Identifying hair traits in a facial image

Alexandra-Natalia Tudorescu

ACM Classification:

I. Computing Methodologies

I.4 IMAGE PROCESSING AND COMPUTER VISION

I.4.7 Feature Measurement

MSC Classification:

68-XX COMPUTER SCIENCE

68Txx Artificial intelligence

68T45 Machine vision and scene understanding

Abstract

Computer Vision problems involving hair have been studied for years, with application purposes in different domains. Andre Walker's hair curl type classification hasn't yet been approached as a computer vision problem yet, though. Hair curl type is a feature whose correct identification in a virtual setting could have applications such as creating a custom hair guide for a person based on their picture, identifying the ethnic or racial background of a person, or be integrated in other beauty-related virtual tools. In this paper, different CNN architectures have been tried for solving this problem, eventually resulting in an 71.69% accuracy.

1. Introduction

Hair traits are a distinguishing feature in the appearance of a human being. They can be used in helping determine a person's background, such as ethnicity and race. They are also useful in beauty industry-oriented applications.

Practicing has always been challenging for students in the beauty industry, especially in the hairstyle branch, as it involves finding hair models or spending a large amount of money on hair dolls. New learners are also prone to mistakes, which are harder to correct in a real environment, as some actions performed on hair cannot immediately be undone (i.e. making a wrong cut). In the time of the pandemic, this issue has become even more prominent, as human contact has been limited and the number of in-person trainings has decreased. A tool allowing the user to perform hair training activities on a virtual model would make for a great learning resource for a beginner, as well as for a more experienced individual who cannot find real models and would like to practice new techniques. The idea of identifying hair traits in an image of a real person and modifying them is useful because it contributes to creating a realistic scenario and it also provides an insight into how the real-life person would look with the modified hair features.

The aim of this research is to provide a foundation for creating a valuable tool for practicing hairstyling techniques, as described above.

2. Related Work

The problem of hair segmentation, as well as of classifying specific traits, has been approached before, from various perspectives and using different implementations.

In [1], the focus is, as the title suggests, on hair segmentation and color classification, from a visagisme perspective. In this approach, the segmentation and classification are done on facial images, face detection playing an important role in identifying the hair region. For the segmentation problem, a U-Net fully convolutional network with two symmetric parts is used. In the first part, represented by the contraction path, the image is down-sampled, using depth-wise convolutions. In the second part, respectively the expansive path, the down-sampling operations are reverted by using transpose layers. A flood-fill algorithm is later used to complete the gaps in the detected hair mask, and an algorithm based on detecting the face outline determines whether the subject is bald or not. For the hair color classification part, a normalized color histogram is fed into a classical ANN, using the RGB, HSV and Lab representations.

S Sarraf focuses, in [2], on the idea that detection of the hair region is of significant importance in automatic human identification for security reasons. Unlike the previously described work, the current paper classifies detected hair into black or non-black. Low-level features (R, G, B, H, S, V) were extracted, a vector normalization was performed and then fed into two types of classifiers: SVM and K-NN. Both had good results on the larger database (460 images), but worse performance on the smaller database (200 images).

An issue that has occurred when trying to solve hair segmentation and classification problems is the relatively limited number of images available in datasets ready to be used. This issue is mentioned or can be noticed in the papers above as well, as the authors had to annotate images themselves or used a relatively small dataset. [3] describes the creation of a large dataset designed specifically for this type of work. For creating the dataset, both an auto-labelling model and humans were used. The resulting masks were used in three types of hair manipulation applications. For hair dyeing, the segmentation masks needed to be predicted corresponding to hair regions. U-Net, Feature Pyramid Network, Pyramid Scene Parsing Network, DeepLabv3+ and Pyramid Attention Network are used for segmentation, and after comparing the results, PAN seems to be producing the most reliable results. K-nearest-neighbor matting and alpha matting are further used to produce accurate segmentation masks. The color is finally changed in a realistic manner by modifying HSV instead of RGB. For hairstyle transfer, a reference-based image-to-image transfer model, StarGAN v2 [5], is used, managing to transfer the hairstyle even when the position or gender of the target subject differs from that of the reference model. For hairstyle classification, ResNet101 [4] produces the best results.

