How can we predict health conditions beyond skin color? "Revisiting skin-tone homogeneity"

Ryo Yokota [1],*, Yusuke Hara[1], Masako Katsuyama[1], Kumiko Kikuchi[1], Motoki Oguri[1], Maki Shirato[1], Mihoshi Yokoo [2], Shinya Iwanaga[2], Chieko Okamura[1]

[1]   MIRAI Technology Institute, Shiseido Co., Ltd.;
[2]   Brand Value R&D Institute, Shiseido Co., Ltd.

1. Introduction

Idiomatic expressions in various languages often associate skin colors and their spatial inhomogeneity with health and mental states. For instance, in English, the terms that represent colors such as "pink" and "pale" sometimes appear in phrases that express positive and negative health states[1].

From both physiology and psychology perspectives, previous studies have demonstrated that facial skin color and skin-tone homogeneity correlate with perceived health status[2-8]. Skin blood prefusion and oxygenation modulate complexion and enhance perceived health[3], and variations in carotenoid-dependent skin yellowness further improve perceived health and attractiveness[5]. Similarly, decreased skin-tone homogeneity is associated with lower perceived health and facial attractiveness, as well as higher perceived age[6].

However, although several studies have reported reproducible link between facial skin colors, skin-tone homogeneity, and perceived health, few have examined correlations with objective health indicators. To my knowledge, only the stage of the menstrual cycle has been reported to be clearly associated with skin colors or skin-tone homogeneity[7,8]. When extended to machine learning-based skin diagnosis, some previous studies have identified specific diseases such as jaundice and anemia from skin colors[9,10]. In contrast, Foo et al. (2017) reported that carotenoid supplementation alters skin color to enhance facial attractiveness and perceived health without affecting actual health conditions such as oxidative stress, innate immunity, or semen quality[5]. Cai et al. also reported no association between facial coloration and salivary immunoglobulin A, a key marker of mucosal immunity[11].

We hypothesized that one reason of this research gap is attributed to the fact that conventional statistical metrics for skin-tone homogeneity primarily focus on perceived age or facial attractiveness. This narrow focus results in a lack of necessity to quantify various spatial patterns for skin-tone homogeneity, potentially explaining why conventional metrics comprise with fewer than ten variables. These metrics are ill-suited to capture the complex associations between diversity of spatial skin color distribution and a wide range of health indicators.

To address the limitations of conventional metrics for skin-tone homogeneity, we applied 93 statistical features in the field of image texture analysis to quantify skin-tone homogeneity in

each color channel. These features can evaluate various characteristics in spatial distribution of pixel values, and widely used in diverse fields of image analysis, including medical images (e.g., CT, MRI)[12] and aerial images[13]. Our approach revealed specific spatial characteristics unique to certain health indicators that conventional metrics overlook, demonstrating the potential of skin-tone homogeneity as model variables for non-invasive diagnosis tool that predict internal health conditions from facial images.

2. Materials and Methods

A total of 150 healthy Japanese women were enrolled as subjects, with 25 individuals in each of the six age groups: 20s, 30s, 40s, 50s, 60s, and 70s. 402 health indicators of this study were derived from the clinical and instrumental evaluations, with a partial list provided in Table 1 (not all measured indicators are listed due to space constraints). Skin color of the cheek was measured using a contact-type spectrophotometer (CM-700d; Konica Minolta Sensing, Tokyo, Japan). Skin topography was evaluated from multiple perspectives using different measurement systems including proprietary measurement devices. In particular, wrinkle area ratio was measured by a three-dimensional analysis of winkles at the eye corner with replicas (Silflo, Flexico Developments Ltd., Potters Bar, UK). The all evaluations were conducted under the following conditions (temperature: 23°C, relative humidity: 45%) after a 15-minute acclimatization period. Prior to enrollment in this study, written informed consent was obtained from all subjects.

Table 1. Partial list of health indicators measured in this study.

