

Analisis Kinerja Model Klasifikasi Pada Multiclass Facial Expression Recognition Berbasis Fitur Eigenface

Syefrida Yulina^{[1]*}, Heni Rachmawati^[2]

Program Studi Sistem Informasi, Jurusan Teknologi Informasi, Politeknik Caltex Riau^{[1],[2]}

Jl. Umbansari, Pekanbaru – Riau, 28265, Indonesia

syefrida@pcr.ac.id^[1], heni@pcr.ac.id^[2]

Abstract— Facial Expression Recognition (FER) is currently widely explored by researchers in the field of Computer Vision. The application of Machine Learning and Deep Learning methods is useful in developing an intelligent system that is accurate in recognizing facial expressions such as emotions. This is inseparable from the type of dataset and classification method used which certainly affects the desired results. To choose the right method, it is necessary to compare the performance of these methods. This study focuses on comparing the performance results of four classification methods namely, Convolutional Neural Network (CNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes Classifier (NBC) on a multiclass dataset for seven classes of facial emotion labels based on Eigenface feature selection uses the Personal Component Analysis (PCA) algorithm. The test parameters used to perform method comparisons are accuracy, recall, precision, f1-score, as well as the Receiving Operating Characteristic (ROC) and Area Under Curve (AUC) curves. The results of the analysis state that the SVM method has the highest accuracy value, while other methods show varying performance based on recall, precision, f1-score, and ROC and AUC analysis. This research was conducted on the FER 2013 dataset which showed that the classification method tested had quite good performance according to the test parameters.

Keywords— Facial expression, FER dataset, Method comparison, Performance analysis

Abstrak— Facial Expression Recognition (FER) saat ini banyak dieksplorasi oleh para peneliti dalam bidang Computer Vision. Penerapan metode Machine Learning dan Deep Learning, bermanfaat dalam mengembangkan sebuah sistem cerdas yang akurat dalam mengenali ekspresi wajah seperti emosi. Hal ini tidak terlepas dari jenis dataset dan metode klasifikasi yang digunakan yang pastinya mempengaruhi hasil yang diinginkan. Untuk pemilihan metode yang tepat, maka diperlukan sebuah perbandingan kinerja terhadap metode-metode tersebut. Penelitian ini berfokus pada perbandingan hasil kinerja empat metode klasifikasi yaitu, Convolutional Neural Network (CNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes Classifier (NBC) pada dataset multiclass untuk tujuh kelas label emosi wajah berdasarkan seleksi fitur Eigenface menggunakan algoritma Personal Component Analysis (PCA). Parameter uji yang digunakan untuk melakukan perbandingan metode adalah seperti nilai akurasi, recall, precision, f1-score, serta kurva Receiving Operating Characteristic (ROC) dan Area Under Curve (AUC). Hasil analisis menyatakan bahwa metode SVM memiliki nilai akurasi tertinggi, sedangkan metode lainnya menunjukkan kinerja bervariasi berdasarkan nilai recall, precision,

f1-score, serta analisis ROC dan AUC. Penelitian ini dilakukan pada dataset FER 2013 yang menunjukkan metode klasifikasi yang diujikan memperoleh kinerja cukup baik sesuai dengan parameter uji.

Kata Kunci— Analisis Performa, Dataset FER, Ekspresi Wajah, Perbandingan Metode

I. PENDAHULUAN

Pengenalan emosi pada wajah menggunakan teknologi Facial Expression Recognition (FER) sedang aktif dieksplorasi dalam penelitian Computer Vision. Banyaknya metode Machine Learning dan Deep Learning, berpotensi untuk membangun sebuah sistem cerdas yang akurat dalam mengenali emosi [1]. Sebagai bagian penting dari Computer Vision, FER semakin menarik perhatian, sehingga penelitian dibidang ini mengadopsi berbagai metode untuk meningkatkan FER tersebut [2-3]. Saat ini terdapat penelitian tentang penerapan serta perbandingan metode klasifikasi pada berbagai jenis dataset [4-5] terutama pada klasifikasi emosi wajah [1][6-8]. Dengan hadirnya library-library yang telah disediakan oleh bahasa pemrograman tertentu seperti Python, sehingga memungkinkan para peneliti dengan mudah dan cepat untuk melakukan klasifikasi terhadap citra wajah. Kinerja berbagai metode klasifikasi perlu diuji pada sebuah dataset agar menjadi salah satu cara dalam penentuan metode yang tepat. Masing-masing dari metode tersebut memiliki kelebihan dan kekurangan, hal ini dapat dilihat pada keunggulan dalam mengolah objek dataset. Gap analisis pada penelitian ini dapat dilihat dari Tabel I.

Metode klasifikasi emosi wajah yang umumnya banyak diterapkan adalah seperti: Convolutional Neural Network (CNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes Classifier (NBC). Metode tersebut merupakan algoritma klasifikasi citra yang efektif untuk masalah dua kelas atau multi kelas [9-10]. Penggunaan sejumlah dataset sangat berpengaruh terhadap akurasi sebuah metode klasifikasi. Hal ini dapat menjadikan acuan bagi peneliti untuk memperoleh nilai akurasi yang bervariasi serta karakteristik yang dimiliki oleh sebuah metode [11].

