

# Применение свёрточных нейронных сетей для установления возраста человека по изображению

А.А. Рудь<sup>1</sup>, С.А. Рудь<sup>1</sup>, М.А. Исаев<sup>1</sup>, Д.А. Савельев<sup>1,2</sup>

<sup>1</sup>Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

<sup>2</sup>Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

**Аннотация.** В работе представлены результаты исследований определения биологического возраста человека по изображению лица. Для решения данной задачи использовались: алгоритм случайного леса с использованием гибридного фильтра Гессе и оператора локальных бинарных шаблонов, а также свёрточная нейронная сеть ResNet50. Приведены использованные наборы данных, проблемы, связанные с их применением, а также точность классификации при выбранном разделении возрастной линии.

## 1. Введение

В настоящее время активно развивается сфера технического зрения [1 – 3]. Обнаружение и распознавание объектов применяется не только в профессиональных специализированных видах деятельности, но и обычными пользователями смартфонов [4]. Задачи компьютерного зрения весьма разнообразны, в частности это может быть распознавание текста, биометрия, видеоаналитика, анализ спутниковых снимков, редактирование изображений, реконструкция объёмных моделей, управление автомобилем [1 – 5]. Целью данной работы является изучение применимости свёрточных нейронных сетей для решения задачи автоматического определения возраста человека на основе изображения.

Решение задачи автоматического определения возраста человека становится всё более и более актуальным вследствие быстрого роста социальных платформ и медиа приложений, а также ввиду маркетинговых исследований, использовании в системах безопасности и других областях, где возможно ограничение допуска по возрасту [6]. Представителям каждого возрастного периода предписаны шаблонные характеристики, требования, обязанности, а также возможные ограничения.

В данной работе мы рассмотрим применение свёрточных нейронных сетей (CNN) [7, 8] для разрешения задачи установления биологического возраста человека по изображению его лица. Несмотря на то, что возрастные рамки современного человека крайне размыты, некоторые конфигурации искусственных сетей делают возможным добиться весьма точных результатов [9].

В первую очередь необходимо обнаружить лицо на фотографии, чтобы отправить соответствующую область на вход нейронной сети для определения возраста. Задачей детектирования является выделение интересующего участка на изображении или видеопотоке

[10]. Существует много методов распознавания лиц [11], и свёрточные нейронные сети являются одними из лучших алгоритмов по распознаванию и классификации изображений [9, 12, 13].

Вообще говоря, обучение такого рода аппроксимации весьма ресурсозатратно. Вследствие этого некоторыми экспериментаторами были предложены конфигурации, решающие конкретные задачи эффективней остальных. ResNet50 [14, 15] представляет одну из таких архитектур, которую в рамках решаемой задачи, мы будем использовать.

Обучение сети будем производить с использованием модифицированного алгоритма градиентного спуска (mini-batch) [16, 17], который делает возможным минимизировать функцию ошибки.

Стоит отметить, что фильтры и сама операция свёртки дают CNN одну из самых существенных специфик – инвариантность к смещениям, т. е. в случае задач обнаружения и идентификации объектов классификации, их местоположение на входном изображении не имеет значения.

## 2. Описание используемых алгоритмов

В данной работе для решения задачи определения биологического возраста человека использовались следующие методы: алгоритм случайного леса с использованием гибридного фильтра Гессе и оператора локальных бинарных шаблонов для предварительной обработки, а также свёрточная нейронная сеть, а именно ResNet50.

Стоит отметить, что в случае со свёрточной нейронной сетью, на вход подавалось изображение. В случае с алгоритмом случайного леса, на вход подавалось не изображение, а признаки, отобранные посредством предобработки фильтром Гессе и последующим использованием метода локальных бинарных шаблонов.

Гибридный фильтр Гессе [18] позволяет обнаружить морщины, вычисляя матрицу Гессе для каждого пикселя входного изображения. Результат его работы представлен на рисунке 1. Оператор локальных бинарных шаблонов представляет собой описание окрестности пикселя изображения в двоичной форме. Восемь пикселей вокруг центрального пикселя принимают значение 0 или 1 в зависимости от порога, которым является значение центрального пикселя. Таким образом получается восьмиразрядный бинарный код, который описывает окрестность пикселя.



**Рисунок 1.** Визуализация работы модификации гибридного фильтра Гессе.

В данной работе для извлечения возрастных признаков, которые в дальнейшем передаются классификатору, используется метод, который одновременно совмещает в себе использование модификации гибридного фильтра Гессе и последующее использование оператора локальных бинарных шаблонов для уменьшения количества данных без потери их ценности в контексте исходной задачи.

