Fast In-the-Wild Hair Segmentation and Color Classification

Tudor Alexandru Ileni[1], Diana Laura Borza[2] and Adrian Sergiu Darabant[1]
[1]Department of Computer Science, Faculty of Mathematics and Computer Science, Babes-Bolyai University, Cluj-Napoca, Romania
[2]Computer Science Department, Technical University of Cluj-Napoca, Romania

Keywords: Hair Segmentation, Hair Color, Fully Convolutional Neural Network, Histograms, Neural Network.

Abstract: In this paper we address the problem of hair segmentation and hair color classification in facial images using a machine learning approach based on both convolutional neural networks and classical neural networks. Hair with its color shades, shape and length represents an important feature of the human face and is used in domains like biometrics, visagisme (the art of aesthetically matching fashion and medical accessories to the face region) , hair styling, fashion, etc. We propose a deep learning method for accurate and fast hair segmentation followed by a histogram feature based classification of the obtained hair region on five color classes. We developed a hair and face annotation tool to enrich the training data. The proposed solutions are trained on publicly available and own annotated databases. The proposed method attained a hair segmentation accuracy of 91.61% and a hair color classification accuracy of 89.6%.

1 INTRODUCTION

Face analysis has received great interest from the computer vision community due to its applications in various domains: behavioral psychology, human-computer interaction, biometrics etc.. However, most of the research conducted in this area focused mainly on internal face features (eyes, eyebrows, lips etc.), while the external features (hair, chin contour) were somewhat neglected.

The hair plays an important role in human face recognition: in (Sinha and Poggio, 2002) it was proved that internal face features are ignored in favor of external face one and the overall head structure. Other studies showed that the facial features are perceived holistically (Sinha et al., 2006) and that the hair line and color are an important recognition cues in cases when the shape features are distorted. In the field of soft biometrics, the hair style is one of the most effective biometric traits (Proença and Neves, 2017).

Nowadays e-commerce and digital interaction with the clients play an important role in the field of modern optometry. Several technologies around virtual/virtual-try on applications were developed, which allow customers to experiment frames and glasses from the comfort of their homes with a similar experience to that in an optical shop. These systems are based on 3D models of real glasses and frames. As the worldly offer of frames and glasses is very large, users (buyers) are seldom in the situation of being overwhelmed by the multitude of choices. One has to physically try thousands of frames to see which one fits him better medically and aesthetically. In this context, aesthetics has seen development of new approaches: visagisme; it is a new subject which allow humans to enhance their appearance by choosing the appropriate accessories that are in harmony with their face. It defines a complex set of rules taking in account facial features like: hair texture and color, face shape, skin tone and texture, location of lips, eyes and facial proportions, etc. Amongst these hair color is one of the decisive factors when choosing the eyeglasses as it covers a major part of the upper side of the head.

However, automatic hair analysis was not intensively studied. First of all, the hairstyle and its color can be easily changed, but in practice, most of the people keep the same hairstyle for a long period of time. Also, numerous hairstyles exist (symmetrical, asymmetrical, curly, bald etc.), so, unlike other face features, it is hard to establish the areas where hair is likely to be present. Also, the hair's color distribution is not uniform (different color for the roots and locks, highlights), making it more difficult to detect.

This paper proposes a hair segmentation and color recognition method targeted for visagisme applicati-

Ileni, T., Borza, D. and Darabant, A.
Fast In-the-Wild Hair Segmentation and Color Classification.
DOI: 10.5220/0007250500590066
In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, pages 59-66
ISBN: 978-989-758-354-4

ons. The hair area is determined using state of the art fully convolutional neural networks (CNN); the detected "hair" pixels are used to construct a color histogram which is further analyzed by an artificial neural network (ANN) to decide on the hair tone.

The remainder of this paper is organized as follows: in Section 2 we review the recent advances in the field of automatic hair analysis. The proposed solution is detailed in Section 3, and the experimental results are reported in Section 4. This work is concluded in Section 5.