Pada penelitian ini melakukan analisis kinerja dari metode-metode klasifikasi tersebut diatas dalam mengelola dataset yang memiliki lebih dari dua label (multiclass) emosi wajah yaitu: bahagia (happiness), tidak bahagia (sadness), terkejut (surprise), marah (anger), takut (fear), jijik (disgust), dan

netral. Dataset yang akan digunakan untuk penelitian ini adalah FER-2013 yang memiliki banyak data citra wajah berisikan tujuh emosi wajah. Tahapan pengenalan emosi wajah (FER) yang dilakukan yaitu, seperti: mendeteksi dan mengakuisi bagian wajah, *pre-processing* data, ekstraksi fitur wajah berupa Eigenface menggunakan algoritma Personal Component Analysis (PCA), dan mengklasifikasikan emosi pada wajah tersebut. Kinerja setiap metode diukur melalui pengujian nilai akurasi, *recall*, *precision*, *f1-score*, serta *Receiving Operating Characteristic* (ROC) dan *Area Under Curve* (AUC) menggunakan dataset yang sama. Dengan adanya perbandingan

beberapa metode klasifikasi yang berbeda, maka dapat dijadikan sebagai acuan dalam menentukan metode yang terbaik.

TABEL I. TABEL GAP ANALISIS PENELITIAN

Tahun	Peneliti	Metode	Dataset	Ekstraksi Fitur	Performa				ROC	AUC
					Akurasi	Presisi	Recall	F-Measure		
2023	Penelitian Sekarang	SVM KNN NBC CNN	FER 2013	Eigenface	✓	✓	✓	✓	✓	✓
2021	Song [4]	CNN	FER 2013	ROI/Gabor	✓	-	-	-	-	-
2019	Canedo & Neves [1]	SVM KNN CNN	CK+/JAFFE/FER 2013	LBP/ROI/Gabor	✓	-	-	-	-	-
2019	Utami et al. [6]	CNN	CK+	LBP/Gabor	✓	-	-	-	-	-
2017	Tarnowski et al. [5]	KNN/MLP - NN	KDEF	AU	✓	-	-	-	-	-

II. PENELITIAN TERKAIT

A. Penelitian Terdahulu

Saat ini terdapat banyak penelitian mengenai penerapan serta perbandingan metode klasifikasi emosi wajah serta penerapannya pada berbagai jenis dataset. Pada tahun 2021, [8] menerapkan metode CNN dalam mengekstraksi fitur penting pada ekspresi wajah serta meningkatkan akurasi pada pengenalan emosi. Tahun 2020, [6] melakukan perbandingan performa beberapa metode klasifikasi pada dataset citra busur panah, dengan hasil penelitian menyatakan bahwa metode SVM memiliki performa yang lebih baik dibandingkan metode lainnya. Kemudian selanjutnya [1] dan [12] ditahun yang sama melakukan studi analisis untuk pengenalan ekspresi wajah dalam berdasarkan perspektif dataset, metode klasifikasi, dan variasi pose wajah, lalu [13] menerapkan metode KNN dan MLP dalam pengenalan emosi wajah untuk model wajah 3 dimensi.

Menurut [9] *face/wajah* merupakan kunci pembeda yang digunakan secara luas sebagai identitas personal yang memiliki keunikan atau ciri-ciri khusus. Bidang kajian *face detection* khususnya *Facial Expression Recognition* telah menjadi perhatian yang dipertimbangkan di dalam pengembangan dan peningkatan interaksi antara manusia dengan mesin atau manusia dengan manusia, dengan pertimbangan dan alasan kealamian serta efisiensi kinerja di dalam berkomunikasi melalui sebuah ekspresi wajah. *Facial Expression Recognition* secara esensial melokasikan dan mengekstraksikan suatu bagian wajah dari latarbelakangnya yang kemudian melakukan pengenalan suatu ekspresi [2][12][14]. Hal ini tampak suatu permasalahan atau tugas yang mudah, tetapi memerlukan suatu teknik/metode dalam penyelesaiannya. Wajah manusia

merupakan suatu objek yang dinamis dan mempunyai derajat kevariasian yang sangat tinggi, sehingga *Facial Expression Recognition* menjadi suatu permasalahan yang sangat rumit di dalam *Computer Vision* [15]. Sebuah citra wajah manusia dapat dijadikan sebagai sampel pada suatu kajian dalam pengklasifikasian emosi wajah manusia.

Facial Expression Recognition terdiri dari beberapa langkah, yaitu: *image acquisition*, *pre-processing*, *feature extraction*, *classification/regression* [1][5]. Berbagai metode/teknik/pendekatan telah banyak dibahas pada suatu kajian ilmiah tentang face detection meliputi: segmentasi warna, citra wajah, eigenface, jaringan syaraf tiruan, dll. Pada *Facial Expression Recognition*, algoritma klasifikasi digunakan untuk memprediksi/mengklasifikasikan label emosi tertentu dengan inputan berupa citra wajah.

