

**LAPORAN KEMAJUAN PENELITIAN YANG DIAJUKAN
KE LEMBAGA PENELITIAN DAN PENGABDIAN KEPADA MASYARAKAT**



**SISTEM PENDETEKSIAN DAN PENGENALAN EKSPRESI WAJAH
DENGAN ALGORITMA YOLO DAN CONVOLUTIONAL NEURAL NETWORK**

Disusun oleh:

Ketua Tim

Dra. Chairisni Lubis, M.Kom (0307096301/10393012)

Anggota:

Novario Jaya Perdana, S.Kom, MT (8856970018/10817002)

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TARUMANAGARA
JAKARTA
TAHUN 2020

**HALAMAN PENGESAHAN
LAPORAN AKHIR PENELITIAN
Semester Genap / Tahun 2019-2020**

1. Judul Penelitian : Sistem Pendeteksian dan Pengenalan Ekspresi Wajah dengan Algoritma Yolo dan Convolutional Neural Network
2. Ketua
- a. Nama dan gelar : Dra. Chairisni Lubis, M.Kom
 - b. NIDN / NIK : 0307096301/10393012
 - c. Jabatan / gol : Ektor Kupaia
 - d. Program Studi : Teknik Informatika
 - e. Fakultas : Teknologi Informasi
 - f. Bidang Keahlian : Artificial Intelligence
 - g. Alamat Kantor : Untar Kampus 1, Gedung R It.10
Jl. S. Parman No.1
 - h. No. HP : 08161388270
3. Anggota Tim Penelitian
- a. Jumlah Anggota Dosen : 1 orang
 - b. Nama Anggota 1/Fak/Keahlian : Novario Jaya Perdana, S.Kom., M.T. / Fakultas Teknologi Informasi / Data Engineering
 - c. Jumlah Mahasiswa : 1 orang
 - d. Nama Mahasiswa/NIM : Nelson Y/535160046
4. Lokasi Kegiatan Penelitian : Laboratorium FTI
5. Luaran yang Dihasilkan : Publikasi Internasional
6. Jangka Waktu Pelaksanaan : Periode 1 / Maret -Agustus 2020
7. Biaya Total
- a. Biaya yang Diajukan ke LPPM : Rp16.500.000,00
 - b. Biaya yang Disetujui LPPM : Rp10.500.000,00

Jakarta, 13 Juli 2020

Mengetahui,
Dekan Fakultas



Prof. Dr. Dyah Emy Herwindiati, M.Si
NIDN/NIK: 0306046301/10389013

Ketua Peneliti



Dra. Chairisni Lubis, M.Kom
NIDN/NIK: 0307096301/10393012

Menyetujui,
Ketua Lembaga Penelitian dan Pengabdian kepada Masyarakat



Jap Tji Beng, Ph.D
NIDN/NIK: 0323085501 / 10381047

RINGKASAN

Penelitian ini menggunakan Convolutional Neural Network (CNN) yang merupakan salah satu arsitektur dalam Deep Learning untuk mendapatkan tingkat keberhasilan pengenalan wajah yang tinggi dari hasil pendeteksian. Proses pendeteksian citra wajah digunakan Algoritma Yolo. Pendeteksian wajah dilakukan terhadap input berupa video hasil perekaman data dari webcam yang memuat ekspresi wajah. Pada proses pelatihan untuk pengenalan, data ekspresi wajah yang sudah diketahui labelnya (data training) dilatih dahulu dengan menggunakan CNN. Setelah itu, hasil pelatihan tersebut digunakan untuk mengenali ekspresi wajah pada data testing. Dari hasil pengujian, Algoritma YOLO berhasil mendeteksi citra wajah dengan sangat baik, dengan tingkat keyakinan hingga 85%. Sedangkan performa CNN kurang bagus, karena hanya berhasil mendeteksi ekspresi wajah hingga 46% saja.

Kata Kunci: *Algoritma Yolo, Convolutional Neural Network, Deep Learning, Wajah*

PRAKATA

Pada kesempatan ini, kami ingin mengajukan laporan akhir penelitian kepada Lembaga Penelitian dan Pengabdian Kepada Masyarakat, Universitas Tarumanagara. Tema dari penelitian kami adalah Pendeteksian dan Pengenalan Ekspresi Wajah Manusia pada video. Metode yang akan digunakan adalah Algoritma Yolo untuk proses pendeteksiannya dan *Convolutional Neural Network* untuk proses pengenalan ekspresi wajahnya. Oleh karena itu, penelitian kami diberi judul “Sistem Pendeteksian dan Pengenalan Ekspresi Wajah dengan Algoritma Yolo dan Convolutional Neural Network”. Semoga penelitian ini dapat memperluas khasanah ilmu pengetahuan, terutama bidang teknologi informasi, khususnya bidang *deep learning*.

Jakarta, Agustus 2020
Peneliti

Dra. Chairisni Lubis, M.Kom

DAFTAR ISI

HALAMAN PENGESAHAN	2
RINGKASAN	3
PRAKATA	4
DAFTAR ISI	5
DAFTAR GAMBAR	7
DAFTAR TABEL	8
BAB 1 PENDAHULUAN.....	9
1.1 Latar Belakang	9
1.2 Tujuan Khusus.....	10
1.3 Pentingnya Penelitian yang direncanakan	10
BAB 2 STUDI PUSTAKA.....	11
2.1 Ekspresi Wajah.....	11
2.2 Algoritma <i>You Only Look Once</i> (YOLO)	12
2.3 Convolution Neural Network	14
2.3.1 Lapisan ekstraksi fitur.....	14
2.3.2 Convolution Layer	14
2.3.3 ReLU (<i>Rectified Linear Units</i>)	14
2.3.4 Lapisan Pooling	15
2.3.5 Lapisan Klasifikasi	15
2.3.6 Softmax.....	16
2.3.7 Residual Neural Network	17
BAB 3 METODE PENELITIAN.....	18
3.1 Metode yang digunakan	18
3.2 Manfaat Penelitian.....	19
3.3 Bentuk Data yang Digunakan.....	22
BAB 4 Hasil dan Pembahasan.....	23
4.1 Metode Pengujian.....	23
4.1.1 Pengujian <i>Black Box Testing</i>	23
4.1.2 Pengujian Algoritma YOLO.....	24
4.1.3 Pengujian Algoritma CNN	24