2 RELATED WORK

Automatic hair color analysis was pioneered by (Yacoob and Davis, 2005): the authors proposed a method for hair segmentation in frontal facial images. The hair area is established using facial proportions, color information and region growing. The work also defined several metrics to describe the hair's properties: length, dominant color, volume, symmetry etc. Paper (Rousset and Coulon, 2008) introduces a novel hair segmentation method by intersecting two image masks computed by frequency and color analysis, respectively. In (Julian et al., 2010) the hair region is segmented in two steps: first, a simple hair shape model is fitted to upper hair region using active shape models. Next, a pixel-wise segmentation is performed based on the appearance parameters (texture, color) learned from the first region. A hair segmentation method tuned for automatic caricature synthesis is described in (Shen et al., 2014). In an off-line training phase, the prior distribution hair's position and color likely-hood are estimated from a labeled data-set of images. Based on this information, the hair is localized through graph-cuts and k-means clustering.

Recently, more robust hair segmentation algorithms were proposed which perform well on images captured in unconstrained environments. In (Proença and Neves, 2017), a two-layer Markov Random Field architecture is proposed: one layer works at pixel level, while the second one operates at object level and guides the algorithm towards possible solutions. The method presented in (Muhammad et al., 2018) constructs a hair probability map from overlapping image patches using a Random Forest classifier and features extracted by a CNN. This rough segmentation is refined using local ternary patterns and support vector machines to perform hair classification at pixel level.

Other works also tackled the problem of hair color classification as a soft biometric trait. The works (Krupka et al., 2014), (Prinosil et al., 2015) propose a hair color analysis method from video sequences. The head area is estimated trough background subtraction and face detection and a face skin mask is computed using flood-fill. The hair area is simply determined as the difference between the head and the skin. The hair color is classified into five distinct tones: white/gray, black, brown, red and blond. In (Sarraf, 2016), the hair color is distinguished only between "black" and "non-black" tones. The values, mean and variance of each channel from the RGB and HSV representation of the images are combined into a feature vector and a machine learning classifier (kNN or SVM) is used to decide on the hair color.

3 PROPOSED SOLUTION

The problem of hair color classification involves two main steps: hair segmentation and color analysis. The segmentation module detects all the pixels from the input image which belong to the "hair" class; this module has a great impact on the color recognition module, as an incorrect segmentation influences the applicability of the color features.

A general outline of the proposed method is depicted in Fig. 1. First, the hair area is extracted using a CNN; as the hair has a uniform structure, an additional post-processing step is applied in order to fill in the (eventual) gaps from the hair pixels.

The hair color classification module analyzes the detected hair pixels to decide on the hair tone. The classification is performed using an artificial neural network which operates on normalized color histograms.

3.1 Hair Segmentation

Segmentation is the process of detecting and highlighting one or many objects of interest in an image. It also can be viewed as a classification problem, by assigning to each pixel of the image a label.

We used a variant of the U-Net fully convolutional network (U-Net FCN) (Ronneberger et al., 2015) to detect the hair pixels. The architecture comprises two symmetric parts.

The first part, the contraction path, iteratively down-samples the original image: at each step a 3×3 pooling operation is applied and the number of output channels is doubled. Different from the implementation in (Ronneberger et al., 2015), the classical convolutions are replaced by depthwise convolutions, in order to reduce the computational cost, but still to benefit from spatial and depthwise information. Such layers are created from a pipeline of operations. First,

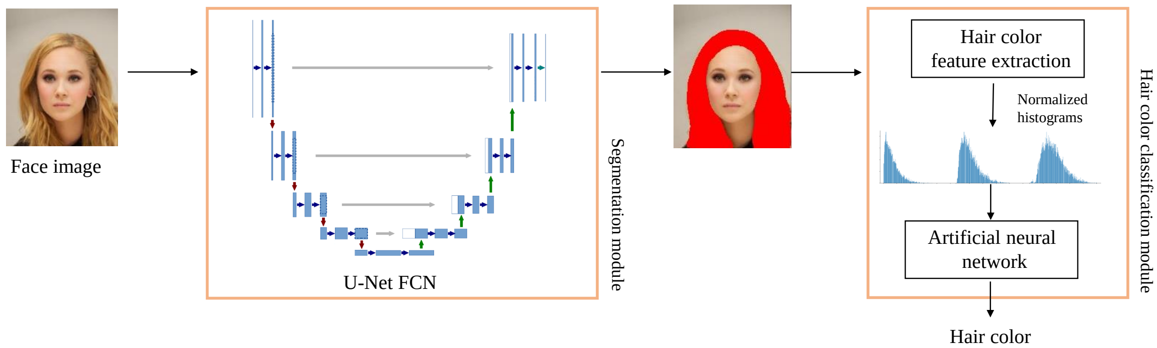


Figure 1: Outline of the proposed hair segmentation and color recognition module.

