

ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ДЛЯ АВТОМАТИЗИРОВАННОЙ РАССТАНОВКИ ЦЕФАЛОМЕТРИЧЕСКИХ ТОЧЕК НА ТЕЛЕРЕНТНОГРАММАХ ГОЛОВЫ В БОКОВОЙ ПРОЕКЦИИ

Мураев А.А.¹, Кибардин И.А.², Оборотистов Н.Ю.³, Иванов С.С.¹
Иванов С.Ю.⁴, Персин Л.С.³

Телерентгенография (ТРГ) является важным методом исследования в ортодонтии и челюстно-лицевой хирургии, позволяющим получить необходимую диагностическую информацию о строении мозгового и лицевого отделов черепа для планирования лечения. Существует множество методов анализа, требующих расстановки цефалометрических точек на телерентгенограмме с последующей их обработкой, что занимает значительное время врача. Современные цифровые технологии с применением искусственного интеллекта позволяют усовершенствовать этот метод исследования и значительно упростить работу врача.

Цель исследования. В настоящем исследовании была разработана конфигурация искусственных нейронных сетей (ИНС), позволяющая с высокой точностью расставлять цефалометрические точки на телерентгенограмме головы в боковой проекции вне зависимости от источника получения ТРГ.

Материалы и методы. В исследовании были использованы 80 телерентгенограмм (ТРГ) головы в боковой проекции.

Результаты. Ошибка ИНС в нашем исследовании составила 2%. Предложенный подход требует в 2-3 раза меньше времени, чем традиционный «ручной» метод расстановки цефалометрических точек, в зависимости от количества точек и сложности цефалометрического анализа.

Ключевые слова: цефалометрический анализ, искусственные нейронные сети, телерентгенография.

Контактный автор: Мураев Александр Александрович, e-mail: ma@visurgery.ru

Для цитирования: Мураев А.А., Кибардин И.А., Оборотистов Н.Ю., Иванов С.С., Иванов С.Ю., Персин Л.С. Использование нейросетевых алгоритмов для автоматизированной расстановки цефалометрических точек на телерентгенограммах головы в боковой проекции. REJR 2018; 8(4):16-22. DOI:10.21569/2222-7415-2018-8-4-16-22.

Статья получена: 13.10.18

Статья принята: 11.11.18

USE OF NEURAL NETWORK ALGORITHMS FOR THE AUTOMATED ARRANGEMENT OF CEPHALOMETRIC MARKERS ON LATERAL CEFALOGRAMS

Muraev A.A.¹, Kibardin I.A.², Oborotistov N.Yu.³, Ivanov S.S.¹, Persin L.S.³

Тhe lateral cephalogram analysis is an important method of a research in orthodontics and maxillofacial surgery that allows to obtain necessary diagnostic information on a structure of brain and facial skull parts for treatment planning. There are numerous techniques of the analysis based on processing of cephalometric points (markers) on the lateral cephalogram. Such approaches take a considerable time for the doctor to arrange cephalometric points. Modern digital technologies with use of artificial intelligence allow to improve this method of a research and to significantly simplify doctor's work.

Purpose. In this paper, we propose a neural network and training strategy that enables to place cephalometric points on the lateral cephalogram with high precision.

1 – ФГАОУ ВО РУДН.
2 – ФГАОУ ВО МФТИ (ГУ)
3 – МГМСУ им.
А.И. Евдокимова.
4 – ФГБОУ ВО «Первый
МГМУ им. И.М.
Сеченова» Минздрава
России (Сеченовский
университет).
г. Москва, Россия.

1 - The Peoples'
Friendship University of
Russia,
2 – Moscow Physics and
Technical University
3 – A.I. Evdokimov
Moscow State Medical
Dentistry University
4 – I.M. Sechenov
Moscow State Medical
University

Materials and methods. The research used 80 lateral cephalograms of the head.

Results. The developed method handles the cephalograms regardless of the source of the image, while error percentage below 2% of the size of the images. The offered approach demands 2-3 times less time than a traditional "manual" method of arrangement of cephalometric points, depending on quantity of points and complexity of the cephalometric analysis.

