Face recognition in Service robotics: Analysis of the padding effect according to people age

Reconhecimento facial em robótica de serviço: análise do efeito de preenchimento de acordo com a idade das pessoas

DOI:10.34117/bjdv7n12-376

Recebimento dos originais: 12/11/2021
Aceitação para publicação: 12/12/2021

Aron Caiuá Viana de Brito
Graduando em Sistemas de Informação
Instituição de atuação atual: Universidade do Estado da Bahia (UNEB)
Endereço :Rua Silveira Martins, 2555, Cabula. Cep 41150-000, Salvador, Ba Brasil
E-mail: aroncaiua@gmail.com

Ana Patrícia Fontes Magalhães Mascarenhas
Doutorado
Instituição de atuação atual: Universidade do Estado da Bahia (UNEB) / Universidade Salvador (UNIFACS)
Endereço :Rua Silveira Martins, 2555, Cabula. Cep 41150-000, Salvador, BA Brasil
E-mail. apmagalhaes@uneb.br

Josemar Rodrigues de Souza
Doutorado (Ph.D)
Instituição de atuação atual: Universidade do Estado da Bahia (UNEB)
Endereço :Rua Silveira Martins, 2555, Cabula. Cep 41150-000, Salvador, BA Brasil
E-mail:josemar@uneb.br

Jorge Alberto Prado de Campos
Doutorado (Ph.D)
Instituição de atuação atual: Universidade do Estado da Bahia (UNEB)/Universidade Salvador (UNIFACS)
Endereço :Rua Silveira Martins, 2555, Cabula. Cep 41150-000, Salvador, BA Brasil
E-mail: jorge@unifacs.br

Marco Antonio Costa Simões
Mestrado
Instituição de atuação atual: Universidade do Estado da Bahia (UNEB)
Endereço :Rua Silveira Martins, 2555, Cabula. Cep 41150-000, Salvador, BA Brasil
E-mail: msimoes@uneb.br

Robson Marinho da Silva
Doutorado
Instituição de atuação atual: Universidade do Estado da Bahia (UNEB)
Endereço :Rua Silveira Martins, 2555, Cabula. Cep 41150-000, Salvador, BA Brasil
E-mail: robsonms@uneb.br

ABSTRACT
Service robots usually perform repetitive tasks such as collecting garbage, cleaning the house, among others. This kind of robot needs different skills to perform its daily tasks, being people´s recognition a critical skill. One of the techniques used to improve face recognition is padding. The padding technique increases, by a given scale factor, the bounding box of a detected face. In previous work, we had presented a comparative analysis of the influence of the padding in the algorithm used for face recognition. This paper extends the previous analysis by considering the effect of various padding scale factors among different life stages (i.e., toddler, children, teenager, adult, senior, and golden oldie). The result of this analysis shows that increasing the bounding box of detected faces is less efficient for middle-aged people than for younger and elderly people.

Keywords:Service Robotics, people recognition, padding technique.

RESUMO
Os robôs de serviço costumam realizar tarefas repetitivas como coletar lixo, limpar a casa, entre outras. Este tipo de robô necessita de diferentes habilidades para realizar suas tarefas diárias, sendo o reconhecimento das pessoas uma habilidade crítica. Uma das técnicas usadas para melhorar o reconhecimento facial é o preenchimento. A técnica de preenchimento aumenta, por um determinado fator de escala, a caixa delimitadora de um rosto detectado. Em trabalho anterior, apresentamos uma análise comparativa da influência do preenchimento no algoritmo usado para reconhecimento facial. Este artigo estende a análise anterior, considerando o efeito de vários fatores de escala de preenchimento entre os diferentes estágios da vida (ou seja, bebês, crianças, adolescentes, adultos, idosos e idosos). O resultado desta análise mostra que aumentar a caixa delimitadora de faces detectadas é menos eficiente para pessoas de meia-idade do que para pessoas mais jovens e idosas.

Palavras-chave: Robótica de Serviço, reconhecimento de pessoas, técnica de enchimento.

