The determining age of a person from an image using convolutional neural networks

Alexandr Rud
Samara National Research University
Samara, Russia
sasha_.96@mail.ru

Sergey Rud
Samara National Research University
Samara, Russia
jin_.96@mail.ru

Michael Isayev
Samara National Research University
Samara, Russia
michailisaev.home@gmail.com

Dmitry Savelyev
Samara National Research University;
Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS
Samara, Russia
dmitrey.savelyev@yandex.ru

***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.**

Keywords*—*Random forest algorithm, Convolutional neural network, ResNet50, IMDB-WIKI, Unfiltered faces for gender and age classification

I. Introduction

Currently, the field of technical vision is actively developing [1 – 5]. Detection and recognition of objects are used not only in professional specialized activities [6, 7] but also by ordinary users of smartphones [8]. The tasks of computer vision are very diverse, in particular, it can be text recognition, biometrics, video analytics, analysis of satellite images, image editing, reconstruction of volumetric models, driving a car [1 – 3, 5, 8, 9]. These and many other tasks can be solved using neural networks. In particular, artificial neural networks are used to solve the problem of image recognition [10] and classification of detected objects [11, 12]. The aim of this work is to study the applicability of convolutional neural networks (CNN) [13, 14] to solve the problem of automatically determining the age of a person based on the image.

The estimating a person's age from facial images is a current research topic that has many applications, such as demographic analysis, visual observation, and age progression. The solution to the problem of automatically determining a person’s age is becoming more and more relevant due to the rapid growth of social platforms and media applications, as well as due to marketing research, use in security systems and other areas where age tolerance is possible [15]. Representatives of each age period are prescribed template characteristics, requirements, responsibilities, as well as possible restrictions. Biometric features of a person are unique for each of them. Identification and verification is becoming an increasingly interesting area of research. Fingerprint, face, voice, iris, retina are widely used for authentication [16, 17]. With increasing age, a person's facial features change, and wrinkles appear. That is why the assessment of a person's age from facial images is a very promising task in the framework of identification and verification of people, in visual observation. Some researchers are performed a search for facial areas, which are preferred by machines and people when assessing the age [18]. The authors found that eye area is vital role both for men and CNN. The importance of the other areas vary.

In this paper, we consider the use of convolutional neural networks to solve the problem of establishing the biological age of a person by the image of their face. Despite the fact that the age framework of modern man is extremely blurred, some configurations of artificial networks make it possible to achieve very accurate results [19].

First of all, it is necessary to detect the face in the photo in order to send the corresponding area to the input of the neural network to determine the age. The detection task is to highlight the area of interest in the image or video stream [20]. There are many methods of face recognition [21], and convolutional neural networks are some of the best algorithms for recognizing and classifying images [6, 7, 19, 22, 23, 24].

Generally speaking, learning this kind of approximation is very resource-intensive. As a result, some experimenters have proposed configurations that solve a specific kind of problem more effectively than others. The authors [25] tested several popular convolutional neural network architectures. They showed that ResNet50 [26, 27] not only gives the best results, but also adapts well to images with rotated faces. This architecture we will use as a part of task being solved.

We will train the network using a modified gradient descent (mini-batch) algorithm [28, 29], which makes it possible to minimize the error function.

It is worth noting that the filters and the convolution operation itself give CNN one of the most significant specifics - displacement invariance, i.e., in the case of detection and identification of classification objects, their location on the input image does not matter.

II. Description of Algorithms Used

In this paper, to solve the problem of determining the biological age of a person, the following methods were used: a random forest algorithm using a Hessian hybrid filter (Fig. 1) and a local binary template operator for pre-processing, as well as a convolutional neural network, namely ResNet50.

It is worth noting that in the case of a convolutional neural network it was using images as an input. In the case of the random forest algorithm, the input was not an image, but signs selected by preprocessing with a Hesse filter and then using the local binary template method.

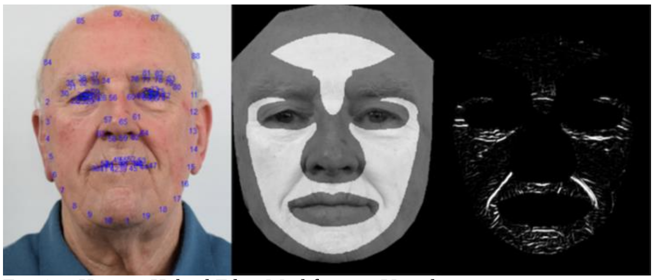

Fig. 1. Hessian Hybrid Filter Modification Visualization.

