Age Classification based on images

Petean Darius-Flaviu*, Ungureanu Mihai-Laurentiu*, and Turcut Stefan-Adrian*

*University of Babes-Bolyai, Cluj-Napoca

12 December 2021

Abstract

Age classification has been a difficult problem to solve, both because it is naturally a difficult problem(it's hard to determine the age of someone just by looking at them) and variations in the image(lighting, pose, expression). In this paper we will discuss different approaches and AI models for this problem, as well as our results. We will compare our results with the results of other already existent approaches.

1 Introduction

With the rise of AI in the technological industry, various types of AI based recognitions and classification have emerged with various needs. One of these algorithms are age classification based on visual images or speech patterns. We will try to solve this problem using visual images, since they are more abundant in knowledge in comparison with the speech aproach one. This problem is however, a difficult one to solve due to the complexity of a human face, which can have various factors that are determined by your age, such as: smootheness, amount of wrinkles, bone re-structuring, maturity of tissue, gravity, and so on(according to [1]).

Age estimation and classification is a tehnique of labelling a human age based on appearance only. The age can be either actual age, appearance age and perceived age. Actual age is the number of years accumulated since birth. The apperance age is the assumption of an age that can should be similar to the actual age, however it cannot always be exact due to the stochastic nature of aging among individuals. Aging is uncontrollable, natural and personalized for each individual. The aging process can be determined by various factors, both internal(genes) and external(life style, working enviroment, health..).

2 Related Work

This topic isn't very sought to achieve in the AI community, thus being very few works for this topic. Usually age recognition is paired with gender/emotion recognition which tend to get more attention from the results, thus overlooking the inaccurate age result.

Based on [2] the previous computation work has been has been carried out in two distinct paradigms. In the first paradigm researchers first extract features such as the eyes, nose, etc., then they relate these features geometrically, and finally they use the geometric relationships to aid in analysis and recognition. The current research has adopted this paradigm of locating features and analyzing them for age classification. The second paradigm treats the complete face image as an input vector and bases analysis and recognition on algebraic transformations of the input space.

3 Overview of the Approach

In this paper we will try a simple supervised learning algorithm based on CNNs, in which we will attempt to achieve better or similar results to the current exis-

tent algorithms. This approach belongs to the second paradigm described earlier, we take the whole image and attempt to train a model to recognise the facial patterns of a certain age type. This however requires a lot of data and will require a lot of training before we can get the needed results.

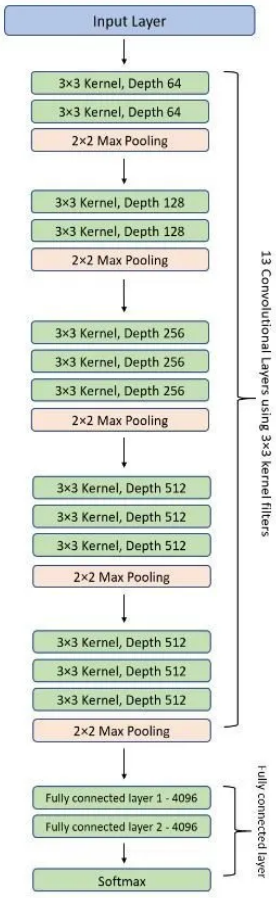

Figure 1: Used Arhitecture

We will start off by using VGG-16 and VGG-19 to see and compare the results on a crawled dataset from the internet. It contains various images of people and their age. The results are presented in the table 1. The results are not good enough for this, however there will be a need to make the dataset bigger and attempt to give it more time to train.

Another approach is to extend the existent pre-trained models, with a few more layers in order to give them more depth that fits for our problem here.

We tried different approaches to this problem, and tried to mix various datasets in order to have a more diverse ecosystem from which we tried to achieve as much diversity as possible, with natural changes to the images, such as the face being obstructed, different lighting , and so on.

For this task we chose to use a combination of UTKFace dataset and WIKI + IMDB faces dataset as described in the [3], we are going the used a crawled dataset from the internet. The dataset contains a large variety of photos with celebrities in various films and other appearances.

Model	Accuracy%	Mean Error
VGG-16	46%	1 interval
VGG-19	48%	1 interval

Table 1: Results

4 Regarding the dataset

We firstly used a dataset found on Kaggle which contained a long list of images with women faces starting from age 18 to 70 years old. We soon discovered that this dataset contained imperfections such as including males or even children disproving the dataset's credibility. Furthermore, this kind of imperfections would later prove to be an impediment in our goal of reaching good and reliable results with our model.

To solve this problem, we later changed the dataset with UTKFace, a large-scale face dataset with long age span starting from 0 to 116 years old, consisting of 20k face images in the wild (only single face in one image), also providing allligned and cropped faces, each one labelled by age, gender and ethnicity.

This dataset [4] is widly known and used in a variety of tasks such as face detection, landmark localization, age progression/regression and so on. This dataset proved to be the best choice by far so we sticked with it throughout the project.