Алгоритм случайного леса [19] сегодня является одним из популярнейших и крайне эффективных методов решения задач машинного обучения, таких как классификация и

регрессия. По эффективности он конкурирует с методом опорных векторов, нейронными сетями и бустингом, хотя, конечно, не лишен своих недостатков [20].

Основным элементом случайного леса является дерево решений. Дерево решений является логической схемой, которая позволяет получить окончательное решение о классификации объекта после ответов на иерархически организованную систему вопросов. Итоговое решение принимается решением голосования по большинству (рисунок 2).

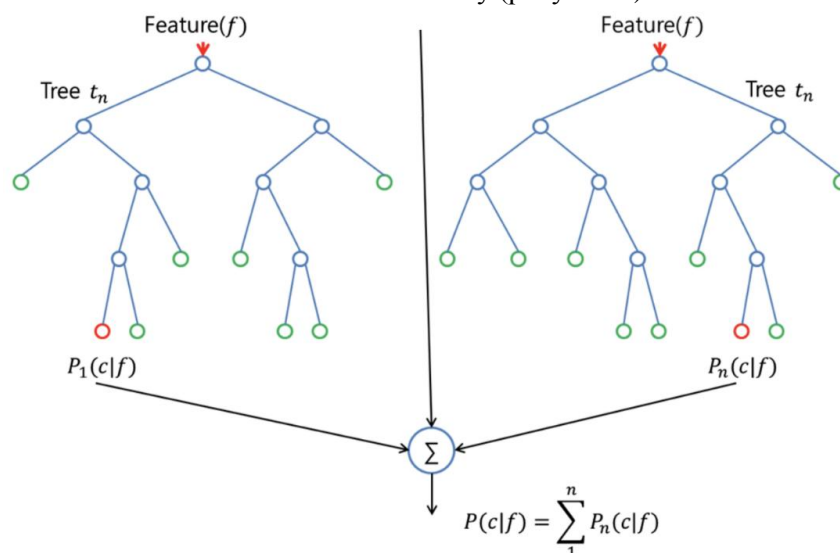


Рисунок 2. Случайный лес.

Каждый лист дерева представляет значение целевой переменной, изменяемой в ходе движения от корня к листу. Каждый внутренний узел соответствует одной из входных переменных. Дерево может быть также «обучено» разделением исходных наборов переменных на подмножества, основанные на тестировании значений атрибутов. Это процесс, который повторяется на каждом из полученных подмножеств. Рекурсия завершается тогда, когда подмножество в узле имеет те же значения целевой переменной, таким образом, оно не добавляет ценности для предсказаний.

ResNet50 [15] является свёрточной нейронной сетью большой глубины. Одной из фундаментальных проблем глубокого обучения является затухающий градиент. Для решения этой проблемы в архитектуре ResNet50 применяется остаточная функция в виде остаточного блока, изображённого на рисунке 3.

Глубокие свёрточные нейронные сети превзошли человеческий уровень классификации изображений. Они извлекают низко-, средне- и высокоуровневые признаки сквозным многослойным способом. Архитектура свёрточных нейронных сетей с остаточными соединениями очень интересна и имеет ряд описанных выше особенностей, в связи с чем было решено использовать архитектуру данной нейронной сети для решения задачи автоматического определения возраста человека.

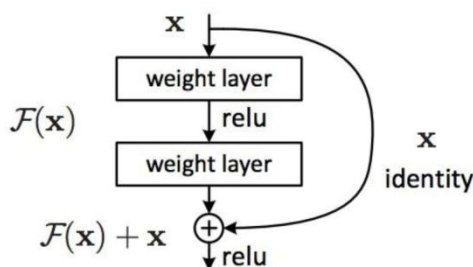
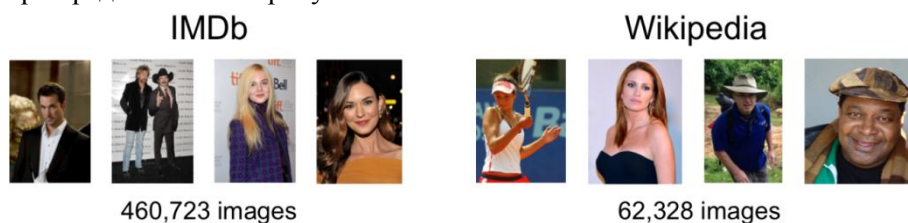


Рисунок 3. Структура остаточного блока архитектуры ResNet50.

В рамках данной работы решается задача классификации, где исходным объектом для классификации является лицо человека. Данные для обучения классификаторов взяты из публичного и доступного на данный момент набора данных изображений лица человека, которые размечены на основании пола и возраста - IMDB-WIKI [21]. Часть примеров из данного набора представлена на рисунке 4.



