https://doi.org/10.2352/ISSN.2470-1173.2020.5.MAAP-082

Beyond Color Correction : Skin Color Estimation In The Wild Through Deep Learning

Robin KIPS, Loïc TRAN, Emmanuel MALHERBE, Matthieu PERROT; L'Oréal Research & Innovation; Clichy, 92110, France

Abstract

Estimating skin color from an uncontrolled facial image is a challenging task. Many factors such as illumination, camera and shading variations directly affect the appearance of skin color in the image. Furthermore, using a color calibration target in order to correct the image pixels leads to a complex user experience. We propose a skin color estimation method from images in the wild, taken with unknown camera, under an unknown lighting, and without a calibration target. While prior methods relied on explicit intermediate steps of color correction of image pixels and skin region segmentation, we propose an end-to-end color regression model named LabNet, in which color correction and skin region segmentation are implicitly learnt by the model. Our method is based on a convolutional neural network trained on a dataset of smartphone images, labeled with $L^*a^*b^*$ measures of skin colors. We compare our method with standard skin color estimation approaches and found that our method over-perform these models while removing the need of color calibration target.

Introduction

Skin color appearance has been studied with attention in several research areas [1], as it is in the center of many applications. In the domain of computer graphics for instance, skin color must be preserved for rendering realistic avatars or AR effects [2]. In medicine, skin diseases diagnostic is largely based on skin color variations [3]. In the cosmetics domain, skin color measurement would for instance lead to accurate foundation shades recommendation or personalized shades creation. From this perspective, L'Oréal is putting research efforts into developing an application for skin color estimation and foundation recommendation based on smartphone pictures.

Most of these topics have been studied in moderately controlled environments. However, bringing such application to consumer hands and smartphones raises new challenging problems. Indeed, skin appearance analysis is a complex real world problem affected by many external factors. For instance, illumination color largely affects the skin color in the image [4], while illuminant orientation is at the sources of shading and specular reflectance effects that locally impact skin color appearance [5]. In addition, the intrinsic properties of the used camera will also affect the colors appearance in the final image [6]. Figure 1 shows example of skin appearance variations for the same subjects under several illumination conditions.

Popular solutions for estimating skin color rely on the use of a color calibration target (figure 3) that allows for color correcting the image pixels [4], [7]. However, such a solution is difficult to implement in practice. Distributing calibration targets to end consumers is costly, and unpleasant in terms of user experience. In addition employing such reference charts in a robust way is complex. Users must avoid occlusions of the calibration target, and ensure that the illumination on the target is consistent with the illumination on the face.

In this paper, we propose to address the problem of skin color estimation in the wild. The objective is to accurately estimate the skin color measurement of a subject given a smartphone self portrait image in an uncontrolled environment with unknown lighting and camera, without the need of a color calibration target.

Our main contributions can be summarized as follow. First, we propose LabNet, a novel end-to-end color regression method based on a convolutional neural network architecture, with measured colors as labels. By using a deep learning model, we intend to learn a specialized color estimation that accounts for complex effects specific to human skin and face geometry such as shading and specularities. Second, we propose to implicitly learn color correction within the skin color estimation model instead of addressing this task as an explicit preprocessing step. We empirically show the ability of our LabNet model to estimate skin color independently of illuminant color without any explicit color correction step.

Related work

In the following section we review related work on skin color. We first detail the different skin color representation choices in the literature, then we describe color correction methods commonly used for skin color estimation. Finally we present traditional skin color estimation models used after color correction.

Skin color representation

Among the research works that focus on estimating the skin color from an image, there is no consensus on how to represent the skin color of a subject. Several research work are based on expert defined, subjective categories of skin color. The field of dermatology has established a seminal scale for skin type classification, the *Fitzpatrick scale* [8], based on visual evaluation of skin reaction to sun exposure. Even if such a scale can be used to differentiate between subjects skin colors, it relies on experts subjectivity. Moreover, other works are based on skin color categories assigned to images by non expert users. For instance, [9] proposes a representation of skin color in two classes, while [10], [11] and [12] are based on a three classes representation. However, these authors stress the lack of objectivity of this approach, as well as the major ambiguities that rise for subjects at the frontier of each class.

