в корпусе [25] в покадровом виде (2048 признаков для ResNet, 4096 — для VGG и 49 — для OpenFace с 30 кадрами/с во всех трех случаях), были прорежены (применен каждый 6-й кадр) с целью уменьшения размерности матрицы признаков и разделены с использованием скриптов, в которых использовались временные метки, полученные ранее при разметке. Таким образом, данные каждой из трех модальностей были предобработаны и представлены в виде наборов признаков согласно временным меткам, полученным ранее при разметке.

Полученный в процессе разметки произнесенный текст использован в качестве данных текстовой модальности. Он был нормализован с применением скриптов путем приведения к нижнему регистру, удаления пун-

¹ Программный инструментарий Faster-Whisper [Электронный ресурс]. Режим доступа: <https://github.com/SYSTRAN/faster-whisper>, свободный. Яз. англ. (дата обращения: 22.08.2025).

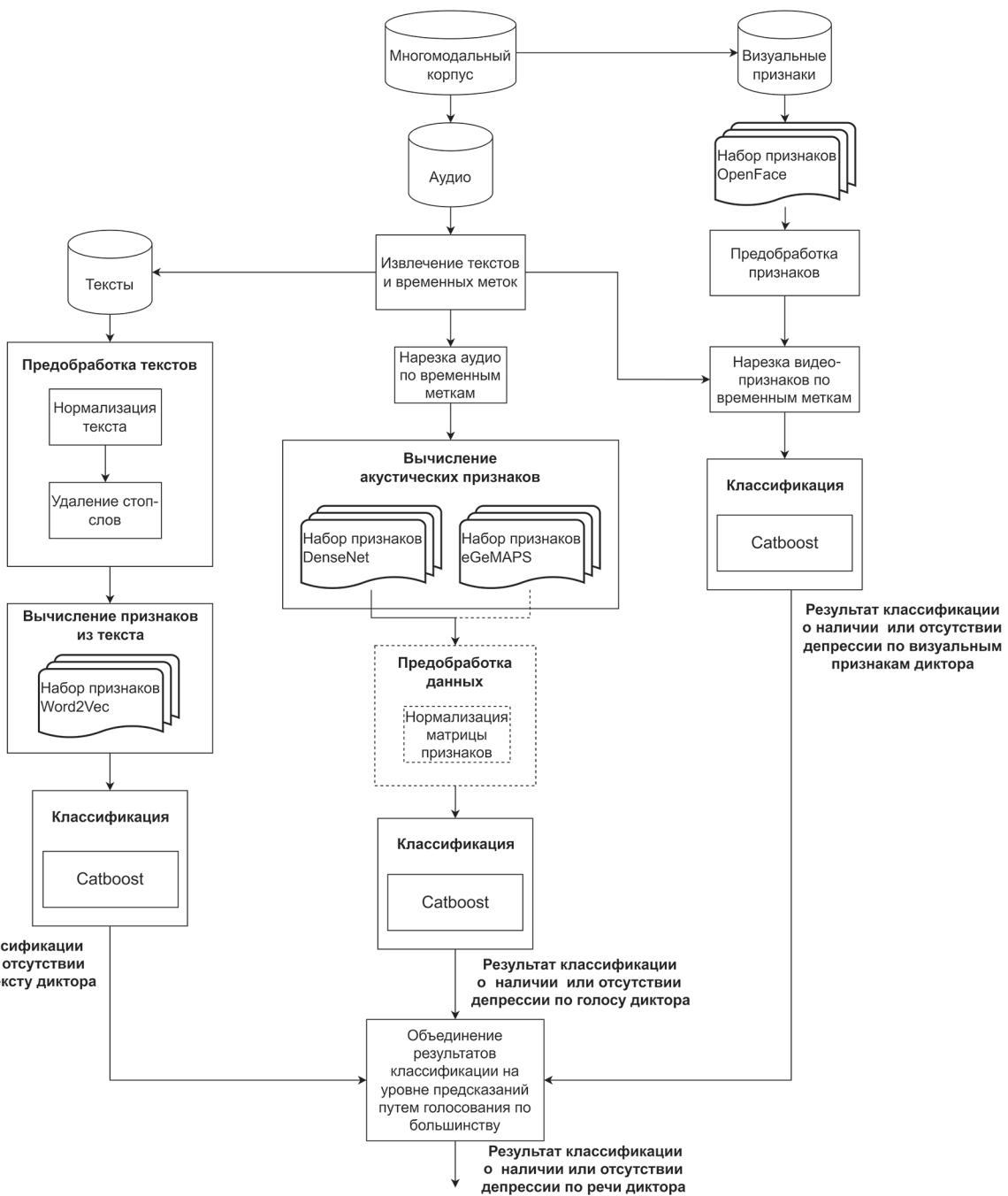


Рисунок. Схематичное изображение подхода к распознаванию депрессии для двух классов (наличие/отсутствие) с использованием многомодальных данных

Figure. A scheme of an approach for depression detection for 2 classes (presence/absence) using multimodal data

ктуации и стоп-слов, после чего с помощью методов векторизации Word2Vec и Term Frequency — Inverse Document Frequency (TF-IDF) вычислены векторные представления, которые являются входными данными классификатора.