4.1.4	Pengujian Keseluruhan Sistem	25
4.2	Hasil dan Pembahasan Pengujian.....	25
4.2.1	Pengujian Algoritma Yolo	25
4.2.2	Hasil Pengujian Algoritma CNN	26
4.2.3	Hasil Pengujian Keseluruhan Sistem.....	26
BAB 5	Kesimpulan dan Saran.....	27
5.1	Kesimpulan.....	27
5.2	Saran.....	27
Daftar Pustaka	28
Lampiran 1	Hasil Black Box Testing	29
Lampiran 2	Hasil Pengujian Algoritma YOLO	33
Lampiran 3	Hasil Pengujian Algoritma CNN	39

DAFTAR GAMBAR

Gambar 1 Ekspresi Wajah Manusia.....	12
Gambar 2 Ilustrasi Algoritma YOLO	13
Gambar 3 Arsitektur Convolution Neural Network	13
Gambar 4 Proses Pooling.....	15
Gambar 5 Perubahan matrik 2 dimensi menjadi vektor	16
Gambar 6 Arsitektur ResNet50.....	17
Gambar 7 Flowchart Sistem Program Pengenalan Ekspresi Wajah	19
Gambar 8 Flowchart metode CNN	20
Gambar 9 Diagram Hirarki Program Aplikasi.....	21
Gambar 10 Tampilan Antarmuka Program Modul Halaman Utama.....	29
Gambar 11 Tampilan Antarmuka Program Modul Register	29
Gambar 12 Tampilan Antarmuka Program Modul Login	30
Gambar 13 Tampilan Antarmuka Program Modul Upload Video	30
Gambar 14 Tampilan Antarmuka Program Modul Edit Video	30
Gambar 15 Tampilan Antarmuka Program Modul My Video	31
Gambar 16 Tampilan Antarmuka Program Modul Search	31
Gambar 17 Tampilan Antarmuka Program Modul Konten Modal PopUp	31
Gambar 18 Tampilan Antarmuka Program Fitur <i>Setting Username</i>	32
Gambar 19 Tampilan Antarmuka Program Fitur Setting Password.....	32

DAFTAR TABEL

Tabel 1 Detil Jumlah Data Untuk Proses Pelatihan.....	24
Tabel 2 Detil Jumlah Data Untuk Proses Pengujian.....	25
Tabel 3. Hasil Pengujian Keseluruhan Sistem.....	26
Tabel 4 Hasil Pengujian YOLO.....	33
Tabel 5 Hasil Pengujian Model CNN.....	39

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Wajah merupakan salah satu instrumen yang dapat digunakan untuk mengidentifikasi seseorang karena setiap orang memiliki wajah yang berbeda-beda. Selain itu, wajah juga dapat digunakan sebagai instrumen untuk membedakan ekspresi seseorang dan dalam ilmu komputer (teknologi informasi) termasuk dalam aplikasi klasifikasi pola (pengenalan). Setiap ekspresi wajah memiliki sebuah respon yang tentunya berbeda-beda. Instrumen wajah yang digunakan untuk membedakan ekspresi wajah adalah dari bentuk dan posisi mata, hidung, dan mulut. Dalam pengenalan ekspresi wajah, ciri-ciri tersebut merupakan komponen-komponen yang digunakan oleh metode-metode pengenalan ekspresi wajah dalam teknologi informasi.

Klasifikasi pola (dalam penelitian ini pola ekspresi wajah) merupakan salah satu aplikasi dari *Deep Learning* (DL). *Deep Learning* merupakan bagian dari *Machine Learning* (ML) yang menggunakan cara baru dalam merepresentasikan pembelajaran dari data dalam bentuk lapisan-lapisan, dimana dengan semakin bertambahnya lapisan dapat meningkatkan representasi yang lebih berarti. Banyaknya lapisan yang digunakan pada model data disebut sebagai kedalaman model. *Deep Learning* (DL) merupakan suatu model ML yang sangat powerful dengan performance yang tinggi dalam mencari solusi yang rumit.

Pada proses klasifikasi dengan menggunakan *Deep Learning*, lapisan terakhir berfungsi untuk memperkuat perbedaan antara kelompok data (input) dan menghilangkan ciri yang tidak penting dengan menggunakan suatu fungsi diskriminan. Dalam pembelajaran yang modern, banyaknya lapisan yang digunakan besarnya puluhan bahkan ratusan lapisan untuk mempelajari data pelatihan secara otomatis.

Convolutional Neural Network (CNN) atau yang biasa dikenal sebagai Convnet merupakan salah satu model *Deep Learning* yang umum digunakan dalam aplikasi *Computer Vision*. CNN menggunakan operasi matematika yang disebut konvolusi yang merupakan salah satu jenis operasi linier khusus. Konvolusi digunakan sebagai pengganti perkalian matriks umum pada sedikitnya satu lapisan. Proses pelatihannya menggunakan *Backpropagation Neural Network* (BPNN).

Sebelum *Deep Learning* melakukan pengenalan ekspresi wajah harus dilakukan pendeteksian pada citra. Pada penelitian ini, digunakan algoritma *You Only Look Once* (YOLO). Algoritma YOLO adalah merupakan metode yang dipakai sebagai *Real Time Object*

Detection yang berguna untuk mendeteksi citra secara langsung. Metode ini pernah diimplementasikan oleh Joshua (NPM: 535160071) mahasiswa Fakultas Teknologi Informasi Universitas Tarumanagara untuk mendeteksi 630 karakter pada tanda nomor kendaraan bermotor dengan nilai akurasi sebesar 80%. Selain mendeteksi karakter, YOLO juga dapat mendeteksi ekspresi wajah manusia. Hasil deteksi dari metode YOLO akan digunakan untuk proses pengenalan ekspresi wajah.