B. Metode Klasifikasi

[9-10] mendeskripsikan algoritma klasifikasi *K-Nearest Neighbor*, *Naive Bayes Classifier*, *Support Vector Machine*, dan *Convolutional Neural Network* merupakan algoritma klasifikasi citra yang efektif untuk masalah dua kelas atau multi kelas.

K-Nearest Neighbor (KNN) adalah sebuah algoritma pembelajaran berbasis *instance* yang menggunakan teknik non-parametrik saat melakukan klasifikasi atau regresi. KNN mencari nilai vektor fitur k dengan sifat termirip, lalu mengelompokkan data baru ke kelompok vektor tersebut. Data pelatihan terdiri dari vektor dalam ruang fitur multidimensi, masing-masing dengan label kelas. Fase pelatihan algoritma hanya terdiri dari penyimpanan vektor fitur ini dan kelas milik mereka. Pada langkah klasifikasi, fitur input atau kumpulan

fitur diprediksi dengan menetapkan kelas yang memiliki fitur terdekat dengan input. Metrik jarak yang umum digunakan untuk menghitung fitur mana yang lebih dekat ke input adalah jarak Euclidean (ED) dan jarak Hamming. Kekuatan metode ini terletak pada implementasi yang sederhana dan pada langkah pelatihan yang cepat.

Naïve Bayes Classifier adalah pengklasifikasi probabilistic berdasarkan penerapan teorema Bayes untuk probabilitas bersyarat. Asumsinya adalah bahwa semua fitur independent dan tidak terkait satu sama lain (*naïve*). Pengklasifikasi dibangun dengan mengalikan probabilitas bersyarat individu dari setiap fitur untuk mendapatkan peluang total suatu kelas. Kemudian kelas dengan tertinggi probabilitas dipilih.

Support Vector Machine (SVM) adalah algoritma *Machine Learning* yang banyak digunakan untuk masalah klasifikasi/regresi. Model SVM adalah representasi fitur dalam ruang, dipetakan sehingga fitur milik setiap kelas dibagi dengan celah yang jelas yang selebar mungkin. Fitur input kemudian dipetakan ke dalam ruang yang sama dan diprediksi menjadi milik kelas berdasarkan sisi celah mana mereka berada. Fase pelatihan membuat peta ini yang digunakan setelah untuk prediksi.

Convolutional Neural Network (CNN) merupakan salah satu jenis jaringan syaraf tiruan yang banyak digunakan dalam *Computer Vision (Deep Learning)*. CNN memiliki kemampuan untuk mendeteksi dan mengidentifikasi pola dasar yang terlalu rumit. Sebuah citra inputan diproses melalui beberapa lapisan tersembunyi CNN yang akan diuraikan menjadi fitur-fitur. Fitur tersebut kemudian digunakan untuk klasifikasi, umumnya melalui fungsi Softmax yang mengambil nilai probabilitas tertinggi dari distribusi probabilitas kelas sebagai kelas yang diprediksi.

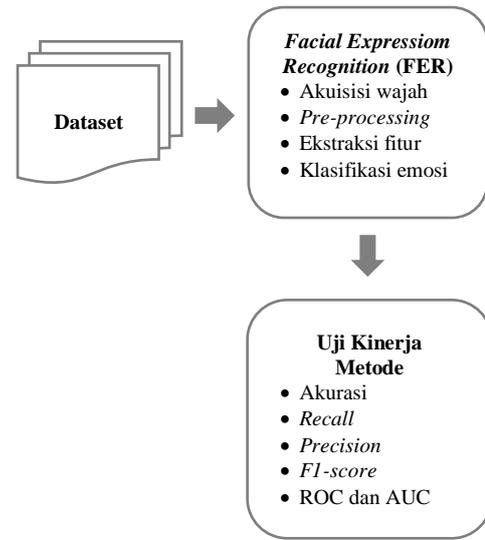
C. Python dan Library Computer Vision

Python merupakan salah satu Bahasa pemrograman yang dapat digunakan untuk berbagai masalah dalam *Artificial Intelligent* dan *Machine Learning*. Python banyak diterapkan pada *Computer Vision (CV)*, yang merupakan bagian dari *Artificial Intelligent* terkait aplikasi-aplikasi CV. Python menyediakan *library-library* yang memudahkan para *developer* untuk mengembangkan aplikasi CV, seperti OpenCV, Keras, Matplotlib, Scikit-Image, dll.