These papers tackle the problem of hair classification from various perspectives, as they have different purposes. The idea of the current research is to try these approaches on a similar, yet different problem, of classifying hair into 12 types based on the curl pattern. This is challenging because of the similarity among close classes and their overall big amount. This is why very high accuracy when it comes to the exact classification is not expected, but it is hoped that a higher accuracy will be obtained if close classifications (one or two classes apart) are also taken into account.

3. Original Contribution

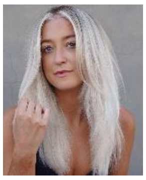
As outlined above, there has been quite some research on the topic of hair detection and classification. One classification that has never been implemented as a computer vision solution, though, is the classification depending on the curl type, namely that defined by Andre Walker, as represented in the table below (section 5). Through the current paper, the aim is to find the best fitted classification method for the problem of curl type, comparing to the methods used in literature for other hair features.

4. Dataset Collection and Annotation

Since hair related computer vision problems have been studied in the past few years, there exist a few datasets consisting of annotated hair images, but they are not all completely suited for the current problem, since they don't contain the exact curl type annotation. K-Hairstyle [18] and Figaro 1K [19] are datasets containing annotated images of real people, and the former is used for several applications described in [3] and mentioned in the "Related Work" chapter, though it censors faces. Hairstyle 30K [20] is generated by a GAN.
However, for the current research, a custom dataset has been created, taking into account the 12 curl types. Most pictures were taken from the Internet, some from the other datasets, and all have been manually annotated to correspond to a class. Variety regarding gender, skin colour, race, hair colour and hair length has been taken into account in order to avoid biased results because of an unbalanced dataset.

5. Proposed Solution

The aim of this paper is to find the most accurate method of classifying hair curl types according to the 12 labels detailed in the table below, from 1A to 4C. It is important to notice that even for the human eye, it is sometimes hard to differentiate among types that are one after another. Thus, the strict classification is not the only one that we follow to evaluate, but we are also interested in how close the prediction label is to the actual label.

	A	B	C
1 (straight)	Fine, straight hair 	Thicker straight hair 	Coarse straight hair 
2 (wavy)	Fine, S-shaped curls 	Looser curl, still S-shaped 	Loose curl 
3 (curly)	Shiny, defined curls 	Combination of curls and coils 	Combination of defined coils 
4 (coily)	Soft, well-defined coils	Shiny, less defined coils	Sensitive, undefined coils

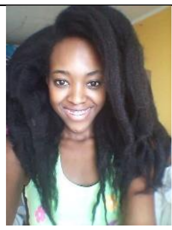
			

Table: Andre Walker Hair Typing System

5.1. Metrics

We are interested in the accuracy of the classification, i.e.
accuracy = (TP + TN) / (TP + TN + FP + FN),
where TP = true positive, TN = true negative, FP = false positive, FN = false negative

5.2. Case Study – Experiments

For classifying hair traits, we follow the approach of using Convolutional Neural Networks, and experiment with different configurations.

1. With the aid of the Tensorflow Keras model, CNNs with different layer configurations are built.

    a. The first configuration has 10 layers, including 5 Conv2D layers, 2 resnet blocks, one Global Average Pooling 2D layer and a Dense layer. As an optimizer, Adam is used, and since it is a classification problem, we use SparseCategoricalCrossentropy as the loss function. The noticed results using 32 batches is unsatisfactory, with a training accuracy of 0.12 after 20 epochs.

    b. Another tried configuration consists of MaxPooling2D and Dense layers as well, and it produces slightly better results faster, respectively a training accuracy of 0.1425 after the 2nd epoch, after which it drops again.

    c. After these unsuccessful experiments, we modify the architecture: Conv2D layer, the number of resnet blocks of various sizes is increased to 12, a GlobalAvgPool2D layer and a Dense layer. After adjusting the activation function on the last layer and adding a dropout layer, satisfactory results are achieved with a 50-epoch training.