Instead of representing skin color in distinct classes, [13] and [7] proposed skin color regression tasks based on a continuous representation of skin color in the $L^*a^*b^*$ color space [14]. Thus, they propose to estimate the $L^*a^*b^*$ triplet of the skin color mea-

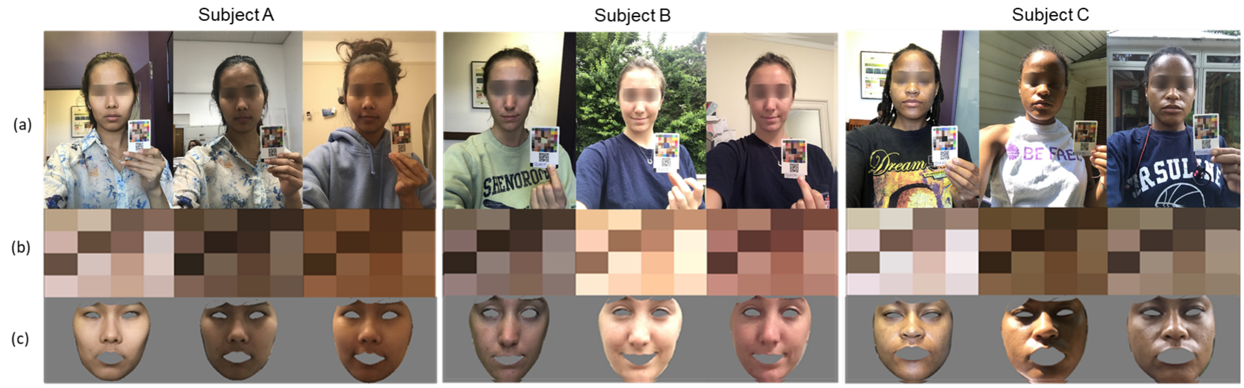


Figure 1: *Example of local and global skin appearance variation between images due to illumination. For each original picture (a), the line below (b) shows the colors extracted from the skin color patches of the color calibration target in the images. The third line (c) shows a segmentation of image skin pixels with a neutral background to emphasize the variation in skin color appeareance bewteen images*

surement of a subject from a camera image.

In order to create an objective typology, [15] proposed skin color classification based on the *Individual Typology Angle (ITA)*. The $ITA \in \mathbf{R}$ measure can be computed as follow, L^* and b^* corresponding to the measurement of skin in the CIE $L^*a^*b^*$ space:

$$ITA = Arctan(\frac{L^* - 50}{b^*}) * \frac{180}{\pi}$$

This representation is useful for establishing a one dimensional scale of skin colors, as well as an objective classification of skin tones. However, it raises problems regarding ambiguity at the border of each defined class, and does not take into account the a^* dimension of the $L^*a^*b^*$ measurement.

Color correction

The main challenge in color estimation problem is the variation in the appearance due to illumination color, as illustrated in figure 1. For this reason, existing research work focus on color correcting the image pixels prior to color estimation. The use of color calibration target for the skin color estimation task was studied by [4]. They propose to use the reference color chart to estimate an affine transformation between pixel values and measured colors. They showed that such a color correction step enables to reduce the impact of both illumination color and camera variations. In addition, [7] showed that this approach can be specialized to skin tones by using skin colored patches in the reference chart, which led to increase performance in skin color estimation.

In order to remove the need of a color chart, other works have attempted to perform color correction using other features infered from the scene, such as illumination color. Indeed, providing the illuminant color, the von Kries model [16] can be used to balance the image colors. For instance, [13] proposed to use the eye sclera and pupil as reference targets to estimate illuminant and derive skin color. This is based on the assumption that both pupil and sclera have a fixed color through all subjects. Meanwhile, in the domain of color constancy other works have proven that deep neural networks can be used to successfully learn illuminant color from a scene in an holistic manner. The color constancy problem consists in estimating the color of the white patch pixels of a color calibration target from an uncontrolled image where the calibration target was hidden. Thus, [17] showed that convolutional neural networks overperform previous statistical methods in this task. Later, [18] confirmed and improved performance using a model with attention mechanisms.

Skin color estimation

While color correction of the pixels is the main focus of existing works, estimating the skin color of the subject from the corrected pixels is not a trivial task. Indeed, the facial skin appearance is heterogeneous with local skin color variations due to specularities and more global effects such as shading, due to light orientation and face geometry.