Экспериментальное исследование предложенного подхода

Экспериментальные исследования были проведены с использованием разбиения данных, представленных в работе [25] (train, dev, test), для возможности последу-

ющего сравнения с аналогичными работами на данном корпусе. Результаты распознавания депрессии по видеоданным приведены в табл. 1, по текстовым данным — в табл. 2. В качестве показателей были выбраны: невзвешенная средняя полнота (Unweighted Average Recall, UAR); полнота для каждого класса (Recall); взвешенная средняя F1-мера (WF1); F-мера для каждого класса. Для бинарной классификации опробованы детерминированные методы машинного обучения: градиентный бустинг на основе деревьев решений (Catboost), метод случайного леса (Random Forest, RF), метод линейного дискриминантного анализа (Linear Discriminant

Таблица 1. Результаты распознавания депрессии на корпусе E-DAIC с использованием видеоданных, %
Table 1. Results of depression detection using E-DAIC corpus data for visual modality, %

Классификатор	UAR	Recall		WF1	F1-мера	
		(0)	(1)		(0)	(1)
VGG						
Catboost	49,1	71,8	27,3	52,8	65,9	31,2
RF	50,1	93,9	6,3	52,2	75,9	10,1
LDA	51,4	76,1	26,6	51,0	68,9	33,8
<i>k</i> -NN	50,4	70,5	30,3	54,0	66,2	33,8
ResNet						
Catboost	46,1	60,8	31,5	49,5	60,1	32,1
RF	50,1	97,0	3,2	49,4	75,9	6,0
LDA	53,6	85,0	22,2	55,9	73,2	30,2
<i>k</i> -NN	50,6	83,2	18,0	53,7	71,4	24,7
OpenFace						
Catboost	52,8	56,7	48,9	54,2	59,9	45,2
RF	55,6	96,1	15,0	56,5	76,8	24,8
LDA	51,2	97,0	5,4	49,8	75,4	10,0
<i>k</i> -NN	52,6	66,8	38,4	55,3	64,8	40,3
BOVW						
Catboost	52,3	64,7	39,4	55,6	65,0	39,7
RF	50,2	70,4	30,0	54,3	66,7	33,1
LDA	49,1	84,0	14,2	52,6	71,7	20,1
<i>k</i> -NN	49,5	67,1	32,0	53,4	64,8	34,0

Analysis, LDA) и метод *k*-ближайших соседей (*k*-Nearest Neighbours, *k*-NN). Жирным выделены лучшие результаты, полученные с использованием классификаторов на различных наборах признаков по соответствующим показателям, а также классификаторы, с использованием которых удалось добиться лучших и/или наиболее сбалансированных результатов по каждому набору

признаков. В круглых скобках указан номер класса: 0 — отсутствие и 1 — наличие депрессии. Результат многомодального распознавания депрессии был получен путем объединения модальностей посредством голосования по большинству между полученными предсказаниями, где на входе имеются три бинарных значения {0, 1}, полученных от трех классификаторов

Таблица 2. Результаты распознавания депрессии на корпусе E-DAIC с использованием текстовых данных, %
Table 2. Results of depression detection using E-DAIC corpus data for text modality, %

Классификатор	UAR	Recall		WF1	F1-мера	
		(0)	(1)		(0)	(1)
Word2Vec						
Catboost (базовые параметры)	51,4	74,3	28,6	55,3	68,4	33,3
RF	50,5	91,4	9,5	52,3	74,4	15,4
LDA	52,9	77,1	28,6	56,7	70,1	34,3
<i>k</i> -NN	47,6	85,7	9,5	50,0	71,4	14,3
Catboost (подобранные параметры)	62,3	62,9	61,9	63,1	67,7	55,3
TF-IDF						
Catboost (базовые параметры)	59,5	100,0	19,0	62,3	80,5	32,0
RF	50,0	100,0	0,0	48,0	77,0	0,0
LDA	50,0	100,0	0,0	48,0	77,0	0,0
<i>k</i> -NN	48,6	68,6	28,6	52,4	64,9	31,6
Catboost (подобранные параметры)	54,3	80,0	28,6	58,1	71,8	35,3

для каждой модальности, а на выходе единственное бинарное значение.

При исследованиях с видеомодальностью было выявлено, что признаки OpenFace в сочетании с методом Catboost оказались наиболее эффективными. В экспериментальных исследованиях с текстовыми данными метод Catboost с признаками Word2Vec показал лучшие результаты, поскольку обеспечил наиболее сбалансированные результаты распознавания между классами.