The Hessian hybrid filter [30] allows you to detect wrinkles by calculating the Hessian matrix for each pixel of the input image. The result of its work is presented in Fig. 1. The local binary template operator is a description of the neighborhood of the image pixel in binary form. The eight pixels around the center pixel take a value of 0 or 1 depending on the threshold, which is the value of the center pixel. This produces an eight-bit binary code that describes the neighborhood of the pixel.

In this paper, to extract age characteristics that are subsequently passed to the classifier, we use a method that simultaneously combines the use of a modification of the Hessian hybrid filter and the subsequent use of the local binary template operator to reduce the amount of data without losing their value in the context of the original task.

The random forest algorithm [31] today is one of the most popular and extremely effective methods for solving machine learning problems, such as classification and regression. In terms of efficiency, it competes with the support vector method, neural networks, and boosting, although, of course, it is not without its drawbacks [32].

The core element of a random forest is a decision tree. The decision tree is a logical scheme that allows you to get the final decision on the classification of the object after answers to a hierarchically organized system of questions. The final decision is made by a majority vote (Fig. 2).

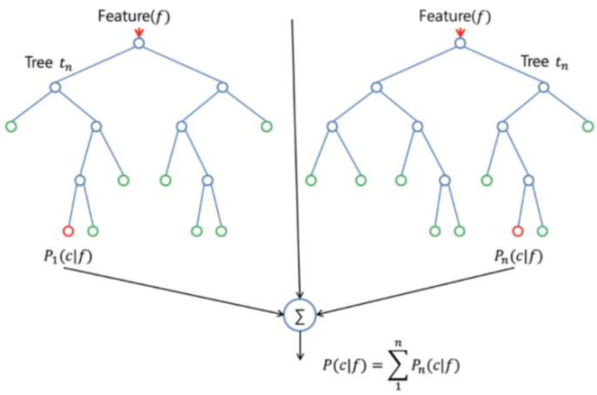

Fig. 2. Random forest.

Each leaf of the tree represents the value of the target variable that changes during the movement from the root to the leaf. Each internal node corresponds to one of the input variables. The tree can also be “trained” by dividing the original sets of variables into subsets based on testing attribute values. This is a process that repeats on each of the resulting subsets. Recursion is completed when the subset in the node has the same values of the target variable, so it does not add value to the predictions.

ResNet50 [27] is a convolutional neural network of great depth. One of the fundamental problems of deep learning is the fading gradient. To solve this problem, the ResNet50 architecture employs a residual function in the form of a residual block, shown in Fig. 3.

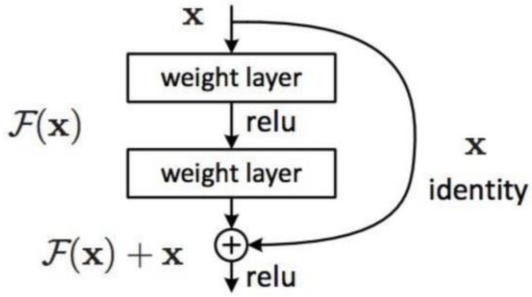

Fig. 3. ResNet50 Architecture Residual Block Structure.

Deep convolutional neural networks have surpassed the human level of image classification. They extract low-, medium-, and high-level features in an end-to-end multilayer manner. The architecture of convolutional neural networks with residual connections is very interesting and has a number of features described above, in connection with which it was decided to use the architecture of this neural network to solve the problem of automatically determining the age of a person.

In the framework of this work, the classification problem is solved, where the person’s face is the initial object for classification. The data for training classifiers are taken from the public and currently available data set of images of a person’s face, which are labeled based on gender and age — IMDB-WIKI [33]. Some examples from this set are presented in Fig. 4.

Fig. 4. Sample images from the IMDB-WIKI dataset.

It should be noted that in the presented data set there may be inaccuracies due to the specifics of the formation of the data set itself. In the process of research, the problem of eliminating the influence of imbalance in the set was solved, which was expressed in the fact that, with a direct prediction of age, the response of the classifier belonged to the group of ages that dominated the total data set.

