

Библиографическое описание согласно ГОСТ Р 7.0.100–2018

Применение компьютерного зрения для определения реперных точек при оценке нарушения осанки / И. Д. Шитоев, В. Н. Никитин, М. Д. Иванова, Г. З. Клоян, С. В. Муравьев. – Текст : непосредственный. – DOI 10.15593/2499-9873/2023.4.06 // Прикладная математика и вопросы управления / Applied Mathematics and Control Sciences. – 2023. – № 4. – С. 94–106.



**ПРИКЛАДНАЯ МАТЕМАТИКА
И ВОПРОСЫ УПРАВЛЕНИЯ**
№ 4, 2023

<https://ered.pstu.ru/index.php/amcs>



Научная статья

DOI: 10.15593/2499-9873/2023.4.06

УДК 004.942



Применение компьютерного зрения для определения реперных точек при оценке нарушения осанки

И.Д. Шитоев^{1,2}, В.Н. Никитин^{1,2}, М.Д. Иванова¹, Г.З. Клоян², С.В. Муравьев^{3,4}

¹Пермский национальный исследовательский политехнический университет, Пермь, Российская Федерация

²ООО «Йорд Тех», Пермь, Российская Федерация

³Пермский государственный медицинский университет имени академика Е.А. Вагнера Пермь, Российская Федерация

⁴Научно-медицинский центр «Родник», Пермь, Российская Федерация

О СТАТЬЕ

Получена: 27 ноября 2023
Одобрена: 14 декабря 2023
Принята к публикации:
15 декабря 2023

Финансирование

Работа выполнена при финансовой поддержке Пермского научно-образовательного центра мирового уровня «Рациональное недропользование».

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Вклад авторов

равноценен.

Ключевые слова:

нейронная сеть, нарушение осанки, реперные точки, обучение, компьютерное зрение, сколиоз, спина.

АННОТАЦИЯ

Применение компьютерного зрения для определения реперных точек на спине человека является относительно новым подходом в медицине. Актуальность данного исследования обусловлена не только необходимостью разработки новых методов и подходов в определении нарушения осанки, но и растущим интересом к развитию и применению искусственного интеллекта в медицине. Настоящее исследование посвящено созданию и обучению нейронной сети для нахождения реперных точек спины при диагностике нарушений осанки. Использована совокупность готовых нейросетевых инструментов, находящихся в свободном доступе. Обучение проводилось на моделях фотографий нейронной сетью достигает 85 %, что свидетельствует о хорошем определении границ объектов и их классификации. Применение компьютерного зрения для определения реперных точек для оценки нарушения осанки может принести значительные преимущества в медицинской практике. Этот инструмент может быть более точным и эффективным, чем традиционные методы диагностики, а также более доступным и удобным для пациента. Кроме того, использование нейронных сетей может ускорить процесс диагностики и снизить затраты на проведение исследования.

© Шитоев Иван Дмитриевич – ассистент кафедры вычислительной математики, механики и биомеханики; генеральный директор, e-mail: ShitoevID@yord.tech, ORCID: 0000-0002-6391-9271.

Никитин Владислав Николаевич – кандидат физико-математических наук, доцент кафедры вычислительной математики, механики и биомеханики; технический директор, e-mail: nikitinvladislav86@gmail.com, ORCID: 0000-0001-9652-9088.

Иванова Мария Дмитриевна – аспирант кафедры вычислительной математики, механики и биомеханики, e-mail: ivanova-mashaa@mail.ru, ORCID: 0000-0002-1471-8326.

Клоян Гаянэ Зурабиевна – биомеханик, e-mail: kloyang@mail.ru, ORCID: 0000-0001-6615-8159.

Муравьев Сергей Владимирович – кандидат медицинских наук, доцент кафедры физической и реабилитационной медицины, ведущий специалист международного отдела управления международного медицинского образования; научный руководитель, e-mail: sergey89.m@mail.ru, ORCID: 0000-0002-3342-4710.



Эта статья доступна в соответствии с условиями лицензии Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Perm Polytech Style: Shitoev I.D., Nikitin V.N., Ivanova M.D., Kloyan G.Z., Muravev S.V. Application of computer vision to identify reference points for assessing postural disorders. *Applied Mathematics and Control Sciences*. 2023, no. 4, pp. 94–106. DOI: 10.15593/2499-9873/2023.4.06

MDPI and ACS Style: Shitoev, I.D.; Nikitin, V.N.; Ivanova, M.D.; Kloyan, G.Z.; Muravev S.V. Application of computer vision to identify reference points for assessing postural disorders. *Appl. Math. Control Sci.* **2023**, 4, 94–106. <https://doi.org/10.15593/2499-9873/2023.4.06>

Chicago/Turabian Style: Shitoev, Ivan D., Vladislav N. Nikitin, Maria D. Ivanova, Gayane Z. Kloyan, and Sergey V. Muravev. 2023. "Application of computer vision to identify reference points for assessing postural disorders". *Appl. Math. Control Sci.* no. 4: 94–106. <https://doi.org/10.15593/2499-9873/2023.4.06>



APPLIED MATHEMATICS
AND CONTROL SCIENCES
№ 4, 2023
<https://ered.pstu.ru/index.php/amcs>



Article

DOI: 10.15593/2499-9873/2023.4.06

UDC 004.942



Application of computer vision to identify reference points for assessing postural disorders

I.D. Shitoev^{1,2}, V.N. Nikitin^{1,2}, M.D. Ivanova¹, G.Z. Kloyan², S.V. Muravev^{3,4}

¹Perm National Research Polytechnic University, Perm, Russian Federation

²LLC "Yord Tech", Perm, Russian Federation

³Perm State Medical University named after Academician E.A. Wagner, Perm, Russian Federation

⁴Scientific and Medical Center "Rodnik", Perm, Russian Federation

ARTICLE INFO

Received: 27 November 2023
Approved: 14 December 2023
Accepted for publication:
15 December 2023

Funding

This research was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation in the framework of the program of activities of the Perm Scientific and Educational Center "Rational Subsoil Use".

