

Нейросетевая модель искусственного интеллекта для реализации аффективных вычислений на основе анализа электроэнцефалограмм

А.Г. Чобан¹, Д.Г. Стадников¹, А.Е. Сулавко¹

¹ ФГАОУ ВО «Омский государственный технический университет» (ОмГТУ),
644050, г. Омск, просп. Мира, д. 11

Аннотация

Статья посвящена анализу возможности оценки эмоционального состояния человека по особенностям работы головного мозга с использованием неинвазивных нейрокомпьютерных интерфейсов. Проведен анализ публикаций последних лет, посвященных применению сигналов электроэнцефалограммы (ЭЭГ) для оценки эмоционального состояния, выявлены актуальные проблемы в этой области. Описаны основные подходы к стимуляции головного мозга для получения информативных сигналов ЭЭГ, а также методы их анализа и распознавания. Предложена архитектура глубокой сверточной нейронной сети для анализа данных ЭЭГ, а также нейросетевая модель искусственного интеллекта для классификации 4 эмоций (страх, счастье, печаль, спокойствие) по шкале валентности-арозальности Рассела на основе двух сверточных нейронных сетей. Проведен эксперимент с участием 50 человек, которые просматривали эмоционально окрашенные видеоролики. Собранные данные ЭЭГ 50 испытуемых, которые использовались для обучения и тестирования нейросетевой модели. Результаты показали высокую точность классификации эмоций (94% ± 3,4%) с использованием беспроводного нейроинтерфейса.

Ключевые слова: распознавание эмоций, электроэнцефалограмма, сверточные нейронные сети, сигнал ЭЭГ, интерфейс мозг-компьютер, нейроинтерфейс, анализ биометрических параметров.

Цитирование: Чобан, А.Г. Нейросетевая модель искусственного интеллекта для реализации аффективных вычислений на основе анализа электроэнцефалограмм / А.Г. Чобан, Д.Г. Стадников, А.Е. Сулавко // Компьютерная оптика. – 2024. – Т. 48, № 5. – С. 782-790. – DOI: 10.18287/2412-6179-CO-1417.

Citation: Choban AG, Stadnikov DG, Sulavko AE. Neural network in an artificial intelligence model for realization of affective computing based on electroencephalogram analysis. Computer Optics 2024; 48(5): 782-790. DOI: 10.18287/2412-6179-CO-1417.

Введение

Аффективные вычисления – это область деятельности, направленная на исследование и разработку систем и устройств, способных распознавать, обрабатывать, интерпретировать и имитировать человеческие эмоции.

Цель аффективных вычислений – обнаружение эмоциональных состояний, возникающих во время взаимодействия человека и компьютера, и синтез ответных реакций, соответствующих обнаруженным эмоциям.

Комплекс задач, связанных с распознаванием эмоций, представляет особый интерес, так как эмоции – важная часть повседневной жизни человека, которая оказывает влияние на принятие решений, восприятие, обучение, рациональное мышление и поведение. Распознавание эмоций может применяться в приложениях, требующих взаимодействия человека и компьютера, от автоматической доставки мультимедийного контента до ментальных и психологических тренировок.

Разметка мультимедийного контента с помощью релевантных тегов применяется в задачах поиска

аудиоинформации. Неявное аффективное маркирование можно отнести к легкому способу создания субъективных и/или эмоциональных тегов. Аффективные характеристики и их маркирование позволяют повысить производительность рекомендательных и поисковых систем [1].

Эмоции могут быть выражены либо вербально с помощью эмоциональной лексики, либо с помощью невербальных сигналов, таких как интонация голоса, мимика и жесты. В настоящее время многие существующие исследования по оценке эмоций сосредоточены на анализе выражения лица и речи для определения эмоционального состояния человека [2, 3]. Однако большинство современных систем взаимодействия человека и компьютера (HCI) не способны идентифицировать и интерпретировать эту информацию, поскольку человек намеренно может скрывать свои эмоции (например, грусть можно скрыть улыбкой).

Сегодня активно развиваются технологии создания нейрокомпьютерных интерфейсов, позволяющих бесконтактно передавать команды устройствам. Большинство нейроинтерфейсов основаны на регистрации и интерпретации электроэнцефалограмм (ЭЭГ), отражающих динамику изменения электриче-

ской активности головного мозга во времени. Идентификационный потенциал электроэнцефалограмм крайне высок – анализ ЭЭГ используется в таких задачах, как мысленное управление механизмами и протезами [4], биометрическая идентификация и аутентификация [5], распознавание «рискованного» поведения [6], оценка функционального (физического, психического) состояния человека [7]. В частности, анализ сигналов ЭЭГ перспективен для задач распознавания эмоций, так как имеет тесную связь с эмоциональным состоянием человека [8].

Целью данного исследования является повышение точности автоматической классификации эмоциональных состояний человека по данным ЭЭГ на базе сверточных глубоких нейронных сетей. Проведен эксперимент по сбору данных сигналов ЭЭГ, в котором в качестве стимулов используются музыкальные видеоклипы. Предложена архитектура сверточной нейронной сети (CNN), на базе которой создан мультиклассификатор эмоций по шкале валентности-арозальности (V-A) Рассела.

1. Анализ достигнутых ранее результатов

Существуют различные дискретные модели для классификации эмоций: шесть базовых эмоций по шкале Экмана [9], древовидная структура эмоций Парротта [10], «колесо» эмоций Плутчика [11], шкала валентности-арозальности Рассела [12]. Последняя широко используется для количественного описания эмоций. В шкале Рассела каждое эмоциональное состояние может быть размещено на двумерной плоскости с осью арозальности (или возбуждения) и осью валентности (рис. 1). Арозальность варьируется от неактивной (неинтересно, скучно) до активной (настороженность, возбуждение). Валентность может быть отрицательной (например, стресс) или положительной (например, радость). Фактически для классификации эмоций достаточно использовать двумерную шкалу (для определения яркости эмоций в модель оценки может быть включена доминантность). Для самооценки по этой шкале используются манекены самооценки (SAM) [13].

Остановимся подробнее на нескольких ключевых работах в области распознавания эмоционального состояния по ЭЭГ.

