

Model *Convolutional Neural Network* untuk Mengukur Kepuasan Pelanggan Berdasarkan Ekspresi Wajah

<http://dx.doi.org/10.28932/jutisi.v8i3.5493>

Riwayat Artikel

Received: 1 Oktober 2022 | Final Revision: 21 Oktober 2022 | Accepted: 3 November 2022

Creative Commons License 4.0 (CC BY – NC)



Daru Prasetyawan^{✉#1}, Rahmadhan Gatra^{#2}

[#] Pusat Teknologi Informasi dan Pangkalan Data, UIN Sunan Kalijaga Yogyakarta
Jl. Laksda Adisucipto Yogyakarta, 55281, Yogyakarta, Indonesia

¹daru.prasetyawan@uin-suka.ac.id

²rahmadhan.gatra@uin-suka.ac.id

[✉]Corresponding author: daru.prasetyawan@uin-suka.ac.id

Abstrak — Kepuasan pelanggan menunjukkan seberapa baik produk atau layanan suatu organisasi secara keseluruhan dalam memenuhi harapan pelanggan. Ekspresi wajah pelanggan dapat menunjukkan kepuasan mereka terhadap layanan yang diberikan. Convolution Neural Network (CNN) adalah jenis algoritma neural network yang dapat digunakan untuk mengenali suatu objek di dalam sebuah gambar. CNN memanfaatkan proses konvolusi untuk menentukan dan membedakan suatu objek di dalam gambar dengan objek lainnya seperti untuk mengenali berbagai ekspresi wajah. Penelitian ini bertujuan untuk mengukur kepuasan pelanggan dengan memanfaatkan model CNN dengan mengenali setiap perubahan ekspresi wajah. Dari hasil pelatihan model CNN diperoleh akurasi sebesar 90,57%. Selanjutnya model yang terbentuk diimplementasikan ke dalam sebuah sistem berbasis web yang melakukan perekaman ekspresi wajah dan melakukan klasifikasi (puas atau tidak puas) terhadap setiap perubahan wajah yang terdeteksi. Ekspresi yang paling dominan merupakan hasil dari pengukuran kepuasan pelanggan.

Kata kunci— *Convolutional Neural Network*; Ekspresi Wajah; Kepuasan Pelanggan.

Convolutional Neural Network Model for Measuring Customer Satisfaction Based on Facial Expressions

Abstract — Customer satisfaction shows how well the product or service of an organization meets customer expectations. Customers' facial expressions can show their satisfaction with the services provided. Convolution Neural Network (CNN) is a type of neural network algorithm that can be used to recognize an object in an image. CNN utilizes the convolution process to determine and distinguish an object in the image from other objects such as to recognize various facial expressions. This study aims to measure customer satisfaction by utilizing the CNN model by recognizing any changes in facial expressions. From the results of the CNN model training, an accuracy of 90.57% was obtained. Furthermore, the formed model is implemented into a web-based system that records facial expressions and performs a classification (satisfied or dissatisfied) on any detected facial changes. The most dominant expression is the result of measuring customer satisfaction.

Keywords— *Convolutional Neural Network*; Customer Satisfaction, Facial Expressions.

I. PENDAHULUAN

Salah satu cara untuk mengukur kualitas layanan suatu organisasi adalah kepuasan pelanggan (*customer*). Kepuasan pelanggan adalah salah satu upaya untuk memenangkan persaingan [1]. Kepuasan pelanggan mewakili sejauh mana harapan pelanggan terhadap produk atau layanan terpenuhi atau dilampaui [2]. Untuk itu, banyak pelaku bisnis yang fokus untuk meningkatkan kepuasan pelanggan. Jika seorang pelanggan merasa puas, mereka akan kembali berkali-kali atau yang biasa disebut dengan pelanggan tetap. Kepuasan pelanggan akan berpengaruh terhadap loyalitas pelanggan tersebut. Hal ini dianggap lebih penting oleh suatu organisasi dibanding mencari pelanggan baru. Loyalitas dari pelanggan sangat diperlukan oleh sebuah organisasi, karena mereka akan menciptakan dampak positif bagi organisasi di masa depan. Selain itu pelanggan yang setia dapat menjadi sarana promosi dengan melakukan rekomendasi kepada orang lain. Menurut Kotler, ada empat cara atau metode untuk mengukur kepuasan pelanggan [3]. Pertama, melalui sistem saran dan keluhan, yaitu dengan menyediakan kotak keluhan dan saran, atau dengan menghubungi pelanggan melalui telepon. Metode ini memiliki beberapa kelemahan, diantaranya adalah keluhan dan saran akan kurang bisa dimengerti oleh organisasi atau pelanggan malas untuk menyampaikan keluhan dan sarannya. Kedua, melalui *ghost shopping*, yaitu salah satu strategi untuk mengetahui dan mendapatkan informasi yang terkait dengan kepuasan pelanggan terhadap kualitas layanan yang diberikan kepada konsumen dengan menugaskan beberapa orang (*ghost shopper*) untuk berperilaku dan bersikap seperti seorang pelanggan. Metode ini memiliki kelemahan karena untuk mendapatkan informasi yang cukup perlu dilakukan berkali-kali dalam beberapa kondisi yang berbeda. Ketiga, melalui *lost customer analysis*, yaitu dengan menghubungi secara langsung pelanggan yang hilang. Kelemahannya adalah suatu organisasi harus menghubungi pelanggan satu per satu, dan tentunya akan mengeluarkan biaya yang cukup banyak. Keempat melalui survey kepuasan. Kelemahannya adalah selain pelanggan malas mengisi survey yang diberikan, pelanggan juga berpotensi menjawab survey dengan tidak jujur.

Penelitian tentang pengembangan sistem kepuasan pelanggan sudah banyak dilakukan di berbagai bidang dan dengan berbagai metode. Di lingkungan pendidikan, pengembangan sistem informasi untuk mengukur kepuasan alumni mengenai layanan yang diberikan oleh sebuah lembaga pendidikan [4]. Metode pengukuran kepuasan yang digunakan adalah *Importance Performance Analysis* dan diujikan kepada alumni dengan mengisi kuesioner tingkat kepuasan. Sistem tersebut akan menunjukkan persentase kepuasan dan prioritas variabel kuesioner yang harus diperbaiki untuk meningkatkan pelayanan kepada pelanggan atau alumni. Penelitian serupa juga pernah dilakukan di lingkungan pemerintahan. Penelitian tersebut menghasilkan aplikasi yang dapat digunakan untuk mengukur kepuasan pelanggan dengan membuat kuesioner yang akan diisi oleh pelanggan [5].