Previous works generally use a skin region segmentation step in order to select the pixels to consider for deriving the subject skin color. [4] uses a face detection algorithm to target the region of interest, and excludes pixels with luminance outside of a specific bound. The objective is to remove pixels saturated by specular reflectance. [7] proposes to filter pixel according to a fixed skin region boundary in the *HSV* color space. The skin color is then estimated using the mean color of the selected pixels. To the best of our knowledge, no neural networks models have been applied to skin color estimation tasks in the litterature.

Problem

In this paper, we propose to address the problem of skin color estimation *in the wild*. The skin color is represented by the measurement of a spectrophotometer in the CIE $L^*a^*b^*$ space. We choose to consider as input images *in the wild* that are as close as possible to real life consumer smartphone images. By *in the wild* we mean that the image is taken with an unknown camera, under an unknown lighting, and without a color calibration target. However, in this study, calibration targets are used in the images for method comparison purposes. For methods without calibration targets, they are masked in the image in order to avoid bias. Compared to previous work, such images rise new problems, as skin appearance will be largely affected by illuminant color changes, intensity changes, and orientation changes, as seen in figure 1. In practice, a method accurately estimating skin color from smartphone images makes possible to develop applications without the need of distributing calibration targets.

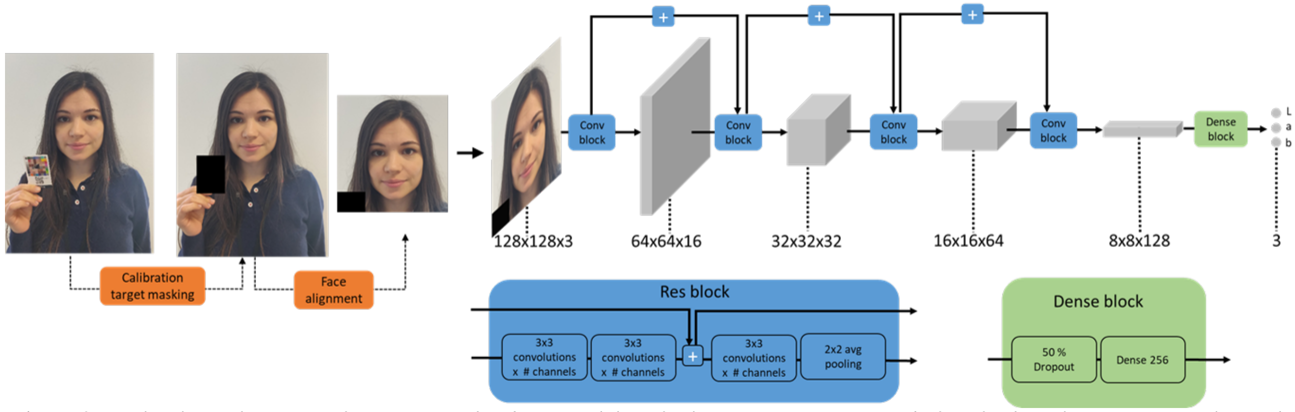


Figure 2: *Right, the architecture of our proposed LabNet model. Left, the preprocessing steps before feeding the images into the model.*

Method

In this section, we describe in details LabNet, the color regression model that we propose for skin color estimation from images *in the wild*. In order to avoid strong assumptions on the color of objects in the image such as in [13], we aim to statistically learn color correction from scene in an holistic manner. In addition, the segmentation of the skin color region in the image is a complex problem that requires a model with a large capacity of representation. Thus, instead of addressing first the surrogate problems of color correction and skin region segmentation as in previous literature, we propose to train an end-to-end color regression model. In such a model, color correction and skin region segmentation are no longer explicitly formulated, but only implicitly learnt within the model. By doing so, we intend to learn a specialized color estimation that accounts for complex effects specific to human skin such as reflection, shading and specularities.