В работе [14] сигналы ЭЭГ используются для классификации положительных и отрицательных эмоций. В исследовании приняло участие 6 человек и использовался 62-канальный электроэнцефалограф. Испытуемому предложено просматривать видеоролики из 12 популярных художественных фильмов. Длительность каждого видеофрагмента составляла 4 минуты, при этом эмоциональный окрас каждого видеоклипа не был заранее определен. После просмотра каждого клипа испытуемому предлагалось заполнить форму SAM для измерения валентности, возбуждения и доминирования. Если оценка валентности была

меньше 5, то эта сессия принадлежала к классу отрицательных эмоций, в противном случае её относили к классу положительных эмоций. С помощью полосовых фильтров выделены частотные диапазоны, соответствующие дельта–(1–4 Гц), тета–(4–8 Гц), альфа–(8–13 Гц), бета–(13–30 Гц) и гамма–(30–40 Гц) ритмам. Затем были сформированы спектрограммы сигнала ЭЭГ. Суммарно получено более 2500 примеров образа ЭЭГ каждого испытуемого. Авторы использовали метод опорных векторов (SVM) с линейным ядром (linear-SVM) для их классификации, по 1 модели на каждый частотный диапазон (всего 5 моделей). Также была обучена 6-я модель SVM на всех частотных диапазонах. Наборы данных для обучения и тестирования разделены в соотношении 70:3. Точность составила 87,53%. Установлено, что полосы высоких частот более информативны при определении эмоций, чем полосы низких частот.



Рис. 1. Двухмерное семантическое пространство эмоций Рассела

В работе [15] проведен анализ сигналов ЭЭГ для распознавания счастья и печали. В эксперименте участвовали 6 человек (3 женщины и 3 мужчины) в возрасте 25 лет. Использовался 62-канальный электроэнцефалограф с частотой дискретизации 1000 Гц. Каждый эксперимент состоял из двух сессий с перерывом в 10 минут. Использовались изображения лиц, представляющие улыбку и слезы. Каждая картинка показывалась в течение 6 секунд. Между испытаниями показывался черный экран (3 секунды). Эмоциональное содержание этих картинок оценивалось с помощью SAM [16] для измерения валентности и возбуждения. Для классификации использовали linear-SVM. Исходный сигнал делился на короткие отрезки с двумя вариантами длины: 3 с и 1 с. Каждый эксперимент состоял из 240 и 720 испытаний для отрезков длительностью 3 с и 1 с соответственно. Наборы данных для обучения и тестирования разделены в соотношении 7:3. Точность классификации составила $93,5\% \pm 6,7\%$ для 3-секундных отрезков, в то время как для 1-секундных – $93,0\% \pm 6,2\%$.

В исследовании [17] для записи ЭЭГ использовался 32-канальный электроэнцефалограф (HD-72, Cognionics, 300 Гц, сухие электроды). Участвовало 4

человека (3 мужчины и 1 женщина), средний возраст испытуемых – 25 лет. Испытуемым демонстрировались 3 эмоционально окрашенных видеоролика по 25 минут. После просмотра испытуемые оценивали валентность и возбуждение, используя SAM с 9-балльной шкалой для каждой сцены [18]. После удаления артефактов были рассчитаны спектральные плотности мощности для тета–(4–7 Гц), альфа–(8–12 Гц), бета–(12–30 Гц) и гамма–(31–50 Гц) ритмов. Вычислены коэффициенты корреляции Спирмена между спектрами и оценками симпатии. Далее использовались комбинации каналов и частотных полос с наибольшими коэффициентами корреляции. Выбрана модель классификации на базе SVM с ядром радиальной базисной функции (RBF), а также регрессионная модель на базе SVR с RBF-ядром для 3-секундных видеоклипов. В результате классификации SVM была достигнута точность 77,6%, модель регрессии показала точность 64,5%.

В работе [19] авторы используют шкалу Рассела для количественного описания эмоционального состояния субъектов по сигналам ЭЭГ. Утверждается, что сигналы ЭЭГ нестабильны во времени и восприимчивы к незначительным изменениям окружающей среды. Сигналы, полученные от одного и того же субъекта в разные моменты времени при одинаковой стимуляции, как правило, различаются. Для удаления помех авторы вычитали из целевого участка ЭЭГ предварительно записанный сигнал субъекта в спокойном состоянии. Артефакты из сигнала удалялись вручную. В качестве алгоритма классификации использованы ансамбли деревьев решений. Точность классификации составила $95,8\% \pm 2,5\%$ на общедоступном наборе данных DEAP и $87,28\% \pm 5,5\%$ на наборе данных DREAMER.

В работе [20] рассматривается подход с выделением дифференциальной энтропии на 4 полосах частот сигнала. Electrodes размещались в соответствии с международной системой 10–20 (рис. 2а) – это стандартная система размещения электродов на поверхности головы (обычно используется 21 электрод), которая рекомендована Международной федерацией электроэнцефалографии и клинической нейрофизиологии. Данные ЭЭГ подавались на вход сверточной нейронной сети. В работе был использован набор данных DEAP. Точность классификации составила 90,24% для возбуждения и 89,45% для валентности.

В исследовании [21] авторы предложили самоконтролируемое обучение (SSL) [22] вместо применения SAM. Использован набор данных AMIGOS. Применялась сверточная нейронная сеть. Достигнуты показатели точности распознавания 84% и 85,8% отдельно для валентности и возбуждения соответственно.

В работе [23] предложен подход для отнесения ЭЭГ к одному из 4 классов в соответствии с квадрантами шкалы Рассела – высокая валентность (V) и арозальность (A) (HVNA), высокая V и низкая A

(HVLA), низкая V и высокая A (LVNA), низкая V и A (LVLA). Исходные сигналы были разделены на эпохи длиной 2 секунды, затем к эпохам была применена Z-нормализация и нормализация от 0 до 1. Обработанные сигналы попадали на вход 1D сверточной нейронной сети с LSTM-ячейками. Показатель точности распознавания составил 97,8%.

