

# Penilaian Rupa Wajah dengan Implementasi Fitur Geometris dan Tekstur Menggunakan Regresi Linear Berganda

Jessica

Informatika, Fakultas Industri Kreatif, Institut Teknologi dan Bisnis Kalbis  
Jalan Pulomas Selatan Kav. 22, Jakarta Timur, 13210  
Email : jessicapda98@gmail.com

**Abstract:** This research aim to scoring the facial look by implementing geometrical and texture features using multiple linear regression on scikit-learn library. Geometrical features calculate the range of facial landmark features, while texture features were calculated using a local binary pattern on scikit-image library. The method used for creating the model was multiple linear regression. This method models the relationship between independent variables (features) and dependent variables (score from humans), so the model can be used to predict new data. The score predicted by the machine learning model was compared to the average score from 13 respondents that were collected by a questionnaire. The result of scoring from the machine learning model and questionnaire were evaluated by mean squared error (MSE). The result showed the model can predict the score of facial look close to scoring from humans with MSE value 0,43. This model is built using a subjective opinion label from 13 respondents. Thus, the conclusion drawn from this research only give the big picture of the dataset used, and not a form of reality validation..

**Keywords:** geometrical features, texture features, facial look scoring, machine learning, multiple linear regression

**Abstrak:** Penelitian ini bertujuan untuk melakukan penilaian rupa wajah dengan mengimplementasikan fitur geometris dan fitur tekstur menggunakan regresi linear berganda pada modul scikit-learn. Fitur geometris menghitung jarak antar fitur landmark wajah, sedangkan fitur tekstur dihitung melalui local binary pattern pada modul scikit-image. Metode yang digunakan dalam pembuatan model adalah regresi linear berganda. Metode ini memodelkan hubungan antara variabel bebas (fitur) dengan variabel tak bebas (skor dari manusia), sehingga model dapat digunakan untuk memprediksi data baru. Hasil akhir penelitian berupa model pembelajaran penilaian kerupaan wajah. Skor dari hasil prediksi model dibandingkan dengan penilaian rata-rata dari 13 responden yang dikumpulkan melalui kuesioner. Skor dari model dan skor dari penilaian kuesioner dianalisis dan dievaluasi menggunakan mean squared error (MSE). Hasil menunjukkan model dapat memprediksi skor kerupaan wajah yang mendekati penilaian dari manusia nilai MSE 0,43. Model ini dibangun menggunakan label opini subjektif dari 13 responden. Sehingga, kesimpulan yang ditarik melalui penelitian ini hanya memberikan gambaran pada data set yang digunakan, dan bukan merupakan validasi realitas.

**Kata kunci:** fitur geometris, fitur tekstur, penilaian kerupaan wajah, pembelajaran mesin, regresi linear berganda

## I. PENDAHULUAN

Wajah merupakan tampilan tubuh pada manusia yang terdiri dari beberapa bagian seperti mata, hidung, mulut. Wajah berfungsi untuk menampilkan ekspresi,

penampilan, mewakili identitas, dan sarana komunikasi manusia. [1, hlm. 3] Selain itu, wajah juga merupakan bagian tubuh yang pertama kali dilihat dan diperhatikan oleh seseorang sehingga menimbulkan kesan pertama mengenai penampilan orang

tersebut. Hal ini mengakibatkan, penilaian mengenai kerupaan wajah dapat meningkatkan atau menjatuhkan kepercayaan diri seseorang [2, hlm. 112], karena manusia sebagai makhluk sosial cenderung memedulikan penilaian dari orang lain.

Persepsi mengenai penilaian rupa wajah merupakan hal yang subjektif, seperti dapat diibaratkan dengan pernyataan “beauty is in the eye of the beholder” (kecantikan ada di mata yang melihatnya). Setiap orang memiliki cara pandang dan penilaian yang berbeda mengenai persepsi kerupaan wajah seseorang sehingga penilaian dinilai subjektif. Namun, dalam skala populasi atau kumpulan berapa orang, terdapat konsistensi dan kemiripan dari penilaian tentang kerupaan wajah. Hal ini membuat kualitas fitur yang didapat dari skala populasi, dapat diaplikasikan dalam skala populasi yang lebih besar, sehingga analisis kerupaan wajah menjadi rasional dan dapat dipertimbangkan. [3, hlm. 123–124]

Adanya teknologi pembelajaran mesin (machine learning), image processing, dan analisis statistik dapat memudahkan dalam mempelajari tren penilaian manusia tentang kerupaan wajah. Sekumpulan data citra wajah serta penilaiannya diproses untuk pembuatan model komputasional penilaian kerupaan wajah. Model ini dapat digunakan untuk memprediksi nilai citra wajah baru yang hasilnya dapat mendekati penilaian dari sekelompok populasi manusia. Analisis dan pengukuran mengenai penilaian kerupaan wajah dapat dimanfaatkan dan diterapkan oleh ilmuwan, dokter, dan artis dalam bidang hiburan, operasi plastik, dan industri kosmetik.