Groups of indicators	Measurement principles	Supplementary information (device, manufacturing company, etc…)
Skin color	Spectrophotometer	CM-700d (Konica Minolta)
Skin topography	Wrinkle analysis, Skin microrelief, etc…	Replica: Silflo (Flexico), VOXELAN(Hamano engineering), etc…
Body composition	Bioelectrical impedance analysis	InBody 770 (InBody Japan Inc.)
Questionnaire	BDHQ, etc…	BDHQ (Brief-type self-administered Diet History Questionnaire): Assessment of Japanese diets, etc…
Viscoelasticity	Negative pressure suction displacement	Cutometer: CT580-2/MPA580D (Courage+Khazaka)
Blood test	Medical checkup items	biochemical factors in blood
Inflammation	Stratum corneum test	In-house developed inspection kit
Glycation	Skin autofluorescence	AGE-Reader mu (DiagnOptics)

The left-facing facial photographs of each subject were taken using a high-resolution digital single-lens reflex camera (EOS Kiss X4; Canon Inc., Tokyo, Japan). The taken RAW image files (cr2 files) were converted to sRGB image of lossless TIFF files using the Genuine software from the camera manufacturer (Digital Photo Professional, version 4.17.0, Canon Inc., Tokyo, Japan). Each photograph had 5184 x 3456 pixels (Width x Height), and the region of the left cheek was set to be the region of interest (ROI).

The ROI was extracted based on the facial landmarks obtained using the Mediapipe python package (version 0.10.14). The pixel values in sRGB coordinates were converted into CIELAB

coordinates using the rgb2lab function in the scikit-image python package (version 1.0.2)[14]. This conversion requires the assumption that a color temperature of a white light source should be controlled to approximately correspond to the CIE standard illuminant such as D65 and D50. However, it is difficult to control a light source in actual operating conditions such as commercial services. Thus, in this study, it was assumed that the light sources of the digital photographs taken in a typical indoor environment was D65 (2-degree observer).

As the image features to evaluate the characteristics of skin-tone homogeneity, 93 features in the field of image texture analysis were extracted from the ROI of each subject using the PyRadiomics python package (version 0.10.14)[12]. Those features are classified into the following 6 classes: first-order statistics, gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), neighboring gray tone difference matrix (NGTDM), gray level dependence matrix (GLDM). First order statistics comprise 19 features that quantify the distribution of pixel intensity within ROI in each color channel, regardless of the spatial relationship between pixels. GLCM has 24 features that evaluate the spatial relationship of pixels based on the empirical observation frequency of pairwise pixels with specific grayscale pixel intensities in a fixed orientation. The inverse difference moment (IDM), which has been utilized in the previous studies as a metric for evaluating skin-tone homogeneity, is one of the GLCM features[15]. GLRLM has 16 features that characterize the run length distribution of consecutive pixels with the same grayscale intensity in a fixed orientation. GLSZM has 16 features that quantify the distribution of the spatial size of consecutive pixels with the same grayscale intensity. NGTDM has 5 features that evaluate the difference between a grayscale intensity and average grayscale intensity of its neighboring pixels within specific distance. GLDM has 14 features that quantify grayscale intensity dependencies which are defined as the number of consecutive pixels within a specific intensity difference and within a specific distance. GLCM, GLRLM, GLSZM, NGTDM, and GLDM belong to image texture features that quantify the complex spatial patterns of pixel intensity from the perspective of inhomogeneity, coarseness, and correlation. The details of each feature should be referred to Ref.[12]. Since the 93 features were derived from the grayscale facial images in each color channel, a total of 279 features were calculated (i.e., 93 features multiplied by 3 color channels). Hereafter, in order to distinguish the existing concept of statistical metrics for skin-tone homogeneity from the concept of our study, the above 279 features are collectively referred to as "visual texture (VT) features".

VT features and the health indicators were standardized using the following formula:

$$mz_i = \frac{x_i - \tilde{x}}{MAD},$$

where x_i, $\tilde{x}$, and MAD are the observation value of the ith subject, the median of observation values and the median of absolute deviation values, i.e., $MAD = median(|x_i - \tilde{x}|)$, respectively.

To confirm how the 402 health indicators correlates with the chronological and the perceived ages of subjects, the correlation coefficients between them were calculated using spearman's correlation with false discovery rate (FDR) post-hoc correction (the statistical significance was set at p<0.05). Moreover, to capture the total tendency of VT features against the chronological

and the perceived age of subjects, the spearman's correlations between their ages and the first principal component of the reduced VT feature dimensions using principal component analysis (scikit-learn version 1.4.2[16]) were also calculated. The perceived age of each subject was assessed by having twenty-three beauty consultants evaluate the subject's facial images, focusing on skin quality, from both frontal and lateral perspectives. The average of these evaluations was then calculated to determine the perceive ages.