K.S. Bhanumathi и др. в своей работе [24] провели обзор существующих алгоритмов по распознаванию эмоций. Установлено, что наибольшую точность дает подход, основанный на использовании глубоких нейронных сетей. Подход может быть улучшен путем применения механизмов внимания для фокусировки на наиболее информативных участках ЭЭГ.

Ключевые выводы:

1. Сигналы ЭЭГ имеют высокую размерность пространства признаков и нелинейные зависимости в данных, поэтому почти невозможно выделить зависимости с помощью линейных алгоритмов с высокой точностью. Методы глубокого обучения сверточных и рекуррентных сетей позволяют получать более адекватные результаты. Преимущества рекуррентных нейронных сетей проявляются, когда объемы данных обучения становятся очень большими.
2. Показательным является следующее: высокая точность наблюдалась при использовании только двух лобных датчиков, что дает возможность применять на практике «сухие» электроды, непосредственно контактирующие с кожей, и использовать нейрогарнитуру в качестве устройства записи сигналов ЭЭГ.
3. Существующие наборы данных ЭЭГ, характеризующих разные эмоциональные состояния человека, как правило, имеют небольшой объем. Недостаток вариативности сигналов может привести к недостаточной обученности (или к переобучению) моделей распознавания эмоций и снижению их обобщающей способности. Также данные могут быть зависимы от оборудования, на котором планируется применять разрабатываемый метод. По этим причинам большинство исследователей формируют собственные наборы данных. В настоящем исследовании решено пойти таким же путем.

2. Формирование набора данных ЭЭГ

В настоящем исследовании в качестве стимулов выбраны музыкальные видеоклипы. Это обусловлено следующими данными. L.A. Schmidt и L.J. Trainor обнаружили [25], что активность сигналов ЭЭГ в левой фронтальной зоне возрастает (рис. 2б) во время прослушивания музыкальных отрывков с положительной эмоциональной окраской (радость и счастье) и в правой фронтальной зоне (рис. 2в) во время представления отрицательно оцененных музыкальных отрывков (страх и печаль). Результаты соответствуют модели лобной акти-

вазии/валентности эмоций, сформулированной Дэвидсоном [26], Фоксом [27] и Хеллером [28], в которой они утверждают, что переживание положительных эмоций локализовано в левой передней области мозга, тогда как переживание отрицательных эмоций локализовано в

правой передней области мозга. Результаты показывают, что эмоции, вызванные музыкой, активируют те же лобные области мозга, что и эмоции, вызванные другими модальностями. Для этих целей может быть использована музыка любых жанров.

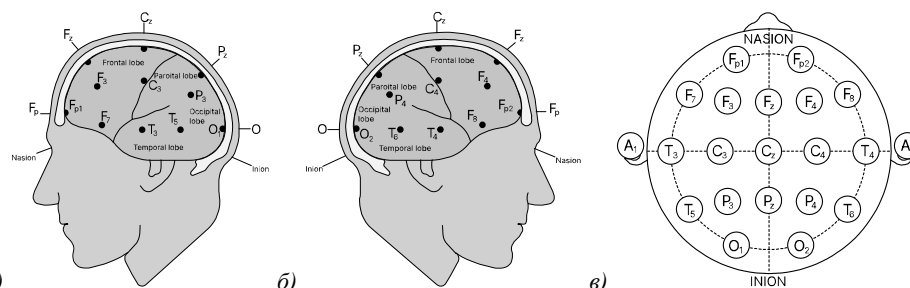


Рис. 2. Расположение электродов по областям мозга в соответствии со схемой 10–20. а) левифронтальная зона, б) правифронтальная зона, в) вид сверху

В ходе предварительного исследования были тщательно отобраны 60 видеоклипов таким образом, чтобы вызвать наиболее яркие эмоции у целевой группы, которая состояла из студентов университета (популярные русскоязычные и иностранные поп- и рок-исполнители).

Пространство валентности-возбуждения можно разделить на четыре квадранта в соответствии со шкалой валентности-арозальности Рассела [12], а именно:

- низкое возбуждение/низкая валентность (LA/LV);
- низкое возбуждение/ высокая валентность (LA/HV);
- высокое возбуждение/низкая валентность (HA/LV);
- высокое возбуждение/высокая валентность (HA/HV).

Испытуемые выбирали те видео, которые вызвали наиболее четкие эмоциональные реакции. Затем из каждого видеостимула выделена одноминутная часть. Испытуемые использовали манекены самооценки [29]. Участники смотрели музыкальные клипы и оценивали их по дискретной 9-балльной валентности, возбуждению и доминированию (рис. 3). Порядок клипов был рандомизирован.

После того, как все 60 видео были оценены 20 добровольцами, чтобы максимизировать силу вызванных эмоций, мы отобрали те видео, которые имели самые высокие оценки. Для каждого видеофрагмента мы рассчитали нормализованный балл возбуждения и валентности, взяв среднее значение. Затем для каждого квадранта мы отобрали 10 видео, которые лежат ближе всего к крайнему углу квадранта. В итоге, из первоначальной коллекции стимульных видеоклипов были отобраны 40 музыкальных видеофрагментов.

В ходе эксперимента ЭЭГ записывалась с помощью нейрогарнитуры NeuroPlay-8Cap (рис. 4). Для предъявления стимулов и записи оценок пользователей использовалось программное обеспечение «NeuroPlayPro» компании Neuroplay. Видеофрагмен-

ты были представлены на широкоформатном экране (1920:1080, 60 Гц), испытуемые находились в 1 м от экрана. ЭЭГ записывалась с помощью нейроинтерфейса NeuroPlay-8Cap (125 Гц, 8 активных AgCl-электродов). Электроды были размещены в соответствии с международной системой 10–20).