Saat ini perkembangan teknologi industri telah sampai pada revolusi industri 4.0, dimana sebagian besar pekerjaan yang sebelumnya dilakukan secara manual digantikan dengan teknologi siber dan teknologi otomatisasi (termasuk di dalamnya adalah kecerdasan buatan dan pembelajaran mesin). Salah satu pemanfaatan kecerdasan buatan dan pembelajaran mesin adalah pengenalan wajah. Kecerdasan buatan dan pembelajaran mesin merupakan alat paling canggih yang dapat digunakan untuk memecahkan dan mengoptimalkan berbagai masalah, dan juga sangat efektif dalam menentukan hubungan kompleks antara beberapa variabel [6]. Pembelajaran mesin adalah jenis kecerdasan buatan yang memungkinkan mesin atau komputer dapat memprediksi suatu hasil tertentu tanpa harus diprogram secara eksplisit. Algoritma pembelajaran mesin bekerja dengan memanfaatkan data historis sebagai input dalam memprediksi nilai output. Pembelajaran mesin meliputi perancangan dan pengembangan algoritma yang membuat sebuah mesin dapat mengembangkan kemampuan untuk belajar secara otomatis berdasarkan data empiris dari pengalaman sebelumnya [7]. Pembelajaran mesin mengidentifikasi pola untuk membuat keputusan melalui proses pembelajaran menggunakan data sebelumnya dengan intervensi manusia yang sangat minimal. Proses pembelajaran tersebut adalah upaya untuk memperoleh kecerdasan melalui tahap pelatihan dan tahap pengujian [8] [9]. Proses pembelajaran dimulai dari mengamati kumpulan data, lalu melatih mesin dengan mengembangkan model menggunakan kumpulan data dan algoritma tertentu untuk membentuk sebuah pola dan menghasilkan keputusan yang lebih baik dengan tujuan agar sebuah mesin dapat belajar secara otomatis tanpa intervensi manusia. Pembelajaran mesin sangat berbeda dengan pemrograman biasa atau biasa dikenal dengan pemrograman tradisional. Dalam pemrograman tradisional, sebuah program dibuat menggunakan suatu algoritma tertentu dan diberikan beberapa input data, sehingga menghasilkan output berdasarkan fungsi-fungsi di dalam algoritma tersebut. Sedangkan di dalam pembelajaran mesin, sebuah mesin melakukan pelatihan dengan menjalankan data, kemudian mesin tersebut membuat programnya sendiri dan dievaluasi melalui tahap pengujian.

Algoritma pembelajaran mesin terbagi menjadi tiga, yaitu algoritma *supervised learning* atau pembelajaran terawasi, *unsupervised learning* atau pembelajaran tak terawasi, dan *reinforcement learning* atau pembelajaran penguatan. *Supervised learning* menggunakan data berlabel, artinya algoritma tersebut mengidentifikasi fitur dan melakukan prediksi atau klasifikasi secara eksplisit. Algoritma *supervised learning* dapat mempelajari dan mengidentifikasi hubungan antar variabel sehingga dapat menghasilkan hasil yang baru apabila menggunakan data yang baru. *Supervised learning* menggunakan model untuk mempelajari pola pemetaan antara variabel input dengan variabel target. Ada dua jenis *supervised learning*, yaitu regresi dan klasifikasi. Regresi digunakan untuk memprediksi nilai numerik dengan mengukur hubungan variabel *dependen* dengan satu atau lebih variabel *independen*. Regresi dapat juga digunakan untuk mengukur kekuatan hubungan antara sebuah variabel

tertentu dengan perkiraan masa depan. Klasifikasi adalah teknik pembelajaran mesin yang digunakan untuk mengkategorikan data ke dalam sejumlah kelas tertentu.

Klasifikasi adalah klasifikasi mengacu pada masalah pemodelan prediktif dimana model akan memprediksi label kelas dari data baru yang diberikan. Klasifikasi merupakan proses menemukan pola untuk memprediksi kelas dari data yang belum memiliki [10]. Algoritma klasifikasi terkenal yaitu *K-Nearest Neighbor* (K-NN), pohon keputusan (*decision tree*), dan jaringan syaraf tiruan (*Artificial Neural Network*). Jaringan Syaraf Tiruan atau sering disebut *neural network* merupakan teknik komputasi yang didasari pada jaringan syaraf biologis yang terdapat di dalam otak manusia. Elemen pemrosesan di dalam jaringan syaraf tiruan yang dikenal dengan *neuron* terhubung ke elemen pemrosesan lainnya. *Neuron-neuron* tersebut diatur dalam lapisan atau vektor, dengan *output* dari satu lapisan berfungsi sebagai *input* ke lapisan berikutnya. *Neuron* pada lapisan tertentu dapat dihubungkan ke semua atau sebagian dari *neuron* pada lapisan berikutnya. Jaringan syaraf tiruan (JST) terdiri dari lapisan *input*, *hidden*, dan *output* dengan *neuron* yang saling terhubung untuk mensimulasikan otak manusia. Seperti jaringan syaraf pada otak manusia, JST juga mempunyai kemampuan untuk beradaptasi terhadap input yang diberikan melalui proses belajar.

Pembelajaran mesin yang lebih tradisional memiliki kemampuan yang terbatas dalam memproses data alami yang masih dalam bentuk mentah [11]. Pembelajaran mendalam (*deep learning*) hadir untuk mengatasi permasalahan tersebut dengan kemampuannya untuk melakukan ekstraksi fitur dan proses pembelajarannya secara bersamaan. Landasan teoritis pembelajaran mendalam berakar dari jaringan syaraf (*neural network*) klasik, tetapi pembelajaran mendalam menggunakan banyak *neuron* tersembunyi dan banyak lapisan yang dikombinasikan dengan paradigma pelatihan baru [12]. Pembelajaran mendalam secara otomatis dapat mempelajari dan meningkatkan kemampuannya sendiri dengan mengevaluasi algoritma yang digunakan. Pembelajaran mendalam memiliki struktur jaringan yang terdiri dari lapisan-lapisan tersembunyi dalam jumlah ratusan atau bahkan ribuan. Pembelajaran mendalam menggunakan banyak lapisan di antara lapisan *input* dan lapisan *output* [13]. Setiap lapisan tersembunyi bertugas melatih sekumpulan fitur berdasarkan *output* dari lapisan sebelumnya. Sistem pembelajaran mendalam membutuhkan perangkat keras yang kuat karena memerlukan jumlah data yang banyak untuk diproses dan melibatkan beberapa perhitungan matematis yang kompleks. Jaringan dikatakan lebih dalam berdasarkan jumlah lapisan yang dimilikinya. Struktur jaringan di dalam pembelajaran mendalam terdiri dari banyak *neuron* yang saling berhubungan seperti *neuron* di dalam otak manusia. Setiap *neuron* menerima sejumlah sinyal *input* akan menghasilkan sinyal output tunggal yang akan ditransmisikan ke banyak cabang, dan berakhir sebagai input bagi *neuron* lain dalam jaringan [14].

Salah satu algoritma *deep learning* yang terkenal adalah jaringan konvolusional atau *Convolutional Neural Network (CNN)*. CNN menjadi sebuah inovasi baru yang digunakan untuk memproses gambar, video, ucapan, dan audio [11]. Jaringan syaraf konvolusional merupakan jenis jaringan syaraf *multilayer* atau arsitektur pembelajaran mendalam yang terinspirasi dari sistem visual makhluk hidup, dan sangat cocok digunakan di dalam bidang visi komputer dan pemrosesan bahasa alami [15]. Di dalam CNN ekstraksi fitur dan reduksi dimensi dapat dilakukan sekaligus di dalam jaringan sehingga dapat mengurangi *preprocessing*. Lapisan konvolusi di dalam CNN yang bertanggung jawab untuk mengekstraksi dan mempelajari fitur. Sedangkan lapisan jaringan syaraf biasa di dalam CNN bertugas melakukan klasifikasi. Sebelum CNN banyak digunakan, pengenalan wajah dilakukan melalui tahapan-tahapan seperti ekstraksi fitur, memperkecil dimensi citra, dan klasifikasi dilakukan secara terpisah [16]. Hal ini menyebabkan kesesuaian algoritma yang digunakan pada setiap tahapan sangat diperlukan. Kinerja sistem klasifikasi seperti ini juga dipengaruhi oleh masing-masing algoritma yang dipilih. Untuk itu diperlukan sebuah ketelitian dalam memilih algoritma yang digunakan pada setiap tahapan. Algoritma *Support Vector machine (SVM)* sering digunakan untuk pengenalan wajah. SVM hanya digunakan sebagai algoritma pengklasifikasi, untuk ekstraksi fitur (seperti mulut dan mata) dilakukan dengan menggunakan algoritma *Gabor Features dan Local Binary Pattern (LBP)*, serta untuk mereduksi dimensi fitur menggunakan *Principal Component Analysis (PCA)* [17].