OpenCV (*Open-Source Computer Vision Library*) adalah library perangkat lunak Machine Learning dan Computer Vision yang bersifat open-source [16]. OpenCV diciptakan untuk menyediakan infrastruktur umum pada aplikasi *Computer Vision* dan untuk mempercepat penggunaan persepsi mesin dalam produk komersial. *Library* ini memiliki lebih dari 2500 algoritma yang dioptimalkan. Algoritma pada OpenCV dapat digunakan untuk mendeteksi wajah, mengidentifikasi objek, klasifikasi, dll. OpenCV-Python memanfaatkan Numpy, yang merupakan perpustakaan yang sangat dioptimalkan untuk operasi numerik dengan sintaks gaya MATLAB. Semua struktur array OpenCV dikonversi ke dan dari array Numpy. Ini juga memudahkan integrasi dengan library lain yang menggunakan Numpy seperti SciPy dan Matplotlib.

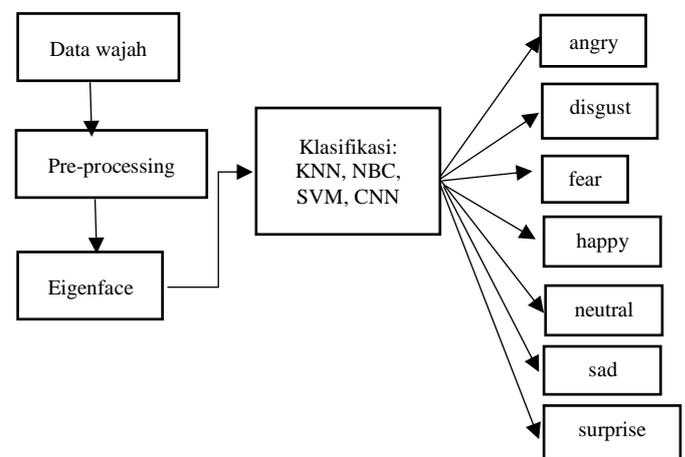
III. METODE PENELITIAN

Metode penelitian terdapat pada Gambar 1. Metode penelitian mencakup tiga tahapan, yaitu dataset citra wajah, *Facial Expression Recognition*, dan uji kinerja metode klasifikasi.



Gambar 1. Metodologi Penelitian

Tahapan *Facial Expression Recognition (FER)* digunakan untuk proses klasifikasi emosi yang dimulai dari proses deteksi dan akuisisi wajah, *pre-processing* data, ekstraksi fitur, kemudian klasifikasi emosi pada wajah menggunakan metode tertentu (Gambar 2). Pada penelitian ini metode yang diterapkan adalah algoritma *k-Nearest Neighbor (KNN)*, *Naive Bayes Classifier (NBC)*, *Support Vector Machine (SVM)*, dan *Convolutional Neural Network (CNN)*. Class yang diklasifikasikan yaitu *angry*, *disgust*, *fear*, *happy*, *neutral*, *sad* dan *surprise*.

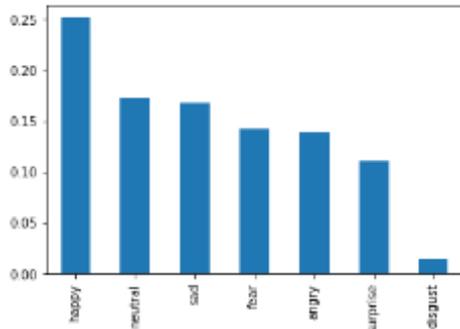


Gambar 2. Tahapan FER

Tahapan uji kinerja setiap metode klasifikasi diukur melalui pengujian nilai akurasi, *recall*, *precision*, *f1-score*, serta

Receiving Operating Characteristic (ROC) dan Area Under Curve (AUC) menggunakan dataset yang sama.

Dataset yang digunakan untuk klasifikasi adalah data FER 2013 yang disediakan oleh Kaggle, dimana data ini terdiri dari berbagai citra ekspresi wajah yang memiliki tujuh label emosi wajah seperti Gambar 3.



Gambar 3. Distribusi Data untuk Emosi Wajah

Pada Gambar 3 setiap emosi wajah memiliki label *multiclass* seperti emosi bahagia (*happiness*), tidak bahagia (*sadness*), terkejut (*surprise*), marah (*anger*), takut (*fear*), jijik (*disgust*), dan netral. Jumlah total ekspresi wajah pada dataset ini sejumlah 28709 wajah dengan distribusi tiap emosi.

IV. HASIL DAN PEMBAHASAN

Tahap pengujian dilakukan pada dataset FER 2013 dengan distribusi data dapat dilihat pada Gambar 3. Data ini telah melewati tahapan akuisisi wajah, *pre-processing*, seleksi fitur wajah, dan klasifikasi ekspresi wajah dengan metode KNN, NBC, SVM dan CNN.

A. Hasil

1) Akuisisi Wajah

Proses deteksi wajah adalah tahapan awal agar proses klasifikasi dalam dilakukan. Untuk penelitian ini, algoritma yang diterapkan pada proses akuisisi wajah adalah Haar Cascade Classifier.