Keywords: cephalometric analysis; artificial neural network; cephalogram.

Corresponding author: Aleksandr A. Muraev; E-mail: ma@visurgery.ru

For citation: Muraev A.A., Kibardin I.A., Oborotistov N.Yu., Ivanov S.S., Persin L.S. Use of neural network algorithms for the automated arrangement of cephalometric markers on lateral cephalograms. REJR 2018; 8 (4):16-22. DOI:10.21569/2222-7415-2018-8-4-16-22.

Received: 13.10.18

Accepted: 11.11.18

Телерентгенография является важным методом исследования в ортодонтии и челюстно-лицевой хирургии, позволяющим получить необходимую информацию о строении мозгового и лицевого отделов черепа, размерах и расположении челюстей в горизонтальном и вертикальном направлениях, наклонах зубов, зубоальвеолярных высотах в переднем и боковых участках челюстей, типе и направлении роста челюстей, соотношении мягких тканей лица с костями лицевого скелета и другое. Для получения данных после проведённого рентгенологического исследования существует множество методов анализа, требующих расстановки цефалометрических точек и ориентиров с последующей их обработкой, что занимает значительное время врача. Современные цифровые технологии с применением искусственного интеллекта позволяют усовершенствовать этот метод исследования и значительно упростить работу врача.

Искусственные нейронные сети могут использоваться для интерпретации медицинских изображений и в качестве клинических систем принятия решений в медицинской диагностике.

Цель исследования: разработка автоматизированного цефалометрического анализа телерентгенограмм (ТРГ) головы в боковой проекции на основе искусственных нейросетевых алгоритмов.

Материалы и методы.

Общая характеристика исследуемого материала. В исследовании были использованы 80 телерентгенограмм (ТРГ) головы в боковой проекции. Снимки случайным образом были выбраны из архивов кафедры ортодонтии МГМСУ им. А.И. Евдокимова и кафедры челюстно-лицевой хирургии и хирургической стоматологии РУДН. Разнородность материала обеспечивала вариативность для обучения искусственных нейронных сетей (ИНС). Файлы были получены из разных источников: сканы рентгеновских плёнок с ТРГ; ТРГ, полученные на ортопантомографе с цефалостатом; ТРГ, по-

строенные программным методом на основе компьютерной томографии головы. Каждую ТРГ импортировали в программу ViSurgery (Россия) и проводили расстановку цефалометрических точек. Точки расставляли на контурах черепа, мягких тканей и зубов. В исследовании были использованы ТРГ из разных источников, поэтому изображения имели ряд особенностей, они могли отличаться четкостью анатомических контуров, яркостью и контрастностью, наличием артефактов. Общее количество точек на каждом снимке составило 62.

Для обучающей выборки искусственной нейронной сети (ИНС) были подготовлены 80 файлов в формате JPG, на которых изображены телерентгенограммы головы в боковой проекции. Из указанных файлов случайным образом были выбраны 50 для обучающей выборки и 10 для валидации во время обучения, 20 для контроля после окончания обучения модели.

Характеристика ИНС и алгоритм обучения.

Искусственная нейронная сеть (ИНС) – математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных связей. Основная задача тренировки ИНС настоящего исследования – обучить ИНС определять правильные координаты цефалометрических точек и расставлять их на ТРГ.

Архитектура ИНС.

Для реализации ИНС был использован фреймворк PyTorch – это современная библиотека глубокого обучения [1]. Она даёт максимальную гибкость и расширяемость при построении ИНС, что удобно для различных исследований.

Так как наша обучающая выборка состояла всего из 50 изображений, обучение нейросети с нуля на этих данных приводило к неудовлетворительной точности и плохой обобщающей способности модели. Чтобы преодолеть эту проблему, был применён метод «переноса обучения» от английского transfer learning. Он поз-

воляет обучить хорошую модель, имея небольшое количество данных. Суть метода заключается в том, чтобы взять готовую ИНС, обученную на большом наборе данных (возможно, решать другую задачу), а затем дообучить её на своих данных и под свою задачу. В качестве такой предобученной модели была взята сеть SE-ResNeXt-50, предобученная на ImageNet – массивной базе данных аннотированных изображений, предназначенной для отработки и тестирования методов распознавания образов и машинного зрения [2]. По состоянию на 2016 год в базу данных было записано около десяти миллионов URL с изображениями из 1000 различных категорий, которые прошли ручную аннотацию для ImageNet.