Ekstraksi fitur yang dipilih untuk pelatihan model adalah fitur geometris dan fitur

tekstur kulit. Fitur-fitur ini dipilih karena adanya pemakaian operasi plastik yang berkaitan dengan proporsi atau geometris wajah, kemudian adanya penggunaan make-up, dan skin care yang berkaitan dengan kulit atau tekstur wajah. Hal ini sangat berkaitan dengan usaha orang untuk mencapai penampilan wajah yang diinginkan atau terlihat cantik, yang dapat menambah nilai kerupaan dari orang lain.

Berbagai studi telah dilakukan untuk meneliti mengenai faktor kerupaan dan persepsi orang mengenai kerupaan. Xie et al. pada tahun 2015 [4] membuat data set 500 citra perempuan Asia dengan rentang skor kecantikan 1-5 yang diperoleh melalui 75 responden. Fitur yang diekstraksi adalah fitur geometris dan fitur tekstur. Fitur geometris dihitung jaraknya menggunakan perhitungan golden ratio. Fitur tekstur diekstrak menggunakan filter Gabor. Nilai kecantikan diprediksi melalui fitur geometris, fitur tekstur, maupun kombinasi fitur geometris dan tekstur. Hasil dari penelitian ini adalah prediksi nilai kecantikan lebih akurat jika menggunakan kombinasi dari dua fitur ini, daripada menggunakan satu fitur saja.

Selanjutnya, Zhang et al. pada tahun 2017 [5] meneliti tentang pembuatan model prediksi nilai kecantikan wajah dan aplikasi untuk mempercantik gambar wajah. Fitur yang diekstraksi yaitu fitur geometris dan fitur tekstur. Fitur geometris dicari fitur terbaiknya menggunakan k-nearest neighbor (KNN) dan dievaluasi menggunakan perhitungan korelasi Pearson. Fitur tekstur diekstrak menggunakan Local Binary Pattern (LBP) dan PCA-net. Model untuk analisis kecantikan dibuatkan dengan membandingkan akurasi dari algoritma KNN, support vector regression (SVR), regresi linear, dan Histogram matching. Hasil dari penelitian ini menunjukkan

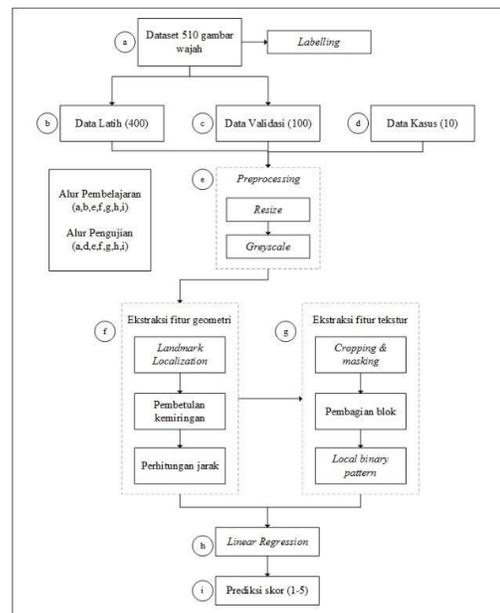
bahwa metode yang lebih baik untuk pembuatan model learning terbaik adalah KNN dan SVR.

Data set yang digunakan dalam penelitian ini sebanyak 510 citra. Pengelompokan data adalah data latih sebanyak 400, data validasi sebanyak 100, dan data uji sebanyak 10. Citra data set pada data latih dan data validasi adalah South China University of Technology – Facial Beauty Prediction (SCUT-FBP) yang dibuat oleh Xie et al. pada tahun 2015. Data set ini terdiri dari 500 citra perempuan Asia. Citra data set pada data uji merupakan citra yang dikumpulkan secara daring. Label skor kerupaan pada data set dikumpulkan secara mandiri dengan menyebarkan kuesioner kepada 13 responden dengan rentang skor kecantikan 1-5, kemudian rata-rata penilaian 13 responden menjadi label skor kerupaan wajah.