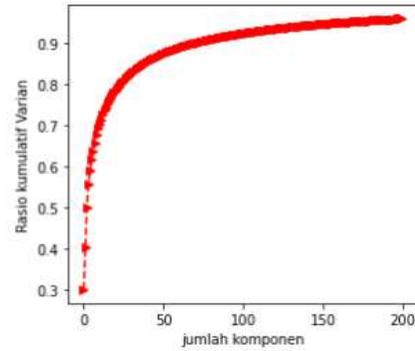
2) Pre-processing Data

Pada tahapan *pre-processing* dilakukan normalisasi data citra wajah untuk menghilangkan serta mengurangi redundansi data. Normalisasi memetakan derajat keabuan citra berdimensi n dengan rentang nilai [min, maks] ke citra dengan rentang nilai [minBaru, maksBaru]. Hasil dari normalisasi ini berupa data array dengan ekstensi file *.npz*. Setelah data dinormalisasi maka kemudian dilanjut proses reduksi dimensi data menggunakan *Principal Component Analysis* (PCA).

3) Ekstraksi Fitur

PCA mengekstraksi fitur wajah yang optimal dengan membentuk suatu arah menurut varian maksimumnya. Algoritma yang digunakan pada PCA ini adalah *Eigenface*, dimana komponen yang berkontribusi dipertahankan. Pemilihan fitur dilakukan sebelum mengekstraksi fitur menggunakan PCA. Proses pemilihan fitur dapat dilihat pada

Gambar 4.



Gambar 4. Seleksi Fitur Wajah Menggunakan *Eigen Face*

Gambar 4 memperlihatkan jumlah komponen yang digunakan pada penelitian ini adalah sebanyak **50 fitur**. Pada citra wajah dapat direpresentasikan sebagai sebuah vektor yang terdiri dari panjang (w) dan lebar (h) piksel. Jumlah komponen dari vektor tersebut adalah $w * h$. Sementara varian mengukur penyebaran nilai komponen yang ada didalam himpunan. Analisis pemilihan fitur berdasarkan jumlah komponen dan jumlah varian tiap data lalu mencari nilai dan vektor eigen matrik kovarians nya. Matriks kovarians ini mengukur kekuatan hubungan antar komponen satu sama lain pada himpunan data.

Pemilihan fitur ini untuk memilih fitur yang akan banyak berpengaruh terhadap nilai n fitur yang ada. Fitur menjadi parameter pembeda atau karakteristik dari wajah. Dataset yang dihasilkan setelah menerapkan PCA adalah sebuah citra *eigenface* dapat dilihat pada Gambar 5.



Gambar 5. Sampel Citra Wajah *Eigenface*

4) Klasifikasi Emosi

Tahapan selanjutnya adalah melakukan proses klasifikasi pada ekspresi wajah. Pada setiap metode klasifikasi, dilakukan perbandingan uji kinerja untuk nilai akurasi, *recall*, *precision*, *f1-score*, dan ROC AUC.

a) Akurasi Metode Klasifikasi

Akurasi setiap metode klasifikasi dilakukan pada dataset

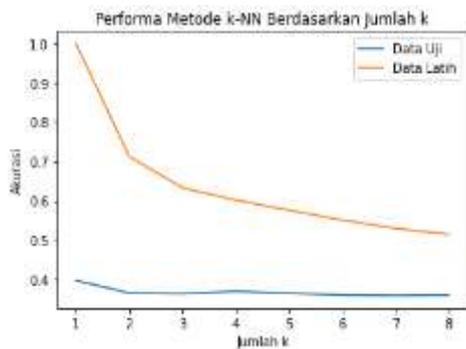
yang sama dengan perbandingan untuk data latih dan data uji adalah 80:20 persen. Rasio perbandingan ini umum digunakan untuk mengetahui kesalahan minimum yang terjadi seperti kondisi *overfitting* maupun *underfitting* pada validasi metode. Hasil akurasi dari setiap metode klasifikasi dapat dilihat pada Tabel II.

TABEL II. AKURASI METODE KLASIFIKASI

Metode	Skor Data Latih (%)	Akurasi Data Uji (%)
KNN	64	37
NBC	34	33
SVM	52	43
CNN	69	62

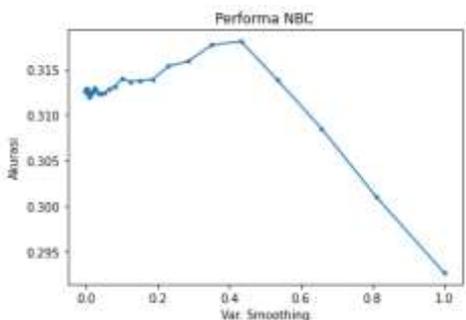
Pada metode KNN didapatkan akurasi data latih sebesar 64% dan data uji sebesar 37%. Pada metode NBC mendapatkan akurasi 34% pada data latih dan 33% pada data uji. Akurasi metode SVM menghasilkan 52% untuk data latih dan 43% untuk data uji. Sedangkan akurasi metode CNN didapatkan 69% untuk data latih dan 62% untuk data uji.

Cross validation dilakukan untuk meningkatkan hasil akurasi pada setiap metode. *Cross validation* yang merupakan evaluasi kinerja metode dimana data dipisahkan menjadi dua subset yaitu data latih dan data uji.