Для того, чтобы адаптировать данную архитектуру для задачи расстановки цефалометрических точек, мы заменили последний слой SE-ResNeXt-50 на четыре полносвязных слоя с использованием Batch Normalization и функции активации ReLU [4], определяемой как $f(x) = \max(0, x)$ [3]. Полученная таким образом архитектура приведена на рисунке 1.

Изображения размером 224x224 подаются в претренированную нейронную сеть SE-ResNeXt-50. Сеть извлекает из каждого изображения вектор из 2048 признаков. Извлеченные признаки проходят через конструкцию из четырех полносвязных слоёв (Fully-Connected layer) с использованием Batch Normalization, функции активации ReLU и Dropout. На выходе получаем вектор из 124-х предсказанных координат, принадлежащих 62-м цефалометрическим точкам (рис. 1).

Так как мы располагали небольшим объёмом данных, использование такой глубокой нейронной сети может привести к её переобучению. Переобучение ИНС – это явление, когда построенная модель хорошо аппроксимирует примеры из обучающей выборки, но относительно плохо работает на примерах из тестовой выборки, не участвовавших в обучении. Это связано с тем, что при построении ИНС в процессе обучения в выборке обнаруживаются некоторые случайные закономерности, которые отсутствуют в генеральной совокупности.

Для преодоления этой проблемы в последних слоях ИНС был использован Dropout, предложенный в работе Srivastava N. [5]. Идея этого метода заключается в том, чтобы при тренировке нейронной сети каждый раз случайным образом занулять выход доли нейронов перед полносвязным слоем. При получении предсказаний для тестового набора изображений, вместо зануления случайной доли выходов нейронов, все выходы уменьшаются в два раза. Этот довольно простой алгоритм хорошо зарекомендовал себя для борьбы с переобучением.

Полученная модель тренируется методом

обратного распространения ошибки. Подробнее о режиме тренировки будет написано ниже.

Подготовка исходных изображений.

Для предупреждения возможных процессов переобучения и для улучшения обобщающей способности модели, к исходным изображениям применялись сильные аугментации. Аугментации – это намеренные искажения и преобразования исходных изображений для того, чтобы искусственно увеличить размер обучающей выборки, что является сильной регуляризацией для модели. То есть ИНС, просматривая при тренировке больше различных изображений, лучше обобщается и меньше переобучается.

Сначала все изображения приводились к размеру 256x256 пикселей с сохранением пропорцией изображения, чтобы не исказить форму черепа пациентов. Чтобы пропорции сохранялись, изображения дополнялись до квадратных серой заливкой. Затем, все изображения проходили предварительные случайные преобразования:

1. Случайное размытие изображения.
2. Случайная коррекция яркости.
3. Отражение изображения относительно вертикальной оси.
4. Изменение масштаба изображения в число раз, выбранное равновероятно из отрезка от 0.6 до 1.4.
5. Поворот на случайный угол от -30° до $+30^\circ$.

После этого из изображений вырезалось случайное окно размером 224x224 пикселей и соответствующим образом рассчитывались координаты цефалометрических точек так, чтобы они соответствовали новым изображениям.

После проведенных преобразований, изображения собирались в большие серии и подавались в ИНС для её обучения.

Тренировка ИНС.

ИНС обучалась оптимизировать стандартную для задач детекции ключевых точек функцию потерь MSE (Mean Squared Error), которая определяется следующим образом:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2,$$

где \hat{Y}_i – это предсказанные координаты цефалометрических точек, а Y_i – это их настоящие значения.

В качестве оптимизатора использовался Adam [6]. Adam (от англ. adaptive moment estimation) – это алгоритм оптимизации, являющийся расширением метода стохастического градиентного спуска.