Fitur geometris diekstraksi melalui perhitungan jarak antar titik proporsi wajah seperti alis, mata, hidung, dan mulut. Selanjutnya fitur tekstur kulit diekstraksi menggunakan algoritma local binary pattern. Kedua fitur yang sudah diekstraksi akan dimasukkan ke dalam algoritma regresi linear berganda untuk menghasilkan model. Kemudian model divalidasi untuk menguji kelayakan model. Pengujian model dilakukan dengan prediksi nilai kerupaan pada data uji. Perhitungan mean squared error dihitung melalui kuadrat selisih antara nilai prediksi model dengan nilai dari manusia, kemudian dibagi dengan jumlah data dengan tujuan untuk mengetahui besarnya kesalahan antara dari prediksi model dengan skor label. Dengan demikian, aplikasi dapat memprediksi skor kerupaan baru yang mendekati persepsi kumpulan manusia.

## II. METODE PENELITIAN

Alur metode penelitian yang digambarkan dalam bentuk *flowchart* dipaparkan pada Gambar 1.



Gambar 1 Alur penelitian dalam bentuk *flowchart*

### A. Data set

Data set yang digunakan adalah citra wajah wanita beserta dengan label penilaian manusia. Citra wajah wanita diambil melalui data set bernama SCUT-FBP yang dibuat oleh Xie et al. pada tahun 2015. [4] Data set ini terdiri dari 500 gambar wanita Asia dan dapat diunduh secara cuma-cuma melalui situs <http://www.hcii-lab.net/data/SCUT-FBP>. Gambar wanita pada data set beresolusi tinggi, wajah menghadap ke arah kamera, berekspresi natural, latar belakang yang simpel, dan tidak menggunakan aksesoris. Untuk keperluan data uji atau data kasus, data set ditambah 10 citra perempuan yang diunduh secara daring. Sehingga total citra

data set yang digunakan dalam penelitian sebanyak 510 citra.

Skor kerupaan setiap 510 gambar dikumpulkan oleh peneliti melalui penyebaran kuesioner kepada 13 responden dengan pengisian form secara daring dengan rentang skor kerupaan 1-5. 13 responden ini terdiri dari 9 laki-laki dan 4 perempuan. 13 responden menunjukkan populasi untuk melihat tren penilaian skor rupa wajah. Kemudian 13 nilai dari responden dirata-rata, dan nilai rata-rata tersebut menjadi label nilai kerupaan wajah untuk satu gambar. Langkah ini dilakukan ke setiap gambar pada data set.

### **B. Pre-processing**

*Pre-processing* perlu dilakukan untuk persiapan ekstraksi fitur. *Pre-processing* yang dilakukan adalah grayscale dan resize. Proses grayscale perlu dilakukan untuk penyederhanaan citra supaya dapat lebih mudah diproses, dan resize perlu dilakukan untuk penyetaraan ukuran resolusi gambar sehingga performa model dapat lebih dimaksimalkan.

### **C. Ekstraksi Fitur Geometris**

Tahap yang dilakukan pada ekstraksi fitur geometris adalah proses deteksi wajah, rotasi citra wajah, lokalisasi landmark wajah, dan perhitungan jarak antar fitur wajah. Deteksi wajah menerapkan konsep histogram of oriented gradients yang mampu mendeteksi objek, atau dalam kasus ini mendeteksi wajah. Kemudian, dilakukan lokalisasi 68 fitur wajah yang mewakili fitur penting wajah seperti mata, alis, dagu, hidung dan bibir. Selanjutnya, wajah citra dirotasi sehingga posisi wajah menjadi tegak lurus dengan acuan garis lurus kedua mata ( $180^\circ$ ). Proses terakhir dari ekstraksi fitur geometris adalah 42 perhitungan jarak antara 68 fitur landmark wajah. Perhitungan jarak ini

terdiri dari jarak mata, alis, dagu, hidung, dan bibir secara vertikal maupun horisontal. Perhitungan ini diambil dari jurnal yang dirujuk [4], [5].

### **D. Ekstraksi Fitur Tekstur**

Citra wajah terlebih dahulu dilakukan face masking untuk mengurangi noise dan fokus pada pemrosesan tekstur kulit bagian tepi dalam wajah saja. Pada face masking, area wajah diambil dengan cara menutupi area wajah dengan mask yang dibuat. Kemudian, setiap piksel pada citra dihitung nilai local binary pattern (LBP) dan dicari nilai histogramnya. Algoritma ini digunakan karena perhitungannya yang cepat dan algoritmanya yang sederhana, dan tetap dapat menghasilkan ciri tekstur yang efisien. Hasil nilai LBP dari setiap piksel dibuat dalam bentuk histogram, kemudian nilai histogram dinormalisasi untuk menunjukkan range intensitas tekstur setiap piksel. Jumlah bar histogram yang mewakili fitur tekstur wajah adalah sebanyak 20 batang.