We define c_i the skin color of the subject i as follow :

$$c_i = (L_i^*, a_i^*, b_i^*) \; ; \; L_i^* \in [0;100], \; a_i^* \text{ and } b_i^* \in [-128;+127]$$

Our objective is to learning a function $\hat{f}$ parametrized by Θ that computes $\hat{c}_i$ the estimate of c_i from a single subject image I_i :

$$\hat{f}_\Theta(I_i) = \hat{c}_i = (\hat{L}_i^*, \hat{a}_i^*, \hat{b}_i^*)$$

In order to learn this function we propose to train a color regression neural network. We call *LabNet* the architecture of our model, which is described in figure 2. As a preprocessing we use a face alignement step. For each image, facial keypoints are computed using an ensemble of regression trees [19], and their coordinates are used to crop an image with aligned eyes. We define the crop scale such that the distance between eyes represents 60 percent of the final cropped image in order to preserve background and lighting information. This preprocessing eases optimization as the model does not need to learn tasks such as localizing the face in the image. The neural network architecture in itself is inspired by the *ResNet* architecture [20], with skip connections used within each block in order to facilitate optimizaton during the learning procedure. The output of the layer is formed by a dense layer of size three, with ReLU activations. This layer directly outputs the estimated skin color $L^*a^*b^*$ of the subject in the image.

The model error is evaluated by computing the distance between ground truth color and estimated color using the color distance metric *CIE* ΔE^* *1976* [14], with the following formula :

$$\Delta E^*(\hat{c}_i, c_i) = \sqrt{(\hat{L}_i^* - L_i^*)^2 + (\hat{a}_i^* - a_i^*)^2 + (\hat{b}_i^* - b_i^*)^2}$$

This measure was build to match human perception of color distance. As we would like to conserve this property we choose not to rescale the $L^*a^*b^*$ values for training. We choose to use ΔE^* *1976* as the ΔE^* *1994* improvements where derived from other industries and are not relevant for human skin color. Furthermore, we need to use a differenciable color distance, which exclude the use of ΔE^* *2000*.

In order to learn the neural network parameters, we minimize the following loss :

$$\min_\Theta \sum_{i=1}^{n} \mathscr{L}(\hat{f}_\Theta(I_i), c_i) = \min_\Theta \sum_{i=1}^{n} \Delta E^*(\hat{f}_\Theta(I_i), c_i)^2$$

Figure 3: *Left, an Xrite CapsureMe used as a color calibration target in our study. Middle, a Xrite Capsure, the spectrophotometer used for measuring panelists skin color. Right, the position of the three measurements on the face*

Data

In order to train the described model, we recruited a set of panelists chosen to be representative of the various skin colors in the United States. For each subject, skin color was measured using an X-rite Capsure spectrophotometer [21] on three regions of the face, as show in figure 3. The obtained $L^*a^*b^*$ measures are averaged to construct the ground truth skin color for each subject. Measurements positions are designed in order to obtain the global face color and avoid local skin redness that might occur on the cheeks.

Model	Color correction using calibration target	Skin region segmentation	Color estimation	ΔE mean ± std
Naive baseline	No	keypoint segmentation	median pixel color	14.05 ± 7.25
Choi et al.	Yes	hsv values range	mean pixel color	9.45 ± 8.13
Choi et al. revisited	Yes	keypoint segmentation	mean pixel color	8.85 ± 6.57
LabNet color corrected	Yes	None	ResNet	3.91 ± 2.53
LabNet in the wild	No	None	ResNet	4.23 ± 2.72

Table 1: Comparisons of skin color estimation performance for different models

Subjects are asked to take images without any makeup, holding an X-rite CapsureMe color calibration target outside of the face region. In order to introduce camera variation in our database, subjects are asked to take pictures with their personal smartphones. Similarly, to ensure lighting variations, they are instructed to take images in various environments : tungsten light, fluorescent light, outdoor light, natural indoor light. In addition, the position of the color calibration target are manually annotated in each image, and pictures where the color checker patches are occluded are discarded. Using this procedure, we collected a dataset of 2795 images taken on 655 different panelists.

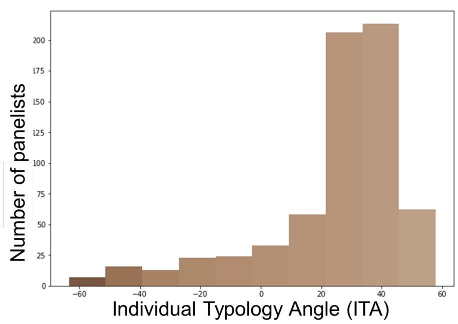


Figure 4: *Distribution of Individual Typology angle of subjects in ours dataset. Each bin of the histogram is colored according to the average observed skin color in the bin*

Experiments and Results

In this section, we detail the different models that we consider in our experiments. Each model is evaluated using a 5 folds cross validation, repeated 5 times with different splits. Since our database has an unbalanced skin tone distribution, as illustrated in figure 4, folds are stratified along the ITA deciles in order to limit bias in the models evaluation. In addition, we use a subject-grouped cross validation, in which each panelist pictures image can only belong to the same fold.