1 INTRODUCTION

Service robots [1] are built to help people in a home environment, perform daily tasks through features such as autonomous navigation, object recognition and manipulation, facial recognition, and voice iteration. Tasks such as hosting people, serving drinks or food according to each person's preferences, throwing out the trash, among others, are tasks performed by a service robot. In this context, this paper deals with the facial recognition ability, which allows the robot to perform actions corresponding to people based on recognized characteristics, such as age, or emotions. For example, the robot may act differently when dealing with an elderly person or a child. In the same way, you can perform actions according to a person's emotions.

For facial recognition tasks, there are already several established methods that can be used, such as [2], Support-Vector [3], and Convolutional Neural Networks [4]. In this paper, we opted for Convolutional Neural Networks, as it is a method widely used by the community and that allows training to be carried out more quickly compared to other methods.

A Convolutional Neural Network [4] is a deep learning algorithm that captures an input and assigns importance as weights and biases to various aspects and characteristics of an image, such as features of a face as gestures, age, features of items or animals, among others. In this way, these characteristics are learned and identified as belonging to a certain object, allowing the algorithm to differentiate that object from others. The main idea is to filter lines, curves, borders, and reduce images for an easier way of processing without losing critical characteristics for good recognition. Soon we have a reduction of image pixels, limiting the characteristics that will be learned. The image pre-processing process has a reduced complexity when compared to other classification algorithms, making their use scalable for a massive set of data. In this process of convolutions, the images may have their pixels reduced faster than is necessary for learning, losing important characteristics for a good forecast. Because of this, Padding is performed, which is a filling process where some pixels are added around the image before convolution. Thus, the dimensionality is maintained even after the image is reduced.

In [5] we presented a controlled experiment in face recognition comparing different padding values in the pre-processing of images processed in a convolutional neural network [4] focusing on how these different values affect the age recognition task. We analyzed the hit rate from the padding perspective and shown only the results of age groups using the best value of padding. In this paper we expand this analysis presenting a discussion of the experiment results considering ranges of ages perspective. We analyze all age groups results using all padding values to see how these age ranges are affected by different padding value. In this direction, we explain the methodology used to build the age model, including training datasets, convolutional neural network architectures, pre-training, pre-processing, how the models were tested and the percentage of correct prediction for each model presented. The age models presented in [6] were used as a reference in choosing the model that best suited our research.

This article is divided into: Introduction where we introduce the subjects that will be addressed in this paper and comment on our objectives; Age Classification Models, that presents the age model adopted in this work; Methodology, that shows how the code for age recognition was structured using padding; Tests and Results, where is commented on how the tests were set up and the results of them, and an analysis is performed of what was observed from these results; Discussions concerning age ranges; and Conclusion and Future Work, concludes the meaning of the presented results obtained in this paper, and discusses about possible future works.

12. AGE CLASSIFICATION MODELS

In order to evaluate the effects of padding on age recognition it is necessary to define a model of face detection and age recognition. This section presents the models adopted for face detection and age recognition. We adopted for face detection, the DeepNeuralNetworks Face Detection [6] module offered by OpenCV [7]. This is a method of good performance and better results compared to other methods such as Haar Cascade [8], HOG [9], Dlib [10], among others.

For age recognition there are different already trained models of convolutional neural networks available in the community that can be reused. Specifically, age recognition models were sought with a classificatory approach. To enable the tests and the evaluation of the padding values, the following classes of age groups were considered: (0-2), (4-6), (8-12), (15-20), (25-32), (38-43), (48-53), (60-100).

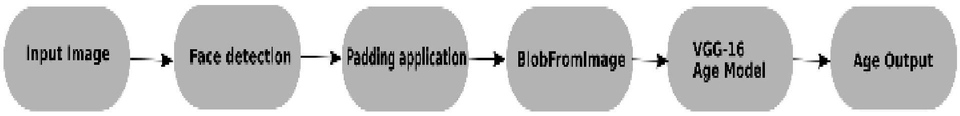
Through the work presented in [6], we had access to different age recognition models, that consider GoogleNet [7], CaffeNet [8], and VGG16 [9] as convolutional neural network architectures. For pre-processing methods is adopted, Unaligned, Landmark-based alignment, and Rotation-based alignment are analyzed. Landmark-based is an alignment created from a reference point, and Rotation-based is a horizontal alignment that gives the image a frontal perspective. Both alignments have the purpose of allowing a greater extraction of characteristics during the processing of the face. Finally, regarding the initialization of weights, that is, pre-training from one or more datasets, we considered the methods Fromscratch, Finetuning, Imdbwiki. These are trained respectively with the datasets: ImageNet [10], ImageNet and ImdbWiki together [11], and ImdbWiki standalone. So, the study presented in [6] demonstrates the performance of each model through its results of success in recognizing gender and age.