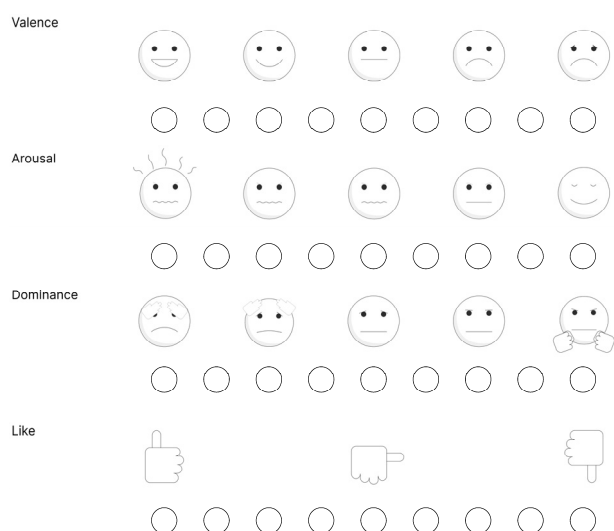


Рис. 3. Манекен самооценки SAM: валентность, возбуждение, доминирование и оценка симпатии



Рис. 4. Нейрогарнитура NeuroPlay-8Cap

В эксперименте приняли участие пятьдесят участников, не имеющих психических или физических патологий в возрасте от 19 до 25 лет. Неврологический статус всех испытуемых до начала эксперимента

оценивался как нормальный. Перед началом эксперимента каждый участник подписал форму согласия и заполнил анкету. Экспериментатор также присутствовал, чтобы ответить на любые вопросы.

Видеоролики были продемонстрированы в 40 испытаниях, каждое из которых состояло из следующих этапов:

1. 2-секундная заставка, отображающая номер текущего испытания, чтобы проинформировать участников об их прогрессе.
2. 5-секундная запись исходного сигнала, чтобы участник мог настроиться на следующее испытание.
3. 1-минутный показ музыкального видеофрагмента.
4. В конце каждого испытания участники проводили самооценку своих уровней возбуждения, валентности, симпатии и доминирования.

Манекены (рис. 3) отображались в центре экрана с цифрами 1–9, напечатанными ниже. Участники перемещали мышшь, чтобы указать свой уровень самооценки. После 20 испытаний участники делали небольшой перерыв.

Таким образом, 50 испытуемых посмотрели 40 музыкальных клипов, сделано $50 \times 40 = 2000$ записей ЭЭГ (4 класса, по 500 файлов на каждое эмоциональное состояние). Итоговая размерность подготовленных данных составила $40 \times 8 \times 8064$ (где 40 – количество видеоклипов, 8 – количество каналов ЭЭГ, 8064 – 60 секунд записи ЭЭГ с частотой дискретизации 128 Гц).

3. Методы анализа данных и архитектура нейросетевой модели искусственного интеллекта для классификации эмоций

Результаты проведенного обзора были учтены при построении нейросетевой модели искусственного интеллекта в настоящем исследовании.

Важным шагом при анализе сигналов ЭЭГ как в клинической, так и в исследовательской практике, является этап предобработки. На данном этапе временные волны данных ЭЭГ проверялись визуально. Записи, сильно загрязненные электромиограммой (ЭМГ), были удалены вручную. Помехи от сети электропитания были устранены наложением режекторного фильтра (общепринятая практика) с частотой среза 45 до 55 Гц. Затем сигналы ЭЭГ были разделены на части длиной 2,5 секунды с помощью окна с перекрывающимся шагом.

Согласно ранее проведенным исследованиям [30, 31, 32] удалось определить, что ЭЭГ содержит информацию как об индивидуальных особенностях работы мозга субъекта, так и о состоянии субъекта, в том числе зависящем от совершаемых действий в реальном времени. В результате анализа амплитудно-частотных характеристик (АЧХ) сигналов ЭЭГ было определено, что на разных устройствах записи ЭЭГ информативность АЧХ сопоставима, а сигналы электроэнцефалограмм почти равноинформативны, при

этом их АЧХ по отдельности малоинформативны, поэтому необходим мощный инструмент обогащения данных, такой как сверточные нейронные сети.

По итогам проведенного систематического обзора научной публицистики решено использовать архитектуры сверточных нейронных сетей для решения задачи распознавания эмоций, так как они дают наиболее высокие результаты при анализе ЭЭГ. Предложено обучать сеть на данных ЭЭГ, записанных субъектом в разные дни, т.к. это позволит значительно повысить надежность идентификации пользователя по данным сигналов ЭЭГ [30].

Все эксперименты реализованы с использованием фреймворка Tensorflow версии 2.11.0 и Python версии 3.10.10. В качестве рабочей станции использовалась GIGABYTE B560 HD3, в состав которой входили Intel i7-11700F (8 CPU @ 2,5 ГГц), NVIDIA GeForce RTX 3060 6 Гб VRAM, 500 Гб SSD, и 32 Гб RAM.

Сформированы две сверточные нейронные сети с идентичной архитектурой, представленной на рис. 5. Отправной точкой для поиска оптимальной архитектуры послужила нейронная сеть EEGNet [33], которая доказала свою эффективность в задачах распознавания моторных стимулов.

Однако в задачах распознавания эмоций следует искать более сложные паттерны в данных, с чем прекрасно справляются глубокие CNN-модели, рассмотренные в работе [34]. Но недостаток обучающей выборки и сложность оптимизации таких сетей [34] не позволяют применять их в задачах с малой выборкой (набор данных из 2000 образов можно назвать относительно малым для таких сетей). Поэтому нами разработана архитектура, основанная на нейросетевых решениях из работ [33–34], которая является значительно более глубокой, чем EEGNet (25 слоёв, против 14), и имеет большее количество фильтров.

Шкалы валентности и арозальности были условно разделены на два интервала равных, представляющих собой области низкой и высокой оценок по соответствующей шкале. Каждая нейронная сеть обучалась задаче бинарной классификации, а именно предсказанию самооценки по одной из размерностей шкалы Рассела: низкая либо высокая валентность и низкая либо высокая арозальность. Обе нейронные сети объединены в мультиклассификатор для эффективного различения 4 состояний в соответствии со шкалой валентности-арозальности Рассела: HVNA, HVLA, LVNA, LVLA (рис. 1), таким образом можно распознать 4 вида эмоций (страх, счастье, печаль, спокойствие).

Использование двух нейронных сетей вместо одной общей позволяет снизить размерность задачи (перейти от задачи классификации 4 классов к двум задачам бинарной классификации), тем самым повысить эффективность обучения на малых выборках (каждая нейронная сеть обучается на одной и той же выборке, но для решения разных задач).