Conflicts of Interest

The authors declare no conflict of interest.

Authors Contributions
equivalent.

Keywords:

neural network, poor posture, reference points, learning, computer vision, scoliosis, back.

ABSTRACT

The use of computer vision to determine reference points on the human back is a relatively new approach in medicine. The relevance of this study is due not only to the need to develop new methods and approaches in determining postural disorders, but also to the growing interest in the development and application of artificial intelligence in medicine. This study is devoted to the creation and training of a neural network to find reference points of the back when diagnosing postural disorders. In this study, we use a set of ready-made neural network tools that are freely available. Training was carried out on models of photographed volunteers to determine the reference points of the back. The results of the study indicate that, as part of the study, a search was made for alternative (reference points) that determine the geometry of the back using developed neural network algorithms. It was shown that the accuracy of point detection by a neural network trained on more than 2400 grouped photographs reaches 85%, which indicates a good determination of object boundaries and their classification. The use of computer vision to identify reference points for assessing postural abnormalities can bring significant benefits to medical practice. This tool may be more accurate and efficient than traditional diagnostic methods, as well as more accessible and convenient for the patient. In addition, the use of neural networks can speed up the diagnostic process and reduce research costs.

© Ivan D. Shitoev – Assistant, Department of Computational Mathematics, Mechanics, Biomechanics; The Chief Executive Officer, e-mail: ShitoevID@yord.tech, ORCID: 0000-0002-6391-9271.

Vladislav N. Nikitin – CSs of Physics and Mathematics Sciences, Associate Professor, Department of computational Mathematics, Mechanics, Biomechanics; A Chief Technical Officer, e-mail: nikitinvladislav86@gmail.com, ORCID: 0000-0001-9652-9088.

Maria D. Ivanova – Ph. D. Student, Department of Computational Mathematics, Mechanics, Biomechanics, e-mail: ivanovamashaa@mail.ru, ORCID: 0000-0002-1471-8326.

Gayane Z. Kloyan – biomechanist, e-mail: kloyang@mail.ru, ORCID: 0000-0001-6615-8159.

Sergey V. Muravev – CSs of Medical Sciences, Associate Professor, Department of Physical and Rehabilitation Medicine, Leading Specialist of the International Department of the International Medical Education; Scientific Director, e-mail: sergey89.m@mail.ru, ORCID: 0000-0002-3342-4710.



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

Введение

Сколиоз – одно из самых распространенных заболеваний позвоночного столба у детей и подростков [1]. С точки зрения пространственной геометрии позвоночник представляет собой кривую, которая в состоянии нормы имеет ряд физиологических изгибов в сагиттальной плоскости – кифозы (грудной и формально крестцовый) и лордоза (поясничный и шейный) [2]. Актуальные инструменты для диагностики позвоночного столба представлены двумя основными группами: лучевые и оптические методы. Данные методы имеют ряд недостатков: при проведении диагностики лучевым методом пациент облучается низкодозовыми рентгеновскими лучами, в то же время принцип «снижения квалификационных требований» к оператору рентгеновских и оптических топографических установок в полной мере применен быть не может, поскольку каждая из них требует участия медицинских кадров высокого уровня профессиональной подготовки [3], которые ограничивают их применение в лечебно-диагностическом процессе. В настоящее время интенсивно развивающееся направление оптической диагностики трехмерных объектов методом фотограмметрии все больше интегрируется в лечебно-диагностический процесс, что позволяет в перспективе использовать этот метод в оценке выраженности деформации позвоночника у детей и подростков.

Биометрические технологии становятся все более популярными в системах компьютерного зрения, представляя инновационные методы, которые используют физические и поведенческие характеристики человека для идентификации и аутентификации [4–6]. Данные технологии являются бесконтактными и неинвазивными, в их основе лежит распознавание биометрических характеристик при помощи компьютерного зрения.

Настоящее исследование посвящено разработке, тестированию и анализу параметров спины человека с использованием нейросетевых алгоритмов. Разработанный продукт по фотографии спины человека определяет реперные точки на ее поверхности на основании созданного нейросетевого алгоритма при диагностике сколиоза.

Цифровой инструмент может стимулировать повышение уровня заинтересованности пользователей в самоконтроле состояния опорно-двигательного аппарата и своевременном обращении за профессиональной медицинской помощью.

Разрабатываемая программная система обладает необходимым набором инструментов для предоставления точного и количественного анализа, способна решить перечисленные проблемы и тем самым усовершенствовать процесс наблюдения и диагностики, а также профилактики.

Цель исследования – улучшение выявления нарушений осанки на более ранних стадиях путем создания и обучения нейронной сети для нахождения реперных точек спины.

Материалы и методы: для обучения нейронной сети использовалась выборка из 3000 размеченных фотографий, находящихся в открытом доступе, на которых изображена спина человека, а также данные учащихся начальных классов, проживающих в ЗАТО «Звездный», Пермский край. Настоящая работа была проведена в рамках мониторингового исследования совокупности испытуемых – школьников младшего школьного возраста (1–3-й классы, средний возраст $8 \pm 0,78$ года) и взрослых (средний возраст 18–40 лет) – представителей европеоидной расы. До проведения исследования были получены информированные согласия на диагностическое вмешательство от всех родителей испытуемых (детей). Разметка фотографий производилась с помощью инструмента графического анализа изображений `labelimg1.8.6`, в котором в режиме графического интерфейса вручную отмечались области нахождения (классы определения) реперных точек. В качестве архитектуры нейронной сети выбрана YOLO 8.

Определение ключевых реперных точек

Современные инструменты в диагностике деформации позвоночника у детей [7] представлены двумя основными группами, в равной степени обладающими недостатками, не позволяющими им быть максимально широко интегрированными в лечебно-диагностический процесс.