Обсуждение результатов

Предложенный подход распознавания депрессии с использованием многомодальных данных, основанный на классификаторе Catboost и нескольких наборах признаков, показал достаточно высокую точность WF1 = 68,0 % (табл. 3), что подтверждает эффективность объединения трех модальностей. Жирным выделены лучшие результаты, полученные при использовании подхода по соответствующим показателям. В круглых скобках указан номер класса: 0 — отсутствие и 1 — наличие депрессии. Полученные результаты сравнения с аналогичными работами показывают, что предложенный подход находится как минимум на мировом уровне, а в ряде случаев превосходит аналогичные подходы, при этом являясь более быстрым и менее ресурсозатратным к оборудованию. Предложенный подход оказался эффективнее одномодальных мето-

дов для аудио-, видео- и многомодальных методов, представленных в работах [36, 37], однако по ряду показателей оказался менее точным, чем подходы, представленные в других работах, посвященных текстовой модальности и многомодальным подходам.

В ходе исследования были использованы данные многомодального корпуса E-DAIC для бинарного распознавания депрессии, проведено исследование эффективности одно- и многомодальных подходов, а также разработан подход многомодального анализа акустических (DenseNet), визуальных (OpenFace) и текстовых (Word2Vec) признаков. Использование полуавтоматической разметки данных позволило получить менее зашумленные данные. Основное внимание удалено поиску оптимальных параметров и типов признаков, а также сравнению классификаторов. Для классификации в работе были использованы детерминированные методы классификации, такие как Catboost, а итоговое решение принималось посредством голосования по большинству, что минимизировало влияние дисбаланса классов и повысило надежность получаемых результатов. Эксперименты с видеоданными показали, что признаки OpenFace в сочетании с деревьями решений, являются наиболее эффективным решением для видеомодальности. В экспериментах с текстовыми данными метод Catboost с признаками Word2Vec оказался наилучшим, поскольку обеспечил сбалансированные результаты распознавания. Итоговые результаты мно-

Таблица 3. Сравнение показателей эффективности подхода к бинарному распознаванию депрессии с современными подходами на корпусе E-DAIC, %

Table 3. A comparison of the proposed approach efficiency with modern approaches for binary depression detection using E-DAIC corpus data, %

Ссылка на источник подхода	UAR	Recall		WF1	F1-мера	
		(0)	(1)		(0)	(1)
Аудиомодальность						
[37]	52,0	—	—	—	—	—
[38]	41,0	—	—	—	—	—
Настоящая работа	61,0	74,3	47,6	63,9	72,2	50,0
Видеомодальность						
[37]	49,0	—	—	—	—	—
[38]	48,0	—	—	—	—	—
Настоящая работа	52,8	56,7	48,9	54,2	59,9	45,2
Текстовая модальность						
[17]	73,9	84,6	63,3	73,0	74,6	71,7
[15]	—	—	—	84,0	—	—
[16]	—	—	—	86,0	—	—
Настоящая работа	62,3	62,9	61,9	63,1	67,7	55,3
Многомодальное распознавание						
[39]	76,9	83,1	70,9	79,5	84,6	68,2
[37]	58,0	—	—	—	—	—
[38]	54,0	—	—	—	—	—
[23]	—	—	—	92,2	—	—
Настоящая работа	64,3	85,7	42,9	68,0	77,9	51,4

гомодального подхода составляют $WF_1 = 68,0\%$ и $UAR = 64,3\%$ (табл. 3) и сопоставимы с современными международными исследованиями, при этом достигается сбалансированная классификация между классами (при отсутствии и наличии депрессии). К сожалению, не представляется возможным полностью исключить влияние данного фактора внутри одного корпуса, так как дисбаланс классов в корпусе E-DAIC достаточно высок и составляет разницу более, чем в два раза: записей, принадлежащих классу 0 (отсутствие депрессии) — 189, а классу 1 (наличие депрессии) — 86. В качестве возможного решения данной проблемы можно представить межкорпусные исследования, с целью повышения количества записей миноритарного класса. В качестве альтернативного решения представляется возможным опробовать такие техники борьбы с дисбалансом классов как синтетическое увеличение миноритарного класса, например, Synthetic Minority Oversampling Technique, успешно примененного в работе [40] для классификации эмоций в видеоданных. Повышение объяснимости представленного подхода и полученных результатов обуславливается тем, что были использованы методы машинной классификации, основанные на деревьях решений, а также преимущественно экспертные признаки, понятные человеку. Интерпретировать результаты, полученные деревьями решений, можно, например, построив гистограммы на основе вычисленных важностей признаков, поскольку современные методы машинного обучения, реализуемые на базе деревьев решений, имеют встроенную функцию вычисления важности признаков, кроме того, возможен вариант графического отображения дерева решений.

Заключение

В ходе данного исследования был разработан автоматический многомодальный подход к определению депрессии, основанный на экспертных признаках и детерминированных методах машинного обучения. В качестве данных для обучения и тестирования использовался многомодальный корпус Extended Distress Analysis Interview Corpus, данные которого были пре-

добротаны полуавтоматическим путем. Полученные экспериментальные результаты показали, что полуавтоматическая разметка данных позволила не только избавиться от шумов в данных, влияющих на качество обучения машинных классификаторов, но и сократить дисбаланс в результатах распознавания между классами, а также повысить качество распознавания депрессии и получить результаты, сопоставимые с известными аналогами, а в ряде случаев превосходящие их.