Penelitian ini merupakan penelitian lanjutan dari penelitian sebelumnya “Sistem Aplikasi Pendeteksian dan Pengenalan Wajah”. Pada penelitian tersebut, sebelum mencari fitur unik wajah terlebih dahulu dilakukan pendeteksian wajah manusia. Pendeteksian dilakukan untuk membedakan bagian yang merupakan wajah dengan bagian yang bukan merupakan wajah. Pada aplikasi tersebut, metode yang akan digunakan untuk mendeteksi wajah manusia adalah metode *Adaptive Boosting* (AdaBoost), sedangkan proses pengenalannya dilakukan dengan menggunakan *K-Nearest Neighbour* (KNN) dengan membandingkan vektor ciri antara data uji dan data latih. Vektor ciri didapat dari proses ekstraksi ciri dengan menggunakan *Gray Level Co-occurrence Matrix* (GLCM). Dari hasil penelitian tersebut didapat hasil pengujian deteksi wajah manusia sebesar 93,79%. Dan hasil pengenalan wajah didapat rata-rata keberhasilan sebesar 73,37%.

Pada penelitian ini, akan dirancang sebuah aplikasi berbasis website untuk mengenali ekspresi wajah manusia secara langsung dengan menggunakan CNN. Untuk pendeteksiannya digunakan metode YOLO.

1.2 Tujuan Khusus

Tujuan penelitian ini adalah untuk mengenali ekspresi wajah dengan menggunakan metode *Convolutional Neural Network* (CNN) yang merupakan salah satu model dalam *Deep Learning*. Sebelum proses pengenalan, dilakukan proses pendeteksian wajah dahulu dengan menggunakan algoritma *You Only Look Once* (YOLO). Pada penelitian ini akan diteliti tingkat keakuratan dari kedua metode tersebut untuk mengetahui apakah metode tersebut cukup baik digunakan dalam aplikasi pengenalan ekspresi wajah manusia.

1.3 Pentingnya Penelitian yang direncanakan

Program dihasilkan dari penelitian ini diharapkan dapat digunakan untuk aplikasi sosial media penayangan video berbasis website yang dapat mengetahui respon para penggunanya secara langsung saat melihat konten yang telah disediakan.

BAB 2

STUDI PUSTAKA

Pada penelitian “Pengenalan Ekspresi Wajah Menggunakan Convolutional Neural Network” ini, Convolutional Neural Network (CNN) digunakan untuk mengenali wajah manusia dari hasil proses pendeteksian wajah dengan menggunakan metode *You Only Look Once* (YOLO). Untuk mendapatkan citra ekspresi wajah, perangkat keras yang digunakan berupa webcam, kemudian citra yang sudah didapat akan diproses oleh aplikasi. Tujuan dari aplikasi yang akan dirancang adalah untuk mendeteksi dan mengenali ekspresi wajah manusia secara langsung.

Pada bab ini akan dijelaskan metode yang digunakan dalam penelitian ini yaitu: metode untuk mendeteksi wajah dengan menggunakan *You Only Look Once* (YOLO) dan Convolutional Neural Network (CNN) untuk pengenalan ekspresi wajah. Dan sedikit teori Ekspresi Wajah yang mendukung penelitian ini.

2.1 Ekspresi Wajah

Ekspresi wajah atau mimik adalah hasil dari satu atau lebih gerakan atau posisi otot pada wajah. Ekspresi wajah merupakan salah satu bentuk komunikasi nonverbal, dan dapat menyampaikan keadaan emosi dari seseorang kepada orang yang mengamatinya. Ekspresi wajah merupakan salah satu cara penting dalam menyampaikan pesan sosial dalam kehidupan manusia.

Manusia dapat mengalami ekspresi wajah tertentu secara sengaja, tetapi umumnya ekspresi wajah dialami secara tidak sengaja akibat perasaan atau emosi manusia tersebut. Biasanya amat sulit untuk menyembunyikan perasaan atau emosi tertentu dari wajah, walaupun banyak orang yang merasa amat ingin melakukannya. Misalnya, orang yang mencoba menyembunyikan perasaan bencinya terhadap seseorang, pada saat tertentu tanpa sengaja akan menunjukkan perasaannya tersebut di wajahnya, walaupun ia berusaha menunjukkan ekspresi netral. Hubungan perasaan dan ekspresi wajah juga dapat berjalan sebaliknya, pengamatan menunjukkan bahwa melakukan ekspresi wajah tertentu dengan sengaja, dapat memengaruhi atau menyebabkan perasaan terkait benar-benar terjadi.