To comprehensively analyze the pairwise relationship between VT features and the 402 health indicators, partial correlation coefficients excluding the chronological age information of the subjects between the VT features and the health indicators were calculated using spearman's correlation with FDR post-hoc correction (the statistical significance was set at $p < 0.05$). All calculations for spearman's correlation were implemented using Pingouin python package (version 0.5.4)[17]. The significant correlations between VT features and health indicators were summarized using a Sankey plot. This plot visualizes the number of items with significant correlations by representing the width of the bands that connect the groups (nodes) to which each VT feature or health indicator belongs. In other words, a wider band indicates a stronger relationship between the groups of VT features and health indicators.

Measures in information theory were employed to quantify the amount of information contained in each variable such as VT features or health indicators, as well as mutual dependencies between different variables. Information entropy, mutual information (MI), and conditional mutual information (CMI) were calculated using the Diffeomorphic information neural estimator (DINE), which is a latent CMI estimation algorithm with continuous random variables[18]. CMI allows us to quantify the dependency between two variables while excluding the influence of specific other variables. Additionally, conditional independent test was performed to confirm the independence between two variables, taking into account the presence of specific other variables[18].

The regression models of the health indicators with the significantly correlated VT features as explanatory variables were trained using Pycaret python library (version 3.3.2)[19]. Considering the problem of multicollinearity, the selection of explanatory variables to identify their optimal subset was conducted based on the spearman's correlation matrix with all the explanatory variables. The best prediction model was determined based on the coefficients of determination (R2) with 5-fold cross-validation from 19 candidate models; the details of the candidate models can be found in the Ref. [19]. The Shapley additive explanation (SHAP) values, which were calculated using SHAP python library (version 0.44.1)[20], were used to evaluate the importance for each selected explanatory variable. The individual conditional expectation (ICE) plots, which can visualize the dependence of the response variable on the changes in specific explanatory variable of interest for the specific subject, were calculated to demonstrate the possibility for the personalized diagnostic services based on the predictions of each costumer's future condition. The partial dependence plots were also calculated to confirm the average relationship between the response variable and the specific explanatory variable of interest. The significance of the increase in the R2 between the different models with different numbers of explanatory variables was confirmed using partial F test based on the calculated R2. The

implementation of partial F test was conducted using the custom-written python codes with the Scipy library (version 1.11.4).

This study was conducted in accordance with the recommendations of the Declaration of Helsinki and was approved by the Ethics Committee of our own organization.

3. Results

Among the 402 health indicators, 204 and 200 indicators showed significant correlations with chronological ages (CA) and perceived ages (PA), respectively. In contrast, among 279 VT features, 252 and 258 VT features showed significant correlations with CA and PA, respectively. Interestingly, significant correlations with CA were observed not only for IDM, which has been widely used as a metric for skin-tone homogeneity, but also for many other VT features. Moreover, as shown in Fig.1, the first principal component (PC1) of VT features calculated by PCA had slightly bigger correlation coefficient of perceived ages (r=0.49; p<0.01) than that of chronological ages (r=0.46; p<0.01).
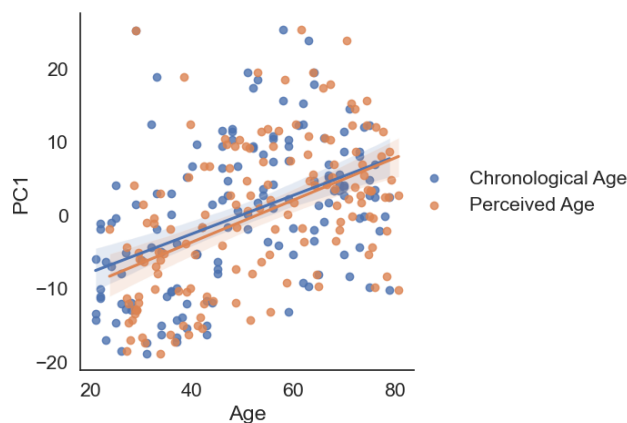

Figure 1. Scatter plot and linear regression of PC1 as a function of age:

Spearman's correlation coefficient with CA is r = 0.46 (p<0.001), and with PA is r =0.49 (p<0.001).