Тренировка ИНС проходила в два этапа. На первом этапе претренированный SE-ResNeXt-50 «замораживался», а тренировались

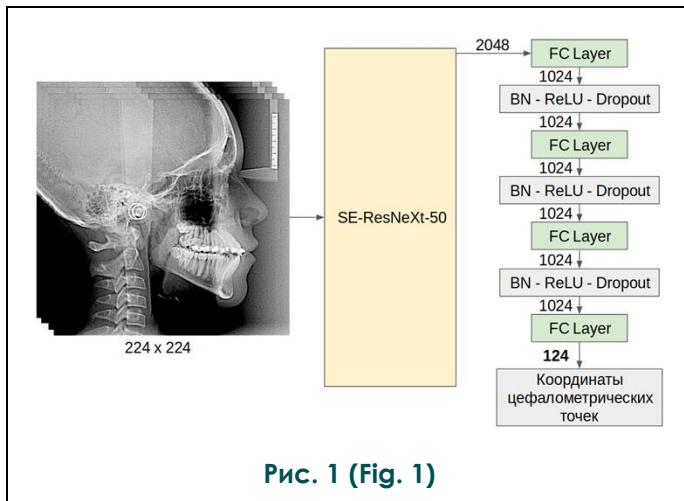


Рис. 1 (Fig. 1)

Рис. 1. Схема.

Итоговая архитектура ИНС.

Fig. 1. Scheme.

Final structure of artificial neural network.

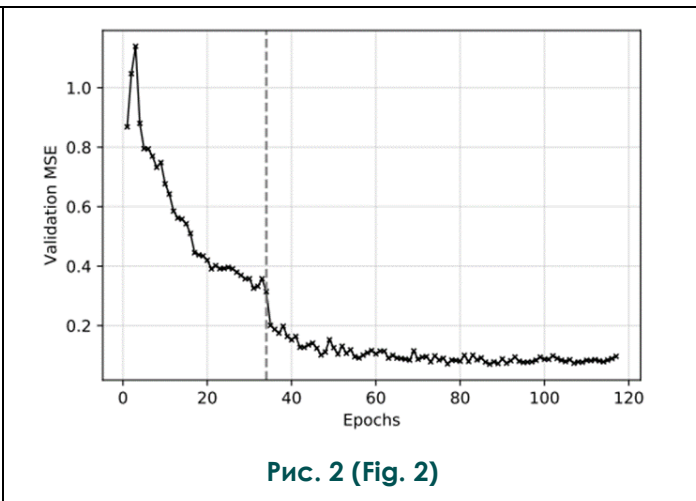


Рис. 2 (Fig. 2)

Рис. 2. График.

Кривая обучения ИНС.

Fig. 2. Graph.

Education line of artificial neural network.

только последние четыре полносвязных слоя. На этом этапе в нейронную сеть изображения отправлялись сериями по 800 штук. Каждую эпоху (эпохой мы называем 100 шагов оптимизатора, на каждом из которых ИНС просматривает серию из 800 изображений), мы измеряли MSE на валидационной выборке из 10 изображений, чтобы контролировать процесс обучения ИНС. Мы использовали следующий режим тренировки: сеть тренировалась с начальным параметром скорости обучения (learning rate) равным 0.0001; каждый раз, когда ошибка MSE на валидационной выборке из 10 изображений не уменьшалась в течение 10 эпох, мы уменьшали параметр скорости обучения в два раза и продолжали обучение. Процедура повторялась до тех пор, пока ошибка на валидационной выборке не перестанет улучшаться.

На втором этапе ИНС тренировалась целиком (вместе с «замороженным» на первом этапе SE-ResNeXt-50) с аналогичным режимом тренировки. Однако на этом этапе изображения отправлялись в ИНС сериями по 185 изображений. Это вызвано необходимостью распространять ошибку сквозь глубокую сеть SE-ResNeXt-50, что накладывает ограничения на размер отдельной серии, связанные с оперативной памятью видеокарт. Соответственно, на этом этапе под эпохой понимается 100 шагов оптимизатора, на каждом из которых ИНС просматривает по 185 изображений.