Naive baseline

In order to illustrate the complexity of the task, we evaluate the performance of a naive baseline. In this model, we do not use any color correction step, and directly segment the skin region in the corrected image. In order to accurately segment skin pixels, we use a keypoints based segmentation method. First, facial landmarks are localized using the method described in section 4, implemented in the dlib package [22]. A spline is then fitted between the chin and eyebrows landmarks to define the skin region. This method is based on the landmarks position only, and does not take into account skin occlusion, such as the presence of hair on the face. For this reason, we chose to exclude the forehead region. Lips and inner eyes are also segmented in order to mask these elements. The figure 1 and 6 show examples results of this skin region segmentation algorithm. Finally, we compute the mean RGB color of the segmented pixel, and convert this value to the $CIEL^*a^*b^*$ color space under the assumption of D65 as white illuminant.

Standard color correction

For comparison purposes, we evaluate the performance of standard color calibration based methods on our dataset. We thus consider the method from [7], based on the use of the skin colored patches of a calibration target. We compute the optimal color transform using least square estimation of a 3*4 affine transformation matrix, between skin colored patches pixels and corresponding measures. In the calibrated image, the skin region in the corrected image is defined using the common boundary for pixel values in the HSV CIE color space : $0 \leq H \leq 50$, $0.2 \leq S \leq 0.68$ and $0.35 \leq V \leq 1$. Finally, the estimated skin color is given by the mean over the pixels of the skin region. We will refer to this model as *Choi et al.*.

However, compared with existing studies our dataset covers a large variety of skin tones. Thus, the fixed color boundary used in this method might not be relevant in our case. For this reason, we propose to replace the original skin region segmentation of *Choi et al.* by the keypoints segmentation method we described earlier. Indeed, the keypoint localization model was trained on the iBUG 300-W face landmark dataset [23] with a large variety of skin tones, and is expected to be more robust to skin tone variation. We will refer to this model as *Choi et al. revisited*.

Furthermore, we chose not to consider color constancy based models in our experiments. Indeed, color constancy aims at estimating the white patch color of the calibration target from the image. By construction, these methods would be less accurate than directly using the ground truth calibration target values, as we do in proposed experiments.

LabNet

Our proposed LabNet model architecture is implemented using the Keras [24] API of the Tensorflow [25] library. To train our model, we use the AdaMax optimizer, with a learning rate of 5.10^{-4} and a $\beta 1$ of 0.9. The batch size is set to 16 and we use data augmentation by randomly applying horizontal flip to the training images.

In order to evaluate our model we perform two different experiments. First, images are color corrected using the calibration target prior to be sent to the neural network. We will refer to this experiment as *LabNet color corrected*. In the second experiment, we propose to evaluate the *LabNet in the wild* model, where images are directly fed to the neural network without any color cor-