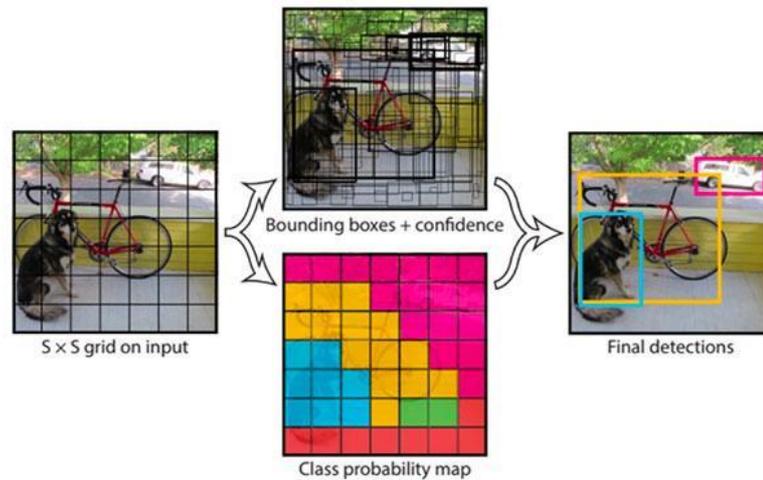
Gambar 1 Ekspresi Wajah Manusia

Sebagian ekspresi wajah dapat diketahui maksudnya dengan mudah, bahkan oleh anggota spesies yang berbeda, misalnya kemarahan dan kepuasan. Namun, beberapa ekspresi lainnya sulit diartikan, misalnya ketakutan dan kejjjikan kadang sulit dibedakan. Selain itu, kadang-kadang suatu wajah dapat disalahartikan mengalami emosi tertentu, karena susunan otot-otot wajah orang tersebut secara alami menyerupai wajah seseorang yang mengalami ekspresi tertentu, misalnya wajah seseorang yang tampak selalu tersenyum. [6]

2.2 Algoritma *You Only Look Once* (YOLO)

Algoritma YOLO adalah metode yang didasarkan dengan regresi yang memprediksi menggunakan *class object* dan *bounding box* untuk menentukan lokasi dari sebuah objek pada citra dalam satu kali algoritma. Algoritma YOLO bekerja dengan cara membagi citra menjadi beberapa sel, setiap sel digunakan untuk memprediksi 5 *bounding box* jika objek pada citra ada lebih dari satu. Prediksi akan menghasilkan nilai prediksi. Hasil prediksi akan dikumpulkan dan *bounding box* dengan probabilitas terkecil akan dihapus. *Bounding box* dengan nilai probabilitas prediksi terbesar akan menjadi hasil akhir.⁷ Proses pada algoritma YOLO dapat dilihat pada **Gambar 2**.

Prediksi pada YOLO menggunakan arsitektur yang mirip seperti *Convolutional Neural Network* (CNN). YOLO hanya menggunakan *convolutional layer* dan *pooling layer*. *Convolutional layer* terakhir pada arsitektur YOLO disesuaikan dengan jumlah kelas dan jumlah kotak prediksi yang ditentukan. Rumus untuk menghitung ukuran keluaran pada *convolutional layer* terakhir dapat dilihat pada rumus (1).



Gambar 2 Ilustrasi Algoritma YOLO

Sumber : Adrian Rosebrock, YOLO object detection with OpenCV, <https://www.pyimagesearch.com/2018/11/12/yolo-object-detection-with-opencv/>

$$Y = S, S, B \times (5 + C) \quad (1)$$

Keterangan:

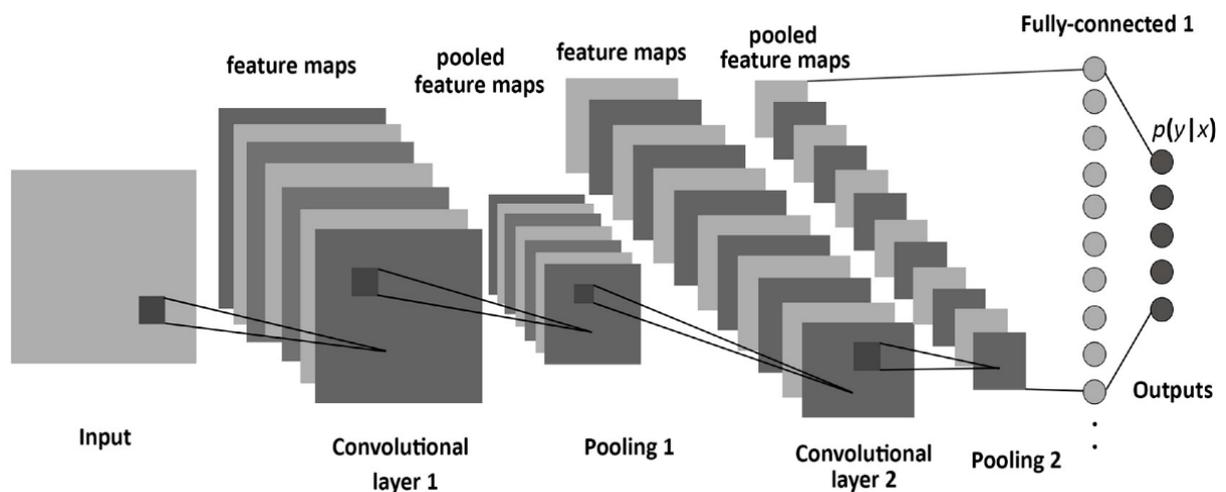
Y= Ukuran keluaran

S= Jumlah baris atau kolom grid

B= Jumlah prediksi yang diinginkan pada setiap grid

C = Jumlah kelas yang ingin diprediksi

Sebelum pendeteksian, algoritma YOLO membutuhkan proses anotasi terlebih dahulu untuk data. Setiap data memiliki nama kelas, titik koordinat X objek, titik koordinat Y objek, panjang *bounding box*, dan lebar *bounding box*.



Gambar 3 Arsitektur Convolution Neural Network

Sumber: Saleh Albelwi, Ausif Mahmood, *A Framework for Designing the Architectures of Deep Convolutional Neural Networks*, <https://www.mdpi.com/1099-4300/19/6/242>

2.3 Convolution Neural Network

Convolutional Neural Network (CNN) atau ConvNet adalah salah satu kelas *Deep Feed-forward Artificial Neural Network* yang banyak digunakan untuk menganalisis suatu citra. CNN terdiri atas satu lapis masukan (*input layer*), satu lapis keluaran (*output layer*) dan sejumlah lapis tersembunyi (*hidden layer*). Pada lapis tersembunyi terdapat beberapa lapisan yang terdiri dari lapisan ekstraksi fitur dan lapisan klasifikasi (Gambar 3).

