Illuminating Hidden Patterns of Skin Aging through AI-Driven Insights

Yizhen Yan*, Qi Zhou, Xueni Lin, Peng Shu*
HuJia Technology Co., Ltd., Shenzhen, China

Abstract

Skin aging is a multifactorial process influenced by genetic and environmental factors, manifested by changes in elasticity, wrinkles, pigmentation, and radiance[1] . We analyzed 396 parameters from 260 Chinese female participants (aged 15–80) to develop predictive models and uncover aging patterns. After data quality control , we used Recursive Feature Elimination to select 20 high-impact features for Random Forest and deep learning models, achieving MSE ≈30.1 and Pearson's r >0.85. Clustering revealed six distinct skin-aging profiles with varied temporal trajectories of pigmentation, wrinkles, and elasticity. Environmental variables, notably UV exposure, modulated aging dynamics. Our findings demonstrate a multidimensional AI-driven approach for identifying personalized anti-aging strategies across climatic regions.

1. Introduction

Skin aging results from intrinsic factors—genetics, hormonal shifts, metabolic changes—and extrinsic stressors such as UV exposure, pollution, and lifestyle habits, which together drive wrinkles, pigmentation changes, and loss of elasticity[2-5]. Classifying skin phenotypes enables targeted analysis of these diverse aging pathways, while predictive "skin age" models harness machine learning to quantify biological versus chronological aging. Together, these approaches advance personalized anti-aging interventions and optimize product development.

Recent AI-driven efforts in skin age estimation leverage imaging data (e.g., high-resolution photographs, 3D scans) and biophysical measurements, employing algorithms such as random forests, convolutional neural networks, and ensemble models to achieve correlation coefficients above 0.8 with chronological age [6,7]. Clustering techniques further enable delineation of distinct aging phenotypes by grouping subjects based on features like elasticity, pigmentation, and wrinkle metrics[8-11]. However, most studies pool data from heterogeneous populations or multiple locations, limiting model generalizability. In this work, we focus on a well-controlled cohort of 260 individuals from a single urban center—each having resided locally for over five years—with broad, uniform age distribution, aiming to refine "skin age" prediction and uncover locally relevant aging subtypes through tailored feature selection and machine learning–based clustering.

2. Materials and Methods

2.1. Participants and Data Collection

A total of 260 healthy female volunteers (age 15–74 years) were recruited in Guangzhou, China. Participants were stratified into twelve 5-year age bins (15–19, 20–24, …, 70–74), with each bin balancing indoor- versus outdoor-based occupations at a 1:1 ratio. All subjects had resided in Guangzhou for at least five years, were free of skin diseases or recent dermatological treatments, and had not used systemic steroids or immunosuppressants within the past month. After obtaining informed consent, each volunteer completed a lifestyle questionnaire (diet, sleep, exercise, occupation) and demographic form; climate data for the testing period (temperature, humidity, pollution levels) were recorded. All measurements took place under controlled temperature and humidity conditions.

2.2. Instrumentation and Measured Parameters

Skin assessments were conducted on four anatomical sites (face including periorbital region, neck, lips, and forearm) using the following devices:

CK‐MPA 580

– Elasticity: R2, R5, R7; firmness: F4

– Barrier function: stratum corneum hydration, transepidermal water loss (TEWL), surface pH

– Sebum content; colorimetry: L*, a*, b*, ITA°; melanin and erythema indices; gloss

PRIMOS CR

– High‐resolution 3D optical imaging for wrinkle depth, density, and roughness (forehead, crow’s feet, cheeks, nasolabial folds, lip lines)

Visia CR 5.0

– 2D multispectral imaging for brown spots, pores, UV spots, porphyrins, red areas

DermaLab® Single (20 MHz ultrasound)

– Epidermal/dermal thickness; echo intensity reflecting collagen–elastin content

Antera 3D (neck only)

– 3D mapping of neck wrinkle volume and depth

VC 20 Plus (neck and forearm)

– Texture analysis parameters: SEr, SEsc, SEw, SEsm

All instruments were calibrated and operated according to standardized SOPs to ensure reproducibility across sites and time points

2.3. Data Quality Control

Data Quality Control involved standardizing and normalizing all features, then using hierarchical clustering heatmaps to identify outlier samples and instruments (e.g., SEr, R2, R5) . Scatter plots and violin plots assessed multi‐sampling consistency, flagging measures with excessive intra‐subject variance. Samples or features exceeding three standard deviations or with >20% missing data were removed. Remaining missing values were imputed using median substitution, yielding a clean, reliable dataset for subsequent analysis.