the convolution kernel is applied for each input channel, and then a pointwise convolution is performed on the resulting matrix.

The input layer has the shape of $224 \times 224 \times$ 3 corresponding to a 3 channel image. The first layer is a classical convolution that outputs 32 filters. In plus, it performs a batch normalization (Ioffe and Szegedy, 2015) and a ReLU activation.

The second part, the expansive path, is symmetrical to the contraction path, and it reverts the dowsampling operations, by using transpose (or fractionally strided convolutions) layers. At each step, a 2×2 deconvolution is applied to increase the feature map size and halve the number output channels.

In addition, as described in (Ronneberger et al., 2015), a cropped part of the corresponding contraction path layer is concatenated to each deconvolution layer. This way, the architecture benefits from a mixture of low and high level features, similar to skip layers, introduced in (Long et al., 2015). Each such concatenation is followed by a depthwise convolutional block, as previously detailed. Finally, a 2D convolution of size one and an upsampling layer are added. The upsampling layer is initialized by bilinear interpolation.

As the proposed method is intended for visagisme applications, where the user is cooperative and has a (near-)frontal position, the network operates on face images. The face is detected in the input image using an off-the-shelf face detector (King, 2009) and the face area is enlarged (both on width and height) with a factor of 1.5; this region of interest is cropped from the input image and used as input for the network.

3.1.1 Segmentation Post-processing

The human hair has a uniform pattern, therefore we apply an additional post-processing step on the hair segmentation mask in order to fill in the eventual gaps within the detected hair area. We used a simple algorithm based on the flood-fill operation. The method is described in the Algorithm 1.

Data: HM - binary hair segmentation mask
Result: RES - hair segmentation mask, gaps filled
Apply back border of size 5 to HM
Select pixel (s_x, s_y) outside the hair area
FLOOD_MASK $\leftarrow$ *floodFill*$(HM, (s_x, s_y))$
FLOOD_MASK $\leftarrow \neg$ FLOOD_MASK
RES $\leftarrow$ HM $\vee$ FLOOD_MASK

Algorithm 1: Hair area post-processing.

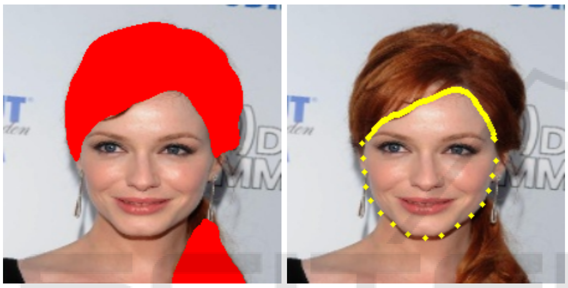
In the above algorithm, $\neg$ and $\vee$ are the *bitwise not* and *or* operators and the *floodFill*$(I, (x, y))$ function performs a flood fill operation on the input image I starting from the seed point (x, y) and returns a binary mask which highlights the pixels modified by this operation.

The algorithm works with a binary mask (0 - background pixel, 255 - hair pixel) and it first applies a border on this image in order to consider the cases where the hair area reaches the borders of the image. Next, a background pixel is selected and flood fill is applied starting from this pixel; as a result we obtain a mask that marks all the background pixels which are not inside the hair contour. This mask is inverted, now all the background pixels within the hair area become white. Finally, a bitwise or operation between the original hair mask and this inverted flood fill mask, is performed in order to fill in the gaps from the initial segmentation.

In the case of bald individuals detecting the hair color does not make sense. Therefore, we need a rule to decide whether the subject has hair before feeding the input image to the hair tone recognition module.