Hasil dari ekstraksi fitur geometris (42 perhitungan jarak), dan ekstraksi fitur tekstur (20 nilai histogram) digabungkan menjadi 62 fitur rupa wajah yang akan digunakan untuk pelatihan model.

### **E. Pembuatan Model**

Setelah seluruh gambar sudah diekstraksi fitur, maka menjadi masukan untuk pembuatan model prediksi skor kerupaan. Sebelum dimasukkan ke dalam algoritma model, data set dibagi ke dalam 2 bagian yaitu data di-split menjadi dua bagian, data latih dan data validasi. Perbandingan jumlah data latih dan data validasi adalah 80:20, sehingga jumlah data latih adalah 400, dan jumlah data validasi adalah 100.

Karena *label* merupakan *multivariate* / variabel kontinu / data numerik (angka 1-

5), maka masalah dapat diselesaikan dengan algoritma regresi. Metode regresi yang digunakan adalah regresi linear berganda. Analisis regresi berganda mempelajari bentuk hubungan antara variabel bebas (X) dengan variabel tak bebas (Y). X adalah fitur, dan Y adalah skor dari manusia. [6, hlm. 474–475]

Setelah model dilatih, selanjutnya prediksi dilakukan terhadap model regresi linear yang telah dibuat menggunakan data validasi untuk menguji validasi model. Hasil prediksi model terhadap data validasi diukur keakuratannya terhadap data realnya yaitu penilaian manusia. Pengujian keakuratan dilakukan menggunakan *Mean Squared Error* (MSE).

Perhitungan MSE merepresentasikan rata-rata kesalahan dari kuadrat bias suatu model. Bias yang dimaksud adalah selisih antara variabel x (nilai prediksi/peramalan) dengan variabel y (nilai sebenarnya). Rumus MSE dipaparkan pada (1). [6, hlm. 278]

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (1)$$

**F. Pengujian Model**

Pengujian dilakukan dengan menerapkan model terhadap 10 data yang dipilih secara acak dari data latih. 10 data terpilih merupakan bagian atau cuplikan data dari data latih. Langkah ini bertujuan untuk melihat hasil prediksi model pada sebagian data latih. Pengujian terhadap data latih secara detil dipaparkan pada Tabel 1. Hasil pada Tabel 1 hanya menunjukkan representasi keseluruhan data yang digunakan, untuk memberikan gambaran pada data set yang digunakan untuk pelatihan model.

Tabel 1 Pengujian pada data hatih (cuplikan dari data set)

No.	Citra Data Latih	Skor Prediksi (x)	Skor Label (y)
1.		2,3077	2,7616
2.		2,7692	3,1122
3.		2,8462	2,8108
4.		3,0000	2,6989
5.		4,0769	3,2826
6.		4,0000	3,1741
7.		3,2308	2,7846

8.		4,3077	3,2861
9.		1,7692	2,0305
10.		2,0000	2,4796

Perhitungan MSE untuk 10 data yang merupakan bagian atau cuplikan data latih pada Tabel 2 dipaparkan pada perhitungan (2). Hasil MSE pada perhitungan (2) menunjukkan nilai kuadrat bias dari representasi keseluruhan data, untuk memberikan gambaran pada data set yang digunakan untuk pelatihan model.

$$MSE = \frac{1}{10}(3,2696) = 0,327 \quad (2)$$

Kemudian, model diterapkan terhadap 10 data uji yang belum dilihat sebelumnya (di luar data uji dan data validasi). 10 data terpilih merupakan bagian atau cuplikan dari data set. Langkah ini bertujuan untuk melihat hasil prediksi model jika diterapkan terhadap data kasus atau data riil. Pengujian terhadap data uji secara detil dipaparkan pada Tabel 2.