В дальнейшем планируется углубленная работа с признаками для каждой модальности. Также будет уделено особое внимание интерпретируемости результатов как при помощи методов машинного обучения, способных предоставить интерпретируемые результаты классификации (что частично было реализовано в настоящей работе), так и методов типа SHapley Additive exPlanations и Local Interpretable Model-agnostic Explanations. Также одним из будущих направлений данной работы является повышение качества распознавания для текстовой модальности и многомодального подхода в том числе с использованием больших языковых моделей.

Системы автоматического многомодального определения депрессии потенциально могут быть использованы специалистами первичного звена здравоохранения в качестве экспертных систем при первичной диагностике тревожно-депрессивных расстройств, а также в ряде комплексных систем, целью которых является автоматическая оценка психоэмоционального состояния человека. Среди ограничений при внедрении предложенного подхода можно выделить: различия в качестве обучающих и данных реального мира (влияние данного фактора может быть уменьшено за счет составления более вариативных наборов обучающих данных, а также добавления методов фильтрации посторонних шумов), относительно низкое качество распознавания (может быть повышенено путем добавления вариантов данных, а также комплексирования моделей и признаков). При этом, при внедрении таких систем стоит учитывать, что вне зависимости от качества распознавания, система является вспомогательным инструментом, а итоговое решение должен принимать специалист.

Литература

- Ушаков И.Б., Бубеев Ю.А., Сыркин Л.Д., Карпов А.А., Поляков А.В., Иванов А.В., Усов В.М. Дистанционное телеконсультирование в первичном звене здравоохранения для скрининга тревожно-депрессивных расстройств с контуром обратной связи от пациента // Системный анализ и управление в биомедицинских системах. 2023. Т. 22. № 4. С. 140–153. <https://doi.org/10.36622/VSTU.2023.22.4.022>
- Depressive disorder WHO (depression). WHO. 2023 [Электронный ресурс]. URL: <https://www.who.int/news-room/fact-sheets/detail/depression> (дата обращения: 22.08.2025)
- Wu W., Zhang C., Woodland P.C. Confidence estimation for automatic detection of depression and Alzheimer's disease based on clinical interviews // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2024. P. 3160–3164. <https://doi.org/10.21437/Interspeech.2024-546>
- Braun F., Bayerl S.P., Perez-Toro P.A., Hoenig F., Lehfeld H., Hillemacher T., Noeth E., Bocklet T., Riedhammer K. Classifying dementia in the presence of depression: a cross-corpus study // Proc.

References

- Ushakov I.B., Bubeev Yu.A., Syrkin L.D., Karpov A.A., Polyakov A.V., Ivanov A.V., Usov V.M. Remote tele-counseling in primary healthcare for screening of anxiety-depressive disorders with a feedback loop from the patient. *System analysis and management in biomedical systems*, 2023, vol. 22, no. 4, pp. 140–153. (in Russian). <https://doi.org/10.36622/VSTU.2023.22.4.022>
- Depressive disorder WHO (depression)*. WHO. 2023. Available at: <https://www.who.int/news-room/fact-sheets/detail/depression> (accessed: 22.08.2025)
- Wu W., Zhang C., Woodland P.C. Confidence estimation for automatic detection of depression and Alzheimer's disease based on clinical interviews. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2024, pp. 3160–3164. <https://doi.org/10.21437/Interspeech.2024-546>
- Braun F., Bayerl S.P., Perez-Toro P.A., Hoenig F., Lehfeld H., Hillemacher T., Noeth E., Bocklet T., Riedhammer K. Classifying dementia in the presence of depression: a cross-corpus study. *Proc. of the Annual Conference of the International Speech Communication*