Since CA appeared to be a confounding factor between VT features and the observed health indicators, partial correlation coefficients that excluded the information of CA were utilized to comprehensively analyze the pairwise relationships between them. Among 112158 possible pairs between 279 VT features and 402 health indicators, 2606 pairs exhibited significant correlations. Specifically, 275 VT features and 83 health indicators were identified as having significant pairs. These health indicators included blood test indicators such as albumin and hematocrit, the metrics of skin topography related to pores, furrows, ridges and wrinkles, and specific questionnaire items such as intakes of egg, frequency of facial cleanser use, degree of cold hands and feet, and sunburn-induced changes. The Sankey plots in Fig.2 visualized the significant relationships between VT features and the health indicators. As shown in Fig.2(a), it may seem obvious that the skin colors measured by spectrophotometer were most strongly related to the VT features and were almost evenly dependent on those derived from the different color channels. However, the other groups of the health indicators unevenly or selectively depended on the VT features derived from the different color channels. For example, the groups of viscoelasticity, inflammation, and saccharification mostly depended on the a* color channel while the groups of skin topography and body composition mainly depends on the b* color channel. Figure 2(b) showed the significant relationships between the groups of VT features and those of health
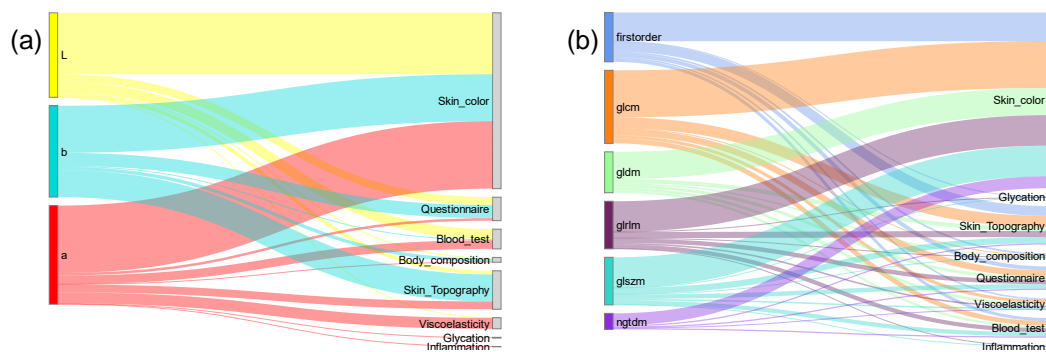
Figure 2. Sankey plots visualizing the number of significant correlations between VT features and health indicators: (a) Connections between the color channels used for VT features and the groups of health indicators; (b) Connections between the groups of VT features and the groups of health indicators.

indicators. Notably, no specific group of VT features was disconnected from any group of health indicators. Interestingly, the indicators from invasive blood test showed significant correlations not only with first-order statistics, which are closely related to color values due to their independence from the spatial distribution of pixel values, but also with other VT features. This suggests that the information regarding the spatial distribution of skin tone provides insight into internal body condition as well as skin colors. In fact, the partial correlation coefficient between serum albumin (Alb) levels and the informational measure correlation 1 (IMC1) derived from GLCM on the b* color channel, excluding the influence of CA, was significant as shown in Fig.3(a). Conversely, the b* skin-color levels did not show a significant partial correlation with Alb levels (Fig.3(b)), which was attributed to the significant correlation between the b* skin-color levels and CA (Fig.3(c)). The differences in the information regarding Alb levels between b* skin-color levels and IMC1 were also confirmed using measures in information theory (Table 2). The amount of information shared between IMC1 and Alb levels remained nearly constant, ranging from 0.062 to 0.058, even after subtracting the information held by CA. In contrast, the amount of information shared between b* skin-color levels and Alb levels drastically decreased from 0.041 to 0.008, when subtracting the information held by CA. The results of conditional independence test further indicated significant mutual dependence between IMC1 and Alb levels, even when CA was conditioned (p<0.01). Conversely, the relationship between b* skin-color levels and Alb levels was independent when CA was conditioned (p>0.1). These results suggested that IMC1 and the b* skin color provide different information about Alb levels, and that the Alb-level information derived from IMC1 may be less related to aging. From a cosmetic perspective, Alb-level information is also important. Figure 3(d) showed a significant partial correlation between the wrinkle area ratio measured using skin surface replicas and Alb levels excluding CA information (r=-0.18, p<0.05).