CNN adalah jenis algoritma pembelajaran mesin untuk mendeteksi dan mengenali suatu objek di dalam sebuah gambar. Untuk mengenali ekspresi wajah seseorang, CNN menggunakan proses konvolusi dengan memindahkan kernel konvolusi (*filter*) dengan ukuran tertentu pada sebuah gambar wajah seseorang. CNN banyak digunakan dalam analisis dan klasifikasi wajah seperti klasifikasi jenis kelamin, usia, dan ekspresi. Untuk klasifikasi jenis kelamin, CNN memiliki kinerja yang sangat baik dengan mendapatkan akurasi sebesar 97% [18]. Dengan arsitektur yang berbeda, CNN dapat memperoleh akurasi hingga 98,75% [16]. Untuk klasifikasi rentang usia berdasarkan wajah seseorang, dengan *multi-class classification* yang terdiri dari 8 kelas, CNN memperoleh akurasi 59,90% [19]. CNN juga digunakan untuk klasifikasi ekspresi wajah. Akurasi model yang didapatkan untuk klasifikasi ekspresi wajah dengan *multi-class classification* sebanyak 7 kelas adalah 60,12% [20]. Dengan menggabungkan CNN dan *Scale Invariant Feature Transform (SIFT)*, model yang dihasilkan mampu memperoleh akurasi hingga 73,4% pada dataset FER-2013 dan 91,1% pada dataset CK+ [21].

Penelitian ini mencoba membuat model CCN untuk mengukur kepuasan pelanggan berdasarkan ekspresi wajah terhadap layanan yang diberikan. Pengenalan ekspresi wajah dapat dimanfaatkan untuk mengukur kualitas layanan yang diberikan oleh suatu organisasi kepada konsumen. Penelitian ini akan mencoba mengembangkan sistem untuk mengukur kepuasan pelanggan berdasarkan ekspresi wajah dengan memanfaatkan teknologi pembelajaran mesin (*machine learning*) untuk mengenali setiap perubahan ekspresi wajah pelanggan. Algoritma yang digunakan adalah *Convolutional Neural Network*

(CNN), yaitu sebuah algoritma klasifikasi yang sering digunakan untuk pengenalan wajah (*face recognition*). Algoritma ini dipilih karena saat ini sedang banyak yang mengembangkan algoritma tersebut, baik dalam implementasi maupun pengembangan kinerja algoritma tersebut.

II. METODE PENELITIAN

A. Data dan Metode Pengumpulan Data

Metode pengumpulan data yang digunakan adalah dengan studi pustaka dan observasi. Langkah awal dalam pengumpulan data adalah dengan melakukan pencarian data dan informasi yang bersumber dari dokumen berupa buku, ensiklopedia, internet, dan penelitian-penelitian sebelumnya. Studi pustaka dilakukan dengan tujuan untuk memperjelas hubungan antara masalah-masalah yang diteliti dengan teori-teori sebelumnya. Observasi merupakan metode pengumpulan data untuk memperoleh informasi dari suatu objek penelitian dan berbagai hal yang terkait dengan objek penelitian tersebut melalui pengamatan secara langsung. Pengamatan dilakukan terhadap perilaku dan ekspresi pelanggan pada saat menyampaikan keluhan mereka.

Data yang dikumpulkan berbentuk citra foto dan video yang menunjukkan ekspresi wajah. Data citra tersebut diperoleh dari pengamatan dan pengambilan foto dan video secara langsung yang dilakukan di Bagian Layanan Pusat Teknologi Informasi dan Pangkalan Data UIN Sunan Kalijaga Yogyakarta. Selain itu data citra wajah juga diperoleh dari internet berupa foto dan video yang menunjukkan kepuasan dan ketidakpuasan seorang pelanggan terhadap sebuah layanan yang diberikan. Data mentah hasil dari pengambilan gambar di lapangan yang akan digunakan dalam pemodelan disajikan pada Gambar 1.



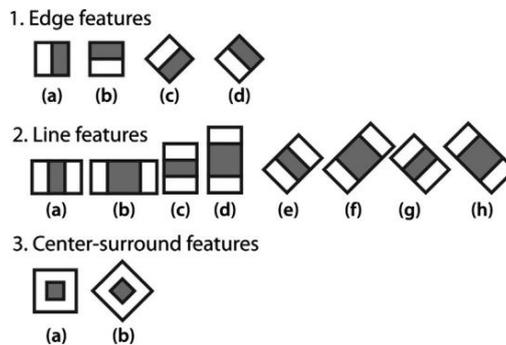
Gambar 1. Gambar mentah hasil pemotretan di lapangan.

Selanjutnya data citra tersebut diberi label “puas” dan “tidak puas”. Label “puas” merujuk pada ekspresi positif yang ditunjukkan dengan ekspresi wajah bahagia atau tersenyum, dan label “tidak puas” merujuk pada ekspresi negatif yang ditunjukkan dengan ekspresi marah, sedih, takut, dan terkejut [22]. Data set yang diperoleh masih dalam bentuk data mentah (*raw data*), sehingga masih harus dilakukan pra pemrosesan agar data tersebut siap digunakan dalam pemodelan. Selanjutnya data diperbanyak dengan teknik augmentasi data, yaitu sebuah teknik untuk meningkatkan jumlah dan keragaman data tanpa mengumpulkan data baru [23] Kemudian data yang akan digunakan untuk pemodelan dibagi menjadi tiga, yaitu data latih (*training set*), data validasi (*validation set*), dan data uji (*testing set*). *Training set* atau data latih diperlukan untuk melatih algoritma *backpropagation* dalam menemukan model yang sesuai. Data validasi digunakan untuk menemukan jumlah unit tersembunyi yang optimal yang akan menentukan titik henti dari algoritma *back-propagation*. Data uji diperlukan untuk membandingkan performa model prediksi yang didapatkan selama proses pelatihan.

B. Pra Pemrosesan Data

Tahap pra pemrosesan data merupakan suatu tahapan untuk mempersiapkan dan membersihkan data mentah yang telah dikumpulkan menjadi data yang siap untuk digunakan di dalam proses selanjutnya. Pra pemrosesan data yang dilakukan antara lain pendeteksian dan *cropping* wajah, penyesuaian ukuran dimensi citra, *grayscale*, *data splitting*. Tahap *grayscale* akan mengubah citra berwarna (RGB) menjadi citra *grayscale* atau citra yang hanya memiliki satu nilai (antara 0 – 255) pada setiap piksel. Citra *grayscale* digunakan karena citra *grayscale* hanya memiliki *channel* tunggal sehingga akan lebih mudah diproses dibandingkan dengan citra berwarna yang memiliki tiga *channel* warna (*red, green, blue*).

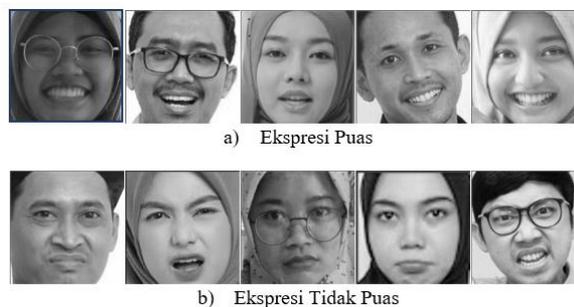
Pendeteksian dan *cropping* wajah dilakukan untuk memperoleh citra wajah. Proses ini menggunakan *Haar Cascade Classifier* [24] diperkenalkan oleh Viola dan Jones pada tahun 2001. Viola dan Jones menggunakan kombinasi fitur *haar-wavelet like* sederhana yang ditunjukkan pada Gambar 2 untuk menganalisis setiap bagian sebuah citra dan mengklasifikasikannya dalam dua kelas, yaitu objek atau bukan objek [25].