We propose a simple, yet efficient method for this task based on the proportion and position of the hair area relative to the face area.

First, an off-the-shelf facial landmark detector (King, 2009) is applied to find 68 facial landmarks on the face. Only the external face landmarks are used to compute the face area; it can be noticed that this algorithm only computes the face contour for the lower face part. In order to estimate the upper region of the face, we scan the upper region of the face, in polar coordinates, starting from the middle eyebrow point (e_x, e_y) with a radius R and with the angle $\theta \in 0, 180$. For each angle, we mark the first pixel on that radius labelled as "hair" as a face contour pixel; if no such pixel exists, we consider that the face boundary for the current angle θ is R pixels away from (e_x, e_y). We heuristically determined that $R = 0.7 \cdot fw$, where fw is the width of the face, is sufficient for most human face shapes. An overview of this process and its result are depicted in Fig. 2. In the figure, the detected landmarks are represented in the lower part of the figure with yellow circles, while the estimated upper face contour is drawn with a yellow curve.



(a) Hair segmentation mask (b) Estimated face area

Figure 2: Face area estimation based on hair segmentation and facial landmarks.

To determine if the person pictured in the input image has hair, we compute the ratio between the hair area and the face area b_r. If the value of b_r is less than 0.15 it is possible that the subject is bald; this threshold value was determined through trial and error experiments. In order not to label persons with (very) short hair as bald, we also add an additional rule regarding the position of the detected hair: only if the detected hair area is split into multiple parts on the sides of the face. We made this assumption based on the fact that human hair loss (androgenic alopecia) follows a similar pattern: the hair starts to fall from above the temples and the calvaria of the scalp (skullcap) and it progressively extends to the side and rear of the head (Asgari and Sinclair, 2011).

3.2 Hair Color Classification

We propose a hair color taxonomy consistent with the natural hair colours (five classes): black, blond, brown, grey/white and red.

To recognize the hair color, only the pixels which were classified as belonging to the "hair" class are analyzed. We compute a normalized color histogram from all the "hair" pixels and we feed this feature vector to an artificial neural network with two hidden layers. The hidden layers contains 4096 neurons each. Of course, the first layer has the size of the feature vector, while the last layer has 5 neurons (the number of classes).

Multiple colorspaces were proposed to encode colors, but none of them can be considered as a "best" representation. In each colorspace the color information is encoded differently, such that colors are more intuitively distinguished or certain computations are more suitable. We tested our method by representing the input image in the most commonly used colorspaces RGB, HSV and Lab; all these experiments are detailed in Section 4.

4 EXPERIMENTAL RESULTS

4.1 Databases

Training data is a crucial aspect for (deep-)learning, as it determines what the classifier learns before being applied to unseen data.

For the segmentation part, we gathered images from two publicly available datasets: Labeled Faces in the Wild (LFW) (Huang et al., 2007) and CelebA (Liu et al., 2015). LFW dataset contains more than 13000 celebrity images captured in uncontrolled scenarios; the only restriction imposed on an image is that the face can be detected using the Viola-Jones face detector. We used an extension of this database, Part Labels Database (Kae et al., 2013), which contains the semantic labelling into Hair-Skin-Background of 2927 image from LFW.

CelebA is a multi face attribute dataset, containing more than 200k images with large pose variations. The database comprises more than 10000 identities and is annotated with 40 binary attributes, like: wearing Hat, Wavy hair, mustache, just to name a few. The annotations also contain information about the hair color: Black_hair, Blond_Hair, Brown_Hair and Gray_Hair. However, it does not include the red hair class and we noticed some inconsistencies in the ground truth annotations for these attributes. Therefore, this labelling cannot be used as it is for hair color classification.

We also developed an application which can be used to manually mark the skin and hair area from facial images; i.e. to create skin and hair mask. We used this application to manually mark the hair area for 2188 additional images from the CelebA dataset.

The total size of the dataset used to training the segmentation network is 5115 images; 20% of these images were used for validation. Data augmentation was also performed, because of the small size of the dataset. We introduced a random rotation of maximum angle 20, a shear deformation of 0.2 and a zoom range of 0.2.