Table 2. Results of information entropies, MIs, and CMIs

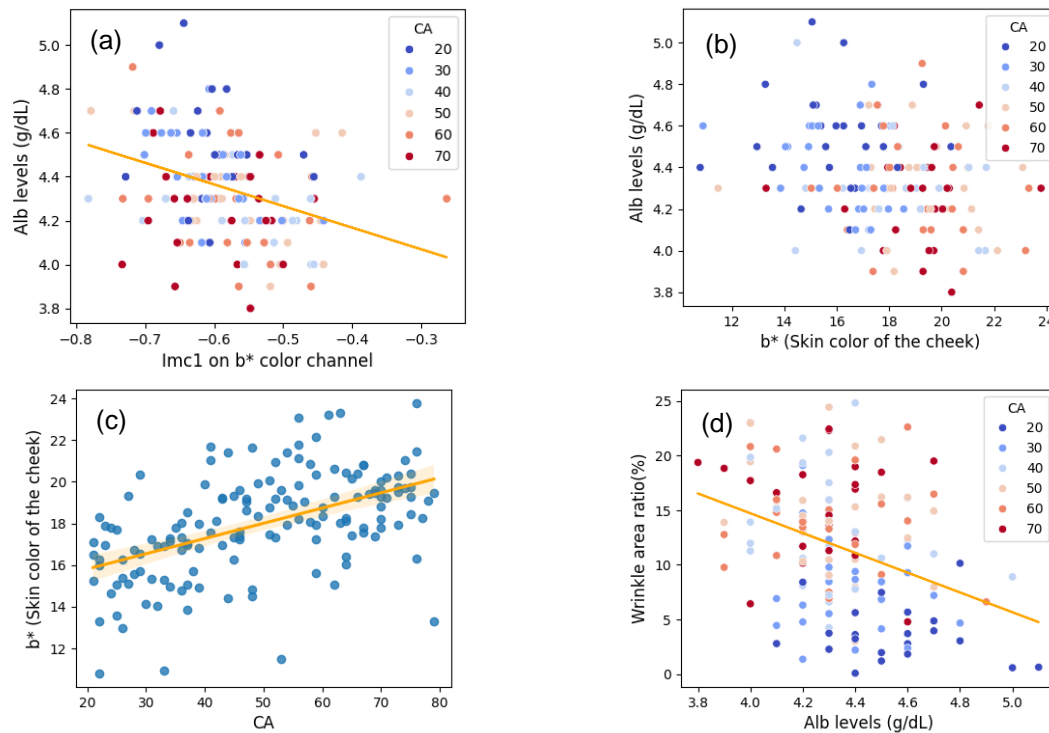| Entropy		Mutual Information		Conditional Mutual Information	
$H[X_{\mathrm{imc1}}]$	4.99	$I[X_{\mathrm{imc1}}; X_{\mathrm{Alb}}]$	0.062	$I[X_{\mathrm{imc1}}; X_{\mathrm{Alb}}|X_{\mathrm{CA}}]$	0.058
$H[X_{\mathrm{Alb}}]$	5.00	$I[X_{\mathrm{b*}}; X_{\mathrm{Alb}}]$	0.041	$I[X_{\mathrm{b*}}; X_{\mathrm{Alb}}|X_{\mathrm{CA}}]$	0.008
$H[X_{\mathrm{b*}}]$	4.99	$I[X_{\mathrm{CA}}; X_{\mathrm{Alb}}]$	0.056		
$H[X_{\mathrm{CA}}]$	4.94	$I[X_{\mathrm{imc1}}; X_{\mathrm{b*}}]$	0.094		

Figure 3. Scatter plots between the variables of interest (Colormap denotes CA):
(a) Relationship between Imc1 on the b* channel and Alb, r = -0.32 (p<0.01);
(b) Relationship between b* levels and Alb levels, r = -0.35 (p<0.01);
(c) Relationship between CA and b* levels, r = -0.35 (p<0.01);
(d) Relationship between Alb levels and wrinkle area ratio on the eye corner, r = -0.18
(p<0.05)