Основная группа методов диагностики, на которую опирались в своем исследовании по определению реперных точек – оптические методы, построенные на анализе изображений дорсальной поверхности тела человека с наложенными теневыми «полосами», которые искажаются в зависимости от особенностей деформации этой поверхности. Передовыми диагностическими установками в этой области являются компьютерная оптическая топография (далее КОТ, КомОТ, ТОДП, ООО «МЕСТОС», г. Новосибирск) [8] и система DIERS (DIERS International GmbH, Schlangenbad, Germany) [9, 10].

Поиск реперных точек осуществляется по образцу КомОТ. Данный метод позволяет дистанционно и бесконтактно определять форму поверхности туловища пациента, которому соответствуют биомеханические нарушения опорно-двигательного аппарата. С помощью этого метода определяются анатомические ориентиры костных структур, на которых основаны графические представления и количественные параметры, описывающие состояние осанки и формы позвоночника. Опираясь на данные результаты, были определены следующие реперные точки спины человека, представленные в табл. 1.

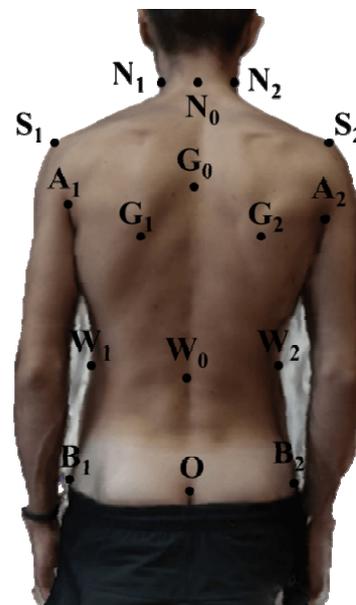


Рис. 1. Реперные точки

Таблица 1

Реперные точки, идентифицируемые на поверхности тела человека

№ п/п	Условное обозначение	Интерпретация	Пояснение
1	N_0	Срединная точка задней поверхности шеи	Соответствует расположению остистого отростка C_{VII}
2	N_1	Левая точка задней поверхности шеи	Соответствует точке, наименее удаленной от срединной линии задней поверхности шеи слева
3	N_2	Правая точка задней поверхности шеи	Соответствует точке, наименее удаленной от срединной линии задней поверхности шеи справа
4	A_1	Точка левой подмышечной впадины	–
5	A_2	Точка правой подмышечной впадины	–
6	S_1	Точка левого плечевого сустава	–
7	S_2	Точка правого плечевого сустава	–
8	G_0	Вершина грудного кифоза	–
9	G_1	Угол левой лопатки	Здесь: «угол» – топографо-анатомическое понятие
10	G_2	Угол правой лопатки	
11	W_0	Вершина поясничного кифоза	–
12	W_1	Левая точка талии	Соответствует точке, наименее удаленной от срединной линии задней поверхности спины, точке в левой поясничной области спины
13	W_2	Правая точка талии	Соответствует точке, наименее удаленной от срединной линии задней поверхности спины, точке в правой поясничной области спины
14	B_1	Левая точка ягодиц	Соответствует точке, наиболее удаленной от срединной линии левой ягодичной области
15	B_2	Правая точка ягодиц	Соответствует точке, наиболее удаленной от срединной линии правой ягодичной области
16	O	Вершина ягодичной складки	Соответствует точке начала координат плоскостей, в том числе при коррекции их положения

В рамках исследования осуществляется поиск 16 основных точек, обозначенных главными латинскими буквами с индексом (рис. 1). Выбор точек зависит от решаемой задачи и для текущего исследования является полным.

Для определения реперных точек был разработан алгоритм по определению специальных точек на трехмерной модели, пространственное положение которых относительно друг друга позволяло сделать вывод о наличии тех или иных нарушений осанки.

Разработка нейросетевого алгоритма

Для качественного поиска реперных точек на плоских снимках использовалась совокупность готовых нейросетевых инструментов, находящихся в свободном доступе. Обучение проводилось на выборке из 2400 размеченных фотографий, находящихся в открытом доступе, на которых изображена спина человека, а также данных обследования учащихся начальных классов, проживающих в ЗАТО «Звездный», Пермский край.

Далее для поиска реперных точек применялась нейросетевая обработка пространственной модели спины. Для создания виртуальной копии трехмерного объекта использовался метод фотограмметрии [11], включающий в себя съемку неподвижного объекта камерой смартфона с разных ракурсов с последующим восстановлением объемной модели по фотографиям на основе их уникальных признаков.

Так как распознавание спины является сложной задачей в силу того, что объемное трехмерное изображение спины содержит большое количество информации, обработка и интерпретация которой делает нейронную сеть дорогостоящей в вычислительном отношении [12], было принято решение проводить поиск реперных точек на плоском снимке, выполненном анфас, и в дальнейшем переносить на трехмерный объект. Также на этом этапе работ был осуществлен переход от готовых решений к разработке и обучению своей нейронной сети.

Для глубокого обучения были использованы популярные наборы инструментов, библиотек и средств разработки с открытым исходным кодом, такие как *keras*, *yolo*, *tenserflow*, предназначенные для распознавания изображений и обнаружения на них объектов [7; 13–15]. К преимуществам *tenserflow*, *keras* относится возможность запуска на большом количестве устройств. Архитектура *yolo* отлично подходит для распознавания объектов в видео [7; 16] и является наилучшим вариантом для обнаружения и распознавания образов (в том числе в реальном времени), так как отличается высокой скоростью и точностью обнаружения объектов, поддерживается на многих операционных системах и платформах, в том числе на *Android* и *Raspberry Pi*.

Создание оригинальной нейросетевой модели

Для создания оригинальной модели алгоритма компьютерного зрения, направленной на поиск реперных точек, использовалась предобученная модель с последующим дообучением на выбранном наборе данных на базе архитектуры YOLO.

Процесс создания модели включал несколько этапов:

1. Сбор и подготовка данных для обучения оригинальной модели. Данные представляют собой набор из 3000 различных фотографий спины школьников младшего школьного возраста и взрослых, представителей европеоидной расы. До проведения исследования были получены информированные согласия на диагностическое вмешательство от всех родителей испытуемых (детей).