To solve this problem, the following was done: first of all, attention was paid to age groups with the least or even zero number of examples. In this regard, it was decided to combine the initial data set with another, which contained examples in those age groups that received the least attention. To enrich the initial data set, the UTKFace data set [34] was selected, the feature of which is a long age line, namely a set for ages from 0 to 116 years. Also, this data set, unlike the original one, has a wide variety of people who belong to different races, which allowed taking into account the person’s race when determining their age, thereby improving the accuracy of determining the person’s age. In addition, the number of copies per class was artificially increased, the number of which is much less than the average number for all classes by copying examples; reduced the number of examples of the dominant class to the average.

III. The Use of Convolutional Neural Networks to Determine the Biological Age

In order to increase the accuracy of classifiers, as well as reduce the immediate time of their training, it was decided to divide the age line from 0 to 100 years into 8 groups: (0, 2), (4, 6), (8, 12), (15, 20), (25, 32), (38, 43), (48, 53), (60, 100).

To assess the accuracy of the classifiers, the following formula was used: $\frac{R}{T} \times 100$, where R corresponds to the number of classifier predictions that match the correct value, and T corresponds to the total amount of data in the test sample.

To solve the initial problem, the ResNet50 architecture was used, with some changes in the configuration, which, in particular, include the addition of a soft maximum layer at the end of the convolutional neural network, the parameters of which correspond to the number of age groups.

To test the architecture of the convolutional neural network, additional training and testing were performed on the Unfiltered faces for gender and age classification [35] (UFFGAAC) dataset containing 26,580 photographs. For some images, age and gender labels are missing. Some examples from this set are presented in Fig. 5.

Fig. 5. Sample images from the Unfiltered faces for gender and age classification dataset.

In addition, part of the photos is too blurry, part contains more than one face. Such instances were excluded. Unlike IMDB-WIKI, UFFGAAC is already divided into age groups. The current dataset is also unbalanced as the previous one. The distribution by age groups is shown in Fig. 6.

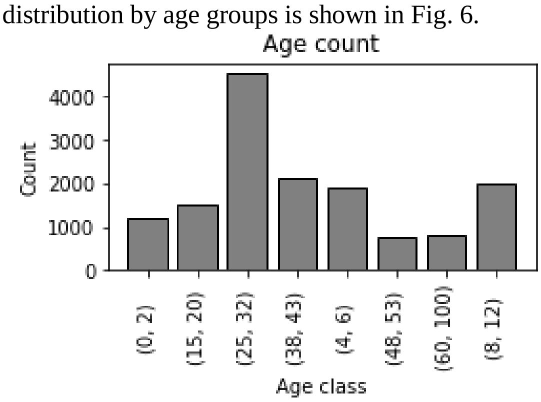

Fig. 6. Age distribution in the UFFGAAC dataset.

The histogram shows that the age group (25, 32) dominates in the data set, and the classes (0, 2), (48, 53), (60, 100) are in excess. In this regard, it was decided to trim the prevailing class, and expand the missing ones from the UTKFace dataset. The final histogram of the distribution of age groups in the dataset is shown in Fig. 7.

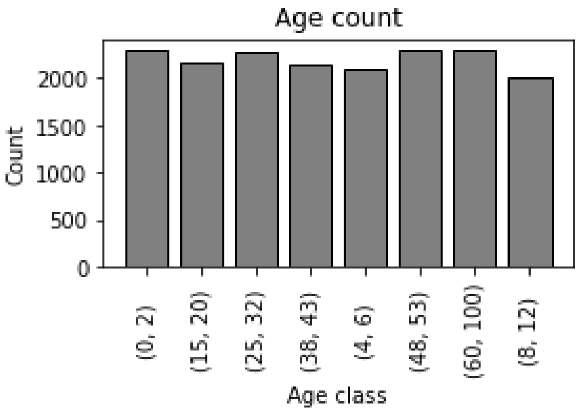

Fig. 7. Balanced age distribution in the UFFGAAC dataset.

It should be noted that the results obtained during testing are comparable with the expected ones. Table 1 presents the results of testing classifiers. Fig. 8 shows several examples of age recognition using a convolutional neural network.

TABLE I. CLASSIFIER TEST RESULTS

Method	**Classification accuracy, %**
RF + HHF (IMDB-WIKI)	81.1
ResNet50 (IMDB-WIKI)	85.7
ResNet50 (UFFGAAC)	82.2

It was assumed that convolutional neural networks would not yield, and even show results better than a random forest, which happened in this case. The explanations for this are as follows: firstly, convolutional neural networks are able to learn to recognize even the most invisible and low-level signs in photographs of not even the best quality, and secondly, convolutional neural networks do not require an unambiguous position of the object in the picture due to the invariance properties.