For the hair color classification module we used images from the CelebA dataset; first, a raw classification was performed based on the binary attributes - Black_hair, Blond_Hair, Brown_Hair and Gray_Hair - provided by the dataset. Starting from this coarse classification, three human labellers classified each image into the following classes: black, blond, brown, grey and red. The annotations are made publicly available. We used more than 20000 images to train the network and 2000 images to evaluate its performance.

4.2 Hair Segmentation

To train the segmentation network we used a stochastic gradient descend optimizer with a learning rate equal to 0.0001. The training time took more than 14h running on two NVIDIA Tesla K40m 12GB GPUs. We run the training 250 epochs, but no major improvements was made after the step 200. This can be view in Fig. 3.

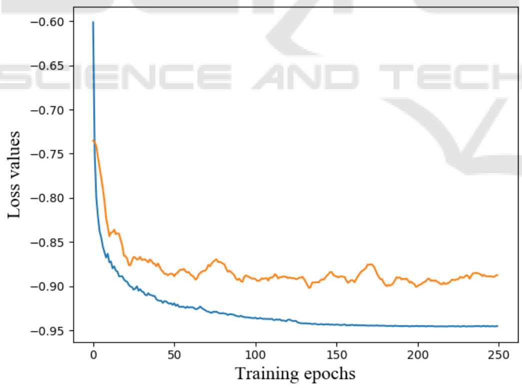


Figure 3: Training (blue curve) and validation loss (orange curve) for hair segmentation task, during 250 epochs.

We evaluated the hair segmentation model using the Intersection Over Union (IoU) metric. It is a scale invariant method, that calculates the similarity between two finite sets, by dividing the size of the intersection by the size of the union. More formally, the metric is defined as:

$$J(A,B) = |A \cap B| / (|A| + |B| - |A \cap B|). \tag{1}$$

We also report two variations of this metric: the Mean IoU and the frequency Weighted IoU. The first one is defined as:

$$\frac{1}{n_{cl}} \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ij} - n_{ii}} \tag{2}$$

and the latter as:

$$(\sum_k t_k)^{-1} \sum_i \frac{t_i n_{ii}}{(t_i + \sum_j n_{ji} - n_{ii})} \tag{3}$$

,

where n_{cl} is the number of segmentation classes, n_{ij} is the number of pixels of class i predicted to be in class j, and t_i the total number of pixels in ground truth segmentation of class i.

We also compute pixel accuracy (or precision)

$$pixelAcc = \sum_i n_{ii} / \sum_i t_i. \tag{4}$$

and mean pixel accuracy:

$$meanPixelAcc = \frac{1}{n_{cl}} \cdot \sum_i n_{ii} / \sum_i t_i. \tag{5}$$

When training the FCN, we tested two configurations for the hyper-parameters of the network:

- *c1* : momentum = 0.9 and batch size = 16 samples
- *c2* : momentum = 0.98 and batch size = 2 samples

In Table 1 we report the performance of the segmentation module, for both training configurations (*c1* and *c2*), on CelebA and Figaro1k databases.

Figaro1k contains 1050 unconstrained image labelled with seven different hair-styles: straight, wavy, curly, kinky, braids, dreadlocks, short. However, not all this images from this dataset can be used to test the proposed solution as the face is not always visible in these samples. The hair segmentation module works with facial images (the first step of the algorithm is to detect and crop the face area). Therefore, we first apply a face detector (King, 2009) and compute the hair segmentation mask only on the samples in which a face was detected. In total, we obtained 171 images from 1050. Although the Figaro1k dataset was envisioned for other purposes (hair texture and hairstyle classification), we tested our solution on this dataset in order to be able to compare with other works published in the literature.

On both datasets, better results are obtained using the second configuration of the network's hyper-parameters. As expected, the segmentation performance decreases on the Figaro1k database. The images contained in this dataset contain unusual hairstyles and shapes. Also, in some images the face is not fully contained and some parts of it are cropped. On average (on both datasets), the mean pixel accuracy for the hair segmentation module is 91.61%.

Some hair color segmentation results are depicted in Fig. 4.