**Рисунок 4.** Примеры изображений из набора данных IMDB-WIKI.

Стоит отметить, что в представленном наборе данных могут присутствовать неточности в связи со спецификой формирования самого набора данных. В процессе исследований решалась задача устранения влияния несбалансированности набора, которая выражалась в том, что при непосредственном предсказании возраста ответ классификатора относился к той группе возрастов, которая доминировала в общем наборе данных.

Для решения данной задачи было сделано следующее: в первую очередь внимание было обращено на группы возрастов с наименьшим или даже нулевым количеством примеров. В этой связи было решено совместить исходный набор данных с другим, который содержал примеры и в тех группах возрастов, которым было уделено наименьшее внимание. Для обогащения исходного набора данных был выбран набор данных UTKFace [22], особенностью которого является длинная возрастная линия, а именно набор для возрастов от 0 до 116 лет. Также данный набор данных в отличие от исходного имеет большое разнообразие людей, которые относятся к разным расам, что позволило учитывать расу человека при определении его возраста, тем самым улучшая точность определения возраста человека. Кроме этого было искусственно увеличено количества экземпляров, принадлежащих на классы, количество которых намного меньше среднего количества по всем классам, посредством копирования примеров; уменьшено количество примеров доминирующего класса до среднего.

### 3. Использование свёрточных нейронных сетей для определения биологического возраста.

В целях увеличения точности классификаторов, а также уменьшения непосредственного времени их обучения было принято решение разделить возрастную линию от 0 до 100 лет на 8 групп: (0, 2), (4, 6), (8, 12), (15, 20), (25, 32), (38, 43), (48, 53), (60, 100).

Для оценки точности классификаторов, использовалась следующая формула:  $\frac{R}{T} \times 100$ , где R – соответствует количеству предсказаний классификатора, которое совпадает с правильным значением, а T соответствует общему количеству данных в тестовой выборке.

Для решения исходной задачи использовалась архитектура ResNet50, с некоторыми изменениями в конфигурации, к которым, в частности, можно отнести добавление слоя мягкого максимума в конце свёрточной нейронной сети, параметры которого соответствуют количеству возрастных групп.

**Таблица 1.** Результаты тестирования классификаторов.

Метод	Точность классификации, %
RF + HNF	81,1
ResNet50	85,7

Стоит отметить, что результаты, полученные при тестировании сопоставимы с ожидаемыми. В таблице 1 представлены результаты тестирования классификаторов. Предполагалось, что свёрточные нейронные сети не уступят, а даже покажут результаты лучше, чем случайный лес, что и произошло в данном случае. Объяснения этому следующие: во-первых, свёрточные

нейронные сети способны научиться распознавать даже самые незаметные и низкоуровневые признаки на фотографиях даже не самого лучшего качества, во-вторых, свёрточные нейронные сети не требуют однозначного положения объекта на снимке ввиду свойств инвариантности.

### 3. Заключение

В данной работе были продемонстрированы возможности использования CNN для определения биологического возраста человека по изображению, было проведено сравнение с алгоритмом случайного леса. Было устранено влияние несбалансированности исходного набора данных IMDB-WIKI путем комбинирования его с набором UTKFace, что позволило учитывать расу человека при определении его возраста, тем самым улучшая точность определения возраста человека. Также было уменьшено количество примеров доминирующего класса до среднего.

Анализ полученных результатов показал, что CNN успешно справляются с задачей автоматического определения биологического возраста человека по его лицу, превосходя алгоритм случайного леса на 5,7% по точности классификации.