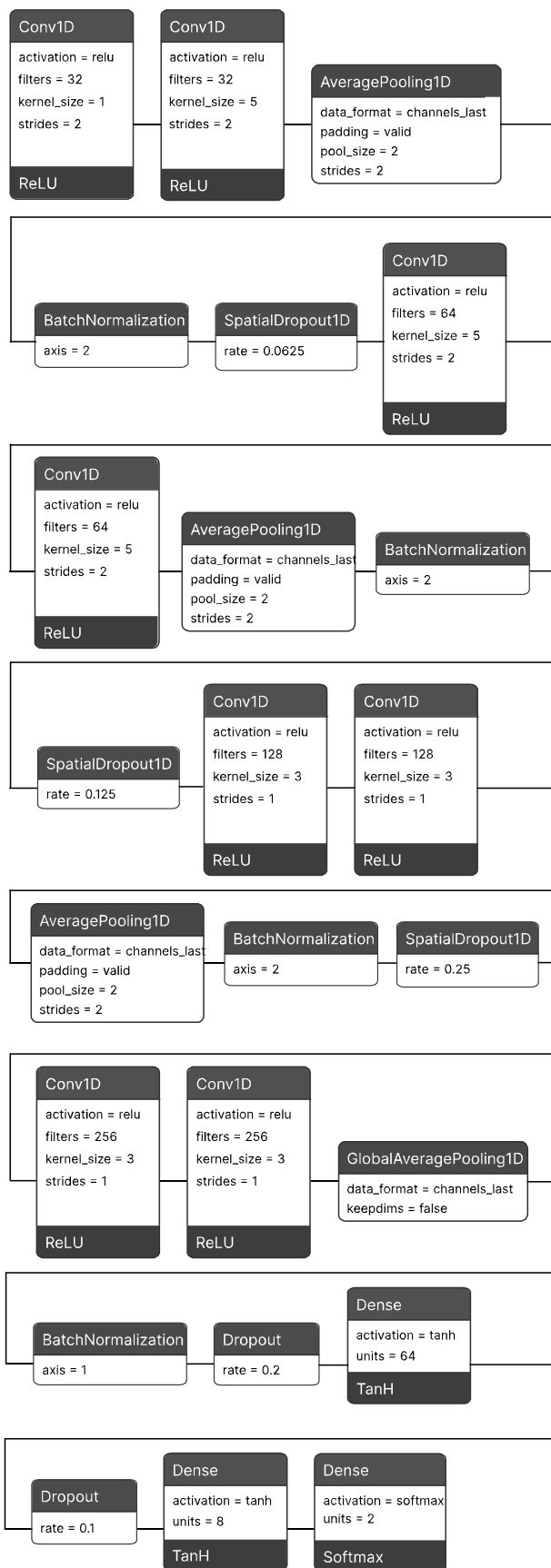


Рис. 5. Архитектура предложенной нейронной сети

Перед классификацией сигналы обрабатывались полосовыми фильтрами с целью выделить дельта- (δ),

тета- (θ), альфа- (α), бета- (β) и гамма- (γ) ритмы, специфичные для обнаружения эмоциональных состояний. Тестовая выборка была сформирована случайным образом и составила 15% от общего объема данных. Обучающая и валидационная выборки формировались путём 10-fold скользящего контроля при обучении каждой модели.

На вход ИНС образы подавались в виде исходных сигналов. Амплитуды образов были нормализованы согласно Z-преобразованию и приведены к виду с нулевым средним и единичным стандартным отклонением. Этот этап выполняется для предотвращения переобучения и повышения точности модели.

4. Обучение и тестирование нейросетевой модели искусственного интеллекта для классификации эмоций

При обучении нейронных сетей окончательно определенный размер минибатча составил 256. В качестве оптимизатора был выбран Adam, а в качестве функции потерь – бинарная перекрестная энтропия. Количество обучаемых параметров представленной нейронной сети составило 425370.

Обучение с промежуточной валидацией для каждой нейронной сети повторялось многократно, каждый раз создавалась новая нейронная сеть, обучаемая независимо, которая проходила процедуру 10-fold скользящего контроля (рис. 6 – 7).

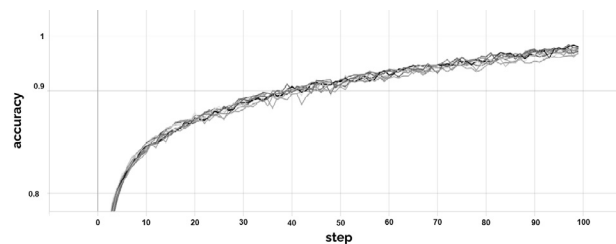


Рис. 6. Динамика изменения функции точности (accuracy) обучения (100 эпох) нейронной сети для классификации эмоций по шкале валентности (положительная или отрицательная) с 10-кратным скользящим контролем (каждая линия ассоциируется с отдельной попыткой обучения и тестирования)

В результате эксперимента модель показала точность классификации эмоций в $94\% \pm 3,4\%$.

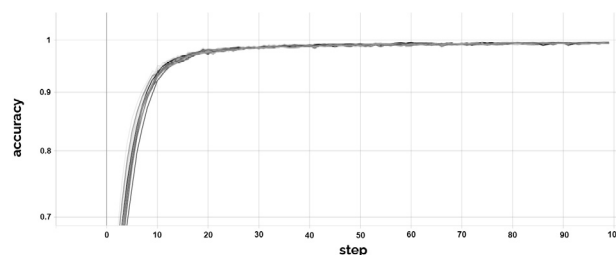


Рис. 7. Динамика изменения функции точности (accuracy) обучения (100 эпох) нейронной сети для классификации эмоций по шкале возбуждения (низкое/высокое) с 10-кратным скользящим контролем (каждая линия ассоциируется с отдельной попыткой обучения и тестирования нейронной сети)

После анализа результатов видно, что разработанная модель распознавания эмоций на основе сверточных нейронных сетей превосходит существующие модели, основанные на классических алгоритмах анализа сигналов, а также другие известные результаты, полученные на базе SVM [35 – 37], Bayesian DNN [38], GCB-Net [39] и др. (табл. 1). Это указывает на то, что предложенная архитектура и нейросетевая модель эффективны.