2. Выбор архитектуры. В качестве архитектуры нейронной сети выбрана YOLO 8. Наибольшим преимуществом YOLO над другими архитектурами является скорость. Мо-

дели обнаружения семейства YOLO (одноэтапные) исключительно быстры и намного превосходят R-CNN (двухэтапные) и другие модели [11], что позволяет добиться обнаружения объектов в режиме реального времени.

3. Обучение модели производилось на 2400 фотографиях из набора собранных данных с размеченными областями нахождения интересующих реперных точек.

4. Оценка и тестирование модели. Тестирование производилось на 600 фотографиях из набора собранных данных и не включенных в группу обучения. Для оценки качества модели использовались различные метрики.

5. Настройка гиперпараметров. Выполнена настройка свойств, управляющих общей архитектурой и настройкой модели, а не непосредственно значением весов или смещений на уровне нейронов. Они устанавливаются до начала обучения и в то же время влияют на процесс обучения. Для достижения лучшей производительности, повышения точности и эффективности модели в процессе обучения выбраны следующие значения гиперпараметров:

а) скорость обучения – параметр, который определяет величину изменения весовой матрицы во время обратного распространения ошибки. Он задает величину, на которую градиент должен быть умножен при обновлении весов. Выбор скорости подбирается, исходя из сложности задачи, и в данном случае оптимальное значение равно 0,001;

б) момент – параметр, который позволяет модели преодолевать локальные минимумы и перемещаться быстрее к глобальному минимуму. Момент добавляет предыдущее изменение весовой матрицы в текущее изменение. Таким образом, ускорение процесса обучения зависит от момента. Значение этого параметра также зависит от задачи и в данном случае равняется 0,9;

в) размер батча отвечает за автоматическую классификацию данных и определяет количество примеров, передаваемых в модель для обновления весов на каждой итерации. Большие размеры батча могут ускорить процесс обучения, однако могут занимать больше ресурсов, потому что большое количество данных передается по сети, а также большой размер батча может увеличить шанс на переобучение. В задаче определения реперных точек использовалась малая выборка и ручная классификация изображений, поэтому размер батча равняется 4;

г) количество эпох обучения – количество полных проходов через обучающую выборку данных при обучении модели. Большее количество эпох может привести к лучшей точности модели, но также и к переобучению. С другой стороны, меньшее количество эпох может привести к недостаточной обученной модели. Оптимальное количество эпох подбиралось экспериментально, путем анализа выходных метрик – 150 эпох;

д) размер входного изображения определяется количеством пикселей изображения по горизонтальной и вертикальной осям и оказывает влияние на точность модели. Если размер изображения слишком большой, то могут возникнуть проблемы с производительностью и временем обучения модели. Оптимальный размер входных данных был определен экспериментально путем постепенного увеличения размера изображения с 640×640 до 1280×1280 и отслеживания влияния на точность модели;

е) функция активации представляет собой нелинейную функцию, которая применяется к результату линейной операции весов и входных данных каждого нейрона. Функция активации определяет выходное значение нейрона и, таким образом, вносит нелинейность в нейронную сеть. В данном исследовании применяется сигмоидная функция активации, которая часто используется в нейронных сетях для задач бинарной классификации, принимает значение от 0 до 1.

Все перечисленные гиперпараметры влияют на процесс обучения и определяют, как быстро и эффективно модель сможет сходиться к минимуму функции потерь и избежать переобучения.

В ходе идентификации обученный алгоритм делит изображения на квадратную сетку. Для каждой ячейки нейронной сети выводятся вероятности определяемого класса. Каждая ячейка сетки отвечает за обнаружение центра объекта внутри ячейки и может предсказать фиксированное количество ограничивающих рамок со значением достоверности (уверенности). Каждая ограничивающая рамка включает в себя пять значений: координата центра по оси абсцисс, координата центра по оси ординат, высота рамки, ширина рамки и значение достоверности. Ячейки, имеющие вероятность класса выше порогового значения, выбираются и используются для определения местоположения объекта на изображении.

Ограничения нейросетевого алгоритма

Для повышения качества поиска было принято решение использовать совокупность готовых нейросетевых инструментов, находящихся в свободном доступе. Обучение проводилось на моделях добровольцев, фотографируемых для определения реперных точек.

Далее для поиска реперных точек применялась нейросетевая обработка фотографий спины. На данном этапе был осуществлен переход от готовых решений к разработке и обучению своей нейронной сети.

В ходе проведенных экспериментов был выявлен ряд недостатков для определения положения ключевых диагностических точек. По результатам обработки более 3000 различных фотографий наиболее распространенными по частоте проблемами являлись неточное определение положения точек: шеи, ягодиц и ягодичной складки. Анализ причин возникновения проблем показал, что зачастую на точность определения точек оказывали влияние такие факторы, как перекрытие волосами исследуемых областей, в том числе одежда, закрывающая область определения нужных точек, а также «нестатичная» фотография в связи с движением модели.

Результаты тестирования нейронной сети

В результате использования обученной нейронной сети на выборке размеченных фотографий удалось получить хорошее совпадение с положением реперных точек, определенных нейронной сетью, с точками, расставленными медицинским специалистом.

Неправильное нахождение реперных точек на фотографиях у нейронной сети связано с определением точек шеи, которые были закрыты волосами на фотографиях. При этом с поиском остальных точек разработанная нейронная сеть справилась в большинстве случаев.

Для оценки полученного результата использовался ряд стандартных метрик, таких как: Box loss, DFL loss, Precision, Recall, mAP50 (средняя точность по IoU 0.5, mAP50-95 (средняя точность по IoU 0.5–0.95) [14; 16].

Для оценки полученного результата использовался ряд стандартных метрик [11; 14; 17], таких как: F-мера, Precision-Confidence Curve, Recall-Confidence Curve, Precision-Recall Curve (рис. 2).