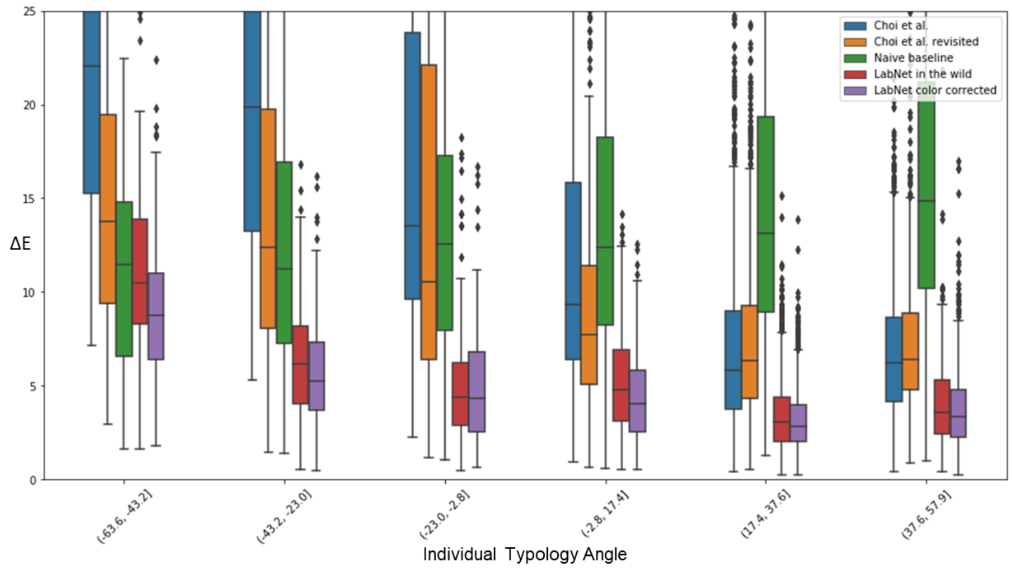


Figure 5: *Distribution of the error by Individual Typology Angle for each model.*

rection step.

Results

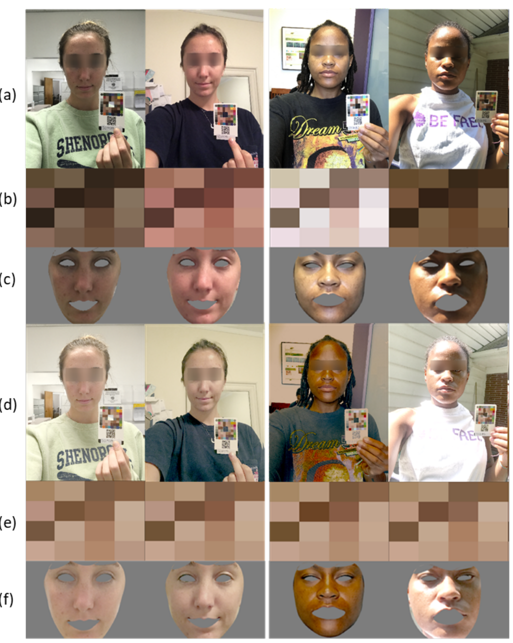
The results of our experiments are reported in table 1. As metric, we consider the average ΔE over all samples as well as the standard deviation of error across samples. The low accuracy ($\Delta E = 14.83$) of the naive baselines illustrates the complexity of the task in our dataset. As a comparison, we observed on average a difference of $\Delta E = 1.47$ between each subject color measure and the corresponding mean aggregated skin color. As expected, in the absence of color correction the naive model has lower accuracy on image under a non-white illuminant, as illustrated in figure 7.

Models based on conventional color calibration, such as *Choi et. al* and its revisited version show an improved performance but still perform poorly, with an accuracy of respectively $\Delta E = 9.45$ and $\Delta E = 8.85$. This could be due to the fact that in practice, due to color calibration target misusage, illuminant lighting the face and the color calibration target may be different, as illustrated in figure 6 example *(3)* and *(4)*, which leads to large errors in color correction.

In addition, on our dataset, we observe that the revisited *Choi et al.* model overperforms the original version of the model in both average and standard deviation of error. Looking at figure 5 confirms that as expected the revisited model based on keypoints segmentations performs better on low ITA, which corresponds to dark skin tones.

Training our *LabNet color corrected* leads to improved performance by a large margin, both in terms of average and standard deviation error, reaching $\Delta E = 3.91$. Such a result seems to show that our model successfully learnt to segment the skin region and model others local appearance variation on the skin such as shading and specularities. In addition, *LabNet in the wild* reach surprisingly good performances $\Delta E = 4.23$, with only a 8.2% increase in error while completely removing the need for a color calibration target. In figure 7, we study the performance of this model under varying illuminant color. The color temperature of the light is approximated using the illuminant color estimated from the white patch of the calibration target under the assumption of a black body radiation. Our model shows constant performance across different temperatures which tends to show that the color correction can be successfully learned from images, even without color calibration target.