Gambar 6. Cross Validation Metode KNN

Pada Gambar 6 memperlihatkan hasil akurasi metode KNN pada nilai k=1 hingga k=jumlah data, nilai akurasi terbaik didapatkan pada nilai k=1 dengan akurasi 39% pada data uji.

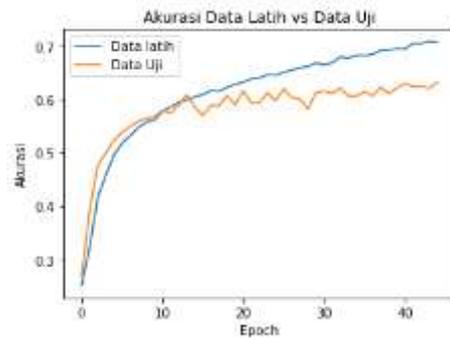


Gambar 7. Distribusi Variabel Smoothing Metode NBC

Pada Gambar 7 menunjukkan hasil klasifikasi metode NBC, dimana nilai akurasi tertinggi adalah 33% pada variabel *smoothing* 0,4. Variabel *smoothing* digunakan untuk

meningkatkan hasil klasifikasi pada NBC sehingga nilai yang dihasilkan lebih besar.

Hasil akurasi pada metode SVM setelah dilakukan *hyper* parameter adalah 99% dengan parameter terbaik adalah C bernilai 1, gamma bernilai 0.1 dan kernel bertipe poly. Uji parameter ini dilakukan dengan memberikan beberapa nilai pada nilai C, nilai gamma dan tipe kernel. C merupakan parameter regularisasi yang bernilai positif, rentang nilai C yang digunakan adalah 1 dan 10. Gamma merupakan nilai coefisien Kernel, pada pengujian ini nilai gamma yang dipakai adalah 0.1. Tipe kernel yang diujikan adalah rbf dan poly.



Gambar 8. Akurasi CNN pada nilai Epoch

Pada Gambar 8 memperlihatkan hasil akurasi pada metode CNN dengan nilai epoch hingga 45 untuk setiap data latih dan data uji. Jumlah layer yang digunakan terdiri dari 32, 64, 128, 512 filter, menggunakan *max pooling* dalam mencari nilai maksimum untuk setiap dimensi vektor. Setelah tahapan convolution dan pooling, maka menghasilkan satu vektor yang kemudian dilewatkan pada *multilayer perceptron (fully connected)* untuk melakukan klasifikasi. Akurasi tertinggi metode CNN adalah 62%.

b) Precision, Recall, F1-Score

Detail hasil *performa precision, recall, dan f1-score* dapat dilihat pada Tabel III. Dalam menguji kinerja metode klasifikasi, nilai *precision* menggambarkan ukuran ketepatan pengklasifikasian untuk setiap kelas, sedangkan *recall* merupakan ukuran kelengkapan metode klasifikasi, dan *f1-score* menunjukkan perbandingan rata-rata *precision* dan *recall* yang dibobotkan.

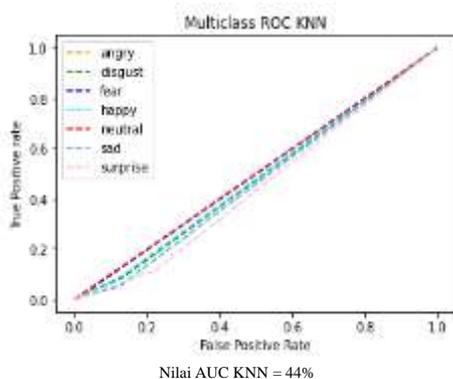
TABEL III. UJI KINERJA METODE KLASIFIKASI

Metode	Variabel	Precision (%)	Recall (%)	F1-Score (%)
KNN	<i>angry</i>	46	45	45
	<i>disgust</i>	34	31	32
	<i>fear</i>	36	41	38
	<i>happy</i>	38	35	36
	<i>neutral</i>	32	31	32
	<i>sad</i>	56	58	57
	<i>surprise</i>	32	38	35
	<i>macro avg</i>	39	40	39
	<i>weighted avg</i>	39	39	39
NBC	<i>angry</i>	38	66	48
	<i>disgust</i>	22	9	13
	<i>fear</i>	5	2	3