### 4. Литература

- [1] Feng, X. Computer vision algorithms and hardware implementations: A survey / X. Feng, Y. Jiang, X. Yang, M. Du, X. Li // *Integration*. – 2019. – Vol. 69. – P. 309-320.
- [2] Voulodimos, A.S. Deep learning for computer vision: A brief review / A.S. Voulodimos, N.D. Doulamis, A.D. Doulamis, E.E. Protopapadakis // *Computational intelligence and neuroscience*. – 2018. – Vol. 2018. – P. 1-13.
- [3] Robie, A.A. Machine vision methods for analyzing social interactions / A.A. Robie, K.M. Seagraves, S.R. Egnor, K. Branson // *Journal of Experimental Biology*. – 2017. – Vol. 220(1). – P. 25-34.
- [4] Billah, M.F.R.M. SMARTLET: A Dynamic Architecture for Real Time Face Recognition in Smartphone Using Cloudlets and Cloud / M.F.R.M. Billah, M.A. Adnan // *Big Data Research*. – 2019. – Vol. 17. – P. 45-55.
- [5] Codevilla, F. On offline evaluation of vision-based driving models / F. Codevilla, A.M. Lopez, V. Koltun, A. Dosovitskiy // *Proceedings of the European Conference on Computer Vision (ECCV)*. – 2018. – P. 236-251.
- [6] Касторнова, В.А. Искусственные нейронные сети как современные средства информатизации / В.А. Касторнова, М.Г. Можаяева // *Информационная среда образования и науки*. – 2012. – № 7. – С. 1-17.
- [7] Le, Q.V. A tutorial on deep learning part 2: Autoencoders, convolutional neural networks and recurrent neural networks / Q.V. Le // *Google Brain*, 2015. – P. 1-20.
- [8] Arandjelovic, R. NetVLAD: CNN architecture for weakly supervised place recognition / R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, J. Sivic // *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016. – P. 5297-5307.
- [9] Sun, Y. Deepid3: Face recognition with very deep neural networks / Y. Sun, D. Liang, X. Wang, X. Tang // *arXiv preprint arXiv: 1502.00873*, 2015.
- [10] Zhang, K. Joint face detection and alignment using multitask cascaded convolutional networks / K. Zhang, Z. Zhang, Z. Li, Y. Qiao // *IEEE Signal Processing Letters*. – 2016. – Vol. 23(10). – P. 1499-1503.
- [11] Журавский, Р.М. Анализ существующих моделей и методов распознавания лиц / Р.М. Журавский // *European Scientific Conference*, 2018. – P. 41-43.
- [12] Wu, H. Simultaneous face detection and pose estimation using convolutional neural network cascade / H. Wu, K. Zhang, G. Tian // *IEEE Access*. – 2018. – Vol. 6. – P. 49563-49575.
- [13] Макаренко, А.А. Методика локализации изображения лица для систем видеоконтроля на основе нейронной сети / А.А. Макаренко, В.Т. Калайда // *Известия Томского политехнического университета. Инжиниринг георесурсов*. – 2006. – Т. 309, № 8. – С. 113-118.

- [14] Krishna, S.T. Deep Learning and Transfer Learning Approaches for Image Classification / S.T. Krishna, H.K. Kalluri // International Journal of Recent Technology and Engineering. – 2019. – Vol. 7. – P. 427-432.
- [15] He, K. Deep residual learning for image recognition / K. He, X. Zhang, S. Ren, J. Sun // Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. – P. 770-778.
- [16] Akiba, T. Extremely large minibatch SGD: training resnet-50 on imagenet in 15 minutes / T. Akiba, S. Suzuki, K. Fukuda // arXiv preprint arXiv: 1711.04325, 2017.
- [17] Yamazaki, M. Yet Another Accelerated SGD: ResNet-50 Training on ImageNet in 74.7 seconds / M. Yamazaki, A. Kasagi, A. Tabuchi, T. Honda, M. Miwa, N. Fukumoto, T. Tabaru, A. Ike, K. Nakashima // arXiv preprint arXiv:1903.12650, 2019.
- [18] Косова, А.Е. Автоматическая посадка малых беспилотных летательных аппаратов с использованием компьютерного зрения / А.Е. Косова, А.М. Кориков // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2017. – Т. 20, № 3. – С. 191-196.
- [19] Biau, G. A random forest guided tour / G. Biau, E. Scornet // Test. – 2016. – Vol. 25(2). – P. 197-227.
- [20] Men, H. Comparison of SVM, RF and ELM on an Electronic Nose for the Intelligent Evaluation of Paraffin Samples / H. Men, S. Fu, J. Yang, M. Cheng, Y. Shi, J. Liu, // Sensors. – 2018. – Vol. 18(1). – P. 285.
- [21] Набор данных IMDB-WIKI [Электронный ресурс]. – Режим доступа: <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/> (29.11.2019).
- [22] Набор данных UTKFace [Электронный ресурс]. – Режим доступа: <https://susanqq.github.io/UTKFace/> (29.11.2019).

## The using convolutional neural networks for determine the age of a person from an image

A.A. Rud<sup>1</sup>, S.A. Rud<sup>1</sup>, M.A. Isayev<sup>1</sup>, D.A. Savelyev<sup>1,2</sup>

<sup>1</sup>Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

<sup>2</sup>Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

**Abstract.** The paper presents the results of research to determine the biological age of a person using the image of the face. To solve this problem, we used a random forest algorithm using a hybrid Hesse filter and a local binary template operator, as well as a convolutional neural network ResNet50. The data sets used, the problems associated with their application, as well as the accuracy of classification in the selected division of the age line are presented.