Кривая обучения нейронной сети на обоих этапах приведена на рисунке 2. На графике показана зависимость функции ошибки MSE на валидационной выборке от числа пройденных эпох. Обучение происходит в два этапа: на первом претренированная SE-ResNeXt-50 не изменяется, обучаются только последние четыре

полносвязных слоя. На втором этапе ИНС обучается целиком. Граница между двумя этапами обозначена пунктирной вертикальной прямой на графике.

Такая процедура тренировки позволила обучить SE-ResNeXt-50 за шесть часов суммарно, используя две современные видеокарты GTX 1080 Ti.

Получение предсказаний.

Для расстановки цефалометрических точек на тестовом наборе данных из изображения вырезается центральная часть размером 224x224 пикселей и отправляется в ИНС. Предсказанные нейросетью координаты пересчитываются обратно, чтобы соответствовать исходному изображению.

Для уточнения предсказаний, мы применили Test Time Augmentation: из исходного изображения вырезаются пять окон размером 224x224 пикселей в центре и по углам, а также эти окна отражаются относительно вертикальной оси. Полученные 10 изображений отправляются в ИНС. Предсказанные координаты соответствующим образом пересчитываются и усредняются арифметическим средним.

Результаты.

Для оценки точности расстановки ЦТ обученной ИНС на контрольной выборке телерентгенограмм сравнивали координаты точек, полученные при их расстановке врачом-ортодонтом и обученной ИНС, рассчитывая среднеквадратичное отклонение, определяемое по формуле

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} = \sqrt{MSE}$$

где \hat{Y}_i – это, аналогично формуле для MSE, предсказанные координаты цефалометрических точек, а Y_i – это их настоящие значения.

Таблица №1. Результаты статистической обработки координат цефалометрических точек после их расстановки обученной ИНС и врачом-ортодонтом.

Метод	RMSE (среднеквадратичная ошибка)
Одно центральное окно	0.0198
Test Time Augmentation	0.0188

Среднеквадратичная ошибка на координатах точек не превышает 2% от размера изображения.