The relationship of VT features and each skin color with Alb levels was further examined from a different perspective, focusing on their contribution to prediction capability using the regression models of Alb levels with the significantly correlated VT features and each skin color as explanatory variables. From the 19 candidate models, extra-tree (ET) regressor was selected as the best prediction model using the function 'compare_models' in Pycaret library. As the explanatory variables, CA, two VT features (IMC1s on the b* and a* channels), and three skin colors (L*, b*, and a*) were selected while considering multicollinearity. In Fig.4, the contributions of each explanatory variable to the prediction performance of the trained ET regression model were quantified using SHAP values. Although CA was the best contributor, the two IMC1s were better than skin colors. Moreover, when compared to using only skin colors and CA as explanatory variables, the addition of two IMC1s as explanatory variables resulted in a significantly higher R2, increasing from 0.085 to 0.254 (p<0.05). These results suggested that some of VT features may be more effective than skin colors in predicting internal health conditions.

Figure 5 showed the PD plot and ICE plots of Alb levels as a function of CA, derived from the trained ET regression model. The bold line and each thin line indicated the PD plot and the ICE plot of each subject, respectively. The downward trends observed in the PD and ICE plots with respect to CA were consistent with the findings reported in the previous cohort studies[21,22]. However, the lower bound of ICE plots was higher than 4.0 g/dL, which exceeded the Alb levels actually observed in the subset of the subjects (Fig.3(d), Fig.5). Because the normal Alb level in healthy adults is $\geq$ 4.0 g/dL[21], part of ICE plots was unable to detect abnormalities in the Alb levels observed in this subset of the subjects.
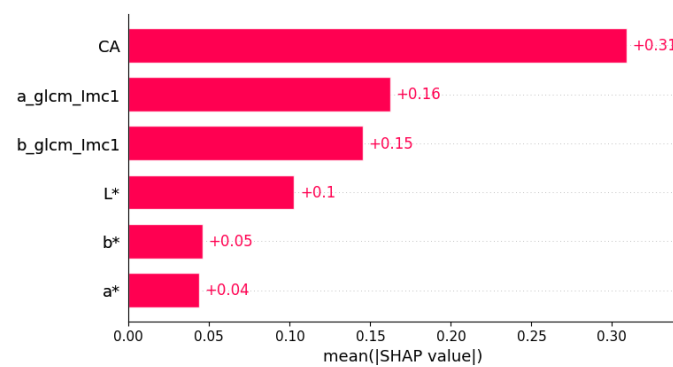
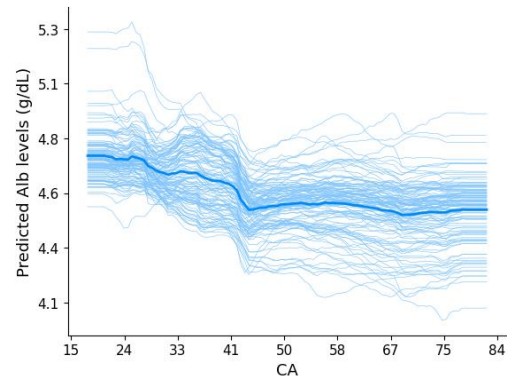**Figure 4.** Feature importance plot based on Shapley (SHAP) values

Figure 5. Partial dependence (PD) plot (bold line) and individual conditional expectation (ICE) plots of all subjects (thin lines).

4. Discussion

In the present study, we applied the image features utilized in the field of image texture analysis to evaluate the characteristics of skin-tone homogeneity, referring to these features as VT features. We demonstrated that VT features exhibited age-dependent properties (Fig.1) and color-channel-dependent relationships with various health indicators (Fig.2(a)), which will be inherited from the properties of skin colors. However, we also confirmed that some of the VT features like IMC1 had significant correlations with internal health indicators like Alb levels (Fig.3(a)), and which was distinct from the relationships between Alb levels and the factors of CA and skin colors (Fig.3(b,c), Table.2). In addition, using the trained ET regression model with CA, skin colors, and two IMC1s as explanatory variables, we demonstrated that IMC1s were more effective contributors to Alb level prediction than skin colors, as indicated by both of SHAP and R^2 values (Fig.4). Taken together of these results, we discovered a potential of VT features as good biomarkers to predict internal health conditions as well as skin colors.