- of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 2308–2312. <https://doi.org/10.21437/Interspeech.2023-1997>
5. Brueckner R., Kwon N., Subramanian V., Blaylock N., O'Connell H. Anxiety and Depression Detection using Vocal Biomarkers. Canaryspeech report. 2025. [Электронный ресурс]. URL: <https://canaryspeech.com/blog/anxiety-and-depression-detection-using-vocal-biomarkers/> (дата обращения: 22.08.2025)
 6. Ji J., Dong W., Li J., Peng J., Feng C., Liu R., Shi C., Ma Y. Depressive and mania mood state detection through voice as a biomarker using machine learning // *Frontiers in Neurology*. 2024. V. 15. P. 1394210. <https://doi.org/10.3389/fneur.2024.1394210>
 7. Ringeval F., Schuller B., Valstar M., Cummins N., Cowie R., Tavabi L., et al. AVEC 2019 Workshop and Challenge: state-of-mind, detecting depression with AI, and cross-cultural affect recognition // Proc. of the 9th International on Audio/Visual Emotion Challenge and Workshop. 2019. P. 3–12. <https://doi.org/10.1145/3347320.3357688>
 8. Gratch J., Artstein R., Lucas G., Stratou G., Scherer S., Nazarian A., et al. The Distress Analysis Interview Corpus of Human and Computer Interviews // Proc. of the 9th International Conference on Language Resources and Evaluation (LREC'14). 2014. P. 3123–3128.
 9. Li Y., Shi S., Yang F., Gao J., Li Y., Tao M., et al. Patterns of comorbidity with anxiety disorders in Chinese women with recurrent major depression // *Psychological Medicine*. 2012. V. 42. N 6. P. 1239–1248. <https://doi.org/10.1017/s003329171100273x>
 10. Zou B., Han J., Wang Y., Liu R., Zhao S., Feng L., Lyu X., Ma H. Semi-structural interview-based Chinese multimodal depression corpus towards automatic preliminary screening of depressive disorders // *IEEE Transactions on Affective Computing*. 2022. V. 14. N 4. P. 2823–2838. <https://doi.org/10.1109/TAFFC.2022.3181210>
 11. Campbell E.L., Dineley J., Conde P., Matcham F., White K.M., Oetzmann C., et al. The RADAR-CNS Consortium. Classifying depression symptom severity: Assessment of speech representations in personalized and generalized machine learning models // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 1738–1742. <https://doi.org/10.21437/Interspeech.2023-1721>
 12. Fara S., Hickey O., Georgescu A., Goria S., Molimpakis E., Cummins N. Bayesian Networks for the robust and unbiased prediction of depression and its symptoms utilizing speech and multimodal data // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 1728–1732. <https://doi.org/10.21437/Interspeech.2023-1709>
 13. Tao F., Esposito A., Vinciarelli A. The androids corpus: a new publicly available benchmark for speech based depression detection // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 4149–4153. <https://doi.org/10.21437/Interspeech.2023-894>
 14. Phukan O.C., Jain S., Singh S., Singh M., Buduru A.B., Sarma R. ComFeAT: Combination of neural and spectral features for improved depression detection // *arXiv*. 2024. arXiv:2406.06774. <https://doi.org/10.48550/arXiv.2406.06774>
 15. Burdisso S., Villatoro-Tello E., Madikeri S., Motlicek P. Node-weighted graph convolutional network for depression detection in transcribed clinical interviews // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 3617–3621. <https://doi.org/10.21437/interspeech.2023-1923>
 16. Zhang X., Li C., Chen W., Zheng J., Li F. Optimizing depression detection in clinical doctor-patient interviews using a multi-instance learning framework // *Scientific Reports*. 2025. V. 15. N 1. P. 6637. <https://doi.org/10.1038/s41598-025-90117-w>
 17. Tank C., Pol S., Katoch V., Meht S., Anand A., Shah R.R. Depression detection and analysis using large language models on textual and audio-visual modalities // *arXiv*. 2024. arXiv:2407.06125. <https://doi.org/10.48550/arXiv.2407.06125>
 18. Zhang W., Mao K., Chen J. A multimodal approach for detection and assessment of depression using text, audio and video // *Phenomics*. 2024. V. 4. N 3. P. 234–249. <https://doi.org/10.1007/s43657-023-00152-8>
 19. Zhang X., Liu H., Xu K., Zhang Q., Liu D., Ahmed B., Epps J. When LLMs meet acoustic landmarks: an efficient approach to integrate speech into large language models for depression detection // Proc. of the Conference on Empirical Methods in Natural Language Processing. 2024. P. 146–158. <https://doi.org/10.18653/v1/2024.emnlp-main.8>