Tabel 2 Pengujian pada data uji (cuplikan dari data set)

No.	Citra Data Uji	Skor Prediksi (x)	Skor Label (y)
1.		3,3846	3,7462
2.		3,9231	3,7063
3.		4,4615	3,9848
4.		3,0000	2,8242
5.		4,3077	3,1469
6.		3,8462	3,3497
7.		4,1538	4,6820

8.		3,5385	3,5419
9.		2,9231	3,5167
10.		4,1538	2,8777

Perhitungan MSE pada 10 data uji yang merupakan bagian atau cuplikan dari data set (di luar data latih dan data validasi) dipaparkan pada perhitungan (3)

$$MSE = \frac{1}{10}(4,2899) = 0,429 \quad (3)$$

**G. Analisis Fitur Kerupaan Wajah**

Untuk mengetahui fitur geometris dan fitur tekstur yang membuat skor menjadi kecil atau besar, maka dibandingkan fitur pada citra wajah yang memiliki nilai prediksi tertinggi dan terendah dari hasil prediksi pada latih dan data uji pada Tabel 1 dan Tabel 2. Citra dengan nilai prediksi tertinggi dan terendah dipaparkan pada Tabel 3.

Tabel 3 Citra wajah dengan nilai tertinggi dan terendah

Citra Wajah		
Nilai	4,682	2,0305

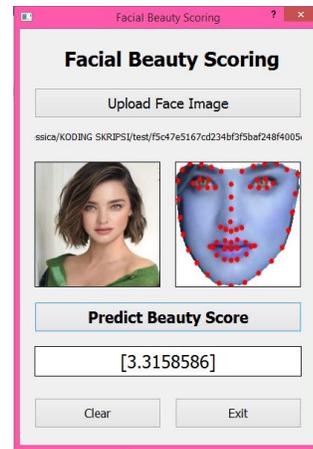
Hasil perbandingan fitur geometris dan fitur tekstur antara citra dengan nilai tertinggi dan terendah dipaparkan pada Gambar 2.

Skor/Fitur	1	2	3	4	5	6	7	8	9	10	11	12
4,682	59	58	28	30	45	31	87	90	31	27	13	26
2,0305	54	58	25	27	49	28	86	87	26	32	16	32
Skor/Fitur	13	14	15	16	17	18	19	20	21	22	23	24
4,682	164	156	145	135	118	95	67	35	54	36	16	33
2,0305	162	159	153	145	131	109	78	41	63	40	17	39
Skor/Fitur	25	26	27	28	29	30	31	32	33	34	35	36
4,682	15	59	21	76	80	16	6	14	43	62	72	25
2,0305	16	46	16	70	70	15	8	14	45	68	75	30
Skor/Fitur	37	38	39	40	41	42	43	44	45	46	47	48
4,682	21	68	11	11	11	10	0,2240	0,0717	0,1315	0,0297	0,1179	0,0174
2,0305	23	57	7	7	7	7	0,1324	0,0596	0,1528	0,0316	0,2313	0,0113
Skor/Fitur	49	50	51	52	53	54	55	56	57	58	59	60
4,682	0,0038	0,0339	0,0146	0,1292	0,0542	0,0047	0,0038	0,0033	0,0033	0,0325	0,0028	0,0207
2,0305	0,0035	0,0241	0,0273	0,1083	0,0162	0,0024	0,0057	0,0015	0,0058	0,0130	0,0041	0,0259
Skor/Fitur	61	62										
4,682	0,0118	0,0891										
2,0305	0,0298	0,1135										

Gambar 2 Perbandingan fitur wajah pada nilai tertinggi dan terendah

**H. Aplikasi Graphic User Interface (GUI)**

Tampilan aplikasi GUI untuk penerapan model penilaian kerupaan wajah dibuat menggunakan bahasa PyQt. Fitur dari aplikasi ini adalah unggah gambar wajah wanita, dan prediksi nilai kerupaan wajah. Hasil dari tampilan aplikasi GUI dipaparkan pada Gambar 4.



Gambar 3 Tampilan Aplikasi GUI

### III. HASIL DAN PEMBAHASAN

Berdasarkan perhitungan dari rumus (1), didapatkan hasil evaluasi model melalui MSE pada sebagian data latih dan data uji adalah 0,327 dan 0,429. Dari hasil tersebut, dapat disimpulkan tingkat kesalahan model dalam memprediksi nilai kerupaan wajah adalah 0,429 dari rata-rata penilaian 13 responden. Untuk meminimalisir persentase kesalahan (supaya model lebih akurat) dapat diatasi dengan memperbanyak data dan memperbaharui ekstraksi fitur.