We confirmed that more than 250 kinds of VT features including IDM significantly correlated with CA and PA, and that the PC1 of VT features using PCA more significantly correlated with PA than CA. This trend was consistent with the previous studies about skin-tone homogeneity[6], suggesting that most of the VT features may be able to capture age-dependent characteristics of skin-tone homogeneity as well as IDM. Because these VT features can capture the spatial distribution characteristics of pixel values which differ from those of IDM[12,13], they will enable us to evaluate specific patterns of skin-tone homogeneity that cannot be captured by conventional evaluation metrics. While the conventional metrics have primarily focused on appearance or attractiveness related to skin-tone homogeneity, our VT features may have the ability to address abnormalities in health indicators beyond these aspects (Fig.3).

Alb levels also had a significant association with the wrinkle area ratio on the eye corner (Fig.3(d)). Meanwhile, prediction of Alb levels is important not only for cosmetic purposes but also from a health maintenance perspective. Some previous cohort studies reported that a decrease in Alb levels, even within the normal range, increased the risks of frailty and diseases such as cardiovascular disease (CVD)[21,22]. In the Japanese cohort study, lower Alb levels had significant linear association with several internal health conditions such as kidney function (estimated glomerular filtration rate; eGFR), frailty measures (handgrip strength), hormones (insulin-like growth factor-1; IGF-1), and trace elements (calcium, magnesium, iron, and zinc)[21]. The longitudinal cohort study in Netherlands also reported that older adults

with a decline in Alb levels over a three-year period had a relatively high risk of CVD[22]. Since each ICE plot (thin line) predicted the personalized age-dependent changes in Alb levels, the trained model could serve as a non-invasive diagnostic tool to encourage improvements in dietary habits of each subject. The reason why the ICE plots overlooked the abnormal Alb values in some subjects is likely due to the small number of subjects with abnormal values in the training dataset (Fig.3(d), Fig.5). In situations where the data points are crucial but limited, it is essential to employ algorithms like SMOTE to resample significant data points or adjust the weighting of these points during training. If the aim of the study is set to detect abnormal values, it will be necessary to consider modifying the training algorithms.

Despite the advantages discussed above, VT features have certain limitations. Firstly, the values of VT features are significantly influenced by color management. In this study, the color temperature was not controlled to align with the CIE standard illuminant such as D65 and D50, because we anticipated challenges in managing light sources under actual operating conditions such as commercial services. However, for academic rigor, VT feature values should ideally be assessed under the correct CIELAB color space, with a strictly controlled color temperature. Secondly, because image resolution also impacts VT feature values, it is important to verify the robustness of significant relationships between VT features and health indicators, even when the resolution of subjects' facial images varies. Moreover, the rationale for using IDM as the optimal indicator of skin-tone homogeneity in various previous studies is grounded in Fink et al. (2001), which identified IDM as the feature most strongly correlated with facial attractiveness among seven features derived from GLCM[15]. However, it is important to note that the resolution of the digital images employed in their study was over seven times lower than that of the high-resolution digital images used in this study. Consequently, it remains uncertain whether IDM continues to serve as an optimal indicator of skin-tone homogeneity when applied to contemporary high-resolution digital images. Thirdly, to predict internal health conditions from VT features, polarized images that capture the subject's face with an optical deflection lens may be more effective. This is because specular reflection light primarily represents skin topography and can obscure the natural skin colors. Although these limitations require point-to-point verification, non-invasive diagnosis tool using VT features has distinct advantages to monitor internal health condition based on facial images.

5. Conclusion

In this study, we proposed new metrics related to skin-tone homogeneity, referred to as VT features. We demonstrated that the VT features had age-dependent properties, similar to conventional metrics for skin-tone homogeneity, while also revealing age-independent significant relationships with various health indicators, such as those obtained from invasive blood tests, including Alb levels. By incorporating two specific VT features, namely IMC1s on the a* and b* color channels, along with CA and skin colors into the regression model for predicting Alb levels, we significantly enhanced its prediction capability. Furthermore, we presented ICE plots to demonstrate the potential for the personalized diagnostic services based on predictions of each costumer's future condition. This represents a first step toward developing non-invasive diagnosis tools that visualize the connections between skin and internal health conditions.

References

[1] Kövecses, Z. Metaphor and Emotion: Language, Culture, and Body in Human Feeling. Cambridge University Press, (2003).

