

УДК 004.891, 004.932.2

ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧИ РАСПОЗНАВАНИЯ ВОЗРАСТА ЧЕЛОВЕКА

© Колодин В.К., Савельев Д.А.

*Самарский национальный исследовательский университет
имени академика С.П. Королева, г. Самара, Российская Федерация*

e-mail: vkkolodin@gmail.com

В настоящее время задача определения возраста человека может быть решена автоматически и применена во многих сферах, например, для ограничения доступа к Интернет-ресурсам на основе возрастного порога, для улучшения работы рекомендательных систем, для уточнения других биометрических данных в системах безопасности [1; 2]. Возраст человека может быть получен рядом методов с различной степенью точности, в том числе с помощью анализа голоса [3] и лица [4; 5]. Также могут применяться методы машинного обучения и нейронные сети [5].

Данная работа рассматривает решение задачи определения возраста человека с помощью сверточных нейронных сетей и исследует влияние различных факторов на результат оценки возраста.

Набор данных для исследования был составлен из нескольких наборов данных, а именно IMDB-WIKI [6], UTKFACE [7], APPA-REAL [8]. Цветные изображения людей были обрезаны до лиц, а их размер изменен до 224 x 224 пикселя. Набор данных был полностью сбалансирован по возрастам. Всего в тренировочной выборке присутствовало 80 000 изображений.

В контексте решения задачи регрессии (определение возраста как вещественного числа) проводилось исследование с помощью архитектуры MobileNetV2 с классифицирующей частью из трех полносвязных слоев со следующим количеством нейронов: 1024, 512, 1, и функциями активации Swish

Первым делом исследовалось влияние предобработки данных на результат оценки возраста, в качестве экспериментальных данных взяты 5 интерпретаций распределения возрастов, а именно в отрезках [1; 80], [-39,5; 39,5], [0; 1], [-1; 1]. Результат работы сетей на тестовой выборке оценивался метрикой MAE (Mean Average Error – средняя абсолютная ошибка), а также оценивалась сбалансированность предсказаний по возрастам с помощью выборочной дисперсии ошибки по возрастам. Результаты представлены в таблице. Оказалось, что лучшего из полученных результатов можно добиться, используя распределение [-1; 1], что можно объяснить удобностью такого промежутка для оптимизации.

Таблица – Результаты исследования влияния отрезка распределения возрастов

Распределение	MAE, лет, тестирование	Выборочная дисперсия ошибки
[1; 80]	8,75	8,38
[-39,5; 39,5]	8,8	7,18
[0; 1]	8,67	8,62
[-1; 1]	8,49	5,41

Далее проводилось исследование по выявлению лучших из следующих рассматриваемых функций активации: Swish, Leaky ReLU, Tanh; базовых архитектур: MobilenetV2, Xception, Resnet50; ф-ий ошибки: MAE, MSE, Huber (0,5). Выявлено, что самыми эффективными по метрике MAE оказались Swish, Xception и MAE соответственно. Используя их, удалось добиться средней абсолютной ошибки в 7,94 года.

Для варианта с определением принадлежности к группе возрастов были обучены нейронные сети с архитектурами Xception и MobilenetV2 с аналогичными варианту с точным определением возраста полносвязными слоями. Обучение и тестирование проводилось на различных разбиениях возрастов на 8 групп.

При смене количества полносвязных слоев или смене архитектуры с MobilenetV2 на Xception (т. е. увеличения количества обучаемых параметров более чем в 7 раз) точность в пределах одного разбиения у разных сетей не менялась больше, чем на 5 %. Было установлено, что разбиение на группы оказывает ключевое влияние на результат. Таким образом, наихудший результат для различных разбиений – 44,3 % точности, наилучший – 65,5 %.

Также для улучшения результата были взяты лучшие из обученных сетей как для классификации, так и для регрессии, и уже обученные модели объединены в одну сеть с последующим обучением 2 новых добавленных слоев со 128 и 1 нейроном соответственно. То есть сети, которые были уже обучены, в обучении не участвовали, а изменялись веса только добавленных слоев. В итоге результат был улучшен на 20 % по сравнению с предыдущей лучшей сетью для регрессии и составил 6,28 года, что можно объяснить тем, что сети, решающие задачу классификации и регрессии во время обучения, обучались обнаружению и обработке различных множеств признаков изображения, а комбинация сетей учитывала оба множества сразу.

В данной работе проведены исследования влияния функции активации, архитектуры сети, распределения возрастов, а также разбиение их на группы на результат распознавания возраста по изображению. Установлено, что разбиение возрастов больше остальных рассмотренных факторов влияет на результат классификации. Подбор лучших параметров позволил добиться улучшения результатов при использовании архитектуры Xception, функции активации Swish и распределения возрастов в [-1; 1]. А при использовании ансамбля сетей возможно уменьшить ошибку распознавания на 20 %.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования в рамках выполнения Государственного задания (проект № FSSS-2023-0006).

Библиографический список

1. Alkaabi S., Yussof S., Al-Khateeb H., Ahmadi-Assalemi G., Epiphaniou G. Deep Convolutional Neural Networks for Forensic Age Estimation: A Review // Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity. Springer, Cham. 2020. P. 375–395.
2. Han H., Otto C., Jain A.K. Age estimation from face images: Human vs. machine performance // IEEE. 2013 International conference on biometrics (ICB), 2013, P. 1–8.
3. Kwasny D., Hemmerling D. Gender and age estimation methods based on speech using deep neural networks // Sensors. 2021. Vol. 21 (14). 4785 p.
4. Benkaddour M.K. CNN based features extraction for age estimation and gender classification // Informatica. 2021. Vol. 45(5). P. 697–703.

5. Castellano G., De Carolis B., Marvulli N., Sciancalerope M., Vessio G. Real-time age estimation from facial images using yolo and efficientnet // Springer, Cham. International Conference on Computer Analysis of Images and Patterns. 2021. P. 275–284.
6. Rothe R., Timofte R., Van Gool L. Deep expectation of real and apparent age from a single image without facial landmarks // International Journal of Computer Vision. 2018. Vol. 126 (2). P. 144–157.
7. Zhang Z., Song Y., Qi H. Age progression/regression by conditional adversarial autoencoder // Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. P. 5810–5818.
8. Clapés A., Anbarjafari G., Bilici O., Temirova D., Avots E., Escalera S. From apparent to real age: gender, age, ethnic, makeup, and expression bias analysis in real age estimation // Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2018. P. 2373–2382.