Table 1: Hair segmentation performance on the CelebA and Figaro-1k databases.

	CelebA		**Figaro-1k**	
Metric	***c1***	***c2***	***c1***	***c2***
mean pixel accuracy	93.84%	94.76%	83.42%	88.46%
mean IoU	88.34%	90.35%	76.28%	81.96%
weighted freq. IoU	92.59%	93.89%	81.91%	86.01%
pixel accuracy	95.99%	96.76%	89.55%	92.13%

Figure 4: Some examples of hair segmentation results.

To evaluate the proposed algorithm for hair vs. no-hair detection (i.e. hair baldness), we selected 100 images (50 bald, 50 non-bald) from the CelebA dataset. The confusion matrix for this test scenario is illustrated in Table 2. The ground truth labels are represented on each row. The accuracy of the proposed algorithm is 91%.

Table 2: Confusion matrix for bald vs hair classification.

	Bald	**Hair**
Bald	44	6
Hair	3	47

4.3 Hair Color Classification

The hair color is classified based on the normalized color histogram of the hair pixels using a classical artificial neural network. To evaluate the classifier performance we selected 2000 (400 for each hair color class) images from the CelebA dataset. The ground truth was obtained by merging the classifications performed by three independent human labelers; in cases of disagreement we used simple voting to obtain the ground truth. We observed that the majority of confusions occurred between the red-brown and blond-brown classes.

In Table 3 we report the performance of the hair color classification module for different colorspaces and sizes for the feature vector. The test samples are balanced, i.e. we have 400 images belonging to each hair color class.

Table 3: Hair color classification performance for different colorspaces.

Colorspace	**Bin size**	**Accuracy**
RGB	1,1,1	0.878
HSV	1,1,1	0.881
LAB	1,1,1	0.883
RGB	8,8,8	0.881
HSV	8,8,8	0.889
LAB	8,8,8	**0.896**

The best results are obtained using the LAB colorspace with a bin size of 8; Table 4 shows the confusion matrix for this configuration. The rows contain the ground truth class.

Table 4: Confusion matrix for the LAB colorspace with bin size of 8.

	Black	**Blond**	**Brown**	**Grey**	**Red**
Black	398	0	1	0	1
Blond	0	398	2	0	0
Brown	0	0	397	0	3
Grey	1	3	2	394	0
Red	0	1	4	0	395

The execution time for the hair color classification method is on average $8 \cdot 10^{-4}$ seconds for a batch of 32 samples, run on the GPU device.

Blond Red Grey Black Brown

Brown as grey Brown as red Blond as grey

Figure 5: Correct and incorrect hair color classification.

Fig. 5 shows some correct and incorrect hair color classification results.

4.4 Comparison with State of the Art

As discussed in Section 2, several works addressed the problem of hair segmentation. However, there is no standardized benchmark for this task and some methods were only tested on internal, non-public datasets. The method (Muhammad et al., 2018) was evaluated on all 1050 images from Figaro1k dataset and the best configuration attained 91.5% segmentation accuracy. The algorithm uses features extracted by a CNN, local ternary patterns, super-pixels and a random forest classifier to segment the hair pixels.

The FCN for hair segmentation proposed in this paper obtained a pixel accuracy of 92.13% on the subset of Figaro 1k database that meets the requirements of our application; i.e. the face must be detected in the input image. Due to this fact (we couldn't use all the images from the dataset), a direct numerical comparison is not relevant. However, our average pixel accuracy on (on all the available test data) is 91.61%, so, we can conclude that our method is at least comparable with (Muhammad et al., 2018).

To the best of our knowledge, only two more papers addressed the problem of hair color classification: (Sarraf, 2016) and (Krupka et al., 2014). In (Sarraf, 2016) the hair tone is distinguished in only two classes: black and non-black, so a direct comparison with this work is not possible. The authors report an accuracy score of 97% in the best case and 55% in the worst scenario. However, we can extrapolate the results from Table 4 and compute the accuracy score for the black vs. non-black hair scenario: 99.85% (it should be noted that the classes are unbalanced in this scenario: 400 black and 1600 non-black).