To strictly assess the impact of padding, our work presented in our paper adopts the VGG16 model, a robust model that obtained better results considering what was analyzed in [6]. Alignments and training were used with both datasets presented in [6]. The VGG16 model has a hit rate of 63% in age recognition. To assess the impact that padding generates on this result, we adopted the same dataset used to test the model we have chosen as a test dataset. These recognition values are our basic reference for the variations, both positive or negative, that the different padding values can generate in these results. Thus, we have a consistent and efficient analysis to carry out our considerations on padding.

3 METHODOLOGY

The methodology used to test the recognition process comprises: structuring our code to receive input image; face detection; padding application; BlobFromImage; VGG16 age model; Age output. Figure 1 illustrates the idea of this structure.

Figure 1. Methodology structure.



Initially, we collect the input data, which can be a video, or an image containing one person per entry, and load the models of face detection and age recognition. The first process is to detect the face at the entrance and draw a bounding box around this detection, visually displaying the detected face. After detecting the face, we apply padding using a formula taken from [12], which is a paper that describes an application for gender and age classification using padding to improve recognition results. The formula is shown in Figure 2 and is basically the addition of pixels in the dimensions of the image, specifically height and width.

Figure 2. Application formula for *Padding* - Using the image format values, we use padding to enlarge the face detection area in width and height.

```
Face = frame [max (0, bbox[1] - padding): min(bbox[3]+ padding, frame.shape[0] - 1), max(0,bbox[0] - padding):
min(bbox[2]  + padding,frame.shape[1]-1]
```

The *frame* variable shown in Figure 2 contains the image's input data without face detection. From this frame, we will reduce it to the size of the detected face with the coordinates of the bounding box (bbox), x1, x2, y1, y2. Then we will apply the padding in the detection coordinates so that the frame is bigger than the bounding box, starting from the padding value. After that we need to subtract the padding from x1 (bbox[1]) and y1 (bbox[0]) and add the padding to x2 (bbox[2]) and y2 (bbox[3]). The "max" and "min" functions are used so that the vertex values are not negative, nor larger than the height and width of the frame. Finally, we will have a new area of interest, larger than the bounding box for face detection, allowing us to extract more features of the face that will be used later. This new area is stored in "face" variable and will be used in a function of the OpenCV [13] library called blobfromImage(), which performs a pre-processing that applies the mean subtraction [14] and scaling techniques [14]. The Mean subtraction technique is the application of RGB values to the image to assist dealing with changes in lighting. After that, the Scaling process is the application of a scale factor on the image after the mean subtraction process. Finally, a resizing of the image resolution is applied. The input image dimension must have the same value used in the training of the model adopted in the recognition.

We use the input data, which has now been applied to padding and pre-processing, in the convolutional neural networks of age model. After passing through the neural networks, the face in question is predicted, and a visual output with this result will be displayed. Every process described so far is repeated for each detected face and at the end, we have an image with the bounding box of all faces, accompanied by a text showing the results of age prediction.

4 CONTROLLED EXPERIMENT

This section presents a controlled experiment to assess age recognition in face detection using different values of padding and their respective results. The experiments is organized based on the guidelines proposed by Wholin [15] and comprises the following steps: experiment planning; operation; and data analysis and interpretation.

4.1 PLANNING - GOAL, QUESTIONS AND METRICS

The experiment aims to analyze the effects of padding on age recognition, to define a range of values in which positive results are obtained, and from which values we have a reduction in the percentage of correct answers.