2.4. Feature Selection and Preprocessing

Features were normalized (z‐score) and scaled (min–max) before applying Recursive Feature Elimination with cross‐validated Random Forests. The RFE pipeline iteratively removed least important predictors, retaining the top 20 high‐impact features (e.g., R5, ITA°, sebum, wrinkle metrics) for modeling . Continuous variables were then re‐standardized to unit variance, ensuring uniform weighting across subsequent analyses.

2.5. Modeling Approaches

Random Forest and a fully connected neural network were trained to predict “skin age” from the 20 selected features. Both models underwent five‐fold cross‐validation with grid‐search hyperparameter tuning (trees, depth; layers, units, dropout) and early stopping for the neural network. Model performance was evaluated by mean squared error and Pearson’s correlation on held‐out folds .

3. **Results**

3.1. Data Quality and Feature Correlation

QC analysis showed that SEr, SEsm, SEw, R2, R5, R7, F4, LEB, Skin Thickness, and Intensity had high intra-subject variability, whereas hydration, TEWL, pH, sebum, gloss/Gloss DSC, colorimetric indices (L*, a*, b*, ITA°), melanin, erythema, and offset were highly consistent (PCC > 0.9) . Variability differed by site, with facial and neck measures more dispersed than forearm. After QC, all 260 samples were retained; age distribution across twelve 5-year bins remained uniform.

3.2. Feature Selection Outcomes

The final feature set was identified via exhaustive grid‐search over all candidate combinations (10–100 features) using RFE and Random Forest performance (neg MSE) as the criterion (Figure1). The optimal 29‐feature panel includes: Antera3D neck wrinkle length and max depth; PRIMOS counts and volumes for forehead, under-eye, eye-bag, nasolabial, and lip wrinkles; Visia percentiles, fractional areas, and intensities for brown spots, UV spots, red zones, visible spots, and wrinkles across left/right and full face; VC 20plus arm and neck SEsm and SEw metrics; MPA 580 sebum (forehead), color b* (left face) and ITA° (arm), elasticity R2/R7 and firmness F4 across sites; and arm ultrasound LEB. This tailored combination maximizes predictive accuracy while minimizing redundancy .

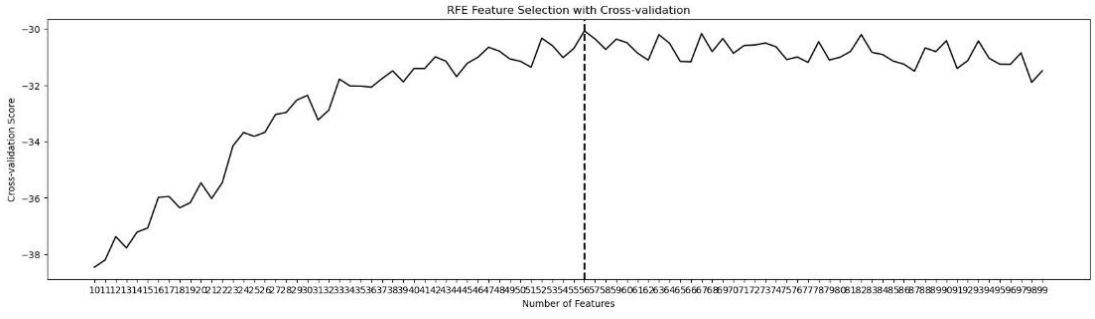


Figure 1. Recursive Feature Elimination (RFE) cross‑validation performance as a function of feature set size. Each point on the curve represents the mean cross‑validation score (negative mean squared error) obtained by training a Random Forest model on the top n features selected by RFE. The vertical dashed line indicates the optimal number of features (29), at which the highest validation score was achieved. Higher (less negative) values correspond to better model performance.

3.3. Predictive Model Performance

The optimized Random Forest model (29 features) yielded an average MSE of 35.17 and Pearson's r of 0.88 over fivefold cross‑validation, with $R^2 = 0.78$ on held‑out data. The neural network using the same feature set achieved MSE = 33.45 and r = 0.90, demonstrating consistent performance (≤4% fold‑to‑fold variance) and robust skin‑age estimation across the cohort (Figure 2) .