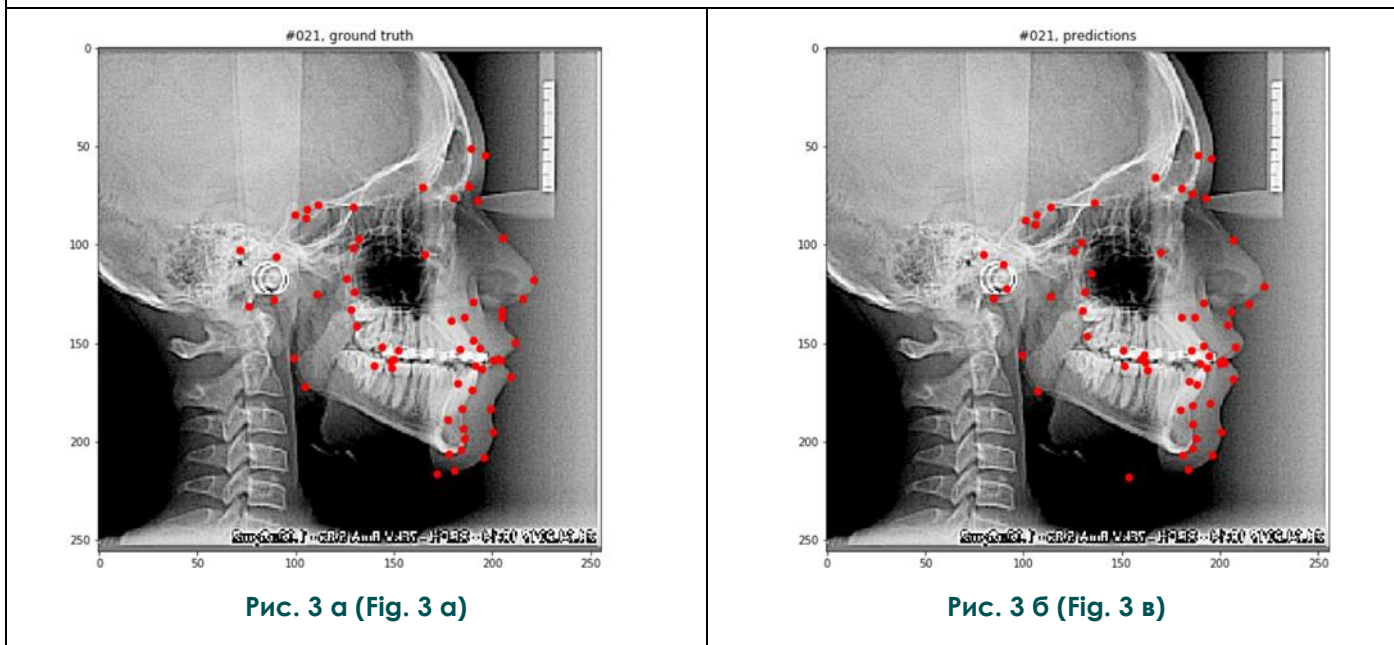


Рис. 3 а (Fig. 3 а)

Рис. 3 б (Fig. 3 в)

Рис. 3. Телерентгенограмма.

Сравнительное представление расстановки цефалометрических точек врачом (слева) и обученной ИНС (справа).

Fig. 3. Teleroentgenogram.

Comparison of cephalometric points made by doctor (left) and educated artificial neural network (right).

Результаты статистической обработки полученных данных представлены в таблице №1 (при подсчете ошибки координаты точек принимают значения от 0 до 1 – края изображения).

На рисунке 3 представлен пример расстановки цефалометрических точек на ТРГ 21 врачом (А – слева) и обученной ИНС (Б – справа).

После расстановки цефалометрических точек (ЦТ) натренированной ИНС, время, затраченное врачом на корректировку положения 62 ЦТ, не превышало 2-3 минут.

Обсуждение.

Задача детекции ключевых точек очень важна для распознавания и анализа человеческих лиц. В связи с этим, для решения этой задачи было предложено множество различных методов. Ранее, большинство из них основывалось на классических методах таких, как случайный лес, градиентный бустинг, SVM [7 - 9]. Однако такие методы сложны в реализации и

не отличаются высокой стабильностью.

В последние годы для решения задач компьютерного зрения большое развитие получили методы, связанные с глубокими сверточными нейронными сетями. Решения, основывающиеся на нейронных сетях, быстро превзошли традиционные методы как по точности, так и по стабильности и обобщающей способности моделей на основных задачах компьютерного зрения. В том числе, нейронные сети были применены для распознавания отдельных областей лица и для детекции ключевых точек [10, 11].

Результаты, полученные в настоящем исследовании, свидетельствуют о принципиальной возможности применения ИНС для цефалометрической разметки ТРГ головы в боковой проекции и проведения на основе этой разметки различных цефалометрических измерений (анализов). Точность расстановки ЦТ посредством ИНС находится в пределах 2% в сравнении с работой человека, в перспективе возмож-

но повышение этой точности при тренировке ИНС на большем количестве ТРГ – около тысячи.

Нами было отмечено, что для обученной ИНС качество изображения ТРГ, на которой происходит расстановка ЦТ не играет решающего значения. В случае отсутствия чётких ориентиров на ТРГ, ИНС ставит ЦТ в усредненном для них положении. Например, как видно на рисунке 3 Б, точка th отмечена ИНС в предполагаемом месте перехода подъязычной области на передней поверхности шеи. То же касается контуров зубов при наложении правой и левой сторон на ТРГ. Такое поведение ИНС может помочь врачу выбирать правильное положение ЦТ.

Расстановка 62 ЦТ у врача занимает в среднем 6-7 минут, обученная ИНС проводит это мгновенно, последующая корректировка врачом не превышает 2-3 минут. То есть можно с уверенностью утверждать, что цефалометрический анализ с использованием ИНС для расстановки ЦТ как минимум в 2 раза быстрее, чем традиционный подход. Повышение точности работы ИНС и программная оптимизация самого анализа могут дополнительно ускорить процесс.