The experiment intended to evaluate the results of age recognition with different padding values, considering the percentage of the correctness of each case, observing the reason why these results occurred in order to estimate an average value to use in the *padding* when dealing with age recognition. The evaluation was performed based on the percentage of the correct answer obtained for each padding value.

Figure 3 presents the objective of the experiment defined according to the Goal Question Metric [16] template.

Figure 3. Goals for the analysis of the padding in the Goal Question Metric format.

Analyze padding values in age recognition
for the porpuse of evaluating the efficiency of recognition
with respect to the rate of sucess and error
from the point of view of the VGG16 age recognition model
in context of the Adience benchmark dataset.

To guide this experiment, the following research questions (RQ) were also defined. For each RQ a null hypothesis (H) and an alternative hypothesis (HA) were also defined.

RQ1: Does the increase in the padding value influence the hit rate? This question aims to get insights that increasing values of padding generates positive results.

H0: The increase in padding value does not influence in better hit rates.

H1: The increase in padding value leads to better hit rate.

RQ2: To what extent can the use of padding be beneficial to the results of age recognition? This question seeks to analyze in which, from what number of padding, the results start to decrease its success rate.

H2: Excessive use of padding affects results negatively.

HA3: There is a range of padding values that the results reach a peak percentage of correctness.

To measure our experiment we defined a metric named *correctness rate*. It represents the correct recognition rate of an age group and is calculated as: the ration between the total number of people in an age group correctly recognized and the total number of people tested. A correctness rate must be calculated for each padding value tested with the same group of images. Following the same idea, the incorrect recognition rate for the age groups, is calculated by the ratio between the total of incorrectly recognized age groups and the total number of people tested, also being calculated for each padding value.

4.2 PLANNING - CONTEXT

The controlled experiment was conducted in a home office environment with remote participation of students and professors from the Computer Architecture and Operating Systems(ACSO) research group. We also had the participation of all the people listed in the dataset [17], as we used their photos to carry out the tests explained in the next subsection.

Among the students and teachers who conducted the tests, one was selected who had a computer with greater computational power to perform the tests. The specifications of this computer are listed below:

- Processor Intel core i7-3612QM CPU @ 2.10ghz x 8
- RAM 7 GB
- Hard disk 983,4 GB
- Integrated video card Intel HD Graphics 4000(IVB GT2)

4.3 PLANNING - EXPERIMENT

The experiment was conducted using the same dataset that the age model used as a test basis: the rotation aligned test-set from the Adience benchmark dataset [17]. This dataset consists of 19370 photos however to facilitate the tests, we selected only 200 photos at random, with 200 different people, 1 person per photo, varying in all the age groups listed in the previous session. The idea then is to define different padding values and analyze the ranges in which there is a difference in percentage, negative or positive, that they generate from our base.

4.4 PLANNING - PADDING SELECTION

In carrying out the experiment it was necessary to decide the values of our independent variable, the padding. So as, we looked for references that indicated typical values used for padding in recognition tasks and selected the values 8, 16, 24 and 32 pixels recommended in [18]. In addition to these values, we selected a minimum value and a maximum value as extreme values, to also observe how they behave. These values were 0 and 64, with no padding and double the highest value also from the first selection of values. The dependent variables that have been defined for this experiment are the correctness of the correct and incorrect recognition rate for each padding value.

4.5 OPERATION

For the operation of the experiment we developed a script responsible for copying the path of each of the photos and run the application. So the tests with 200 photos were performed one at a time, recording the result obtained in a file. After the execution, the results obtained in each photo were manually checked from the correct ages according to [17]. The experiment was carried out on 19 of February 2021 in the environment specified in the subsection "Planning - Context".

4.6 Results - Data collect

The data collected after the experiment were used for the analysis of the experiment. Through the Figure 4, the results obtained for the selected padding values were displayed. Each bar represents the results obtained as a percentage of each padding value, which were initially: 0; 8; 16; 32; 64, respectively from left to right. Figure 5 shows the exact number of hits and errors for each padding value and finally the percentage of hits for each. The left column displays the correct answers, the middle column displays the errors, and the right column displays the percentage of correctness.

Figure 4. Age Padding Results - Using the finetuning _mixed _vgg16 model, and varying the padding values between 0.8,16,32 and 64. we obtain these results.