Точность определяет отношение числа корректно обнаруженных объектов ко всему количеству обнаруженных объектов и вычисляется по формуле:

$$\text{Точность} = \frac{TP}{TP + FN}.$$

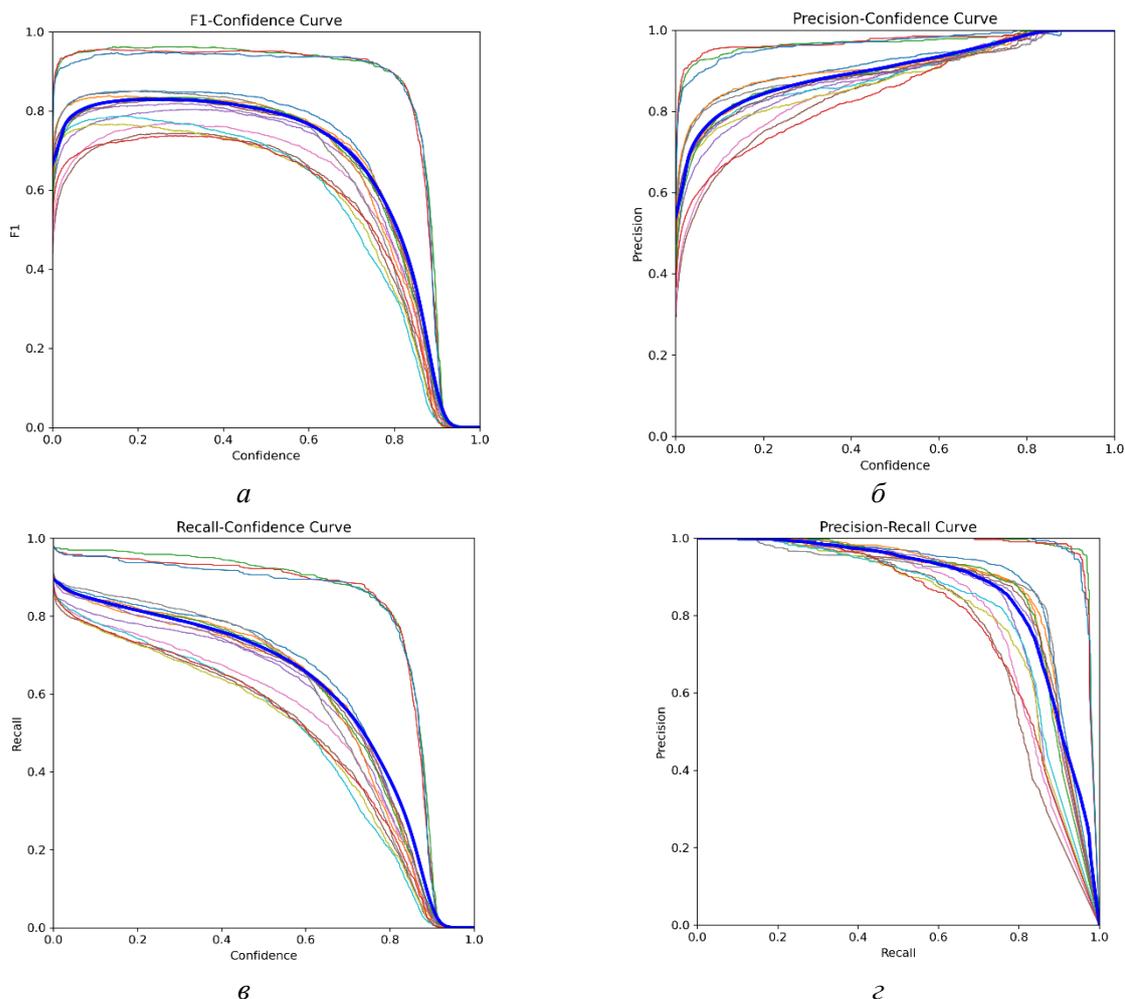


Рис. 2. Метрики оценки работы нейронной сети: a – F-мера (0,83–0,243); b – Precision-Confidence Curve (1–0,94); c – Recall-Confidence Curve (0,89–0); d – Precision-Recall Curve (0,867)

Полнота определяет отношение числа корректно обнаруженных объектов ко всем объектам класса и вычисляется по формуле:

$$\text{Полнота} = \frac{TP}{TP + FN}.$$

F-мера позволяет сравнить две модели, одновременно оценив полноту и точность, и вычисляется по формуле:

$$F\text{-мера} = \frac{2}{\frac{1}{\text{Точность}} + \frac{1}{\text{Полнота}}}.$$

Есть всего четыре возможных исхода при выделении ограничивающих рамок и обнаружении объекта (табл. 2):

- истинно-положительное решение (TP) – количество объектов, которые были правильно классифицированы как положительные;
- истинно-отрицательное решение (TN) – количество объектов, которые были правильно классифицированы как отрицательные;
- ложно-положительное решение (FP) – количество объектов, которые были неправильно классифицированы как положительные;

– ложно-отрицательное решение (FN) – количество объектов, которые были неправильно классифицированы как отрицательные [3; 11].

Таблица 2

Исходы обнаружения объекта

Соотношение	Реальность	
Прогноз	TP	FP
	FN	TN

На представленных далее результатах видно, что точность идентификации большинства реперных точек достигла 1,0.

Графики, автоматически генерируемые при работе нейронной сети YOLO, описывают процесс обучения и результаты сети в процессе тестирования (рис. 3). В частности, график train/boxloss отображает значение функции потерь для каждой эпохи обучения, которое связано с ошибками распознавания объектов и их позиционирования (определением их координат рамок bounding boxes). Если модель ошибается в нахождении координат рамок, это ведет к высокому значению box_loss. Val/cls_loss отвечает за потерю, связанную с определением классов объектов на изображении. Если модель ошибается в определении класса объекта, это ведет к высокому значению cls_loss. Соответственно val/boxloss, val/dflloss – отображают значение ошибок (loss) для каждого параметра – бокса, и локализации. Dfl_loss (дифференциальная функция потерь по Focal Loss'у) – это улучшенный вариант cls_loss, разработанный для борьбы с проблемой неравных значений классов, когда на изображении есть редко встречающиеся объекты. Dfl_loss помогает модели лучше справляться с такими объектами и снижать общий cls_loss.