 20. Association Interspeech, 2023, pp. 2308–2312. <https://doi.org/10.21437/Interspeech.2023-1997>
 21. Brueckner R., Kwon N., Subramanian V., Blaylock N., O'Connell H. *Anxiety and Depression Detection using Vocal Biomarkers*. Canaryspeech report. 2025. Available at <https://canaryspeech.com/blog/anxiety-and-depression-detection-using-vocal-biomarkers/> (accessed: 22.08.2025)
 22. Ji J., Dong W., Li J., Peng J., Feng C., Liu R., Shi C., Ma Y. Depressive and mania mood state detection through voice as a biomarker using machine learning // *Frontiers in Neurology*. 2024, vol. 15, pp. 1394210. <https://doi.org/10.3389/fneur.2024.1394210>
 23. Ringeval F., Schuller B., Valstar M., Cummins N., Cowie R., Tavabi L., et al. AVEC 2019 Workshop and Challenge: state-of-mind, detecting depression with AI, and cross-cultural affect recognition. *Proc. of the 9th International on Audio/Visual Emotion Challenge and Workshop*, 2019, pp. 3–12. <https://doi.org/10.1145/3347320.3357688>
 24. Gratch J., Artstein R., Lucas G., Stratou G., Scherer S., Nazarian A., et al. The Distress Analysis Interview Corpus of Human and Computer Interviews. *Proc. of the 9th International Conference on Language Resources and Evaluation (LREC'14)*, 2014, pp. 3123–3128.
 25. Li Y., Shi S., Yang F., Gao J., Li Y., Tao M., et al. Patterns of comorbidity with anxiety disorders in Chinese women with recurrent major depression. *Psychological Medicine*, 2012, vol. 42, no. 6, pp. 1239–1248. <https://doi.org/10.1017/s003329171100273x>
 26. Zou B., Han J., Wang Y., Liu R., Zhao S., Feng L., Lyu X., Ma H. Semi-structural interview-based Chinese multimodal depression corpus towards automatic preliminary screening of depressive disorders. *IEEE Transactions on Affective Computing*, 2022, vol. 14, no. 4, pp. 2823–2838. <https://doi.org/10.1109/TAFFC.2022.3181210>
 27. Campbell E.L., Dineley J., Conde P., Matcham F., White K.M., Oetzmann C., et al. The RADAR-CNS Consortium. Classifying depression symptom severity: Assessment of speech representations in personalized and generalized machine learning models. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 1738–1742. <https://doi.org/10.21437/Interspeech.2023-1721>
 28. Fara S., Hickey O., Georgescu A., Goria S., Molimpakis E., Cummins N. Bayesian Networks for the robust and unbiased prediction of depression and its symptoms utilizing speech and multimodal data. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 1728–1732. <https://doi.org/10.21437/Interspeech.2023-1709>
 29. Tao F., Esposito A., Vinciarelli A. The androids corpus: a new publicly available benchmark for speech based depression detection. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 4149–4153. <https://doi.org/10.21437/Interspeech.2023-894>
 30. Phukan O.C., Jain S., Singh S., Singh M., Buduru A.B., Sarma R. ComFeAT: Combination of neural and spectral features for improved depression detection. *arXiv*, 2024, arXiv:2406.06774. <https://doi.org/10.48550/arXiv.2406.06774>
 31. Burdisso S., Villatoro-Tello E., Madikeri S., Motlicek P. Node-weighted graph convolutional network for depression detection in transcribed clinical interviews. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 3617–3621. <https://doi.org/10.21437/interspeech.2023-1923>
 32. Zhang X., Li C., Chen W., Zheng J., Li F. Optimizing depression detection in clinical doctor-patient interviews using a multi-instance learning framework. *Scientific Reports*, 2025, vol. 15, no. 1, pp. 6637. <https://doi.org/10.1038/s41598-025-90117-w>
 33. Tank C., Pol S., Katoch V., Meht S., Anand A., Shah R.R. Depression detection and analysis using large language models on textual and audio-visual modalities. *arXiv*, 2024, arXiv:2407.06125. <https://doi.org/10.48550/arXiv.2407.06125>
 34. Zhang W., Mao K., Chen J. A multimodal approach for detection and assessment of depression using text, audio and video. *Phenomics*, 2024, vol. 4, no. 3, pp. 234–249. <https://doi.org/10.1007/s43657-023-00152-8>
 35. Zhang X., Liu H., Xu K., Zhang Q., Liu D., Ahmed B., Epps J. When LLMs meet acoustic landmarks: an efficient approach to integrate speech into large language models for depression detection. *Proc. of the Conference on Empirical Methods in Natural Language Processing*, 2024, pp. 146–158. <https://doi.org/10.18653/v1/2024.emnlp-main.8>