Табл. 1. Сопоставительные данные по эффективности методов распознавания эмоционального состояния образов ЭЭГ субъектов (*V* – валентность, *A* – арозальность)

Методы	Точность	Набор данных
Свёрточная нейронная сеть [20]	90,24 % (<i>V</i>), 89,45 % (<i>A</i>)	DEAP, 32 человека
Сверточная нейронная сеть [21]	84 % (<i>V</i>), 85,8 % (<i>A</i>)	AMIGOS, 40 человек
1D сверточная нейронная сеть с LSTM ячейками [23]	97,8 % – 4 класса	DEAP, 32 человека
Статистические признаки и SVM [35]	67 % (<i>V</i>), 68,8 % (<i>A</i>)	AMIGOS, 40 человек
Преобразование Фурье и SVM с радиальной базисной функцией [36]	83,9 % (<i>V</i>), 65 % (<i>A</i>)	AMIGOS, 40 человек
Статистические признаки и SVM [37]	61,84 % (<i>V</i>), 63,32 % (<i>A</i>)	AMIGOS, 40 человек
Байесовская глубокая сеть (Bayesian DNN) [38]	90 % (<i>V</i>)	AMIGOS, 40 человек
Свёрточная широкая сеть графов (GCB-Net) [39]	86,99% (<i>V</i>), 89,32% (<i>A</i>)	DREAMER, 23 человека
Свёрточная нейронная сеть с динамическим графом (DGCNN) [40]	86,23 % (<i>V</i>), 84,54 % (<i>A</i>)	DREAMER, 23 человека
Свёрточная нейронная сеть [41]	84 % (<i>V</i>), 85,8 % (<i>A</i>)	AMIGOS, 40 человек
Рекуррентная нейронная сеть с LSTM-ячейкам (LSTM-RNN) [42]	79,4 % (<i>V</i>), 83,3 % (<i>A</i>)	AMIGOS, 40 человек
Свёрточная нейронная сеть (3D-CNN) [43]	99,7 % – 4 класса	AMIGOS, 40 человек
Свёрточная нейронная сеть CNN-VGG16 [44]	55,56 %	DREAMER, 23 человека
Рекуррентная нейронная сеть (LSTM + CNN) [45]	90,8 % – DREAMER 99 % – AMIGOS	DREAMER, 23 человека, AMIGOS, 40 человек
Разработанная нейросетевая модель на базе двух CNN	94 % ± 3,4 %	50 человек

Заключение

В ходе проведенных исследований разработана методика проведения эксперимента по сбору данных ЭЭГ пользователей. Собрана база данных для анализа эмоционального состояния человека по ЭЭГ. В экс-

перименте приняло участие 50 человек, каждый из которых просматривал 40 эмоционально окрашенных видеороликов длительностью в 1 минуту.

Разработана нейросетевая модель классификации эмоционального состояния, основанная на двух глубоких сверточных нейронных сетях с идентичной архитектурой, предложенной в настоящей работе. В результате эксперимента модель показала точность классификации эмоций в 94 % ± 3,4 %, что соответствует мировому уровню и превосходит большинство достигнутых ранее результатов.

В будущем планируется провести оптимизацию предложенной модели, применив методы сжатия нейронных сетей, с целью ускорения процесса обучения и снижения потребляемых вычислительных ресурсов, что особенно важно для использования модели на низкопроизводительных устройствах. Возможности использования сжатой нейронной сети на мобильных устройствах значительно расширяют области применения и доступность технологии анализа эмоций на основе ЭЭГ, делая её более удобной для конечных пользователей. Также планируется расширить нейросетевую модель классификации с помощью третьей нейронной сети, которая оценивает силу эмоции по шкале доминантности.

Благодарности

Работа выполнена ОмГТУ в рамках государственного задания Минобрнауки России на 2023 – 2025 годы (FSGF-2023-0004).

References

- [1] Shan MK, Kuo FF, Chiang MF, Lee SY. Emotion-based music recommendation by affinity discovery from film music. *Expert Syst Appl* 2009; 36(4): 7666-7674.
- [2] Anderson K, McOwan PW. A real-time automated system for the recognition of human facial expression. *IEEE Trans Syst Man Cybern B: Cybern* 2006; 36(1): 96-105.
- [3] Ang J, Dhillon R, Krupski A, Shriberg E, Stolcke A. Prosody-based automatic detection of annoyance and frustration in human-computer dialog. *Int Conf on Spoken Language Processing* 2002: 2037-2039.
- [4] Craik A. Deep learning for electroencephalogram (EEG) classification tasks: a review, *J Neur Eng* 2019; 3(16). DOI: 10.1088/1741-2552/ab0ab5.
- [5] Sulavko AE, Kuprik MA, Starkov AI, Stadnikov DG. Analysis of human pattern recognition methods based on electroencephalogram features (review) [In Russian], *Information Protection Issues* 2018; 4: 36-46.
- [6] Vance A, Anderson BB, Kirwan BC, Eargle D. Using measures of risk perception to predict information security behavior: Insights from electroencephalography (EEG). *J Assoc Inf Syst* 2014; 15(10): 679-722. DOI: 10.17705/1jais.00375.
- [7] Nigray AA. Methods of automatic assessment of human psychophysiological state by parameters of electroencephalograms (review) [In Russian]. *Biomedical Radioelectronics* 2020; 5: 5-18.
- [8] Khushaba RN, Greenacre L, Kodagoda S, Louviere J, Burke S, Dissanayake G. Choice modeling and the