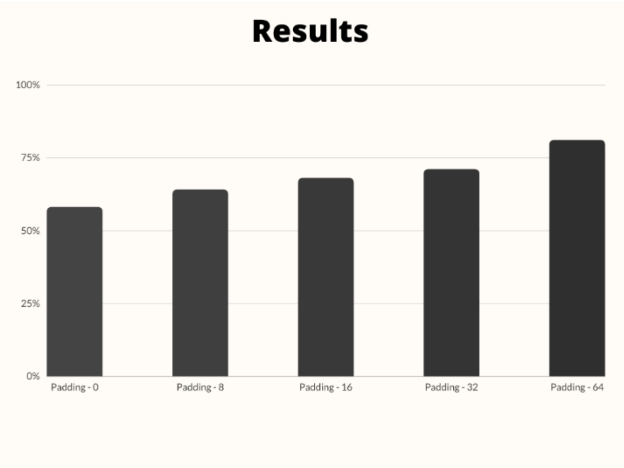


Figure 5. Age Padding Results table - Using the finetuning _mixed _vgg16 model, and varying the padding values between 0.8,16,32 and 64. Number of hits and errors for each padding value, and percentage of hits.

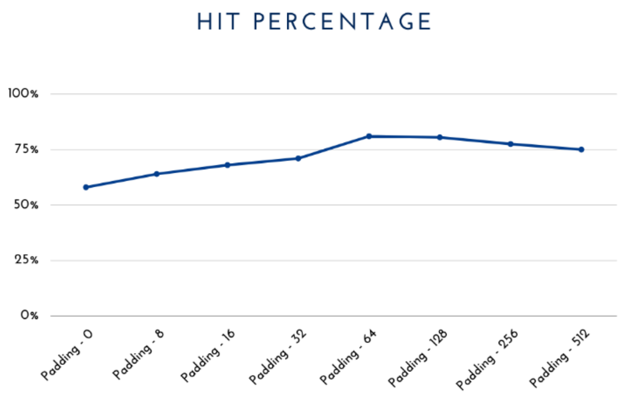
RESULTS

	Hit	Error	Hit rate
Padding - 0	116	84	58%
Padding - 8	128	72	64%
Padding - 16	136	64	68%
Padding - 32	142	58	71%
Padding - 64	162	38	81%

4.6 DATA INTERPRETATION AND ANALYSIS

With the results shown in figures 4 and 5 it was observed how the addition of higher values of padding allowed an increase in the percentage of correct answers. For example, between the values 0 to 64 pixels, we observed 23% of increase on correct answers. Besides, more tests were performed using higher padding values in order to evaluate if the hit rate continued rising, started to decline or remained at the same rate. So, we tested the values 128, 256, and 512. To cover a long range of value we defined that the new tested value should be double the previous value. The new results are shown in the Figure 6.

Figure 6. Age Padding hit percentage - The graph shows the 23% growth curve in the hit rate.



The hit rate growth occurred due to of the increase in the region of interested (i.e. the use of padding) for each submitted face. It happens because in the convolution process less characteristics were lost facilitating the recognition of the correct class for each group. However, the excessive increase in the padding values may expand the region of interest more than what is adequated including characteristics that are not related to the face. As a consequence, it can generate errors in the age range forecast.

As we can see in Figure 6, we got 80.5% using padding = 128, 77.5% using padding = 256 and 75% with padding = 512. The tested values greater than 64 pixels show that the hit rate starts to decline gradually. For example, between the 64 padding values to 512, we computed a total of 6% loss in the hit rate concerning the highest achieved value.

To demonstrate how each of the age group classes was affected by these padding values, we computed the percentage of the correctness of all classes using padding 0 and padding 64, with the lowest overall accuracy rate and the highest general accuracy rate, respectively. This percentage of correct answers for each class refers to the number of correct answers for the class divided by the number of times the class was recognized. The results for padding 0 were: (0-2) =57%, (4-6)=60%, (8-12)=37%, (15-20)=41%, (25-32)=66%, (38-43)=66%, (48-53)=58%, (60-100)=72%. The results for padding 64 were: (0-2) =100%, (4-6)=100%, (8-12)=87%, (15-20)=73%, (25-32)=72%, (38-43)=87%, (48-53)=86%, (60-100)=95%.