20. Dumpala S.H., Dikaios K., Nunes A., Rudzicz F., Uher R., Oore S. Self-supervised embeddings for detecting individual symptoms of depression // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2024. P. 1450–1454. <https://doi.org/10.21437/Interspeech.2024-2344>
21. Sadeghi M., Richer R., Egger B., Schindler-Gmelch L., Rupp L.H., Rahimi F., Berking M., Eskofier B.M. Harnessing multimodal approaches for depression detection using large language models and facial expressions // npj Mental Health Research. 2024. V. 3. N 1. P. 66. <https://doi.org/10.1038/s44184-024-00112-8>
22. Wang J., Ravi V., Flint J., Alwan A. Speechformer-CTC: Sequential modeling of depression detection with speech temporal classification // Speech Communication. 2024. V. 163. P. 103106. <https://doi.org/10.1016/j.specom.2024.103106>
23. Jin N., Ye R., Li P. Diagnosis of depression based on facial multimodal data // Frontiers in Psychiatry. 2025. V. 16. P. 1508772. <https://doi.org/10.3389/fpsyg.2025.1508772>
24. Zhou L., Liu Z., Shangguan Z., Yuan X., Li Y., Hu B. JAMFN: Joint attention multi-scale fusion network for depression detection // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 3417–3421. <https://doi.org/10.21437/Interspeech.2023-183>
25. Величко А.Н., Карпов А.А. Подход к распознаванию депрессии по речи человека с использованием полуавтоматической разметки данных // Информационно-управляющие системы. 2024. № 4(131). С. 2–11. <https://doi.org/10.31799/1684-8853-2024-4-2-11>
26. Litvinova T., Ryzhkova E. RusNeuroPsych: open corpus for study relations between author demographic, personality traits, lateral preferences and affect in text // International journal of Open Information Technologies. 2018. V. 6. N 3. P. 32–36.
27. Stankevich M., Ignatiev N., Smirnov I. Predicting depression with social media images // Proc. of the 9th International Conference on Pattern Recognition Applications and Methods ICPRAM. 2020. V. 1. P. 235–240. <https://doi.org/10.5220/0009168602350240>
28. Stankevich M.A., Smirnov I.V., Kuznetsova Y.M., Kiselnikova N.V., Enikolopov S.N. Predicting depression from essays in Russian // Proc. of the International Conference “Dialogue 2019”. 2019. P. 647–657.
29. Stankevich M., Smirnov I., Kiselnikova N., Ushakova A. Depression detection from social media profiles // Communications in Computer and Information Science. 2020. V. 1223. P. 181–194. https://doi.org/10.1007/978-3-030-51913-1_12
30. Stepanov D., Smirnov A., Ivanov E., Smirnov I., Stankevich M., Danina M. Detection of health-preserving behavior Among VK.com users based on the analysis of graphic, text and numerical data // Lecture Notes in Networks and Systems. 2022. V. 296. P. 574–587. https://doi.org/10.1007/978-3-030-82199-9_39
31. Кисельникова Н., Станкевич М., Данина М., Куминская Е., Лаврова Е. Выявление информативных параметров поведения пользователей социальной сети ВКонтакте как признаков депрессии // Психология. Журнал Высшей школы экономики. 2020. Т. 17. № 1. С. 73–88. <https://doi.org/10.17323/1813-8918-2020-1-73-88>
32. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely connected convolutional networks // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
33. Baltrusaitis T., Zadeh A., Lim Y.C., Morency L.-P. OpenFace 2.0: facial behavior analysis toolkit // Proc. of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). 2018. P. 59–66. <https://doi.org/10.1109/FG.2018.00019>
34. Mikolov T., Sutskever I., Chen K., Corrado G.S., Dean J. Distributed representations of words and phrases and their compositionality // Advances in Neural Information Processing Systems. 2013. V. 26. P. 1–9.
35. Boersma P. Praat, a system for doing phonetics by computer // Glot International. 2001. N 5. P. 341–345.
36. Величко А.Н., Карпов А.А. Методика и программная система интегрального анализа деструктивных паралингвистических явлений в разговорной речи // Информационно-управляющие системы. 2023. № 4 (125). С. 2–11. <https://doi.org/10.31799/1684-8853-2023-4-2-11>
37. Gimeno-Gómez D., Bucur A.M., Cosma A., Martínez-Hinarejos C.D., Rosso P. Reading between the frames: multi-modal depression detection in videos from non-verbal cues // Lecture Notes in Computer Science. 2024. V. 14608. P. 191–209. https://doi.org/10.1007/978-3-031-56027-9_12
20. Dumpala S.H., Dikaios K., Nunes A., Rudzicz F., Uher R., Oore S. Self-supervised embeddings for detecting individual symptoms of depression. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2024, pp. 1450–1454. <https://doi.org/10.21437/Interspeech.2024-2344>
21. Sadeghi M., Richer R., Egger B., Schindler-Gmelch L., Rupp L.H., Rahimi F., Berking M., Eskofier B.M. Harnessing multimodal approaches for depression detection using large language models and facial expressions. *npj Mental Health Research*, 2024, vol. 3, no. 1, pp. 66. <https://doi.org/10.1038/s44184-024-00112-8>
22. Wang J., Ravi V., Flint J., Alwan A. Speechformer-CTC: Sequential modeling of depression detection with speech temporal classification. *Speech Communication*, 2024, vol. 163, pp. 103106. <https://doi.org/10.1016/j.specom.2024.103106>
23. Jin N., Ye R., Li P. Diagnosis of depression based on facial multimodal data. *Frontiers in Psychiatry*, 2025, vol. 16, pp. 1508772. <https://doi.org/10.3389/fpsyg.2025.1508772>
24. Zhou L., Liu Z., Shangguan Z., Yuan X., Li Y., Hu B. JAMFN: Joint attention multi-scale fusion network for depression detection. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 3417–3421. <https://doi.org/10.21437/Interspeech.2023-183>
25. Velichko A.N., Karpov A.A. AN Approach to depression detection in speech using a semi-automatic data annotation. *Information and Control Systems*, 2024, no. 4 (131), pp. 2–11. (in Russian). <https://doi.org/10.31799/1684-8853-2024-4-2-11>
26. Litvinova T., Ryzhkova E. RusNeuroPsych: open corpus for study relations between author demographic, personality traits, lateral preferences and affect in text. *International journal of Open Information Technologies*, 2018, vol. 6, no. 3, pp. 32–36.
27. Stankevich M., Ignatiev N., Smirnov I. Predicting depression with social media images. *Proc. of the 9th International Conference on Pattern Recognition Applications and Methods ICPRAM*, 2020, vol. 1, pp. 235–240. <https://doi.org/10.5220/0009168602350240>
28. Stankevich M.A., Smirnov I.V., Kuznetsova Y.M., Kiselnikova N.V., Enikolopov S.N. Predicting depression from essays in Russian. *Proc. of the International Conference “Dialogue 2019”*, 2019, pp. 647–657.
29. Stankevich M., Smirnov I., Kiselnikova N., Ushakova A. Depression detection from social media profiles. *Communications in Computer and Information Science*, 2020, vol. 1223, pp. 181–194. https://doi.org/10.1007/978-3-030-51913-1_12
30. Stepanov D., Smirnov A., Ivanov E., Smirnov I., Stankevich M., Danina M. Detection of health-preserving behavior Among VK.com users based on the analysis of graphic, text and numerical data. *Lecture Notes in Networks and Systems*, 2022, vol. 296, pp. 574–587. https://doi.org/10.1007/978-3-030-82199-9_39
31. Kiselnikova N., Stankevich M., Danina M., Kuminskaya E., Lavrova E. Identification of informative behavior parameters in users of VKontakte social network as markers of depression. *Psychology. Journal of Higher School of Economics*, 2020, vol. 17, no. 1, pp. 73–88. (in Russian). <https://doi.org/10.17323/1813-8918-2020-1-73-88>
32. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely connected convolutional networks. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
33. Baltrusaitis T., Zadeh A., Lim Y.C., Morency L.-P. OpenFace 2.0: facial behavior analysis toolkit. *Proc. of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 59–66. <https://doi.org/10.1109/FG.2018.00019>
34. Mikolov T., Sutskever I., Chen K., Corrado G.S., Dean J. Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 2013, vol. 26, pp. 1–9.
35. Boersma P. Praat, a system for doing phonetics by computer. *Glot International*, 2001, no. 5, pp. 341–345.
36. Velichko A.N., Karpov A.A. Methods and a software system for integrative analysis of destructive paralinguistic phenomena in colloquial speech. *Information and Control Systems*, 2023, no. 4 (125), pp. 2–11. (in Russian). <https://doi.org/10.31799/1684-8853-2023-4-2-11>
37. Gimeno-Gómez D., Bucur A.M., Cosma A., Martínez-Hinarejos C.D., Rosso P. Reading between the frames: multi-modal depression detection in videos from non-verbal cues. *Lecture Notes in Computer Science*, 2024, vol. 14608, pp. 191–209. https://doi.org/10.1007/978-3-031-56027-9_12