Gambar 2. *Haar-like features* yang digunakan oleh Viola dan Jones untuk mendeteksi objek [25][26].

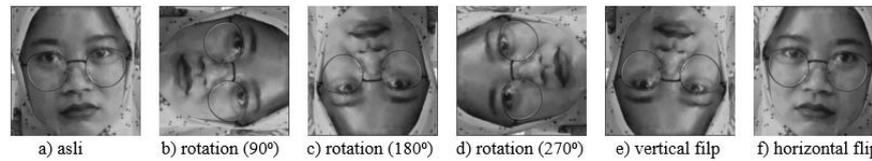
Sebutan *Haar* mengacu pada fungsi matematika yang memiliki bentuk persegi bernama *Haar Wavelet* [21]. Kelebihan algoritma *Haar Cascade Classifier* adalah dapat melakukan komputasi dengan cepat karena hanya melakukan komputasi terhadap piksel yang ada di dalam persegi pada sebuah citra [26]. OpenCV menyediakan banyak pengklasifikasi pra-terlatih untuk mendeteksi wajah (seperti mendeteksi wajah seutuhnya, mata, mulut, senyuman, dan sebagainya) yang disimpan dalam bentuk file XML. Dalam penelitian ini jenis pengklasifikasi yang digunakan adalah *frontal face (haarcascade_frontalface_alt2.xml)*. Untuk menggunakan pengklasifikasi tersebut, OpenCV menyediakan fungsi *cv2.CascadeClassifier()*. Untuk menemukan objek di dalam sebuah citra, OpenCV menyediakan fungsi *cv2.CascadeClassifier.detectMultiScale()*. Parameter penting yang diperlukan fungsi tersebut adalah *image*, *scaleFactor*, dan *minNeighbors*. Parameter *image* merupakan citra dalam bentuk matriks yang berisi objek yang akan dideteksi. Parameter *scaleFactor* adalah parameter yang menentukan seberapa besar ukuran citra diperkecil pada setiap skala citra. Sedangkan parameter *minNeighbors* adalah parameter yang menentukan seberapa banyak jumlah tetangga kandidat kotak yang harus dipertahankan. Parameter *minNeighbors* akan berpengaruh terhadap kualitas objek yang terdeteksi. Semakin tinggi nilai *minNeighbors* mengakibatkan objek yang terdeteksi semakin sedikit tetapi memiliki kualitas yang tinggi [26]. Parameter tersebut seperti nilai ambang batas bawah respon yang dihasilkan, suatu objek akan dianggap sebagai wajah yang valid apabila memiliki respon lebih dari yang ditentukan. Pada penelitian ini, nilai *minNeighbors* yang digunakan adalah 4. Nilai tersebut ditentukan setelah melakukan percobaan dengan nilai yang berbeda. Jika nilai lebih rendah ditetapkan, banyak objek yang sebenarnya bukan wajah (seperti lekukan baju, bros, dan motif pada pakaian) terkadang terdeteksi sebagai wajah. Akan tetapi, jika nilai yang ditetapkan lebih tinggi mengakibatkan objek wajah akan sulit ditemukan di dalam sebuah citra yang memiliki objek wajah yang kurang jelas.

Jika objek wajah terdeteksi di dalam sebuah citra, fungsi *cv2.CascadeClassifier.detectMultiScale()* akan mengembalikan posisi dan ukuran objek wajah tersebut sebagai objek $Rect(x,y,w,y)$, dimana x,y merupakan koordinat objek tersebut berada, dan w,h menunjukkan lebar dan tinggi objek tersebut. Dari objek yang terdeteksi tersebut, dapat dibentuk sebuah citra baru yang hanya berisi wajah secara utuh. Contoh hasil deteksi dan *cropping* wajah disajikan pada Gambar 3.



Gambar 3. Contoh hasil deteksi dan *cropping* wajah disajikan

Di dalam penelitian ini, dimensi citra wajah perlu diubah menjadi ukuran 128 x 128 pixel agar semua data yang digunakan untuk pemodelan memiliki dimensi yang sama. Teknik augmentasi data digunakan untuk memperbanyak jumlah dan variasi data. Augmentasi yang digunakan dalam penelitian ini antara lain *rotation* (90°, 180°, dan 270°) dan *flip* (*vertical* dan *horizontal*).



Gambar 4. Hasil augmentasi data

Setelah dilakukan pra pemrosesan diperoleh data sebanyak 5400 citra wajah, 4000 data digunakan melatih model dan 1400 digunakan sebagai data uji.