[2] Stephen, I. D., Law Smith, M. J., Stirrat, M. R., Facial Skin Coloration Affects Perceived Health of Human Faces. *Int. J. Primatol.* **30**, 845–857 (2009).

[3] Stephen, I. D., Coetzee, V., Law Smith, M., et al., Skin Blood Perfusion and Oxygenation Colour Affect Perceived Human Health. *PLoS ONE* **4**(4): e5083 (2009).

[4] Lu, Y., Yang, J., Xiao, K., et al., Skin Coloration is a Culturally Specific Cue for Attractiveness, Healthiness, and Youthfulness in Observers of Chinese and Western European Descent. *PLoS ONE* **16**(10): e0259276 (2021).

[5] Foo, Y. Z., Rhodes, G., Simmons, L. W., The Carotenoid Beta-Carotene Enhances Facial Color, Attractiveness and Perceived Health, but Not Actual Health, in Humans. *Behav. Ecol.* **28**(2), 570–578 (2017).

[6] Matts, P. J., Fink, B., Grammer, K., et al., Color Homogeneity and Visual Perception of Age, Health, and Attractiveness of Female Facial Skin. J. Am. Acad. Dermatol. 57, 977-984 (2007).

[7] Burriss, R. P., Troscianko, J., Lovell, P. G., et al., Changes in Women's Facial Skin Color over the Ovulatory Cycle are Not Detectable by the Human Visual System. *PLoS ONE* **10**(7): e0130093 (2015).

[8] Oberzaucher, E., Katina, S., Schmehl, S. F., et al., The Myth of Hidden Ovulation: Shape and Texture Changes in the Face During the Menstrual Cycle. *J. Evol. Psychol.* **10**(4), 163-175 (2012).

[9] Zhang, A., Lou, J., Pan, Z., et al., Prediction of Anemia Using Facial Images and Deep Learning Technology in the Emergency Department. *Front. Public Health* **10**, 964385 (2022).

[10] Khanam, F. T. Z., Al-Naji, A., Perera, A. G., et al., Non-invasive and non-contact automatic jaundice detection of infants based on random forest. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, **11**(6), 2516–2529 (2023).

[11] Cai, Z., Hahn, A. C., Zhang, W., et al., No evidence that facial attractiveness, femininity, averageness, or coloration are cues to susceptibility to infectious illnesses in a university sample of young adult women. *Evolution and Human Behavior*, **40**(2) (2019).

[12] van Griethuysen, J. J. M., Fedorov, A., Parmar, C., et al., Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, **77**(21), e104–e107 (2017).

[13] Erdem, F., & Bayrak, O. C., Evaluating the effects of texture features on Pinus sylvestris classification using high-resolution aerial imagery. *Ecological Informatics*, **78**, 102389 (2023).

[14] van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al., scikit-image: Image processing in Python. *PeerJ* **2**:e453 (2014).

[15] Fink, B., Grammer, K., Thornhill, R., Human (Homo sapiens), facial attractiveness in relation to skin texture and color. *J Comp Psychol.* **115**(1), 92-9 (2001).

[16] Pedregosa, F., Varoquaux, G., Gramfort, A., et al., Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825-2830 (2011).

[17] Vallat, R., Pingouin: statistics in Python. *Journal of Open Source Software*, **3**(31), 1026 (2018).

[18] Duong, B., & Nguyen, T., Diffeomorphic Information Neural Estimation. *Proceedings of the AAAI Conference on Artificial Intelligence*, **37**(6), 7468-7475 (2023).

[19] Ali, M., PyCaret: An open source, low-code machine learning library in Python. PyCaret version 1.0.0 (2020). Available at: *https://www.pycaret.org*

[20] Lundberg, S. M., & Lee, S.-I., A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems* **30**, 4768–4777 (2017).

[21] Yamamoto, M., Adachi, H., Enomoto, M., et al., Lower albumin levels are associated with frailty measures, trace elements, and an inflammation marker in a cross-sectional study in Tanushimaru. *Environ Health Prev Med* **26**, 25 (2021).

[22] Schalk, B. W. M., Visser, M., Bremmer, M. A., et al., Change of Serum Albumin and Risk of Cardiovascular Disease and All-Cause Mortality: Longitudinal Aging Study Amsterdam. *American Journal of Epidemiology*, **164**(10), 969–977 (2006).