2.3.1 Lapisan ekstraksi fitur

Lapisan ekstraksi fitur citra terletak pada awal arsitektur yang tersusun atas beberapa lapisan dan setiap lapisan tersusun atas neuron yang terkoneksi pada daerah lokal (*local region*) lapisan sebelumnya. Lapisan pertama adalah lapisan konvolusi (*convolution layer*) dan lapisan kedua adalah lapisan pooling (*Pooling Layer*). Pada keluaran setiap lapisan digunakan suatu fungsi aktivasi.

2.3.2 Convolution Layer

Tahap awal dalam CNN adalah memproses input yang akan menghasilkan peta fitur yang banyak. Dari input awal berupa citra akan dilakukan proses konvolusi dengan menggunakan filter, dimana filter yang digunakan tergantung citra input yang dimasukkan sebagai input, misalnya : 3x3, 5x5, 7x7. Setelah menentukan filter yang digunakan maka akan ditentukan beberapa banyak piksel perpindahan dari filter yang sudah ditentukan.

2.3.3 ReLU (*Rectified Linear Units*)

Fungsi aktivasi ReLU digunakan untuk menghilangkan linearitas yang ada pada hasil Lapisan Konvolusi (*Convolution Layer*). Fungsi aktivasi ini digunakan untuk menghasilkan keluaran dari Lapisan Konvolusi dan Lapisan Pooling. Rumus untuk fungsi aktivasi ini adalah sebagai berikut:

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (2)$$

Keterangan:

$f(x)$ = fungsi ReLU

x = nilai piksel

Fungsi aktivasi ReLU ini bertujuan mengubah nilai yang berada di bawah 0 menjadi 0 dan yang di atas 0 nilainya tidak berubah. Kelebihan dari ReLU yaitu:

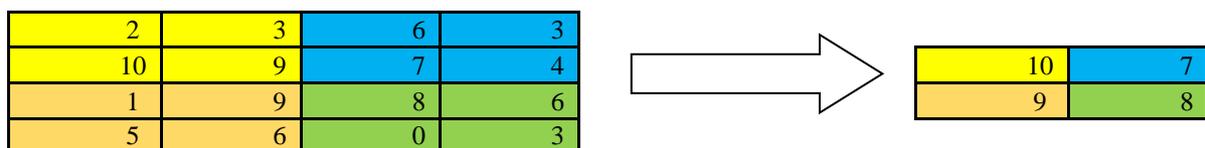
1. Bisa mempercepat gradient stokastik dibandingkan dengan fungsi sigmoid / tanh karena ReLU berbentuk linear.
2. Tidak menggunakan operasi eksponensial seperti sigmoid / tanh, sehingga bisa melakukan dengan pembuatan matriks aktivasi saat ambang batas berada pada nilai 0.

Kekurangan dari ReLU yaitu bisa rapuh saat pelatihan dan mati karena gradient besar yang mengalir melalui ReLU akan menyebabkan update bobot, sehingga neuron tidak aktif pada data point lagi.

2.3.4 Lapisan Pooling

Pada Lapisan Pooling (*Pooling Layer*). peta-peta fitur yang telah diproses melalui ReLU akan dikecilkan dengan memilih ukuran *window*. Dalam Lapisan Pooling, data akan menjadi kecil, mudah dikelola, dan mudah mengontrol *overfitting*.

Cara mengecilkan peta fitur yaitu dengan menggunakan fungsi MaxPooling. MaxPooling yaitu fungsi yang memilih nilai maksimum dari daerah *window* tersebut lalu di representasikan sebagai piksel baru. MaxPooling bergerak dari pojok kiri atas kekanan dengan ukuran Pooling yang telah ditentukan dan pergeseran Pooling yang telah ditentukan.



Gambar 4 Proses Pooling

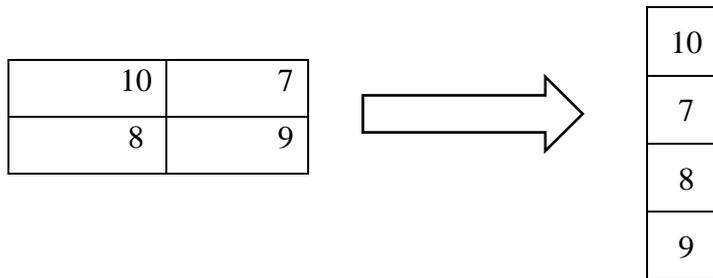
2.3.5 Lapisan Klasifikasi

Lapisan Klasifikasi (*Classification Layer*) digunakan untuk mengklasifikasi tiap neuron yang telah diekstraksi fitur pada lapisan ekstraksi fitur. Lapisan ini terdiri dari *Flatten* dan *Fully-Connected Layer*.

a. Flatten

Flatten mengubah nilai dari sebuah matriks dua dimensi menjadi matriks satu dimensi atau biasa disebut vektor. Array ini disusun dari atas kiri matriks ke kanan

setelah itu pindah ke baris selanjutnya matriks dimulai dari kiri dan seterusnya. Vektor ini akan digunakan sebagai input dari *fully-connected layer*.



Gambar 5 Perubahan matrik 2 dimensi menjadi vektor

b. *Fully Connected Layer*

Fully Connected Layer menggabungkan proses dari *Convolution*, *ReLU*, *Pooling*, dan *Flatten*. Hasil dari *flatten* dimasukan satu persatu, kemudian akan dikalikan dengan bobot w dimana nilai w adalah nilai random yang dinilainya berada di antara 0-1.

$$\hat{y} = \sum_{i=1}^m W_i X_i \quad (3)$$

Keterangan:

\hat{y} = hasil akhir

W_i = bobot ke-i

X_i = nilai input ke-i

2.3.6 Softmax

Softmax berfungsi untuk menghitung probabilitas dari setiap kelas target atas semua kelas target yang memungkinkan dan akan membantu untuk menentukan kelas target untuk input yang diberikan. Dari perhitungan tahap *Fully-Connected Layer* didapatkan beberapa nilai \hat{y} yang akan diperjelas hasilnya dengan melakukan proses *softmax*.