2. In [3], the most efficient approach for hairstyle classification was using a ResNet101. Indeed, using a pre-trained model has been observed to perform well on smaller datasets. This is reflected in the current case as well, since in a dataset with 12 classes and approximately 50 input pictures/class, ResNet101 achieves a training accuracy of 0.84 after only 10 epochs, an actually good accuracy, especially compared to the approach described at 1.a. and 1.b.

5.3. Results

	Last layer activation function	Other information about architecture	Training Loss	Validation Loss	Training Accuracy	Validation Accuracy
Fig. 1	ReLU		2.4849	2.4849	0.0947	0.1132
Fig. 2	softmax		0.0001	8.3255	1.0	0.6792
Fig. 3	softmax	Contains dropout (0.2)	1.0868	14.6009	0.9573	0.7169

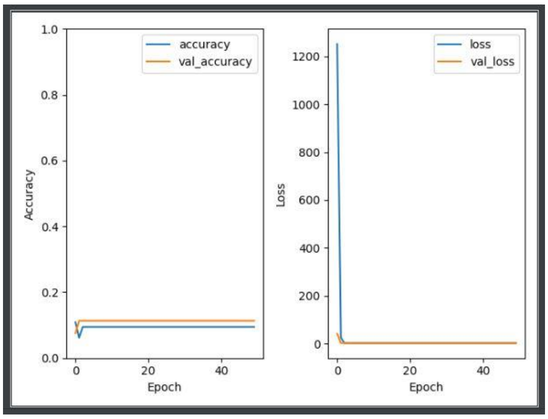


Fig. 1. CNN 5.2.1.c, relu activation

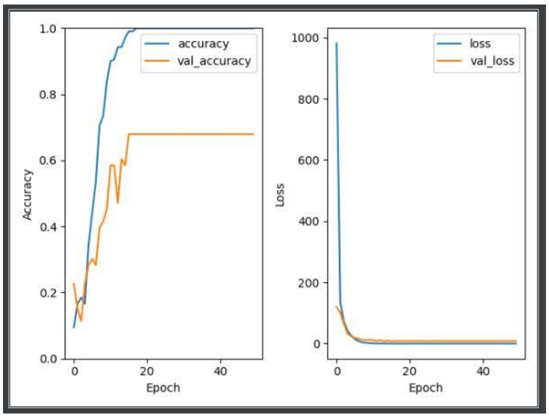


Fig. 2. CNN 5.2.1.c, softmax activation

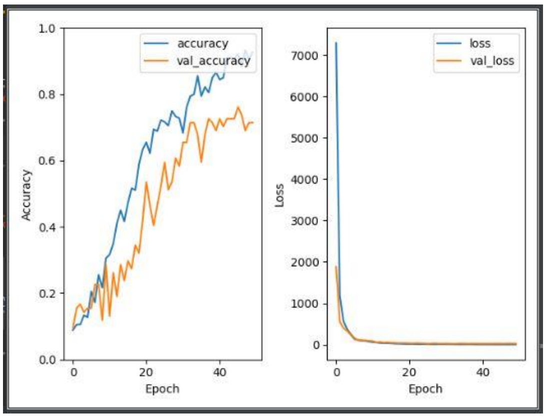


Fig. 3. CNN 5.2.1.c, softmax activation, dropout layer rate 0.2

For training the CNN models described at 5.2.1.c, we use 64 batches, learning rate of 1.00E-04. It can be noticed that while configuration [Fig. 1] yields extremely bad results (0.0561 accuracy), modifying the activation function of the output dense layer of the architecture from relu to sigmoid results into an accuracy of over 0.60. However, sigmoid should not be used for multiple label classification in which the probabilities should sum up to 1. Instead, as an alternative non-linear function, softmax is an option to try. Indeed, this improves the results dramatically [Fig. 2], but unfortunately we can see that overfitting occurs, with training accuracy of 100%, while validation accuracy stays much lower. After adding one dropout layer to the configuration, the results are satisfactory [Fig. 3], with a training accuracy of 95.73% and a validation accuracy of 71.69%.