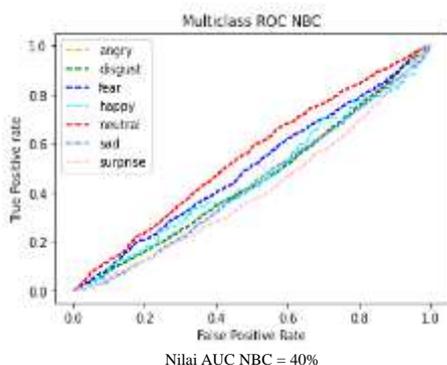
	<i>happy</i>	20	14	17
	<i>neutral</i>	27	24	25
	<i>sad</i>	45	40	42
	<i>surprise</i>	30	27	28
	<i>macro avg</i>	27	26	25
	<i>weighted avg</i>	30	33	30
SVM	<i>angry</i>	45	75	56
	<i>disgust</i>	36	16	22
	<i>fear</i>	100	3	6
	<i>happy</i>	36	18	24
	<i>neutral</i>	34	36	35
	<i>sad</i>	58	48	53
	<i>surprise</i>	37	38	38
	<i>macro avg</i>	49	33	33
	<i>weighted avg</i>	41	42	39
	CNN	<i>angry</i>	16	16
<i>disgust</i>		0	0	0
<i>fear</i>		14	9	11
<i>happy</i>		27	24	25
<i>neutral</i>		18	22	20
<i>sad</i>		17	22	19
<i>surprise</i>		13	13	13
<i>macro avg</i>		15	15	15
<i>weighted avg</i>		18	18	18

c) ROC dan AUC

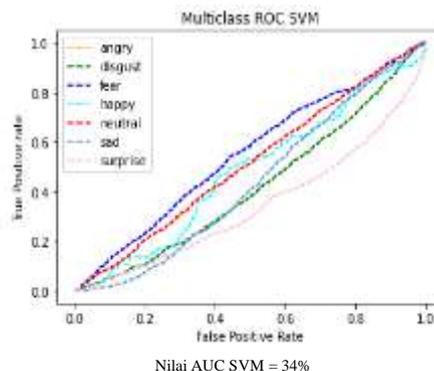
Hasil visualisasi kurva ROC dan nilai AUC untuk metode KNN, NBC, SVM dan CNN dapat dilihat pada Gambar 9.a hingga 9.d. *Receiving Operation Characteristic* (ROC) merupakan kurva probabilitas sedangkan *Area Under Curve* (AUC) mewakili derajat/ukuran keterpisahan. Ini memberitahukan seberapa besar metode mampu membedakan antar kelas.



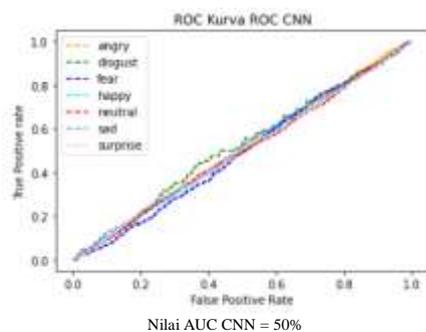
Gambar 9.a. Kurva ROC Metode KNN



Gambar 9.b. Kurva ROC Metode NBC



Gambar 9.c. Kurva ROC Metode SVM



Gambar 9.d. Kurva ROC Metode CNN

B. Pembahasan

Akurasi pada setiap metode memperoleh nilai yang bervariasi. Hasil dari akurasi didapatkan dari nilai akurasi data training dan data testing dari dataset FER-2013. Analisis berdasarkan nilai akurasi ini adalah terdapatnya kondisi overfitting pada metode KNN yang cukup signifikan dimana akurasi data training lebih besar daripada akurasi data testing. Sementara untuk metode NBC, SVM, dan CNN tidak memiliki perbedaan akurasi yang signifikan. Uji validasi akurasi terbaik dilakukan *cross validation* untuk setiap metode. Maka didapatkan nilai akurasi terbaik untuk KNN = 39% dan SVM = 99%, untuk NBC dan CNN tidak memiliki perbedaan akurasi setelah *cross validation*.

Berdasarkan hasil yang diperoleh dari nilai *precision*, *recall*, dan *f1-score* untuk setiap klasifikasi emosi menghasilkan nilai *macro avg* dan *weighted avg*. Penelitian ini menghasilkan nilai yang berbeda pada setiap metode klasifikasi untuk *precision*, *recall*, dan *f1-score*. Nilai *macro avg* dan *weighted avg* juga dihitung untuk tiap metode klasifikasi.

Berdasarkan kurva ROC yang dihasilkan oleh metode KNN untuk klasifikasi ketujuh emosi wajah, memperlihatkan bahwa semua nilai berada dibawah baseline yang mendekati keadaan *false positive rate* dengan nilai AUC = 44%, sedangkan metode NBC, SVM, CNN menunjukkan kurva ROC yang bervariasi untuk setiap emosi yang diklasifikasikan, dimana nilai AUC NBC = 40%, AUC SVM = 34%, dan AUC CNN = 50%.