$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}} \quad (4)$$

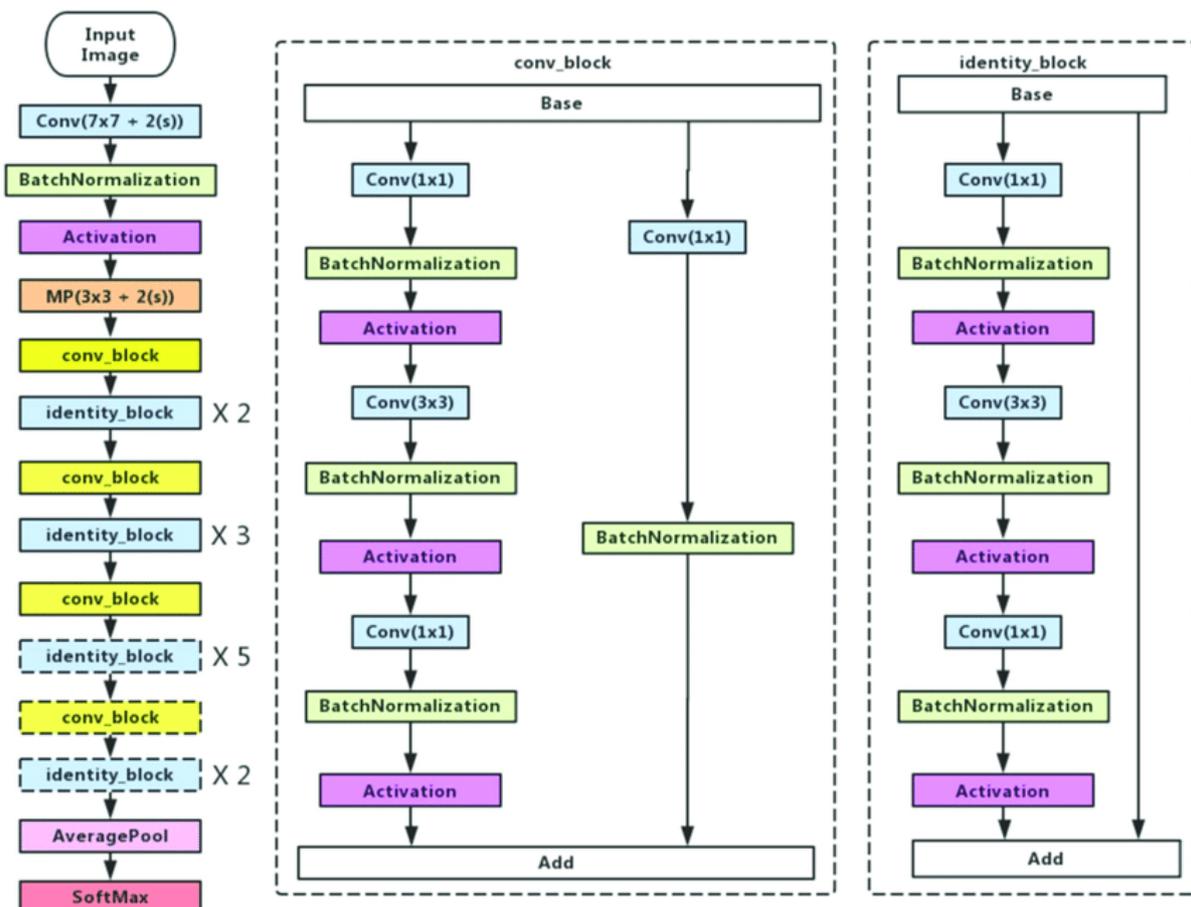
Keterangan:

$f_j(z)$ = fungsi *SoftMax*

e_j = nilai \hat{y} ke- j

2.3.7 Residual Neural Network

Residual Neural Network (ResNet) adalah sebuah arsitektur *Convolutional Neural Network* yang diperkenalkan oleh Kaiming He dalam perlombaan *ImageNet Large Scale Visual Recognition Competition* (ILSVRC) pada tahun 2015.¹⁴ ResNet memiliki tingkat akurasi yang baik karena memiliki *residual block* yang berfungsi untuk mencegah *vanishing gradient problem* yang mengakibatkan nilai *gradient* yang dicari menjadi kecil sehingga menjadi tidak akurat. Arsitektur dari ResNet50 dapat dilihat pada **Gambar 6**.



Gambar 6 Arsitektur ResNet50

Sumber : Qingge Ji, Optimized Deep Convolutional Neural Networks for Identification of Macular Diseases from Optical Coherence Tomography Images, https://www.researchgate.net/figure/Left-ResNet50-architecture-Blocks-with-dotted-line-represents-modules-that-might-be_fig3_331364877

BAB 3

METODE PENELITIAN

3.1 Metode yang digunakan

Pada penelitian ini, akan dibuat sebuah program aplikasi yang akan melakukan operasi pendeteksian dan pengenalan citra ekspresi wajah manusia dengan algoritma *You Only Look Once* (YOLO) dan *Convolutional Neural Network* (CNN) berbasis website. Mula-mula sistem akan dilatih menggunakan citra latih ekspresi wajah yang sudah ada labelnya. Pelatihan dilakukan supaya sistem dapat mendeteksi ekspresi wajah pada saat pengujian. Citra latih ekspresi wajah tersebut diinputkan ke program aplikasi. Selanjutnya program aplikasi akan melakukan operasi *grayscale*, normalisasi ukuran supaya citra yang dimasukan sesuai dengan syarat ukuran citra. Algoritma Yolo akan mendeteksi wajah secara langsung lalu citra ekspresi wajah akan dikenali menggunakan CNN. Flowchart Perancangan Aplikasi Sistem Pengenalan Wajah untuk mengenali ekspresi wajah ini dapat dilihat pada gambar 7.