4.7 HYPOTHESES EVALUATION

Based on the results presented in the previous subsections, we answered the research questions and evaluated the hypotheses. Concerning to RQ1: although the addition of pixels in the image filling generates a significant increase in the results, its excess ends up impairing the prediction. Therefore, the hit rate does not always grouth with the increase of number of padding. This is because by expanding the region of interest of the face too much, it is possible to process characteristics beyond the face that are not relevant for the recognition of age. The answer to that question fits to evaluate our first hypothesis HA1. The increase in padding does generate better results, such as the 23% increase between values 0 to 64, however, this has a limit because from this value a decrease in the percentage of correctness was noticed. Related to RQ2: With the results obtained, the percentage of correctness started to decline with padding values greater than 64 pixels, exactly 6% reduction compared to the highest percentage reached (81%). The two best results are in the range of 32 to 64 pixels, where we got about a 20% increase in results compared to not using padding. With that answer, we could also answer our remaining two hypotheses. Hypothesis H2 states that the excessive use of padding generates negative values in the results, and with the percentage of the correctness of each of the padding values, it was noticed a decrease of 6% of correctness for the values greater than 64 pixels, therefore the hypothesis proved to be correct. Finally, Hypothesis HA3, which stated that there is a range of values where we have a peak of growth in the results. RQ2 also helps with this assessment. Different padding values and their results were analyzed, so we could define a range of values that obtained better results. So, with the answer to the research questions, we observed the range from 32 to 64 pixels for age recognition divided into age groups, with photos containing only 1 person. Values greater than this range showed a rate of decrease and we believe that from that point on, the higher the value of the padding is, greater the rate of decreases. Values lower than this range, despite generating positive results in comparison to not using padding, have still inferior results to the defined range, which has 20% growth in results. It is worth mentioning that this is an initial analysis based on the context used, on the configurations adopted and on the results obtained.

4.8 TEST VALIDITY

It is still necessary to expand these texts to other contexts in other to understand whether these results apply in general to age recognition. Therefore, it is necessary to carry out further tests with different configurations to arrive at a more complete analysis and possibly define a more extensive range of values.

In the context of 200 images, one person per photo, we had promising initial results, computing 23% increase in results using 64 pixels of padding in comparison to not using padding. In [6], about 56.5% of correct answers were obtained in the age recognition without the use of padding, therefore we had approximately 25.5% increase in the correct rate in conjunction with this study, however, the value of 56.5% obtained in [6] was computed using a much larger test database, which highlights the need to increase our test dataset to validate the results using padding values.

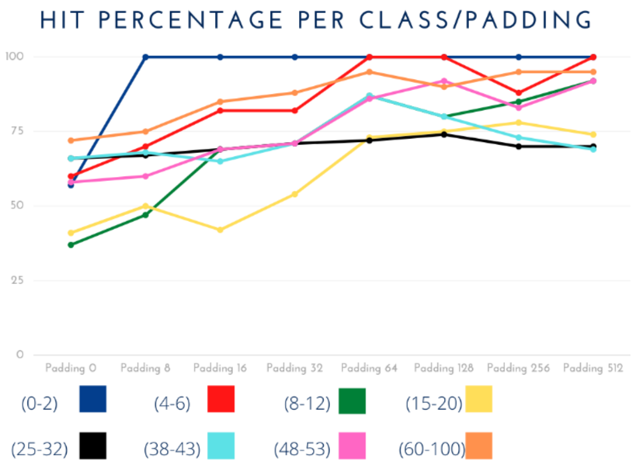
Considering the results by class with the use of padding, we obtained respectively 53%, 40%, 50%, 41%, 6%, 21%, 28%, 23% increase in the hit rate for the classes (0-2), (4-6), (8-12), (15-20), (25-32), (38-43), (48-53), (60-100). This increase in the hit rate was computed using padding in the value of 64, which was the value that obtained more hits considering all classes. As mentioned in the Data interpretation and analysis section, padding enhanced the results of all classes, being more impactful on classes that have characteristics that are easier to be distinguished and not confused. However, there is a need to perform more tests to validate this idea, as the test dataset does not have the number of photos equally divided for all classes, which may favor some age groups in favor of others.