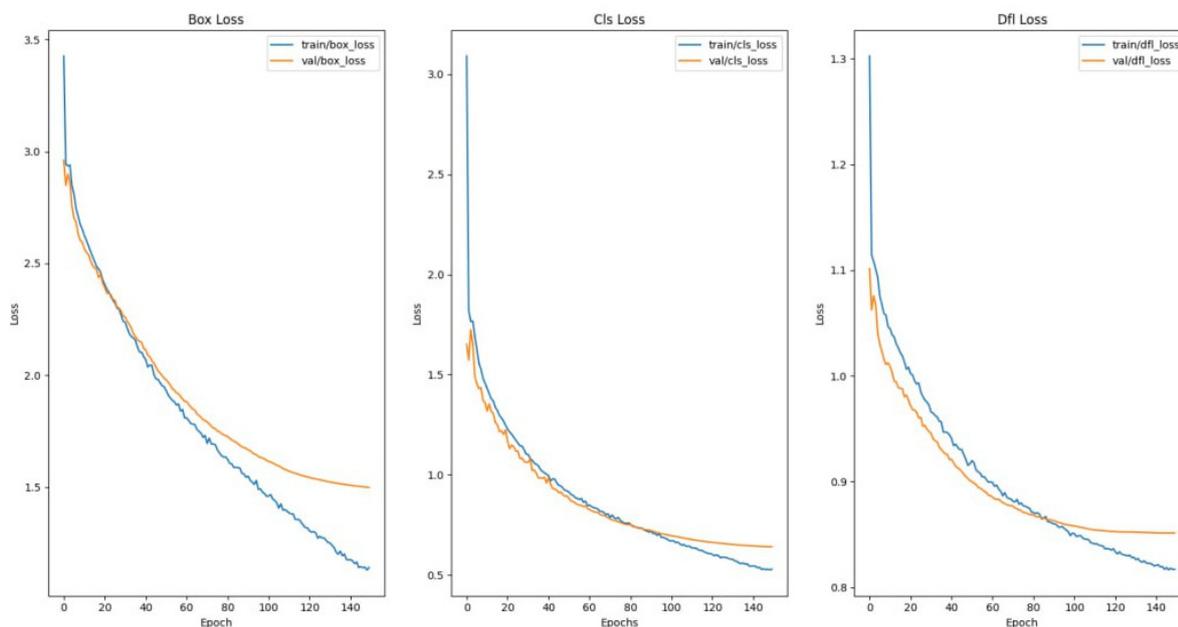


Рис. 3. График результатов обучения модели YOLO на 150 эпохах

Метрика precision (B) (рис. 4) показывает точность (precision) обнаружения объектов на изображениях. rain/dflloss отображает значение функции потерь (ошибки) детектора спин на этапе обучения. metrics/recall показывает, насколько полно и точно модель находит объекты, определяя, сколько объектов было найдено во всех изображениях, разделенных на общее ко-

личество объектов в этих изображениях. metrics/mAP50 и metrics/mAP50-95 показывают среднюю точность обнаружения объектов в зависимости от порога обнаружения. mAP50 показывает точность при использовании порога обнаружения в 50 %, а mAP50-95 – от 50 до 85 %.

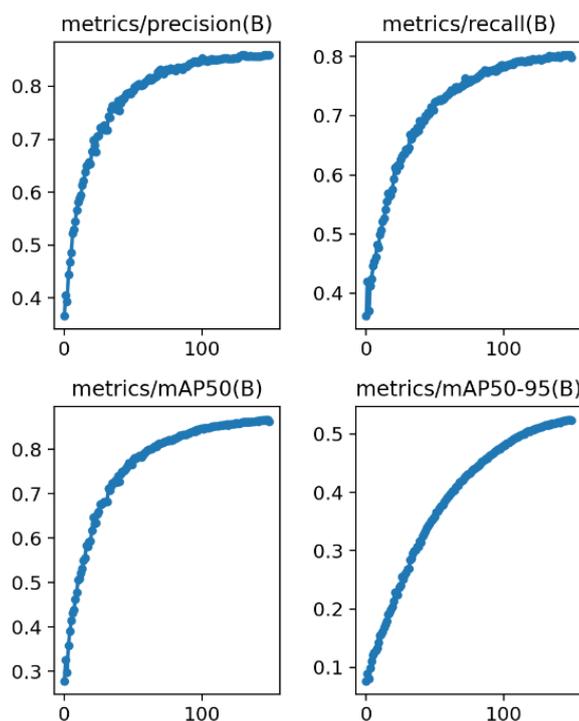


Рис. 4. Метрика precision (B)

Результаты показали, что mAP нейронной сети составил 85 %. Видно, что `box_loss`, `obj_loss`, а значит и `cls_loss`, во время обучения модели постоянно уменьшались. Средняя точность при `mAP_0,5` постоянно увеличивалась. `mAP_0,5` близок к 0,85, что говорит о хорошем эффекте обучения на представленном наборе данных.

Представленные на рис. 3 графики включают две линии – линию обучения и линию проверки. Линия обучения (`train`) отображает показатели производительности модели на каждой эпохе обучения на обучающих данных. Линия проверки (`val`) свидетельствует, как модель проходит проверку на проверочном наборе данных, который она не видела во время обучения. При первоначальном обучении линии `train` и `val` были близки. Однако при переобучении на обучающих данных линия обучения продолжала уменьшаться при увеличении разброса данных линии проверки. Таким образом, уменьшение промежутка между `train` и `val` линиями говорит о лучшей генерализации модели на новые данные. Наиболее оптимальная модель выбирается с тем, чтобы она обладала свойством объективности по отношению к новым и обучающим данным. Таким образом, ошибка при обучении будет снижаться, в то время как на валидационной выборке не должна расти. После дообучения модели удалось снизить количество ошибок, связанных со следующими факторами: распознавание и позиционирование – до 0,8, определение классов – до 0,4, редко встречающиеся объекты – до 0,9.