Seperti dipaparkan pada Gambar 2, dibandingkan fitur wajah antara citra wajah dengan nilai rupa wajah tertinggi dan nilai rupa wajah terendah. Sehingga dapat dibahas, fitur wajah yang dapat menaikkan nilai kerupaan wajah adalah :

1. Alis yang panjang (fitur 1 dan 2)
2. Jarak antar alis panjang (fitur 3)
3. Mata yang panjang dan lebar atau belo (fitur 4, 6, dan 39-42)
4. Jarak antar mata panjang (fitur 5, 7 dan 8)
5. Jarak mata ke tepi wajah pendek (fitur 9 dan 10)
6. Hidung tidak lebar atau pesek (fitur 11 dan 12)
7. Rahang tidak lebar atau rahang tirus (fitur 13-20)
8. Bibir kecil dan pendek, tetapi tebal (fitur 21-25, 31-32, dan 37)
9. Terdapat jarak yang cukup (panjang) antar vertikal alis, mata, hidung, dan bibir (fitur 27-30)
10. Jarak yang pendek antara bibir dan dagu (fitur 33, 34, 36)

Rentang skor dari manusia pada label data latih berada pada rentang 1.615 – 4.308. Hal ini dikarenakan perhitungan skor diambil dari rata-rata penilaian 13

orang saja. Jika jumlah responden diperbanyak, maka rentang skor dapat menjadi lebih panjang dan variatif. Adapun rentang skor pada label data latih mempengaruhi rentang skor prediksi. Sehingga skor kerupaan yang dihasilkan berada pada rentang 2-4 saja.

Kemudian, skor satu orang tidak selalu sama. Gambar yang berbeda meskipun orangnya sama, dapat menghasilkan skor yang berbeda bergantung dari posisi wajah dan tekstur pada foto tersebut. Jika ada luka, jerawat, minyak, atau hal lainnya yang mengganggu pada kulit wajah dapat mempengaruhi skor kerupaan wajah.

Hasil evaluasi model dalam memprediksi skor diperoleh melalui data set citra wajah SCUT-FBP dan label opini 13 responden. Sehingga, kesimpulan yang ditarik melalui penelitian ini menunjukkan akurasi subjektivitas 13 orang melalui label data set pada model. Hasil penelitian hanya memberikan gambaran pada data set yang digunakan, dan bukan merupakan validasi realitas.

### IV. SIMPULAN

Berdasarkan penelitian yang dilakukan, dapat disimpulkan :

1. Fitur geometris wajah yaitu perhitungan jarak antar fitur wajah (mata, alis, hidung, bibir, dagu) dan fitur tekstur wajah melalui perhitungan local binary pattern dapat dijadikan salah satu faktor penilaian rupa wajah.
2. Regresi linear berganda dapat digunakan untuk pembuatan model dan memprediksi nilai rupa wajah.
3. Pada tahap pengujian, model diuji terhadap 10 data uji / kasus. Nilai MSE atau tingkat kesalahan model regresi

dalam memprediksi nilai kerupaan wajah adalah 0,43.

4. Fitur geometris yang dapat menaikkan nilai kerupaan wajah adalah alis yang panjang, mata yang panjang dan belo, hidung yang tidak pesek, rahang atau dagu yang tirus, bibir yang kecil namun tebal, dan terdapat jarak yang cukup (panjang) atau tidak sempit antar fitur wajah tersebut.

## DAFTAR RUJUKAN

- [1] D. Zhang, F. Chen, dan Y. Xu, *Computer Models for Facial Beauty Analysis*. Switzerland: Springer, 2016.
- [2] Faizah, U. Rahma, dan Y. P. Dara, *Psikologi Pendidikan: Aplikasi Teori di Indonesia*. Malang: UB Media, 2017.
- [3] Z. David, Z. Bob, dan Z. Qijun, *Facial Multi-characteristics And Applications*. Beijing: Higher Education Press, 2018.
- [4] D. Xie, L. Liang, L. Jin, J. Xu, dan M. Li, "SCUT-FBP: A Benchmark Dataset for Facial Beauty Perception," *Ieee-Smc*, hlm. 1821–1826, 2015.
- [5] B. Zhang, X. Xiao, dan G. Lu, "Facial Beauty Analysis Based on Features Prediction and Beautification Models," *Pattern Anal. Appl.*, vol. 21, no. 2, hlm. 529–542, 2018, doi: 10.1007/s10044-017-0647-2.
- [6] K. F. Weaver, V. C. Morales, S. L. Dunn, K. Godde, dan P. F. Weaver, *An Introduction to Statistical Analysis in Research: With Applications in the Biological and Life Sciences*. Hoboken: Wiley, 2017.