Учитывая универсальность подхода и гибкость ИНС, предложенный алгоритм можно применять для обучения новых ИНС для реализации схожих целей – разметки любых медицинских изображений и последующих антропометрических измерений. Наиболее близкой задачей является цефалометрический анализ

ТРГ в прямой проекции. В нашем исследовании мы использовали 62 ЦТ, однако это количество ничем не ограничено. Поэтому предложенный алгоритм можно применить для разметки медицинских изображений любым количеством реперных точек.

Особую актуальность данный метод получает при проведении 3D-цефалометрии по данным компьютерной томографии или сканам лица, так как в настоящее время это наиболее трудоемкий метод обследования в ортодонтии и челюстно-лицевой хирургии.

Выводы.

ИНС можно использовать для автоматизированной расстановки цефалометрических точек на телерентгенограммах и любых плоскостных рентгеновских снимках и фотографиях лица. Разработана конфигурация ИНС, позволяющая с высокой точностью расставлять цефалометрические точки на ТРГ головы в боковой проекции вне зависимости от источника получения ТРГ. Ошибка ИНС в нашем исследовании составила 2%. Предложенный подход требует в 2-3 раза меньше времени, чем традиционный «ручной» метод расстановки цефалометрических точек, в зависимости от количества точек и сложности цефалометрического анализа.

Источник финансирования и конфликт интересов.

Авторы данной статьи подтвердили отсутствие финансовой поддержки исследования и конфликта интересов, о которых необходимо сообщить.

Список литературы:

1. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*. 2015; 521 (7553): 436.
2. Hu J., Shen L., Sun G. Squeeze-and-excitation networks. *arXiv preprint arXiv*. 2017; 7.
3. Ioffe S., Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv*. 2015.
4. Nair V., Hinton G. E. Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010; 807-814.
5. Srivastava N. et al. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*. 2014; 15 (1): 1929-1958.
6. Kingma D. P., Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv*. 2014.
7. Amberg B., Vetter T. Optimal landmark detection using shape

References:

1. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*. 2015; 521 (7553): 436.
2. Hu J., Shen L., Sun G. Squeeze-and-excitation networks. *arXiv preprint arXiv*. 2017; 7.
3. Ioffe S., Szegedy C. Batch normalization: Accelerating deep

models and branch and bound. *Computer Vision (ICCV)*. 2011 *IEEE International Conference on*. 2011: 455-462.

8. Liang L. et al. Face alignment via component-based discriminative search. *European conference on computer vision*. Springer, Berlin, Heidelberg. 2008: 72-85.

9. Belhumeur P. N. et al. Localizing parts of faces using a consensus of exemplars. *IEEE transactions on pattern analysis and machine intelligence*. 2013; 35 (12): 2930-2940.

10. Luo P., Wang X., Tang X. Hierarchical face parsing via deep learning. *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012: 2480-2487.

11. Sun Y., Wang X., Tang X. Deep convolutional network cascade for facial point detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013: 3476-3483.

network training by reducing internal covariate shift. *arXiv preprint arXiv*. 2015.

4. Nair V., Hinton G. E. Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010; 807-814.

5. Srivastava N. et al. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*. 2014; 15 (1): 1929-1958.
6. Kingma D. P., Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv*. 2014.
7. Amberg B., Vetter T. Optimal landmark detection using shape models and branch and bound. *Computer Vision (ICCV)*. 2011 *IEEE International Conference on*. 2011: 455-462.
8. Liang L. et al. Face alignment via component-based discriminative search. *European conference on computer vision*. Springer, Berlin, Heidelberg. 2008: 72-85.
9. Belhumeur P. N. et al. Localizing parts of faces using a consensus of exemplars. *IEEE transactions on pattern analysis and machine intelligence*. 2013; 35 (12): 2930-2940.
10. Luo P., Wang X., Tang X. Hierarchical face parsing via deep learning. *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012: 2480-2487.
11. Sun Y., Wang X., Tang X. Deep convolutional network cascade for facial point detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013: 3476-3483.