- brain: A study on the Electroencephalogram (EEG) of preference. *Expert Syst Appl* 2012; 39(16): 12378-12388.
- [9] Ekman P, Friesen WV, O'Sullivan M, Chan A, DiacoyanniTarlatis I, Heider K, Krause R, LeCompte WA, Pitcairn T, Ricci-Bitti PE. Universals and cultural differences in the judgments of facial expressions of emotion. *J Pers Soc Psychol* 1987; 53(4): 712-717.
- [10] Parrott WG. *Emotions in social psychology: Essential readings*. Psychology Press; 2001.
- [11] Plutchik R. The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *Am Sci* 2001; 89: 344.
- [12] Russell JA. A circumplex model of affect. *J Pers Soc Psychol* 1980; 39(6): 1161-1178.
- [13] Bradley MM, Lang PJ. Measuring emotion: The self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry* 1994; 25(1): 49-59.
- [14] Nie D, Wang X-W, Shi L-C, Lu B-L. EEG-based emotion recognition during watching movies. *5th Int IEEE EMBS Conf on Neural Engineering* 2011: 667-670. DOI: 10.1109/NER.2011.5910636.
- [15] Li M, Lu BL. Emotion classification based on gamma-band EEG. *IEEE Int Conf Engineering in Medicine and Biology Society* 2009; 1223-1226.
- [16] Wu W, Gao X, Hong B, Gao S. Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). *IEEE Trans Biomed Eng* 2008; 55(6): 1733-1743.
- [17] Terasawa N, Tanaka H, Sakti S, Nakamura S. Tracking liking state in brain activity while watching multiple movies. *19th ACM Int Conf on Multimodal Interaction* 2017: 321-325. DOI: 10.1145/3136755.3136772.
- [18] Shi LC, Lu BL. Off-line and on-line vigilance Estimation based on linear dynamical system and manifold learning. *Proc 32nd Int Conf of the IEEE Engineering in Medicine and Biology Society* 2010: 6587-6590.
- [19] Cheng J, Chen M, Li C, Cheng J. Emotion recognition from multi-channel EEG via deep forest. *IEEE J Biomed Health Inform* 2021; 25(2): 453-464.
- [20] Yang Y, Wu Q, Fu Y, Chen X. Continuous convolutional neural network with 3D input for EEG-based emotion recognition. In Book: Cheng L, Leung ACS, Ozawa S, eds. *Neural information processing. Pt VII*. Cham: Springer Nature Switzerland AG; 2018: 433-443.
- [21] Sarkar P. Self-supervised learning for ecg-based emotion recognition. *IEEE Int Conf on Acoustics, Speech and Signal Processing (ICASSP)* 2020: 3217-3221.
- [22] Balestrieri R, Ibrahim M, Sobal V, Morcos A, Shekhar S, Goldstein T, Bordes F, Bardes A, Mialon G, Tian Y, Schwarzschild A, Wilson AG, Geiping J, Garrido Q, Fernandez P, Bar A, Pirsivash H, LeCun Y, Goldblum M. A cookbook of self-supervised learning. *arXiv Preprint*. 2023. Source: <<https://arxiv.org/abs/2304.12210>>. DOI: 10.48550/arXiv.2304.12210.
- [23] Farhad Z, Retno W. Emotion classification using 1D-CNN and RNN based On DEAP Dataset. *10th Int Conf on Natural Language Processing* 2021: 363-378. DOI: 10.5121/csit.2021.112328.
- [24] Bhanumathi KS, Jayadevappa D, Tunga S. Feedback artificial shuffled shepherd optimization-based deep maxout network for human emotion recognition using EEG signals. *Int J Telemed Appl* 2022; 2022: 3749413. DOI: 10.1155/2022/3749413.
- [25] Schmidt LA, Trainor LJ. Frontal brain electrical activity distinguishes valence and intensity of musical emotions. *Cogn Emot* 2001; 15(4): 487-500.
- [26] Davidson RJ. The neuropsychology of emotion and affective style. In Book: Lewis M, Haviland JM, eds. *Handbook of emotion*. The Guilford Press; 1993: 143-154.
- [27] Fox NA. If it's not left, it's right: Electroencephalograph asymmetry and the development of emotion. *Am Psychol* 1991; 46: 863-872.
- [28] Heller W. Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychology* 1993; 7: 476-489.
- [29] Morris JD. SAM: The self-assessment manikin an efficient cross-cultural measurement of emotional response. *J Advert Res* 1995; 35(8): 63-68.
- [30] Sulavko AE, Lozhnikov PS, Choban AG, Stadnikov DG, Nigrey AA, Inivatov DP. Evaluation of EEG identification potential using statistical approach and convolutional neural networks. *Information and Control Systems* 2020; 6(109): 37-49. DOI: 10.31799/1684-8853-2020-6-37-49.
- [31] Bragin AD, Spitsyn VG. Motor imagery recognition in electroencephalograms using convolutional neural networks. *Computer Optics* 2020; 44(3): 482-487. DOI: 10.18287/2412-6179-CO-669.
- [32] Hodashinsky IA, Sarin KS, Bardamova MB, Svetlakov MO, Slezkin AO, Koryshev NP. Biometric data and machine learning methods in the diagnosis and monitoring of neurodegenerative diseases: a review. *Computer Optics* 2022; 46(6): 988-1019. DOI: 10.18287/2412-6179-CO-1134.
- [33] Lawhern VJ, Solon AJ, Waytowich NR, Gordon SM, Hung CP, Lance BJ. EEGNet: A compact convolutional network for EEG-based Brain-Computer Interfaces. *J Neural Eng* 2018; 15(5): 056013. DOI: 10.1088/1741-2552/aace8c.
- [34] Osherov E, Lindenbaum M. Increasing cnn robustness to occlusions by reducing filter support. *IEEE Int Conf on Computer Vision* 2017: 550-561. DOI: 10.1109/ICCV.2017.67.
- [35] Yang HC, Lee CC. An attribute-invariant variational learning for emotion recognition using physiology. *IEEE Int Conf on Acoustics, Speech and Signal Processing (ICASSP)* 2019: 1184-1188.
- [36] Shukla J, Barreda-Angeles M, Oliver J, Nandi G. Feature extraction and selection for emotion recognition from electrodermal activity. *IEEE Trans Affect Comput* 2021; 12(4): 857-869. DOI: 10.1109/TAFFC.2019.2901673.
- [37] Katsigiannis S. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE J Biomed Health Inform* 2017; 22: 98-107.
- [38] Harper R, Southern J. A Bayesian deep learning framework for end-to-end prediction of emotion from heartbeat. *arXiv Preprint*. 2020. Source: <<https://arxiv.org/abs/1902.03043>>.
- [39] Zhang T, Wang X, Xu X, Chen C. GCB-Net: Graph convolutional broad network and its application in emotion recognition. *IEEE Trans Affect Comput* 2022; 13(1): 379-388. DOI: 10.1109/TAFFC.2019.2937768.
- [40] Song T, Zheng W, Song P, Cui Z. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Trans Affect Comput* 2020; 11(3): 532-541. DOI: 10.1109/TAFFC.2018.2817622.
- [41] Candra H, Yuwono M, Handojoseno A, Chai R, Su S. Recognizing emotions from EEG subbands using wavelet