Заключение

Обученная нейросеть создана для оптимизации процесса диагностики нарушений осанки, которая может служить для использования в клинической медицине, специалисты в области которой могут быть заинтересованы в появлении новых и валидных инструментов для диагностики деформации позвоночного столба.

В рамках проведенного исследования выполнен поиск специальных (реперных точек), определяющих геометрию спины при помощи разработанных нейросетевых алгоритмов. Было показано, что точность поиска точек нейронной сети, обученной на более 3000 сгруппированных фотографиях, достигает 85 %, что свидетельствует о хорошем определении рамки объектов и классифицирует их. Полученные результаты свидетельствуют об эффективности обучения нейронной сети.

Список литературы

1. Отдельные нейрофизиологические аспекты этиопатогенеза юношеского идиопатического сколиоза / С.В. Муравьев, В.Г. Черкасова, П.Н. Чайников, О.О. Мехоношина, М.А. Ковалев, М.О. Гушин // Пермский медицинский журнал. – 2019. – Т. XXXV, № 4. – С. 39–45. DOI: 10.17816/pmj36439%45
2. Синельников Р.Д., Синельников Я.Р., Синельников А.Я. Атлас анатомии человека: в 4 т. – Т. 1. Учение о костях, соединениях костей и мышцах. – М.: Новая волна, 2021. – 488 с.
3. The effect of simulating leg length inequality on spinal posture and pelvic position: a dynamic rasterstereographic analysis / M. Betsch, M. Wild, B. Große, W. Rapp, T. Horstmann // European Spine Journal. – 2012. – Vol. 21, no 4. – P. 691–697. DOI: 10.1007/s00586-011-1912-5
4. Meena G., Mohbey K.K., Kumar S. Sentiment analysis on images using convolutional neural networks-based Inception-V3 transfer learning approach // International journal of information management data insights. – 2023. – No 3. – 13 p. DOI: 10.1016/j.jjime.2023.100174
5. Rahouma K.H., Mahfouz A.Z. Design and implementation of a face recognition system based on API mobile vision and normalized features of still images // 18th International Learning & Technology Conference. – 2021. – P. 32–44. DOI: 10.1016/j.procs.2021.10.057
6. Akhgar B., Yates S. Strategic Intelligence Management. – Butterworth-Heinemann, 2013. – 340 p.
7. Geometric analysis and YOLO algorithm for automatic face detection system in a security setting / F.E. Ayo, A.M. Mustapha, J.A. Braimah, D.A. Aina // Journal of Physics: Conference Series. – 2022. – Vol. 2199. – No. 1. – 10 p. DOI: 10.1088/1742-6596/2199/1/012010
8. Шнайдер Л.С., Сарнадский В.Н., Павлов В.В. Лучевой и оптический методы оценки позвоночно-тазовых взаимоотношений у пациентов с врожденным вывихом бедра // Хирургия позвоночника. – 2009. – Т. 16, № 1. – С. 63–69. DOI: 10.14531/ss2019.1.63-69
9. Факторы риска развития и прогрессирования дегенеративно-дистрофических заболеваний позвоночника по результатам скринингового обследования жителей Санкт-Петербурга / М.В. Авдеева, Ю.А. Кренева, В.П. Панов, В.Н. Филатов, А.В. Мельцер, Л.А. Карасаева // Анализ риска здоровью. – 2019. – № 1. – С. 125–134. DOI: 10.21668/health.risk/2019.1.14
10. Колесников В.Н., Шандыбина Н.Д., Эриум С.С. Экология человека: сбережение нации как стратегия успешного развития // Управленческое консультирование – 2018. – № 2 (110). – С. 73–79. DOI: 10.22394/1726-1139-2018-2-73-79
11. Цифровые модели и инструменты для диагностики трехмерной деформации позвоночника на основе данных стереофотограмметрии / И.Д. Шитоев, В.Ю. Столбов, С.В. Муравьев, В.Н. Никитин, Г.З. Клоян // Прикладная математика и вопросы управления. – 2022. – № 4. – С. 60–82. DOI: 10.15593/2499–9873/2022.4.04
12. Bhople A.R., Shrivastava A.M., Prakash S. Point cloud based deep convolutional neural network for 3D face recognition // Multimedia Tools and Applications. – 2020. – 21 p. DOI: 10.1007/s11042-020-09008-z

13. Paniego S., Sharma V., Cañas J.M. Open-source assessment of deep learning visual object detection // *Sensors*. – 2022. – Vol. 22, no 12. – 16 p. DOI:10.3390/s22124575
14. Sanchez S.A., Romero H.J., Morales A.D. A review: comparison of performance metrics of pretrained models for object detection using the TensorFlow framework // *IOP Conf. Series: Materials Science and Engineering*. – 2020. – 15 p. DOI: 10.1088/1757-899X/844/1/012024
15. A convolutional neural network based on TensorFlow for face recognition / L. Yuan, Z. Qu, Y. Zhao, H. Zhang, Q. Nian // *IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. – 2017. – P. 525–529. DOI: 10.1109/IAEAC.2017.8054070
16. Weed25: A deep learning dataset for weed identification / P. Wang, Y. Tang, F. Luo, L. Wang, Li. Chengsong, Q. Niu, H. Li // *Frontiers in Plant Science*. – 2022. – 14 p. DOI: 10.3389/fpls.2022.1053329
17. Дудченко П.В. Метрики оценки классификаторов в задачах медицинской диагностики // *Молодежь и современные информационные технологии: сб. тр. XVI Междунар. науч.-практ. конф. студентов, аспирантов и молодых ученых / Томский политехнический университет*. – Томск, 2019. – С. 164–165