6. Conclusion

From the experimented approaches, it can be inferred that the classification of hair curl type needs to use a CNN with a large number of layers of different types, because of the quite large number of classes and the similitude among them, the labels being situated on a scale rather than representing completely different characteristics altogether. For a small dataset, 12 classes and the similitude among them, the results obtained so far are satisfactory. It should be noted that results might be improved in the future by using a larger dataset. This would involve annotating a larger amount of data, but this process might now be automated by the use of the best discovered classifier yet, and later fine-tuning by hand to obtain an accurate ground truth.

Bibliography

1. TA Ileni, DL Borza, AS Darabant, “Fast In-the-Wild Hair Segmentation and Color Classification”, VISIGRAPP (4: VISAPP), 2019, pp. 59-66
2. S Sarraf, “Hair Color Classification in Face Recognition using Machine Learning Algorithms”, American Academic Scientific Research Journal for Engineering, Technology, and Sciences, 2016
3. T Kim, C Chung, S Park1, G Gu, K Nam, W Choe, J Lee, J Choo, “K-Hairstyle: A Large-scale Korean Hairstyle Dataset for Virtual Hair Editing and Hairstyle Classification”, IEEE International Conference on Image Processing (ICIP), 2021
4. K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016
5. Y. Choi, Y. Uh, J. Yoo, and J. Ha, StarGAN v2: diverse image synthesis for multiple domains, 2020
6. Zheng Ji, Bao-Liang Lu, and Xiao-Chen Lian, “Gender classification by information fusion of hair and face”, INTECH Open Access Publisher, 2009
7. M. Svanera, U. R. Muhammad, R. Leonardi, and S. Benini, “Figaro, hair detection and segmentation in the wild,” in Proc. of the IEEE International Conference on Image Processing (ICIP), 2016
8. Weidong Yin, Yanwei Fu, Yiqing Ma, Yu-Gang Jiang, Tao Xiang, and Xiangyang Xue, “Learning to generate and edit hairstyles”, Association for Computing Machinery, 2017
9. Linda Shapiro, The University of Washington, George Stockman, Department of Computer Science, Michigan State University, “Computer Vision”, 2000
10. Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang, “Deep learning face attributes in the wild,” Computing Research Repository (CoRR), 2014
11. Cloete E., Khumalo N.P., Ngoepe M.N., “The what, why and how of curly hair: A review”, Proc. Math. Phys. Eng. Sci., 2019
12. Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” in Proc. of the European Conference on Computer Vision (ECCV), 2018, pp. 801–818
13. A. I. Marinescu, A. S. Darabant and T. A. Ileni, "A Fast and Robust, Forehead-Augmented 3D Face Reconstruction from Multiple Images using Geometrical Methods," 2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 2020, pp. 1-6
14. A. Ion Marinescu, T. Alexandru Ileni, and A. Sergiu Darabant, “A versatile 3d face reconstruction from multiple images for face shape classification,” in 2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 2019, pp. 1-6
15. Yacoob, Yaser, and Larry Davis, "Detection, analysis and matching of hair,", Tenth IEEE International Conference on Computer Vision (ICCV), Volume 1, vol. 1, pp. 741-748. IEEE, 2005
16. Proenca, H. and Neves, J. C., 2017, “Soft biometrics: Globally coherent solutions for hair segmentation and style recognition based on hierarchical mrfs”, IEEE Transactions on Information Forensics and Security, 2017, pp. 1637–1645

17. Julian, P., Dehais, C., Lauze, F., Charvillat, V., Bartoli, A., and Choukroun, A., "Automatic hair detection in the wild" International Conference on Pattern Recognition (ICPR), 2010, pp. 4617– 4620

Related Datasets:

18. K-Hairstyle: A Large-scale Korean Hairstyle Dataset for Virtual Hair Editing and Hairstyle Classification https://paperswithcode.com/dataset/k-hairstyle
19. Figaro 1K http://projects.i-ctm.eu/it/progetto/figaro-1k
20. Hairstyle 30K https://yanweifu.github.io/papers/hairstyle_v_14_weidong.pdf