- analysis. Annual Int Conf of the IEEE Engineering in Medicine and Biology Society (EMBC) 2015: 6030-6033.
- [42] Li C, Bao Z, Li L, Zhao Z. Exploring temporal representations by leveraging attention-based bidirectional LSTM-RNNs for multi-modal emotion recognition. *Inf Process Manag* 2020; 57: 102185. DOI: 10.1016/j.ipm.2019.102185.
- [43] Zhao Y, Cao X, Lin J, Yu D, Cao X. Multimodal emotion recognition model using physiological signals. In Book: Liu D, Xie S, Li Y, Zhao D, El-Alfy E-LM, eds. *Neural Information Processing. Pt IV*. Cham: Springer International Publishing AG; 2017: 811-819. DOI: 10.1007/978-3-319-70093-9_86.
- [44] Siddharth S, Jung T-P, Sejnowski TJ. Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing. *IEEE Trans Affect Comput* 2022; 13(1): 96-107. DOI: 10.1109/TAFFC.2019.2916015.
- [45] Dar MN, Akram MU, Khawaja SG, Pujari AN. CNN and LSTM-based emotion charting using physiological signals. *Sensors* 2020; 20(16): 4551. DOI: 10.3390/s20164551.

Сведения об авторах

Чобан Адиль Гаврилович, 1998 года рождения, в 2016 году окончил Омский государственный технический университет по направлению «Информационная безопасность», аспирант кафедры комплексной защиты информации Омского государственного технического университета по направлению «Когнитивное моделирование». Область научных интересов: биометрия, распознавание образов ЭЭГ. E-mail: adil_choban@mail.ru

Стадников Денис Геннадьевич, 1998 года рождения, в 2016 году окончил Омский государственный технический университет по направлению «Информационная безопасность», аспирант кафедры комплексной защиты информации Омского государственного технического университета по направлению «Искусственный интеллект и машинное обучение». Область научных интересов: биометрия, распознавание образов ЭЭГ. E-mail: sdg250598@inbox.ru

Сулавко Алексей Евгеньевич, 1986 года рождения, в 2009 году окончил Сибирскую государственную автомобильно-дорожную академию по специальности «Комплексное обеспечение информационной безопасности автоматизированных систем», кандидат технических наук, доцент кафедры комплексной защиты информации Омского государственного технического университета. Область научных интересов: биометрия, распознавание образов, машинное обучение, искусственный интеллект, защита информации. E-mail: sulavich@mail.ru

ГРНТИ: 28.23.25

Поступила в редакцию 29 августа 2023 г. Окончательный вариант – 26 января 2024 г.

Neural network in an artificial intelligence model for realization of affective computing based on electroencephalogram analysis

A.G. Choban¹, D.G. Stadnikov¹, A.E. Sulavko¹

¹ Omsk State Technical University, 644050, Omsk, Russia, Mira 11

Abstract

This paper analyzes the possibility of assessing the emotional state of a person by analyzing features of their brain activity using non-invasive neuro-computer interfaces. The analysis of recent publications dealing with the use of electroencephalogram (EEG) signals for assessing the emotional state is carried out and topical problems in this area are revealed. The main approaches to brain stimulation for obtaining informative EEG signals, as well as describing methods for their analysis and recognition. The architecture of a deep convolutional neural network for EEG data analysis is proposed, as well as a neural network model for classifying 4 emotions (fear, happiness, sadness, calmness) according to the Russell valency-arozality scale based on two convolutional neural networks. An experiment was conducted with 50 participants who watched emotion-laden videos. EEG data from 50 subjects were collected and used to train and test the neural network model. The results showed a high emotion classification accuracy of $94\% \pm 3.4\%$ while using a wireless neural interface.

Keywords: emotion recognition, electroencephalogram, convolutional neural networks, EEG signal, brain-computer interface, neural interface, analysis of biometric parameters.

Citation: Choban AG, Stadnikov DG, Sulavko AE. Neural network in an artificial intelligence model for realization of affective computing based on electroencephalogram analysis. *Computer Optics* 2024; 48(5): 782-790. DOI: 10.18287/2412-6179-CO-1417.

Acknowledgements: The research was financially supported by the Ministry of Science and Higher Education of the Russian Federation (theme No. FSGF-2023-0004).

Authors' information

Adil G. Choban (b. 1998), graduated from Omsk State Technical University in 2016 in the direction of "Information Security", postgraduate student of Integrated Information Protection department of Omsk State Technical University in the direction of "Cognitive Modeling". Research interests: biometrics, EEG pattern recognition. E-mail: adil_choban@mail.ru

Denis G. Stadnikov (b. 1998), graduated from Omsk State Technical University in 2016 in the direction of "Information Security", postgraduate student of Integrated Information Protection department of Omsk State Technical University in the direction of "Artificial Intelligence and Machine Learning". Research interests: biometrics, EEG pattern recognition. E-mail: sdg250598@inbox.ru

Alexey E. Sulavko (b. 1986) in 2009 graduated from the Siberian State Automobile and Highway Academy with a degree in Integrated Information Security of Automated Systems. Candidate of Technical Sciences, Assistant Professor, Complex Information Protection department of Omsk State Technical University. Research interests are biometry, pattern recognition, machine learning, artificial intelligence, information security. E-mail: sulavich@mail.ru

Code of State Categories Scientific and Technical Information (in Russian – GRNTI): 28.23.25

Received August 29, 2023. The final version – January 26, 2024.