IV. Conclusion

In this paper, the possibilities of using CNN to determine the biological age of a person from an image were demonstrated, a comparison was made with the random forest algorithm. The influence of the imbalance of the original Unfiltered faces for gender and age classification and IMDB-WIKI datasets was eliminated by combining them with the UTKFace set, which allowed taking into account the person's race when determining their age, thereby improving the accuracy of determining the person's age. The number of examples of the dominant class was also reduced to average.

An analysis of the results showed that CNN successfully copes with the task of automatically determining the biological age of a person by their face, surpassing the random forest algorithm by 5.7% in terms of classification accuracy.

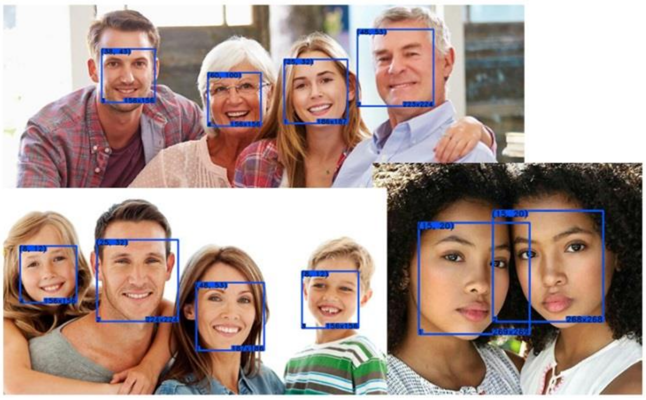

Fig. 8. Samples of recognition.

REFERENCES

[1] X. Feng, Y. Jiang, X. Yang, M. Du and X. Li, “Computer vision algorithms and hardware implementations: A survey,” Integration, vol. 69, pp. 309-320, 2019.

[2] A.S. Voulodimos, N.D. Doulamis, A.D. Doulamis and E.E. Protopapadakis, “Deep learning for computer vision: A brief review,” Computational intelligence and neuroscience, vol. 2018, pp. 1-13, 2018.

[3] A.A. Robie, K.M. Seagraves, S.R. Egnor and K. Branson, “Machine vision methods for analyzing social interactions,” Journal of Experimental Biology, vol. 220, pp. 25-34, 2017.

[4] V.S. Gorbatsevich, A.S. Moiseenko and Y.V. Vizilter, “FaceDetectNet: face detection via fully-convolutional network,” Computer Optics, vol. 43, no. 1, pp. 63-71, 2019. DOI: 10.18287/2412-6179-2019-43-1-63-71.

[5] B.R. Petrovich and I.Y. Zakharova, “Algorithm for tracking people on video sequences using convolutional neural networks for indoor video surveillance,” Computer Optics, vol. 44, no. 1, pp. 109-116, 2020. DOI: 10.18287/2412-6179-CO-565.

[6] Yu.D. Agafonova, A.V. Gaidel, P.M. Zelter and A.V. Kapishnikov, “Efficiency of machine learning algorithms and convolutional neural network for detection of pathological changes in MR images of the brain,” Computer Optics, vol. 44, no. 2, pp. 266-273, 2020. DOI: 10.18287/2412-6179-CO-671.

[7] R.P. Bohush and I.Y. Zakharava, “Person tracking algorithm based on convolutional neural network for indoor video surveillance,” Computer Optics, vol. 44, no. 1, pp. 109-116, 2020. DOI: 10.18287/2412-6179-CO-565.

[8] M.F.R.M. Billah and M.A. Adnan, “SMARTLET: A Dynamic Architecture for Real Time Face Recognition in Smartphone Using Cloudlets and Cloud,” Big Data Research, vol. 17, pp. 45-55, 2019.

[9] F. Codevilla, A.M. Lopez, V. Koltun and A. Dosovitskiy, “On offline evaluation of vision-based driving models,” Proceedings of the European Conference on Computer Vision (ECCV), pp. 236-251, 2018.

[10] J. Zhang, K. Shao and X. Luo, “Small sample image recognition using improved convolutional neural network,” Journal of Visual Communication and Image Representation, vol. 55, pp. 640-647, 2018.

[11] Yu.V. Vizilter, V.S. Gorbatsevich and S.Y. Zheltov, “Structure-functional analysis and synthesis of deep convolutional neural networks,” Computer Optics, vol. 43, no. 5, pp. 886-900, 2019. DOI: 10.18287/2412-6179-CO-565.