The work (Sarraf, 2016) uses the same hair color taxonomy as the one presented in this paper. Their accuracy is 88.66% (value computed from the confusion matrix). However, the test data from (Sarraf, 2016) is not balanced: the red hair class is represented by only 3 samples, while the black hair class contains 30 samples.

Our method attains a hair color classification accuracy of 89.6%, so it can be concluded that the proposed classification module achieves better results than the other works presented in the literature.

5 CONCLUSION

This paper presented an automatic skin tone analysis system targeted for (on-line) eyeglasses virtual try on applications. Using a simple consumer camera and a virtual reality application, the user can perceive his/her appearance with different type of eyeglasses. Our method intervenes in the virtual eyeglasses display strategy: as the available dataset of 3D glasses is large, the assets should be displayed to the user such that the most suitable glasses for his/her appearance show up first. For this purpose a new field of study, visagisme was developed to help users choose the appropriate accessories based on their physical appearances. Hair color is one of the most important visagisme attribute in the choice of eyeglasses. Our method analyses an input image and outputs the hair color of the user: black, blond, brown, grey or red.

The proposed method involves two main steps: segmentation and color analysis. First, the hair area is determined using a state of the art fully CNN; additional morphological operators are applied to the hair mask in order to fill in the eventual gaps in the hair area. The hair pixels are further analysed by a classical artificial neural network in order to determine the hair color. To train and test the proposed algorithm, we annotated more than 4000 images from an existing database with the hair color.

The experiments we performed and the reported results (a hair segmentation accuracy of 91.61% and a hair color classification accuracy of 89.6% demonstrate the effectiveness of the proposed solution.

As a future work, we plan to add more classes to the hair tone taxonomy in order to be able to also recognize un-natural, dyed hair colors: blue, violet, green or hair with highligths.

REFERENCES

Asgari, A. and Sinclair, R. (2011). Male pattern androgenetic alopecia.

Huang, G. B., Ramesh, M., Berg, T., and Learned-Miller, E. (2007). Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst.

Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.

Julian, P., Dehais, C., Lauze, F., Charvillat, V., Bartoli, A., and Choukroun, A. (2010). Automatic hair detection in the wild. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 4617–4620. IEEE.

Kae, A., Sohn, K., Lee, H., and Learned-Miller, E. (2013). Augmenting CRFs with Boltzmann machine shape priors for image labeling. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2019–2026.

King, D. E. (2009). Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10(Jul):1755–1758.

Krupka, A., Prinosil, J., Riha, K., Minar, J., and Dutta, M. (2014). Hair segmentation for color estimation in surveillance systems. In *Proc. 6th Int. Conf. Adv. Multimedia*, pages 102–107.

Liu, Z., Luo, P., Wang, X., and Tang, X. (2015). Deep learning face attributes in the wild. In *Proceedings of International Conference on Computer Vision (ICCV)*.

Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440.

Muhammad, U. R., Svanera, M., Leonardi, R., and Benini, S. (2018). Hair detection, segmentation, and hairstyle classification in the wild. *Image and Vision Computing*, 71:25–37.

Prinosil, J., Krupka, A., Riha, K., Dutta, M. K., and Singh, A. (2015). Automatic hair color de-identification. In *Green Computing and Internet of Things (ICGCIoT), 2015 International Conference on*, pages 732–736. IEEE.

Proença, 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*, 12(7):1637–1645.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.

Rousset, C. and Coulon, P.-Y. (2008). Frequential and color analysis for hair mask segmentation. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pages 2276–2279. IEEE.

Sarraf, S. (2016). Hair color classification in face recognition using machine learning algorithms. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, 26(3):317–334.

Shen, Y., Peng, Z., and Zhang, Y. (2014). Image based hair segmentation algorithm for the application of automatic facial caricature synthesis. *The Scientific World Journal*, 2014.

Sinha, P., Balas, B., Ostrovsky, Y., and Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, 94(11):1948–1962.

Sinha, P. and Poggio, T. (2002). 'united'we stand. *Perception*, 31(1):133.

Yacoob, Y. and Davis, L. (2005). Detection, analysis and matching of hair. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 741–748. IEEE.