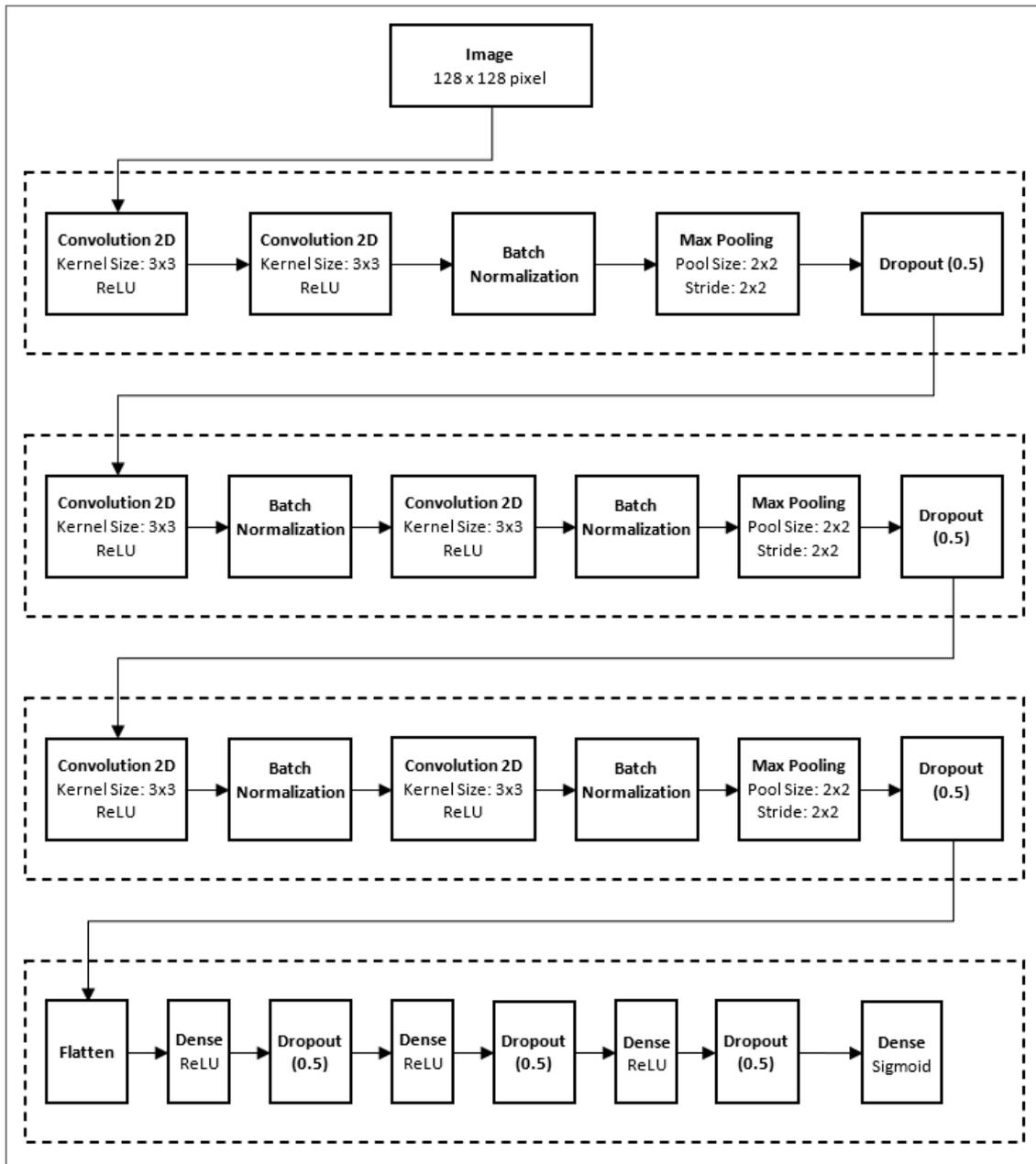
C. Pemodelan

Arsitektur jaringan yang dibangun berisi beberapa lapisan, antara lain lapisan konvolusi (*convolution layers*), lapisan *batch normalization*, lapisan *pooling*, lapisan *dropout*, *flatten*, dan lapisan *dense* yang terbagi menjadi blok ekstraksi fitur dan blok klasifikasi. Lapisan konvolusi menjadi dasar dari arsitektur yang digunakan dalam membuat model CNN. Lapisan konvolusi menerima input berupa matriks dua dimensi kemudian akan melakukan operasi konvolusi dengan kernel yang sudah ditentukan. Lapisan tersebut menggunakan fungsi aktivasi ReLU untuk mendefinisikan nilai output dari setiap *neuron*. Pemodelan dilakukan dengan memanfaatkan pustaka *TensorFlow* yang memiliki berbagai fungsi untuk analisis citra, audio, dan video. Untuk operasi konvolusi terhadap citra, *TensorFlow* menyediakan fungsi *tf.keras.layers.Conv2D*. Fungsi tersebut membuat kernel konvolusi yang dililitkan dengan input lapisan untuk menghasilkan tensor output.

Lapisan normalisasi *batch* atau *batch normalization* digunakan untuk membuat pelatihan jaringan syaraf tiruan menjadi lebih cepat dan stabil melalui normalisasi input dari setiap lapisan, dengan melakukan pemusatan dan penskalaan ulang. Lapisan tersebut mempercepat proses pelatihan di dalam jaringan dan membantu untuk mendapatkan stabilitas [29] [27]. Normalisasi *batch* bertujuan untuk mengurangi pergeseran kovariat internal dengan mencoba menjaga distribusi output dari lapisan sebelumnya tidak berubah [28]. Pergeseran kovariat internal adalah perubahan distribusi aktivasi pada jaringan yang disebabkan oleh perubahan parameter jaringan selama proses pelatihan [28].

Lapisan *pooling* diletakkan setelah lapisan konvolusi, tepatnya setelah fungsi *non-linear* yang diterapkan pada lapisan konvolusi. Lapisan *pooling* melibatkan operasi *pooling* menggunakan filter yang diterapkan untuk operasi konvolusi. Filter yang digunakan lebih kecil dari pada filter yang digunakan untuk operasi konvolusi. Selanjutnya filter tersebut bergeser pada keseluruhan *feature map* dengan langkah atau *stride* yang telah ditentukan. Metode *max pooling* digunakan di dalam arsitektur ini karena tidak ada parameter yang perlu disetel [29]. *Max Pooling* memastikan bahwa fitur yang dapat dipelajari tetap terbatas jumlahnya sekaligus mempertahankan fitur utama dari gambar apa pun [30]. Hal tersebut dilakukan dengan mengambil nilai tertinggi dari setiap sub matriks pada peta fitur dan membentuk matriks yang terpisah. *Max pooling* biasanya menggunakan filter berukuran 2 x 2 [12]. Arsitektur ini menggunakan filter berukuran 2x2 dengan *stride* 2x2. Dengan demikian, lapisan *pooling* akan membagi setiap dimensi *feature map* menjadi dua, sehingga operasi ini akan mengurangi jumlah piksel atau nilai di setiap *feature map* menjadi seperempat ukuran awal. Jenis *pooling* yang dipilih adalah *max pooling*, artinya nilai tertinggi yang terdapat pada filter diambil sebagai nilai baru. Lapisan terakhir di dalam setiap blok ekstraksi fitur adalah *dropout*. *Dropout* merupakan pendekatan yang digunakan untuk regularisasi di dalam jaringan saraf untuk mengurangi ketergantungan antar *neuron*. *Dropout* bekerja dengan mengeluarkan sebagian *hidden layers* dan *visible layers* dari jaringan syaraf [31]. *Dropout* dapat diimplementasikan bersama dengan sebagian besar lapisan di dalam jaringan syaraf, seperti lapisan konvolusi, *lapisan dense*, atau lapisan berulang pada jaringan LSTM.

Pada blok klasifikasi terdapat lapisan *flatten* dan *dense*. Lapisan *flatten* akan mengubah *array* multi dimensi yang dihasilkan *feature map* menjadi bentuk *single vector* yang panjang untuk digunakan sebagai input lapisan jaringan syaraf biasa (*dense layers* atau *fully-connected layers*). *Flattening* diperlukan karena *output* dari lapisan konvolusional berbentuk multi dimensi, sedangkan lapisan *dense* menyaratkan *input* dalam bentuk satu dimensi. Lapisan *dense* merupakan lapisan sederhana dimana setiap neuron akan menerima *input* dari semua *neuron* pada lapisan sebelumnya. Oleh karena itu, lapisan ini juga disebut *fully-connected layers*. Lapisan *dense* melakukan klasifikasi citra berdasarkan keluaran yang dihasilkan dari lapisan konvolusi. Arsitektur CNN yang dikembangkan disajikan pada Gambar 5.



Gambar 5. Arsitektur CNN

D. Pengujian dan Evaluasi Model

Jumlah populasi yang digunakan untuk pengujian model tidak diketahui, sehingga jumlah sampel yang digunakan untuk pengujian diperoleh mengikuti rumus Paul Leedy [35] seperti pada Persamaan 1.

$$n = \left(\frac{Z}{e}\right)^2 (P)(P - 1) \quad (1)$$

Dimana n merupakan ukuran sampel minimum, Z merupakan standard score untuk α yang dipilih, e adalah *sampling error* yang dipergunakan, dan P adalah Proporsi harus dalam populasi.

Pengukuran akurasi model merupakan salah satu cara menilai kinerja suatu model pembelajaran mesin. Perhitungan nilai akurasi bertujuan untuk memperkirakan efektifitas suatu algoritma dengan menunjukkan nilai aktual terhadap keseluruhan label kelas [36]. Dengan demikian, nilai akurasi suatu model dapat menunjukkan keefektifan sebuah algoritma secara keseluruhan. Perhitungan akurasi dapat dilakukan menggunakan Persamaan 2 [36].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

True positive (TP) adalah jumlah sampel yang sebenarnya positif dan diprediksikan positif. *True negative (TN)* adalah jumlah sampel negatif yang diklasifikasikan negatif juga. *False positive (FP)* adalah jumlah sampel yang diprediksi negatif tetapi sebenarnya positif. *False negative (FN)* adalah jumlah sampel negatif yang diprediksi positif.

Selain akurasi, untuk mengukur kinerja model klasifikasi dapat juga menggunakan *Predictive Positive Value (PPV)* atau *precision* dan *True Positive Rate (TPR)* atau *recall*. *Precision* diperoleh dengan membagi jumlah sampel positif yang diklasifikasikan positif dengan jumlah seluruh sampel yang diklasifikasikan sebagai positif (baik sampel positif maupun negatif) seperti pada Persamaan 3 [36].