V. KESIMPULAN

Pada penelitian ini membandingkan metode klasifikasi emosi wajah: KNN, NBC, SVM, dan CNN pada dataset FER-2013 telah melewati tahapan preprocessing, ekstraksi fitur dan klasifikasi. Penelitian ini juga menerapkan algoritma PCA dan Eigenface dalam menentukan fitur wajah. Analisis kinerja metode-metode klasifikasi berdasarkan nilai akurasi, *precision*, *recall*, *f1-score*, dan *AUC*. Akurasi tertinggi dihasilkan oleh metode SVM sebesar 99% setelah dilakukan *cross validation* dalam mengklasifikasikan emosi seperti bahagia (*happiness*), tidak bahagia (*sadness*), terkejut (*surprise*), marah (*anger*), takut (*fear*), jijik (*disgust*), dan netral. Masing-masing kinerja metode klasifikasi memiliki karakteristik nilai optimal yang berbeda-beda. *Cross validation* dilakukan untuk menganalisis setiap metode agar dapat meningkatkan akurasi terhadap klasifikasi dengan membagi data menjadi data latih dan data uji. ROC dan AUC dilakukan untuk mengukur kinerja metode klasifikasi pada berbagai ambang batas (*threshold*), semakin tinggi AUC, semakin baik model dalam memprediksi. Nilai AUC tertinggi dihasilkan oleh metode CNN sebesar 50%, kemudian metode KNN sebesar 44%, metode NBC sebesar 40% dan nilai AUC terendah dihasilkan oleh metode SVM sebesar 34%. Meskipun hasil analisis FER untuk citra wajah telah diperoleh berdasarkan kinerja masing-masing metode, penelitian selanjutnya diharapkan dapat menggabungkan dengan metode lainnya yang saling terintegrasi sehingga dapat memberikan informasi tambahan tentang kehandalan metode. Kemudian untuk kompresi data patut dipertimbangkan agar informasi tentang fitur wajah yang tidak relevan dapat dikurangi agar memperoleh akurasi yang optimal.

DAFTAR PUSTAKA

- [1] D. Canedo and A. J. R. Neves, "Facial Expression Recognition Using Computer Vision: A Systematic Review," *Appl. Sci.*, vol. 9, no. 21, p. 4678, Nov. 2019, doi: 10.3390/app9214678.
- [2] H. Y. Cao and C. Qi, "Facial Expression Study Based on 3D Facial Emotion Recognition," in *2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS)*, Dec. 2021, pp. 375–381. doi: 10.1109/IUCC-CIT-DSCI-SmartCNS55181.2021.00067.
- [3] Y. Huang, F. Chen, S. Lv, and X. Wang, "Facial Expression Recognition: A Survey," *Symmetry (Basel)*, vol. 11, no. 10, p. 1189, Sep. 2019, doi: 10.3390/sym11101189.
- [4] H. Azis, F. Tangguh Admojo, and E. Susanti, "Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah," *Techno.Com*, vol. 19, no. 3, pp. 286–294, 2020, doi: 10.33633/tc.v19i3.3646.
- [5] V. Cherepanova, S. Reich, S. Dooley, H. Souri, M. Goldblum, and T. Goldstein, "A Deep Dive into Dataset Imbalance and Bias in Face Identification," Mar. 2022, [Online]. Available: <http://arxiv.org/abs/2203.08235>
- [6] Z. Song, "Facial Expression Emotion Recognition Model Integrating Philosophy and Machine Learning Theory," *Front. Psychol.*, vol. 12, no. September, Sep. 2021, doi: 10.3389/fpsyg.2021.759485.
- [7] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, "Emotion recognition using facial expressions Emotion recognition using facial expressions," *Procedia Comput. Sci.*, vol. 108, pp. 1175–1184, 2017, doi: 10.1016/j.procs.2017.05.025.
- [8] P. Utami, R. Hartanto, and I. Soesanti, "A Study on Facial Expression Recognition in Assessing Teaching Skills: Datasets and Methods," *Procedia Comput. Sci.*, vol. 161, pp. 544–552, 2019, doi: 10.1016/j.procs.2019.11.154.
- [9] J. Erik Solem, *Programming Computer Vision with Python*. O'Reilly, 2012.
- [10] J. Wira Gotama Putra, "Pengenalan Konsep Pembelajaran Mesin dan Deep Learning," 2020.
- [11] I. M. Revina and W. R. S. Emmanuel, "A Survey on Human Face Expression Recognition Techniques," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 6, pp. 619–628, Jul. 2021, doi: 10.1016/j.jksuci.2018.09.002.
- [12] S. Bayrakdar, D. Akgün, and İ. Yücedağ, "A Survey on Automatic Analysis of Facial Expressions," *Res. Artic.*, vol. 20, no. 2, p. 383, Dec. 2016, doi: 10.16984/saufenbilder.92940.
- [13] F. Fadlisyah, *Computer Vision dan Pengolahan Citra*. ANDI Yogyakarta, 2007.
- [14] Y.-L. Tian, T. Kanade, and J. F. Cohn, "Facial Expression Analysis," in *Handbook of Face Recognition*, New York: Springer-Verlag, 2005, pp. 247–275. doi: 10.1007/0-387-27257-7_12.
- [15] B. K. Durga and D. V Rajesh, "Review of Facial Emotion Recognition System," *Int. J. Pharm. Res.*, vol. 10, no. 03, Jul. 2018, doi: 10.31838/ijpr/2018.10.03.056.
- [16] "Introduction to OpenCV." https://docs.opencv.org/4.x/da/df6/tutorial_py_table_of_contents_setup.html