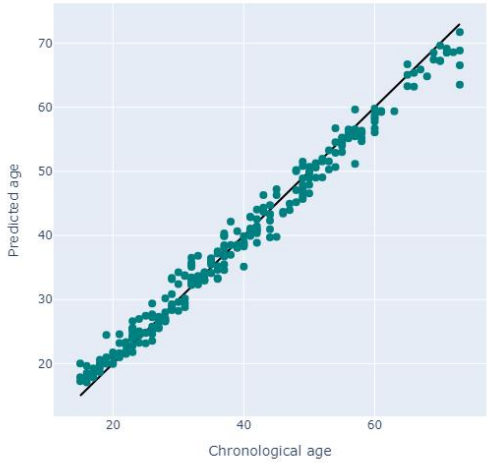


Figure 2. Predicted versus chronological age for the study cohort. Each green dot represents an individual sample's predicted "skin age" using the optimized Random Forest model, plotted against its true age; the black diagonal line denotes perfect prediction (y = x).

3.4. Clustering-Derived Aging Profiles

Six aging phenotypes emerged from K-means clustering (k=6), each characterized by distinct biomarker signatures (Figure 3) :

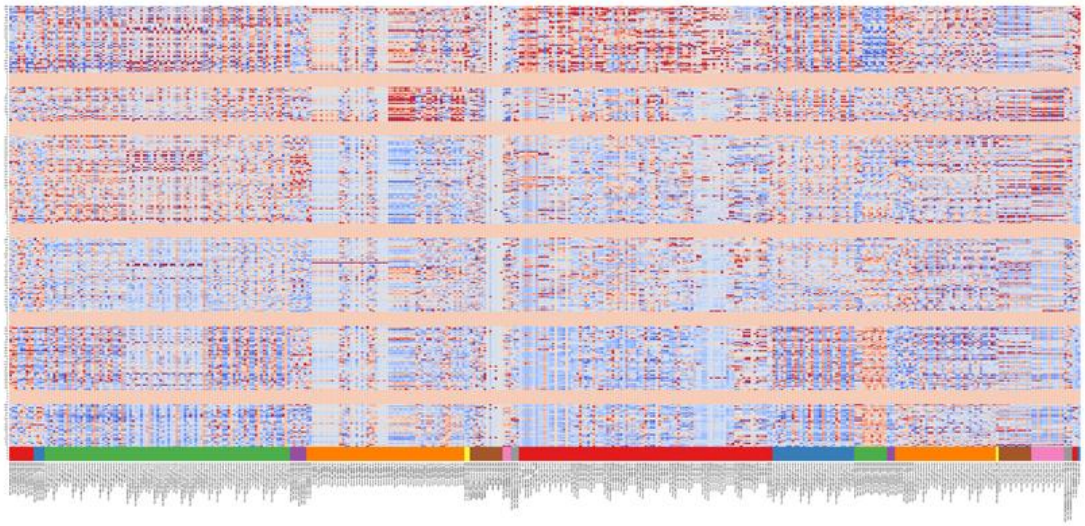
-Cluster 1 : "Dark–Rough" – Deep skin tone, high melanin/spots, pronounced roughness (SEr), low elasticity (R7), severe erythema, prominent wrinkles, and under-eye bags; seeks radiance-enhancing, smoothing treatments.

- Cluster 2 : "Oily–Red" – Natural tone with abundant sebum and porphyrins, high erythema and pore density; demands oil-control and calming formulas.

- Cluster 3 : "Moderate–Spotty" – Mid-tone, moderate spot counts and slight redness; balanced roughness and elasticity.

- Cluster 4 : “Deep–Healthy”– Darker tone, minimal spots and roughness, good elasticity; overall resilient skin.
- Cluster 5: “Light–Tight” – Very fair, few spots, smooth, elastic, slight redness due to high baseline brightness; values non-greasy, brightening actives.
- Cluster 6 : “Natural–Smooth”– Natural tone, low spots/pore counts, smooth texture, moderate elasticity; prefers lightweight, non-sticky formulations.