(a)
(b)
(c)
(d)
(e)
(f)



(1) (2) (3) (4)

Figure 6: *Example of color correction using a calibration target. The original images (a) are shown with the extracted color of the skin colored patches (b) and the segmented skin pixels (c). After color correction (d) the target patches have the same color by construction (e), and the skin pixel show less variation in appereance. Example (3) and (4) shows example failures due to difference in illumination of the face and the color calibration target.*

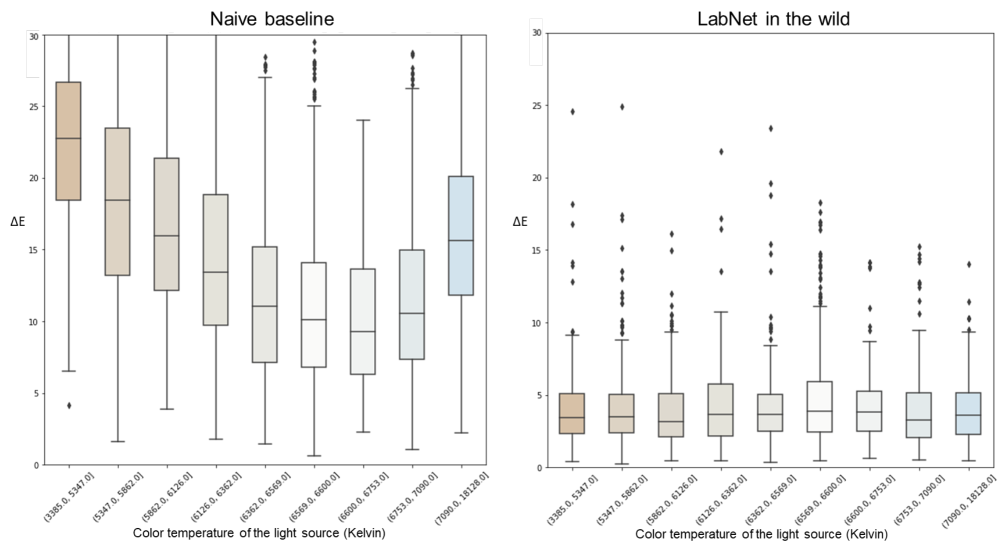


Figure 7: *Performance of the naive baseline (left) and LabNet in the wild (right), across light color temperature deciles. While the naive model has lower accuracy under non white illuminant, our model is able to learn color correction, and shows robust performance across light color temperatures. Boxplots are colored according to the average light color in each decile.*

Finally, we evaluate the performance of the LabNet model across the range of skin tones. Figure 5 indicates that for both *LabNet color corrected* and *in the wild*, errors of the model are higher for skin tones with low ITA. This is a limitation of our model : as it is learning based, its performance is closely related with the distribution of ITA in the training data. Looking at figure 4 confirms that ITA deciles with higher errors are ITA for which we have the fewer data. A potential solution could be to conduct new data studies targeting underrepresented skin tones. In spite of this limitation, our LabNet models still overperfoms other models for all skin tones.

However, one limitation of our approach is that such models are highly specialized compared to conventional color correction procedures. Indeed, color correction using a calibration target provide color information for the entire scene. Thus, while color correction can be use to address multiple color based tasks, such as analysing hair or garnments color, our skin color estimation model could not be leverage to solve another closely related problem.

Conclusion

Appearance of human skin in images in the wild is complex and affected by many sources of variations such as lighting color and direction. While prior methods were dividing this problem into several surrogate tasks such as color correction and skin region segmentation, we propose to learn a specialized skin color estimation within a neural network model in an end-to-end manner.

We show that our model over-performs conventional techniques for skin color estimation by a large margin. This suggests that our model is able to learn a representation that accounts for complex effects specific to human skin such as shading and specularities. Moreover, we show that our model can be trained on images without any color correction with a minor increase in error. Compare to previous methods this removes the need for a color calibration target. We empirically show that the accuracy of our model is independent of the light color temperature, which tends to demonstrate that our model is able to accurately apply color correction internally. In addition, we showed that our model is accurate over the range of skin tones, even though the model performance is closely related to the skin tone distribution in our database.

These results suggest that, using deep learning models, color estimation problems can be addressed in the absence of explicit color correction and calibration target.

Acknowledgments

We would like to thank all the contributors and collegues who participated to the development of the project in France and in the United States. Special thanks to Kelsey Norwood and Joseph Badami for organizing the challenging data collection that supported this research. Our gratitude goes to all our volunteer colleagues who tested our algorithm, as well as the team members who carrefully annotated the data.