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Kebalikan dari *precision* adalah *Negative Prediction Value (NPV)*. NPV merupakan rasio dari jumlah sampel negatif yang diklasifikasikan negatif dengan jumlah seluruh sampel yang diklasifikasikan negatif, seperti pada Persamaan 4 [36].

$$NPV = \frac{TN}{TN + FN} \quad (4)$$

Recall atau *True Positive Rate (TPR)* diperoleh dengan membagi jumlah sampel positif yang terklasifikasi positif dibagi dengan jumlah seluruh sampel positif seperti dalam Persamaan 5 [36].

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

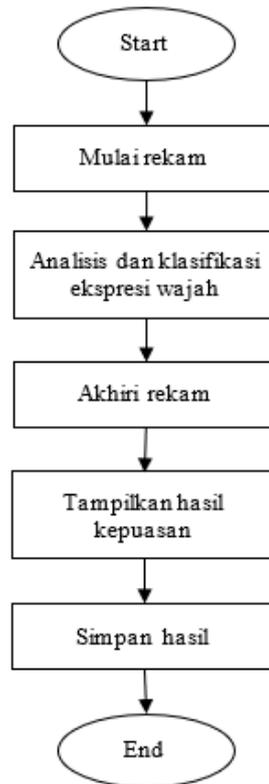
Sedangkan *True Negative Rate (TNR)* diperoleh dengan membagi jumlah sampel negatif yang diklasifikasikan negatif dibagi dengan jumlah sampel negatif seperti pada Persamaan 6 [36].

$$TNR = \frac{TN}{TN + FP} \quad (6)$$

E. Pengembangan Sistem

Software Development Life Cycle (SDLC) digunakan untuk merancang, mengembangkan, dan menguji Sistem Pengukuran Kepuasan Pelanggan Berdasarkan Ekspresi Wajah. Tahapan yang dilalui dalam pengembangan sistem tersebut yaitu analisis, desain, pengkodean, testing, *deployment*. Tahap analisis bertujuan untuk mengumpulkan sekaligus mendefinisikan kebutuhan apa saja yang harus dicapai oleh sistem. Kebutuhan yang harus dipenuhi oleh sistem yang dikembangkan yaitu mampu memberikan informasi tentang kepuasan (puas/tidak puas) pelanggan terhadap layanan yang diberikan dengan menganalisis setiap perubahan ekspresi wajah pada saat tertentu.

Tahap perancangan/desain bertujuan untuk memperkirakan sistem yang akan dikembangkan. Pada tahap ini akan menentukan arsitektur sistem, alur sistem, *interface* sistem, dan *platform* yang digunakan. Alur kerja sistem dimulai dengan merekam *customer* yang datang. Sistem akan membaca rekaman tersebut secara *live*. Pada saat proses perekaman berlangsung sistem akan melakukan klasifikasi terhadap ekspresi wajah pada setiap *frame* menggunakan model yang telah dibuat sebelumnya. Selanjutnya, sistem akan menghitung jumlah hasil klasifikasi untuk setiap kelas ekspresi, kelas dengan jumlah terbanyak merupakan hasil dari pengukuran kepuasan dari *customer* tersebut. Kemudian *customer service* dapat menyimpan hasil pengukuran kepuasan tersebut ke dalam *database*. Alur kerja sistem disajikan pada Gambar 6.



Gambar 6. Alur kerja sistem

III. HASIL DAN PEMBAHASAN

A. Pemodelan

Cross-validation sering digunakan dalam pembelajaran mesin untuk mengatasi permasalahan yang sering terjadi dalam pemodelan prediktif. *Cross-validation* diterapkan dalam proses pemodelan ini dengan tujuan untuk mengevaluasi model supaya mampu mengenali data baru dengan baik. *Cross-validation* menggunakan parameter k yang mewakili jumlah partisi dari pembagian kumpulan data yang digunakan selama pelatihan model. Ketika nilai spesifik untuk k dipilih, nilai tersebut dapat digunakan sebagai pengganti k dalam referensi model. Dalam pelatihan model ini, nilai k yang dipilih adalah 8, artinya proses validasi silang dilakukan sebanyak delapan kali. Nilai 8 dipilih agar proses pelatihan model tidak terlalu lama. Sebanyak 7 partisi digunakan sebagai data pelatihan dan satu partisi atau 12,5% dari jumlah data digunakan sebagai validasi setiap proses pelatihan. Proses ini dilakukan sebanyak delapan kali, sehingga semua partisi secara bergantian dapat berperan sebagai data latih dan data validasi.

Untuk melatih model, *epoch* yang digunakan sebanyak 100. Satu *epoch* berarti melatih jaringan syaraf dengan semua data pelatihan untuk satu siklus. Dalam satu *epoch*, semua data digunakan tepat satu kali, artinya algoritma *neural network* akan melatih model dan melakukan validasi menggunakan keseluruhan keseluruhan data sejumlah *epoch* yang ditentukan. Ketepatan dalam menentukan jumlah *epoch* menjadi hal yang harus diperhatikan. Proses pelatihan akan berhenti meskipun model belum mencapai kondisi optimal apabila *epoch* yang dipilih terlalu sedikit, sehingga model yang terbentuk kurang baik dalam mengklasifikasikan data yang diberikan. Sedangkan jumlah *epoch* yang terlalu banyak menyebabkan model yang terbentuk tidak dapat mengenali data baru yang diberikan. Permasalahan tersebut dapat diselesaikan dengan menerapkan teknik *early stopping*, dimana proses pelatihan akan dihentikan meskipun jumlah *epoch* yang diberikan belum terpenuhi. Proses pelatihan akan dihentikan apabila dari hasil validasi setiap *epoch* tidak mengalami perbaikan. Dalam proses pelatihan model, rata-rata *loss* yang terjadi sebesar 0,2444 dan rata-rata akurasi yang dihasilkan sebesar 0,9063 seperti disajikan dalam Tabel 1.

TABEL 1
LOSS DAN AKURASI YANG DIHASILKAN

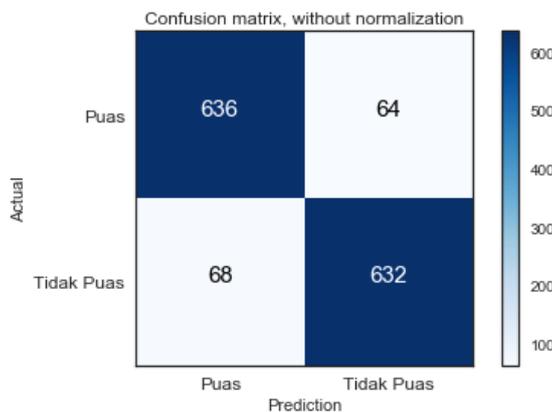
Fold	Loss	Accuracy
Fold-1	0,2883	0,8873
Fold-2	0,2471	0,9073
Fold-3	0,2704	0,8877
Fold-4	0,2233	0,9178
Fold-5	0,2129	0,9242
Fold-6	0,2203	0,9165
Fold-7	0,2397	0,9064
Fold-8	0,2531	0,9028
Average	0,2444	0,9063

B. Evaluasi dan Pengujian Model

Mengevaluasi dan menguji sebuah model pembelajaran mesin diperlukan untuk memastikan model tersebut memiliki kinerja yang baik. Di dalam penelitian ini pengujian terhadap model yang terbentuk dilakukan menggunakan data uji sebanyak 1.400 data, yang terbagi menjadi dua kelas, yaitu puas dan tidak puas, masing-masing sejumlah 700 sampel data. Penentuan jumlah sampel tersebut didasarkan pada penentuan sampel data jika jumlah populasi tidak diketahui sesuai Persamaan 1. Dengan tingkat kesalahan 3% dan *confident level* 97% diperlukan jumlah sampel minimal 1.309 sampel.