(**a**)



(**b**)

Figure 3. (a) Heatmap of z-score–normalized values for all 396 features across the six K-means clusters. Rows are ordered from top to bottom as Cluster 1 (Dark–Rough), Cluster 2 (Oily–Red), Cluster 3 (Moderate–Spotty), Cluster 4 (Deep–Healthy), Cluster 5 (Light–Tight), and Cluster 6 (Natural–Smooth). Red horizontal divider lines separate each cluster. Columns are grouped by measurement modality (elasticity, texture, colorimetry, wrinkles, spots, ultrasound), with blue indicating below-mean and red above-mean values. (b) Representative Visia CR facial images for each cluster, illustrating typical phenotypic characteristics.

4. **Discussion**

The clustering-derived phenotypes reveal coherent aging patterns that align with known pathophysiology and consumer needs. For example, the “Dark–Rough” group’s high melanin, roughness, and low elasticity suggest targeted interventions with brightening and firming actives, while the “Oily–Red” cluster would benefit from oil-control and anti-inflammatory formulations .
Our optimized “skin age” models—Random Forest (MSE=35.17, r=0.88) and neural network (MSE=33.45, r=0.90)—demonstrate strong predictive power using a compact 29-feature panel, outperforming previous probe-based models (MSE≈75) . The rigorous QC process ensured reliability by removing high-variance metrics and imputing missing data, yielding a robust dataset for modeling .
However, the single-city cohort may limit external generalizability; environmental and ethnic factors could influence both measured parameters and aging trajectories. Future work should validate these findings in multi-center studies, integrate longitudinal data to assess temporal dynamics, and explore causal links between lifestyle variables and skin aging, ultimately refining personalized skincare strategies.

5. **Conclusion**

Our study demonstrates that a carefully curated 29-feature panel—spanning high-resolution wrinkle metrics, multispectral imaging, biophysical measurements, and ultrasound—enables robust “skin age” prediction (RF: MSE=35.17, r=0.88; NN: MSE=33.45, r=0.90) in a well-controlled urban cohort. Six distinct aging phenotypes uncovered through K-means clustering offer actionable insights for targeted skincare formulations. Rigorous data quality control and grid-search–driven feature selection minimized noise and redundancy, enhancing model reliability. While single-city sampling limits broader applicability, our framework provides a scalable blueprint for personalized anti-aging strategies. Future multi-center validation and longitudinal analyses will further refine predictive accuracy and support dynamic, individualized intervention design.

Reference:

1. Farage, Miranda A., et al. "Structural characteristics of the aging skin: a review." *Cutaneous and ocular toxicology* 26.4 (2007): 343-357.
2. Kim, Miri, and Hyun Jeong Park. "Molecular mechanisms of skin aging and rejuvenation." *Molecular mechanisms of the aging process and rejuvenation* 450 (2016).
3. Mohiuddin, Abdul Kader. "Skin aging & modern age anti-aging strategies." *Int. J. Clin. Dermatol. Res* 7 (2019): 209-240.
4. Makrantonaki, Evgenia, and Christos C. Zouboulis. "Characteristics and pathomechanisms of endogenously aged skin." *Dermatology* 214.4 (2007): 352-360.
5. Li, Tian‐Hao, et al. "Artificial intelligence analysis of over a million Chinese men

and women reveals level of dark circle in the facial skin aging process." *Skin Research and Technology* 29.11 (2023): e13492.
6. Nurtiwi, Nurtiwi, Ruliana Ruliana, and Zulkifli Rais. "Convolutional Neural Network (CNN) Method for Classification of Images by Age." *JINAV: Journal of Information and Visualization* 3.2 (2022): 126-130.
7. Meyer, David H., and Björn Schumacher. "Aging clocks based on accumulating stochastic variation." *Nature Aging* 4.6 (2024): 871-885.
8. Nam, Gae Won, et al. "Differences in skin properties of Korean Women at the initial aging phase." *Journal of Cosmetics, Dermatological Sciences and Applications* 2014 (2014).
9. Rew, Jehyeok, et al. "Skin aging estimation scheme based on lifestyle and dermoscopy image analysis." *Applied Sciences* 9.6 (2019): 1228.
10. Lee, Seongwoo, et al. "Identifying patterns behind the changes in skin pores using 3‐dimensional measurements and K‐means clustering." *Skin Research and Technology* 28.1 (2022): 3-9.
11. Suppa, M., et al. "The determinants of periorbital skin ageing in participants of a melanoma case–control study in the UK." *British Journal of Dermatology* 165.5 (2011): 1011-1021.