From 64 pixels to 512 pixels we had a loss 6% of the hit rate in comparison to the highest percentage. For a more consistent and broad analysis to define a range of padding values suitable for age recognition, it is still necessary to carry out tests with more photos, mainly with more than one person per photo. The increase in the region of interest of a face, may end up colliding with the region of interest on another side of the image that are close. So, high padding values can generate positive results for photos with only one person, but negative results with two or more people.

5 DISCUSSIONS CONCERNING AGE RANGES

With the worst and best padding value results presented in Section 4, it is interesting to see how the results for each padding value vary across classes, to assess whether some classes are more impacted by padding values than others. To facilitate the visualization of the results, we provide the chart below showing the variation of the results for each class with all padding values. Each line represents a class, the X-axis represents the padding values and the Y-axis represents the hit percentage.

Figure 7. Age Padding hit percentage by class - The graph shows the variance in percentage of all age groups



According to the results shown in Figure 7, we can see how some classes are more affected by the padding change than others. The classes that reached the highest percentage of correct answers were (0-2), (4-6), and (60-100), being respectively 100%, 100%, 95%. In the initial results presented in [6] these classes already had the best results compared to the others and this is due to the characteristics of these age groups being easier to differentiate compared to others. We can note that the application of padding further leveraged these results, generating a variation of 53%, 40%, and 20%, this result being the subtraction of the padding that obtained the best result by the result of the initial padding value. This variation can also be attributed to characteristics being more specific and more difficult to be confused with characteristics of other classes. The age groups (15-20), (25-32), (38-43) reached their maximum 78%, 74%, and 87%, with a variation of 37%, 8%, and 21% respectively, considering the padding that got the best result and the result of the initial padding value. Also from [6] it is reported that these 3 classes have a lower hit rate compared to the others since their characteristics are more relative. This means that it is more common for people from one of these 3 ranges to have characteristics of the other ranges and this makes it difficult for the algorithm to be accurate, requiring greater attention to the training of these 3 classes. The use of padding reflects this pattern, as despite generating a significant increase in results, the variation tends to be smaller than in other classes. Finally, we have classes (8-12) and (48-53) that both reached 92% maximum accuracy, with a variation of 55% and 34%.

The padding use generates this increase in all classes because the expansion that is generated in the image, simply allows more characteristics of the classes to be recognized on the faces, so, for example, if the characteristics of the class (0-2) are easier to be identified, padding allows it to identify even more of these characteristics, which ends up causing an increase in the hit rate and justifies why padding impacted even more in this class.

6 CONCLUSION AND FUTURE WORKS

This article describes an analysis of the use of different padding values applied to age recognition considering different age range groups, noting how their use can impact in different ways each age group depending on the chosen padding value. In this article we expand the results from [5] showing the results in the age group perspective. According to the results and analysis of our experiment we could observe how the increase of padding value in pixels, expands the region of interest of the face, allowing important characteristics not to be lost in the convolution process. Consequently, there is an increase in the percentage of correctness of the model. Therefore, the excessive increase in padding can impair the predictions. The region of interest starts to capture characteristics not related to the age group, decreasing the percentage of correctness of the model. We can note that age groups with younger and elderly ages have more efficient results and the padding use increase even more these results. Therefore in the middle-age classes that have more difficult to predict correctly, the padding use increase the results but with less impact. In fact the classes that characteristics are easier to identify, padding allow it to identify even more because its just getting a bigger region of interest from the face.

It is important to consider these results as an initial analysis. To obtain a more consistent analysis, it is necessary to increase the number of photos in the test dataset, and also observe how the increase in padding values affects in photos where there are two or more people. In this way, it may be possible to find a padding value which generates an increase in the correctness rate of predictions using photos with any number of people.