38. Jaegle A., Gimeno F., Brock A., Zisserman A., Vinyals O., Carreira J. Perceiver: General perception with iterative attention // Proc. of the 38th International Conference on Machine Learning. 2021. V. 139. P. 4651–4664.
39. Li Y., Yang X., Zhao M., Wang Z., Yao Y., Qian W., Qi Sh. FPT-Former: A flexible parallel transformer of recognizing depression by using audiovisual expert-knowledge-based multimodal measures // International Journal of Intelligent Systems. 2024. V. 1564574. P. 1–13. <https://doi.org/10.1155/2024/1564574>
40. Рюмина Е.В., Карпов А.А. Сравнительный анализ методов устранения дисбаланса классов эмоций в видеоданных выражений лиц // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 5. С. 683–691. <https://doi.org/10.17586/2226-1494-2020-20-5-683-691>
38. Jaegle A., Gimeno F., Brock A., Zisserman A., Vinyals O., Carreira J. Perceiver: General perception with iterative attention. *Proc. of the 38th International Conference on Machine Learning*, 2021, vol. 139, pp. 4651–4664.
39. Li Y., Yang X., Zhao M., Wang Z., Yao Y., Qian W., Qi Sh. FPT-Former: A flexible parallel transformer of recognizing depression by using audiovisual expert-knowledge-based multimodal measures. *International Journal of Intelligent Systems*, 2024, vol. 1564574, pp. 1–13. <https://doi.org/10.1155/2024/1564574>
40. Ryumina E.V., Karpov A.A. Comparative analysis of methods for imbalance elimination of emotion classes in video data of facial expressions. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 5, pp. 683–691. (in Russian). <https://doi.org/10.17586/2226-1494-2020-20-5-683-691>

Авторы

Величко Алёна Николаевна — кандидат технических наук, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57203962694](#), <https://orcid.org/0000-0002-8503-8512>, velichko.a@iias.spb.su

Карпов Алексей Анатольевич — доктор технических наук, профессор, руководитель лаборатории, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57219469958](#), <https://orcid.org/0000-0003-3424-652X>, karpov@iias.spb.su

Authors

Alena N. Velichko — PhD, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, [sc 57203962694](#), <https://orcid.org/0000-0002-8503-8512>, velichko.a@iias.spb.su

Alexey A. Karpov — D.Sc., Professor, Head of Laboratory, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, [sc 57219469958](#), <https://orcid.org/0000-0003-3424-652X>, karpov@iias.spb.su

Статья поступила в редакцию 26.08.2025
Одобрена после рецензирования 21.09.2025
Принята к печати 26.11.2025

Received 26.08.2025
Approved after reviewing 21.09.2025
Accepted 26.11.2025



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»