[12] R. Magdeev and A.G. Tashlinskii, “Efficiency of object identification for binary images,” Computer Optics, vol. 43, no. 2, pp. 277-281, 2019. DOI: 10.18287/2412-6179-2019-43-2-277-281.

[13] Q.V. Le, “A tutorial on deep learning part 2: Autoencoders, convolutional neural networks and recurrent neural networks,” Google Brain, pp. 1-20, 2015.

[14] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla and J. Sivic, “NetVLAD: CNN architecture for weakly supervised place recognition,” Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5297-5307, 2016.

[15] V.A. Castronova and M.G. Mozhayeva, “Artificial neural networks as modern means of informatization,” Information environment of education and science, pp. 1-17, 2012.

[16] E.D. Rymkevich and E.V. Sholomitskaya, “Identification/ authentication using biometric data,” Belarusian National Technical University, 2019.

[17] E.N. Barashko, E.V. Vernigorova and E.A. Gurin, “Information security technologies. New biometric and identification systems,” Advanced innovative developments. Prospects and experience of use, problems of introduction into production, pp. 131-134, 2019.

[18] J. Alarifi, J. Fry, D. Dancey and M.H. Yap, “Understanding Face Age Estimation: humans and machine,” International Conference on Computer, Information and Telecommunication Systems (CITS), IEEE, pp. 1-5, 2019.

[19] Y. Sun, D. Liang, X. Wang and X. Tang, “Deepid3: Face recognition with very deep neural networks,” arXiv preprint: 1502.00873, 2015.

[20] K. Zhang, Z. Zhang, Z. Li and Y. Qiao, “Joint face detection and alignment using multitask cascaded convolutional networks,” IEEE Signal Processing Letters, vol. 23, pp. 1499-1503, 2016.

[21] R.M. Zhuravsky, “Analysis of existing models and methods of face recognition,” European Scientific Conference, pp. 41-43, 2018.

[22] H. Wu, K. Zhang and G. Tian, “Simultaneous face detection and pose estimation using convolutional neural network cascade,” IEEE Access, vol. 6, pp. 49563-49575, 2018.

[23] A.A. Makarenko and V.T. Kalaida, “Facial image localization technique for video monitoring systems based on a neural network,” News of Tomsk Polytechnic University. Geo-resource engineering, vol. 309, pp. 113-118, 2006.

[24] C.S.N. Koushik, S.B. Choubey, A. Choubey, D. Naresh and N. Reddy, “Determination of age, gender, dress color and type of a person by convolutional neural network (CNN),” Determination of Age, Gender, Dress Color and Type of a Person by Convolutional Neural Network (CNN), 2020.

[25] D.V. Pakulich, S.A. Yakimov and S.A. Alyamkin, “Age recognition from facial images using convolutional neural networks,” Optoelectronics, Instrumentation and Data Processing, vol. 55, pp. 255-262, 2019.

[26] S.T. Krishna and H.K. Kalluri, “Deep Learning and Transfer Learning Approaches for Image Classification,” International Journal of Recent Technology and Engineering, vol. 7, pp. 427-432, 2019.

[27] K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition,” Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.

[28] T. Akiba, S. Suzuki and K. Fukuda, “Extremely large minibatch SGD: training resnet-50 on imagenet in 15 minutes,” arXiv preprint: 1711.04325, 2017.

[29] M. Yamazaki, “Yet Another Accelerated SGD: ResNet-50 Training on ImageNet in 74.7 seconds,” arXiv preprint: 1903.12650, 2019.

[30] E.A. Kosova and A.M. Korikov, “Automatic landing of small unmanned aerial vehicles using computer vision,” Reports of Tomsk state University of control systems and radioelectronics, vol. 20, pp. 191-196, 2017.

[31] G. Biau and E. Scornet, “A random forest guided tour,” Test, vol. 25, pp. 197-227, 2016.

[32] H. Men, “Comparison of SVM, RF and ELM on an Electronic Nose for the Intelligent Evaluation of Paraffin Samples,” Sensors, vol. 18, pp. 285, 2018.

[33] IMDB-WIKI – 500k+ face images with age and gender labels [Online]. URL: https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki.

[34] UTKFace [Online]. URL: https://susanqq.github.io/UTKFace.

[35] Unfiltered faces for gender and age classification, [Online]. URL: https://talhassner.github.io/home/projects/Adience/Adience-data.html.