Teknik yang sering digunakan untuk mengukur kinerja model klasifikasi adalah *confusion matrix*. Dalam klasifikasi biner, *confusion matrix* memiliki empat nilai dasar, yaitu *True Positive (TP)*, *False Negative (FN)*, *True Negative (TN)*, dan *False Postive (FP)*. Dalam hal ini, *True Positive (TP)* adalah sampel yang sebenarnya termasuk kedalam kelas puas dan juga diklasifikasikan ke dalam kelas “puas”. *False Negative (FN)* adalah sampel yang sebenarnya termasuk ke dalam kelas “puas” tetapi model mengklasifikasikannya ke dalam kelas “tidak puas”. *True Negative (TN)* adalah sampel dengan kelas “tidak puas” diklasifikasikan ke dalam kelas “tidak puas” juga. *False Positive (FP)* adalah sampel dengan kelas “tidak puas” tetapi diklasifikasikan ke dalam kelas “puas”.

Dari hasil pengujian diperoleh jumlah *true positif (TP)* sebanyak 636 sampel, *false negative (FN)* sebanyak 64 sampel, *true negative (TN)* sebanyak 632 sampel, dan *false positive (FP)* sebanyak 68 sampel seperti yang ditunjukkan pada Gambar 7.



Gambar 7. Confusion matrix pengujian model

Berdasarkan hasil pengujian yang ditunjukkan oleh *confusion matrix* seperti pada Gambar 5, akurasi model CNN dapat diketahui dengan memasukan nilai-nilai pada *confusion matrix* ke dalam persamaan 2. Akurasi menunjukkan persentase sampel data yang diklasifikasikan benar oleh model klasifikasi. Akurasi model yang terbentuk ditunjukkan pada persamaan 7.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% = \frac{636 + 632}{636 + 632 + 64 + 68} \times 100\% = 90,57\% \quad (7)$$

Precision menunjukkan persentase menunjukkan sampel data “puas” dari keseluruhan sampel data yang diklasifikasikan “puas”. Nilai *precision* yang diperoleh dapat dilihat pada persamaan 9.

$$Precision = \frac{TP}{TP + FP} \times 100\% = \frac{636}{636 + 68} \times 100\% = 90,36\% \quad (8)$$

Negative Predictive Value (NPV) menunjukkan persentase menunjukkan sampel data “tidak puas” dari keseluruhan sampel data yang diklasifikasikan “tidak puas”. NPV yang diperoleh dapat dilihat pada persamaan 9:

$$NPV = \frac{TN}{TN + FN} \times 100\% = \frac{632}{632 + 64} \times 100\% = 90,8\% \quad (9)$$

Recall menunjukkan persentase sampel data yang terklasifikasi “puas” dari keseluruhan sampel data yang sebenarnya “puas”. Nilai *recall* yang diperoleh dapat dilihat pada persamaan 10.

$$Recall = \frac{TP}{TP + FN} \times 100\% = \frac{636}{636 + 64} \times 100\% = 90,86\% \quad (10)$$

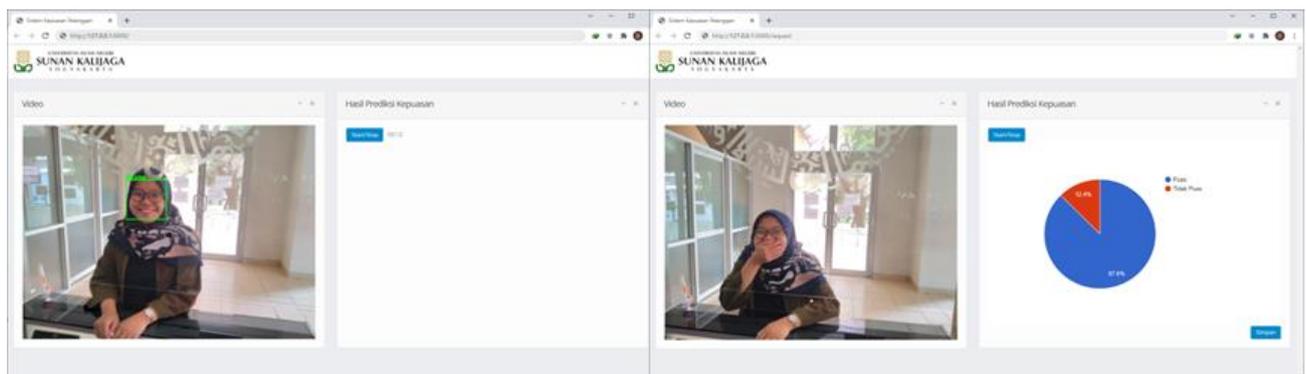
True Negative Rate (TNR) menunjukkan persentase sampel data yang terklasifikasi “tidak puas” dari keseluruhan sampel data yang sebenarnya “tidak puas”. Nilai TNR yang diperoleh dapat dilihat pada persamaan 11:

$$TNR = \frac{TN}{TN + FP} \times 100\% = \frac{632}{632 + 68} \times 100\% = 90,28\% \quad (11)$$

Dengan demikian, model yang dibangun memiliki akurasi 90,57% dengan *precision* dan *recall* masing-masing sebesar 90,34% dan 90,86%.

C. Pengembangan Sistem

Sistem Pengukuran Kepuasan Pelanggan merupakan sistem berbasis web, sehingga untuk menjalankan sistem ini diperlukan sebuah *web browser* seperti *Firefox* atau *Google Chrome*. Sistem Pengukuran Kepuasan Pelanggan diimplementasikan berbasis web menggunakan bahasa pemrograman Python. *Flask* dipilih sebagai framework untuk mempermudah dalam implementasi model CNN. Untuk menyimpan data, sistem tersebut menggunakan database MySQL. Sistem Pengukuran Kepuasan Pelanggan dapat dijalankan melalui protokol HTTP menggunakan *web browser*. Tampilan aplikasi pada saat proses deteksi dan klasifikasi kepuasan disajikan pada Gambar 8.



Gambar 8. Tampilan hasil pengukuran kepuasan pelanggan

Proses pengenalan dimulai saat perekaman dimulai. Model CNN melakukan pengenalan dan klasifikasi pada wajah yang terdeteksi pada setiap *frame*. Misalnya jika video yang diambil memiliki *frame rate* 30 fps, artinya dalam satu detik terdapat

30 *frame* yang harus dikenali. Tetapi tidak semua *frame* dapat terdeteksi objek wajah, sehingga klasifikasi hanya dilakukan pada *frame* yang di dalamnya terdeteksi objek wajah. Hal ini dapat terjadi ketika seorang pelanggan tersebut sedang menunduk atau menoleh. Dalam hal ini, sebisa mungkin kamera dapat merekam wajah pelanggan yang menghadap ke depan. Selanjutnya sistem akan menghitung jumlah kelas yang terdeteksi selama proses perekaman. Jumlah kelas yang terbanyak merupakan hasil dari pengukuran kepuasan tersebut.

IV. SIMPULAN

Model CNN dapat diimplementasikan pada sebuah sistem pengukur kepuasan pelanggan dengan melakukan pengenalan ekspresi wajah pelanggan dan melakukan klasifikasi terhadap ekspresi wajah yang dikenali. Dalam pemodelan CNN, digunakan citra wajah dengan posisi menghadap ke depan yang telah memiliki label puas dan tidak puas, sebanyak 4000 citra ekspresi wajah digunakan sebagai data latih, dan sebanyak 1400 citra wajah digunakan sebagai data uji. Setelah dilakukan pengujian, model CNN yang terbentuk memiliki akurasi mencapai 90.57%. Sistem pengukuran kepuasan pelanggan mendeteksi setiap perubahan ekspresi wajah pelanggan dan mengklasifikasikan kedalam ekspresi puas atau tidak puas kemudian menghitung jenis ekspresi terbanyak yang merupakan hasil dari pengukuran kepuasan pelanggan tersebut.