For future work we aims to expand test scope increasing the number of photos used as well as varying the number of people in these photos. For that, we now are searching for different datasets that vary in this characteristic and have the age group data of each person. In addition, we can also perform the same procedure described in this article with another type of recognition, which can be recognition of gestures, emotions, among others, in order to observe and verify whether the behavior patterns of the padding application are the same in relation to age groups. It is worth mentioning that the use of padding is a way to help in the processing, being able to improve the results in small percentage ranges. The main results are due to the structure and training of the CNN models for each task and padding comes in as an aid to these models, but it cannot generate major changes in models that were not well trained or not well structured.

ACKNOWLEDGEMENT

This project was supported by FAPESB.

REFERENCES

[1] D. Belanche, L. V. Casal´o, C. Flavi´an, and J. Schepers, "Servicerobot implementation: a theoretical framework and research agenda,"The Service Industries Journal, vol. 40, no. 3-4, pp. 203–225, 2020.[Online]. Available: https://doi.org/10.1080/02642069.2019.1672666

[2] M. Anggo and La Arapu, "Face Recognition Using Fisherface Method,"Journal of Physics: Conference Series, vol. 1028, p. 012119, Jun. 2018.[Online].Available:https://iopscience.iop.org/article/10.1088/1742-6596/1028/1/012119

[3] Guodong Guo, S. Z. Li, and Kapluk Chan, "Face recognition by supportvector machines," pp. 196–201, 2000.

[4] S. Khan, M. H. Javed, E. Ahmed, S. A. A. Shah, and S. U. Ali, "Facialrecognition using convolutional neural networks and implementation onsmart glasses," pp. 1–6, 2019.

[5] A. Brito, V. Martins, L. Queiroz, D. Barbosa, T. Silva, A. P.Magalh˜aes, J. Souza, and J. Campos, "Analysis of the paddingeffecton ageclassificationappliedto servicerobotics."IVBrazilian Humanoid Robot Workshop (BRAHUR) and the V BrazilianWorkshop on Service Robotics (BRASERO), 2021. [Online]. Avail-able: https://www.even3.com.br/anais/brahurbrasero/384165-analysis-of-the-padding-effect-on-age-classification-applied-to-service-robotics/

[6] W. Samek, A. Binder, S. Lapuschkin, and K.-R. Muller, "Understandingand Comparing Deep Neural Networks for Age and GenderClassification," in2017 IEEE International Conference on ComputerVision Workshops (ICCVW). Venice: IEEE, Oct. 2017, pp. 1629–1638.[Online]. Available: http://ieeexplore.ieee.org/document/8265401/

[7] "GoogLeNet - Going Deeper with Convolutions," p. 35.

[8] "Caffe|DeepLearningFramework."[Online].Available:https://caffe.berkeleyvision.org/

[9] V. Khandelwal, "The Architecture and Implementation of VGG-16," Aug. 2020. [Online]. Available: https://pub.towardsai.net/the-architecture-and-implementation-of-vgg-16-b050e5a5920b

[10] "ImageNet." [Online]. Available: http://www.image-net.org/

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

[12] "Gender&AgeClassificationusingOpenCVDeepLearning(C++/Python),"Feb.2019.[Online].Avail-able:https://learnopencv.com/age-gender-classification-using-opencv-deep-learning-c-python/

[13] “Opencv,”OpenCV. [Online]. Available: https://opencv.org/

[14] “Deep learning: How OpenCV’s blobFromImage works,” Nov. 2017.[Online]. Available: https://www.pyimagesearch.com/2017/11/06/deep-learning-opencvs-blobfromimage-works/

[15] C. Wohlin, P. Runeson, M. H¨ost, M. C. Ohlsson, and B. Regnell,Experimentation in Software Engineering.Springer, 2012.

[16] V. R. Basili, G. Caldiera, and D. H. Rombach,The Goal Question MetricApproach. John Wiley & Sons, 1994, vol. I.

[17] “FaceImageProject-Data.”[Online].Available:https://talhassner.github.io/home/projects/Adience/Adience-data.html

[18] “HOGdetectMultiScaleparametersexplained,”Nov.2015,section:ImageDescriptors.[Online].Avail-able: https://www.pyimagesearch.com/2015/11/16/hog-detectmultiscale-parameters-explained/