References

1. Muravev S.V., Cherkasova V.G., Chainikov P.N., Mekhonoshina O.O., Kovalev M.A., Gushchin M.O. Otdel'nye neurofiziologicheskie aspekty etiopatogeneza iunosheskogo idiopatičeskogo skolioza [Particular neurophysiological aspects of etiopathogenesis of juvenile idiopathic scoliosis]. *Perm Medical Journal*, 2019, vol. XXXV, no. 4, pp. 39-45. DOI: 10.17816/pmj36439%45
2. Sinelnikov R.D., Sinelnikov Ia.R., Sinelnikov A.Ia.: Atlas anatomii cheloveka. V 4-kh tomakh. Tom 1. Uchenie o kostiakh, soedineniakh kostei i myshtsakh [Atlas of Human Anatomy. In 4 volumes. Volume 1. The doctrine of bones, joints of bones and muscles]. Moscow, Novaia volna, 2021, 488 p.
3. Betsch M., Wild M., Große B., Rapp W., Horstmann T. The effect of simulating leg length inequality on spinal posture and pelvic position: a dynamic rasterstereographic analysis. *European Spine Journal*, 2012, vol. 21, no 4, pp. 691-697. DOI: 10.1007/s00586-011-1912-5
4. Meena G., Mohbey K.K., Kumar S. Sentiment analysis on images using convolutional neural networks-based Inception-V3 transfer learning approach. *International Journal of Information Management Data Insights*, 2023, no 3, 13 p. DOI: 10.1016/j.ijime.2023.100174
5. Rahouma K.H., Mahfouz A.Z. Design and implementation of a face recognition system based on API mobile vision and normalized features of still images. 18th International Learning & Technology Conference, 2021, pp. 32-44. DOI: 10.1016/j.procs.2021.10.057
6. Akhgar B., Yates S. Strategic Intelligence Management. Butterworth-Heinemann, 2013, 340 p.
7. Ayo F.E., Mustapha A.M., Braimah J.A., Aina D.A. Geometric analysis and YOLO algorithm for automatic face detection system in a security setting. *Journal of Physics: Conference Series*, 2022, vol. 2199, no 1, 10 p. DOI:10.1088/1742-6596/2199/1/012010
8. Shnaider L.S., Sarnadskii V.N., Pavlov V.V. Luchevoi i opticheskie metody otsenki pozvonочно-tazovykh vzaimootnoshenii u patsientov s vrozhdennym vyvikhom bedra [Radiological and optical methods for assessing the spino-pelvic relationships in patients with congenital hip dislocation]. *Hirurgia Pozvonochnika*, 2009, vol. 16, no. 1, pp. 63-69. DOI: 10.14531/ss2019.1.63-69

9. Avdeeva M.V., Kreneva Iu.A., Panov V.P., Filatov V.N., Meltser A.V., Karasaeva L.A. Faktory riska razvitiia i progressirovaniia degenerativno-distroficheskikh zabolevanii pozvonochnika po rezul'tatam skringovogo obsledovaniia zhitelei Sankt-Peterburga [Risk factors for the development and progression of degenerative-dystrophic diseases of the spine according to the results of a screening examination of residents of St. Petersburg]. *Health Risk Analysis*, 2019, no. 1, pp. 125-134. DOI: 10.21668/health.risk/2019.1.14
10. Kolesnikov V.N., Shandybina N.D., Erium S.S. Ekologiya cheloveka: sberezhenie natsii kak strategiiia uspeshnogo razvitiia. [Human ecology: saving the nation as a strategy for successful development]. *Upravlencheskoe konsul'tirovanie*, 2018, no. 2 (110), pp. 73-79. DOI: 10.22394/1726-1139-2018-2-73-79
11. Shitoev I.D., Stolbov V.Iu., Murav'ev S.V., Nikitin V.N., Kloian G.Z. Tsifrovye modeli i instrumenty dlia diagnostiki trekhmernoi deformatsii pozvonochnika na osnove dannykh stereofotogrammetrii [Digital models and tools for diagnosing three-dimensional spinal deformity based on stereophotogrammetry data]. *Applied Mathematics and Control Sciences*, 2022, no. 4, pp. 60–82. DOI: 10.15593/2499–9873/2022.4.04
12. Bhople A.R., Shrivastava A.M., Prakash S. Point cloud based deep convolutional neural network for 3D face recognition. *Multimedia Tools and Applications*, 2020, 21 p. DOI: 10.1007/s11042-020-09008-z
13. Paniego S. Sharma V. Cañas J.M. Open-source assessment of deep learning visual object detection. *Sensors*, 2022, vol. 22, no 12, 16 p. DOI: 10.3390/s22124575
14. Sanchez S.A Romero H.J Morales A.D. A review: comparison of performance metrics of pretrained models for object detection using the TensorFlow framework. *IOP Conference Series: Materials Science and Engineering*, 2020, 15 p. DOI: 10.1088/1757-899X/844/1/012024
15. Yuan L., Qu Z., Zhao Y., Zhang H., Nian Q. A convolutional neural network based on TensorFlow for face recognition. *IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2017, pp. 525-529. DOI: 10.1109/IAEAC.2017.8054070
16. Wang P., Tang Y., Luo F., Wang L., Chengsong Li., Niu Q., Li H. Weed25: A deep learning dataset for weed identification. *Frontiers in Plant Science*, 2022, 14 p. DOI: 10.3389/fpls.2022.1053329
17. Dudchenko P.V. Metriki otsenki klassifikatorov v zadachakh meditsinskoii diagnostiki [Metrics for classifier evaluation in medical diagnostic tasks]. *Molodezh i sovremennye informatsionnye tekhnologii: sbornik trudov XVI Mezhdunarodnoi nauchno-prakticheskoi konferentsii studentov, aspirantov i molodykh uchenykh. Tomskii politekhnicheskii universitet, Tomsk*, 2019, pp. 164–165.