UCAPAN TERIMA KASIH

Terimakasih diucapkan kepada LPPM UIN Sunan Kalijaga Yogyakarta atas bantuan yang diberikan sehingga penelitian ini dapat berjalan lancar. Artikel ini didasarkan pada penelitian yang didanai oleh Dana BOPTN UIN Sunan Kalijaga Yogyakarta.

DAFTAR PUSTAKA

- [1] D. R. Elisabeth, A. Nasir and J. Suyono, "The Effect of Service Quality on Customer Satisfaction at Koperasi Karyawan PT. Lotus Indah Textile Industry," *Int. J. Entrep. Bus. Dev.*, vol. 2, no. 2, p. 172–178, 2019.
- [2] N. Anisor, D. Adela-Eliza and C. Luciana, "The method used for measuring the customers' satisfaction," *Int. Conf. Educ. Educ. Technol. - Proc.*, p. 197–200, 2010.
- [3] T. S. Ruslim and M. Rahardjo, "Identifikasi Kepuasan Konsumen Ditinjau Dari Segi Harga dan Kualitas ada Restoran Abuba Steak di Greenville," *J. Ilm. Manaj. dan Bisnis*, vol. 2, no. 1, p. 51–65, 2016.
- [4] Z. Ikrawan, Y. Ariyanto and B. Harijanto, "Sistem Informasi Pengukuran Kepuasan Pelanggan Menggunakan Metode Importance Performance Analysis Pada Program Studi Manajemen Informatika," *J. Inform. Polinema*, vol. 1, no. 4, p. 48, 2015.
- [5] A. N. F. Islami, I. Aknuranda and A. R. Perdanakusuma, "Pengembangan Sistem Survei Kepuasan Pelanggan Pada PDAM Surya Sembada Kota Surabaya," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 4, p. 1690–1697, 2018.
- [6] S. Alamdari, M. H. Basiri, A. Mousavi and A. Soofastaei, "Application of Machine Learning Techniques to Predict Haul Truck Fuel Consumption in Open-Pit Mines," *J. Min. Environ.*, vol. 13, no. 1, p. 69–85, 2022.
- [7] D. Prasetyawan and S. 'Uyun, "Penentuan Emosi pada Video dengan Convolutional Neural Network," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 5, no. 1, p. 23–35, 2020.
- [8] G. B. Huang, Q. Y. Zhu and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1-3, p. 489–501, 2006.
- [9] A. Roihan, P. A. Sunarya and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, p. 75–82, 2020.
- [10] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, 3rd ed, Massachusetts: Morgan Kaufmann, 2001.
- [11] Y. Lecun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 2014, 436–444.
- [12] D. R. e. al., "Deep Learning for Health Informatics," *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 1, p. 4–21, 2018.
- [13] M. H. Diponegoro, S. S. Kusumawardani and I. Hidayah, "Tinjauan Pustaka Sistematis: Implementasi Metode Deep Learning pada Prediksi Kinerja Murid," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, p. 131–138, 2021.
- [14] T. Bikku, "Multi-layered deep learning perceptron approach for health risk prediction," *J. Big Data*, vol. 7, no. 1, 2020.
- [15] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti and D. De, "Fundamental concepts of convolutional neural network," 2019. [Online]. Available: https://scholar.google.co.id/scholar?hl=en&as_sdt=0%2C5&as_vis=1&q=A.+Ghosh%2C+A.+Sufian%2C+F.+Sultana%2C+A.+Chakrabarti+and+D.+De%2C+%E2%80%9CFundamental+concepts+of+convolutional+neural+network%2C&btnG=.
- [16] S. S. Liew, M. Khalil-Hani, S. A. Radzi and R. Bakhteri, "Gender classification: A convolutional neural network approach," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 24, no. 3, p. 1248–1264.
- [17] P. .. Vasanth and K. .. Nataraj, "Facial Expression Recognition Using SVM Classifier," *Indones. J. Electr. Eng. Informatics*, vol. 3, no. 1, p. 16–20, 2015.
- [18] K. Khan, M. Attique, I. Syed and A. Gul, "Automatic gender classification through face segmentation," *Symmetry (Basel)*, vol. 11, no. 6, p. 1–14, 2019.
- [19] Z. Qawaqneh, A. A. Mallouh and B. D. Barkana, "Deep Convolutional Neural Network for Age Estimation based on VGG-Face Model," *ArXiv*, vol. 1, 2017.

- [20] I. Talegaonkar, K. Joshi, S. Valunj, R. Kohok and A. Kulkarni, "Real Time Facial Expression Recognition using Deep Learning," *SSRN Electron. J.*, 2019.
- [21] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*, 1st ed, O'Reilly Media, 2008.
- [22] Q. X. e. al., "Review Article Negative and Positive Bias for Emotional Faces : Evidence from the Attention and Working Memory Paradigms," *Hidawi*, 2021.
- [23] J. Sanjaya and M. Ayub, "Augmentasi Data Pengenalan Citra Mobil Menggunakan Pendekatan Random Crop , Rotate , dan Mixup," *Justisi*, vol. 6, no. 2, p. 311–323, 2020.
- [24] P. Viola and J. Michael, "Robust Real-Time Face Detection," *Int. J. Comput. Vis.*, vol. 57, p. 137–154, 2004.
- [25] C. Maas and J. Schmalzl, "Using pattern recognition to automatically localize reflection hyperbolas in data from ground penetrating radar," *Computers and Geosciences*, vol. 58, p. 116–125, 2013.
- [26] S. Abidin, "Deteksi Wajah Menggunakan Metode Haar Cascade Classifier Berbasis Webcam Pada Matlab," *J. Teknol. Elekterika*, vol. 15, no. 1, p. 21, 2018.
- [27] A. Biswas and M. S. Islam, "An Efficient CNN Model for Automated Digital Handwritten Digit Classification," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 7, no. 1, p. 42, 2021.
- [28] Y. Huang and Y. Yu, "An internal covariate shift bounding algorithm for deep neural networks by unitizing layers' outputs," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2020.
- [29] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *32nd Int. Conf. Mach. Learn. ICML 2015*, 2015.
- [30] A. Z. e. al., "A Comparison of Pooling Methods for Convolutional Neural Networks," *Appl. Sci.*, vol. 12, no. 17, p. 1–21, 2022.
- [31] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, vol. 15, p. 1929–1958, 2014.
- [32] K. Hong, "Object Detection : Face Detection Using Haar Cascade Classifiers," 2020. [Online]. Available: https://www.bogotobogo.com/python/OpenCV_Python/python_opencv3_Image_Object_Detection_Face_Detection_Haar_Cascade_Classifiers.php
- [33] A. Ajit, K. Acharya and A. Samanta, "A Review of Convolutional Neural Networks," in *Int. Conf. Emerg. Trends Inf. Technol. Eng. ic-ETITE 2020*, 2020.
- [34] P. D. Leedy and J. E. Ormrod, *Practical Research: Planning and Design Research*, 8th ed, Pearson Merrill Prentice Hall, 2005.
- [35] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, p. 427–437, 2009.
- [36] R. Bousaid, M. E. Hajji and Y. Es-Saady, "Facial Expression Recognition Using a Hybrid CNN - SIFT Aggregator," *Lect. Notes Bus. Inf. Process.*, vol. 449, p. 61–70, 2022.