References

[1] T. Igarashi, K. Nishino, S. K. Nayar *et al.*, "The appearance of human skin: A survey," *Foundations and Trends® in Computer Graphics and Vision*, vol. 3, no. 1, pp. 1–95, 2007.

[2] I.-S. Jang, J. W. Kim, J.-Y. You, and J. S. Kim, "Spectrum-based color reproduction algorithm for makeup simulation of 3d facial avatar," *ETRI Journal*, vol. 35, no. 6, pp. 969–979, 2013.

[3] T. P. Habif, *Clinical Dermatology E-Book: A Color Guide to Diagnosis and Therapy*. Elsevier Health Sciences, 2015.

[4] M. Harville, H. Baker, N. Bhatti, and S. Susstrunk, "Consistent image-based measurement and classification of skin color," in *IEEE International Conference on Image Processing 2005*, vol. 2. IEEE, 2005, pp. II–374.

[5] B. Khanal and D. Sidibé, "Efficient skin detection under severe illumination changes and shadows," in *International Conference on Intelligent Robotics and Applications*. Springer, 2011, pp. 609–618.

[6] J. Marguier, N. Bhatti, H. Baker, M. Harville, and S. Süsstrunk, "Assessing human skin color from uncalibrated images," *International Journal of Imaging Systems and Technology*, vol. 17, no. 3, pp. 143–151, 2007.

[7] H. Choi, K. Choi, and H.-J. Suk, "Performance of the 14 skin-colored patches in accurately estimating human skin color," *Electronic Imaging*, vol. 2017, no. 17, pp. 62–65, 2017.

[8] T. B. Fitzpatrick, "The validity and practicality of sun-reactive skin types i through vi," *Archives of dermatology*, vol. 124, no. 6, pp. 869–871, 1988.

[9] M. Jmal, W. S. Mseddi, R. Attia, and A. Youssef, "Classification of human skin color and its application to face recognition," in *The Sixth International Conferences on Advances in Multimedia MMEDIA 2014*. Citeseer, 2014.

[10] I. Boaventura, V. Volpe, I. da Silva, and A. Gonzaga, "Fuzzy classification of human skin color in color images," in *2006 IEEE International Conference on Systems, Man and Cybernetics*, vol. 6. IEEE, 2006, pp. 5071–5075.

[11] D. Borza, S. C. Nistor, and A. S. Darabant, "Towards automatic skin tone classification in facial images," in *International Conference on Image Analysis and Processing*. Springer, 2017, pp. 299–309.

[12] D. Borza, A. S. Darabant, and R. Danescu, "Automatic skin tone extraction for visagism applications." in *VISIGRAPP (4: VISAPP)*, 2018, pp. 466–473.

[13] H. Choi, K. Choi, and H.-J. Suk, "The human sclera and pupil as the calibration targets," *Electronic Imaging*, vol. 2017, no. 17, pp. 200–203, 2017.

[14] K. McLaren, "Xiii—the development of the cie 1976 (l* a* b*) uniform colour space and colour-difference formula," *Journal of the Society of Dyers and Colourists*, vol. 92, no. 9, pp. 338–341, 1976.

[15] A. Chardon, I. Cretois, and C. Hourseau, "Skin colour typology and suntanning pathways," *International journal of cosmetic science*, vol. 13, no. 4, pp. 191–208, 1991.

[16] J. von Kries, "Theoretische studien über die umstimmung des sehorgans," *Festschrift der Albrecht-Ludwigs-Universität*, pp. 145–158, 1902.

[17] S. Bianco, C. Cusano, and R. Schettini, "Color constancy using cnns," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 81–89.

[18] Y. Hu, B. Wang, and S. Lin, "Fc4: Fully convolutional color constancy with confidence-weighted pooling," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4085–4094.

[19] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1867–1874.

[20] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[21] X-Rite, Inc., "Capsure specifications," https://www.xrite.com/categories/portable-spectrophotometers/capsure-rm200, Last accessed on 2019-09-11.

[22] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, no. Jul, pp. 1755–1758, 2009.

[23] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: Database and results," *Image and vision computing*, vol. 47, pp. 3–18, 2016.

[24] F. Chollet *et al.*, "Keras," 2015.

[25] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.