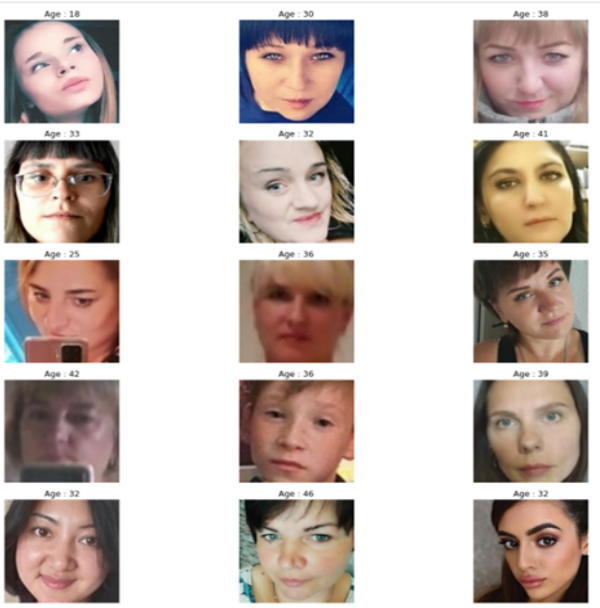

Figure 2: Kaggle Dataset Entries

Figure 3: UTKFace Dataset Entries

5 Regarding the state of the art

When comparing our approach with the state of the art it is clear that there are certain tricks that can be employed to make further progress. Other approaches claim to have reached accuracies over 90% at detecting ages, but they often take into consideration multiple factors. We found the distinction between the biological and apparent age being discussed in the literature and it is clear to us that there is a lot of way for improvement.

Although we are happy to say that our model got better now than it was at our first demo, these numbers can be pushed further with time and further experimentation.

6 Conclusion

As a paper:

This paper tests the VGG-16 architecture as a solution for the task of facial age estimation. Using the UTKFace dataset we tried to find ways to maximize the accuracy of this model by employing different artifices on the original dataset.

As a school project:

This project was a great opportunity for the three of us to learn more about computer vision while working on a project that we were interested about, solving problems as a team. We chose a rather difficult task and that pushed us to research the subject and come up with ideas, discuss and choose the ones we deemed to be the best.

After that, the comes the implementation, then observing and analyzing the results we got. The implementation step was at times a difficult task of its own. For example, while trying to optimize the training process we found that tensorflow wasn't registering the GPU's for some of us, and that sent us searching for alternative solutions and we familiarized ourselves with services like google cloud and kaggle for computing power.

We also found and experimented with datasets we didn't know about, like UTKFace and the IMDB faces sets, and some more obscure ones found on kaggle. Looking back, we had our fair share of speed-

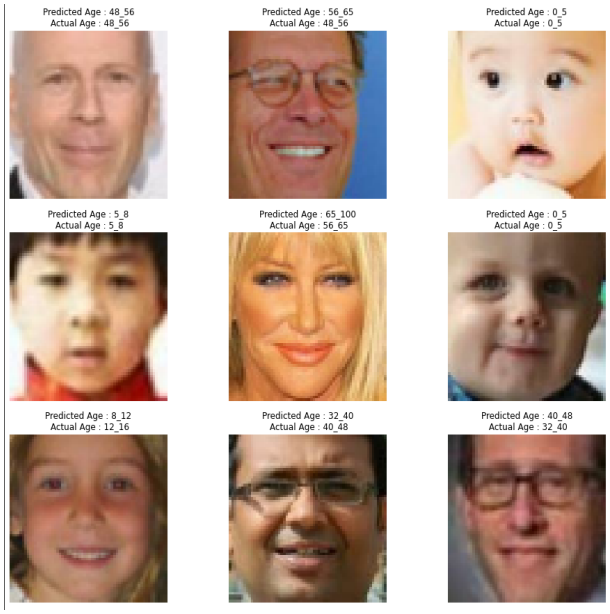

Figure 4: Experimental Results

bumps on our road, but we got through it all stronger than ever, as a team. When it comes to the numbers, as it was stated earlier, we are not claiming to have beaten the state of the art, but we believe that the journey was more important than the destination itself and we are happy that our project got better and better with the time and we found ways to solve problems we have never seen before, so while it is not a technological push, it is definitely a win for us all.

References

[1] R. Angulu, J. R. Tapamo, and A. O. Adewumi, "Age estimation via face images: a survey," *EURASIP Journal on Image and Video Processing*, vol. 2018, p. 42, Jun 2018.

[2] Y. H. Kwon and N. da Vitoria Lobo, "Age classification from facial images," *Computer Vision and Image Understanding*, vol. 74, no. 1, pp. 1–21, 1999.

[3] R. Rothe, R. Timofte, and L. V. Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *International Journal of Computer Vision*, vol. 126, no. 2-4, pp. 144–157, 2018.

[4] S. Y. Zhang, Zhifei and H. Qi, "Age progression/regression by conditional adversarial autoencoder," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2017.