Pada proses pengenalan dari hasil pendeteksian wajah, Citra yang digunakan dalam tahapan ini adalah hasil dari proses sebelumnya. Adapun langkah-langkah dalam proses ini adalah:

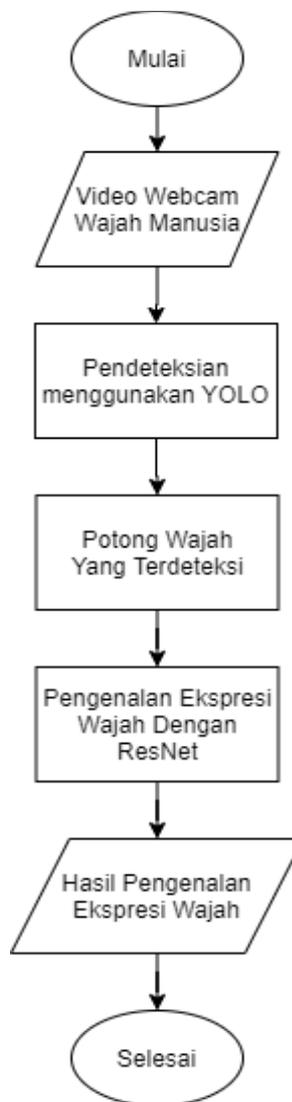
1. Normalisasi ukuran citra.
2. Menghitung boot pada *convolution layer*.
3. Menghilangkan linearitas yang ada pada hasil *convolution layer*.
4. Mengecilkan peta fitur (*polling*).
5. Mengubah nilai dari sebuah matriks dua dimensi menjadi matriks satu dimensi (*flatten*).
6. Menggabungkan semua proses diatas (*Fully Connected Layer*).
7. Untuk proses pelatihan, data hasil metode *Fully Connected Layer* akan disimpan dalam basis data. Data tersebut akan digunakan sebagai basis data pada proses pengujian.
8. Untuk proses pengujian, data hasil metode *Fully Connected Layer* akan dilanjutkan dengan proses *softmax*, hingga akhirnya didapatkan data hasil pengenalan ekspresi wajah.

Diagram Hirarki yang bertujuan untuk memberikan gambaran umum mengenai program aplikasi untuk mengenali ekspresi wajah ini dapat dilihat pada gambar 9.

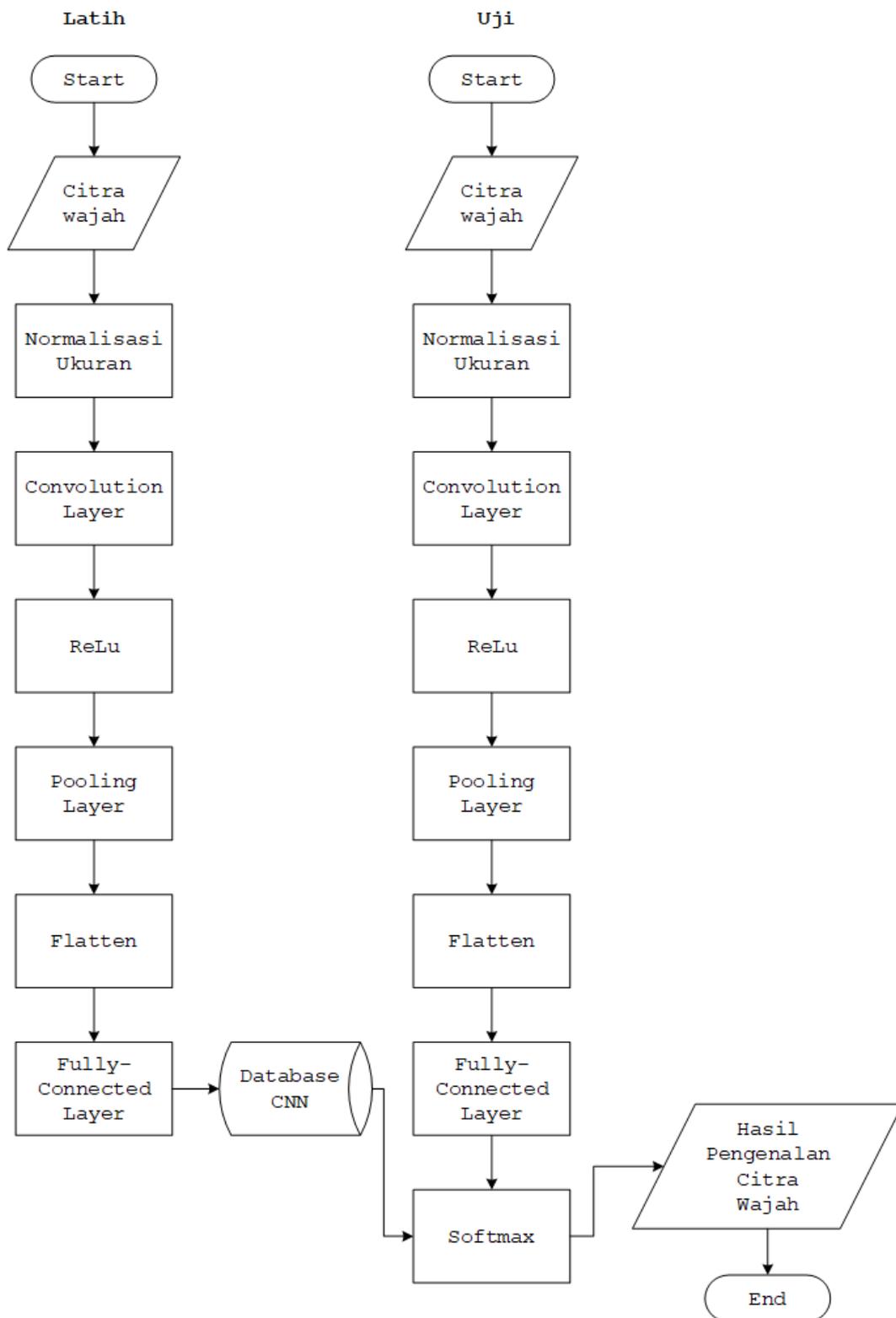
3.2 Manfaat Penelitian

Hasil Penelitian “Sistem Pendeteksian dan Pengenalan Ekspresi Wajah dengan Algoritma Yolo dan Convolutional Neural Network” ini, memiliki manfaat antara lain:

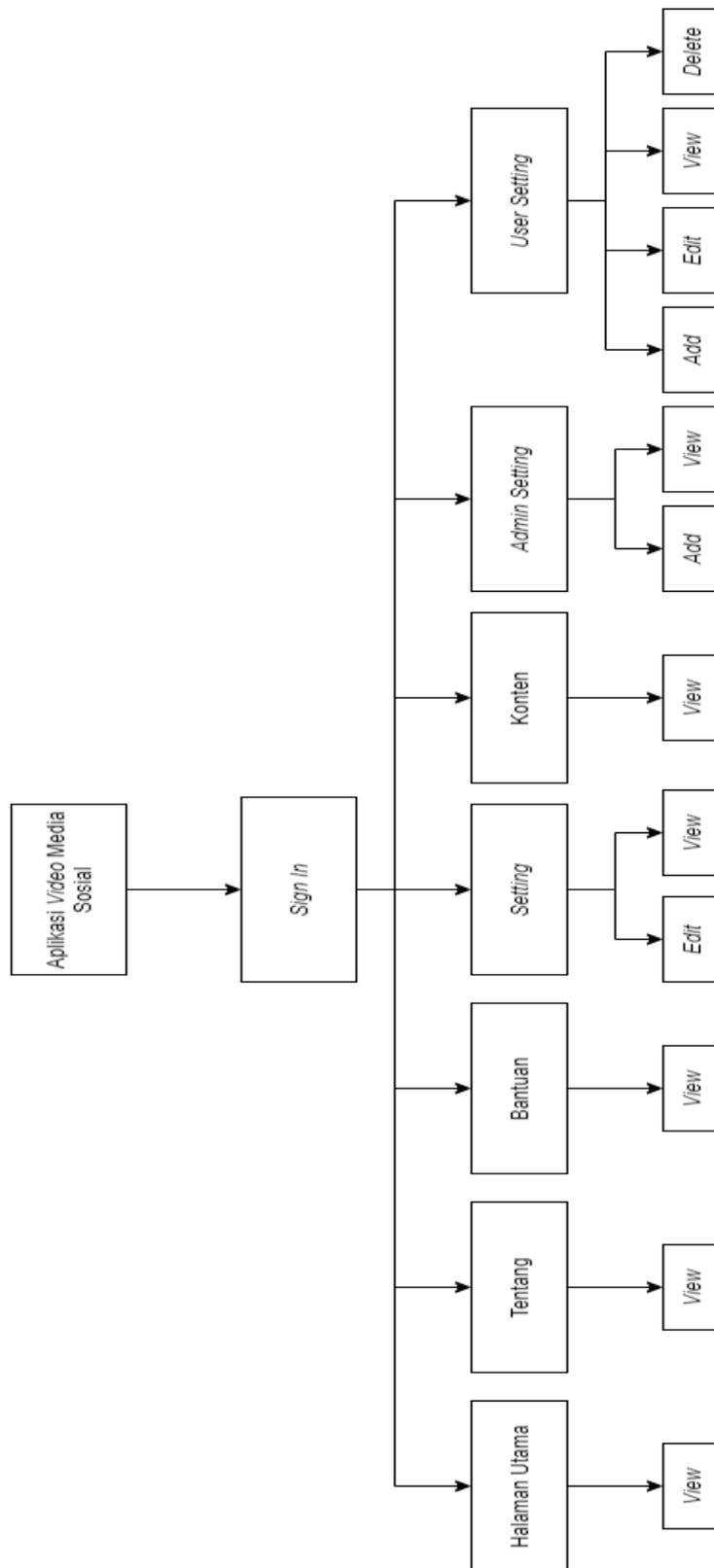
- Program Aplikasi Sistem Pendeteksian dan Pengenalan Ekspresi Wajah dengan Algoritma Yolo dan Convolutional Neural Network yang dihasilkan dari penelitian ini dapat dimanfaatkan untuk mengenali ekspresi wajah seseorang dari suatu video secara cepat dan tepat. Misalnya, untuk pengenalan ekspresi wajah pada penentuan respon saat melihat youtube.
- Hasil penelitian ini dapat dipublikasikan pada Seminar atau Jurnal Nasional/ Internasional.



Gambar 7 Flowchart Sistem Program Pengenalan Ekspresi Wajah



Gambar 8 Flowchart metode CNN



Gambar 9 Diagram Hirarki Program Aplikasi

3.3 Bentuk Data yang Digunakan

Data yang digunakan oleh aplikasi berupa data dalam bentuk citra dua dimensi. Data tersebut dibagi menjadi 2 yaitu citra latih dan citra validasi. Citra latih adalah citra yang digunakan untuk melatih model dalam algoritma YOLO dan untuk melatih model CNN. Citra validasi adalah citra yang digunakan untuk melakukan evaluasi model CNN yang telah dilatih.

Citra latih untuk algoritma YOLO berupa citra wajah manusia yang sudah diberi anotasi. Citra latih dan validasi untuk metode CNN berupa citra ekspresi wajah manusia yang terdiri dari 7 kelas dan setiap kelas memiliki jumlah yang berbeda-beda. Data citra didapatkan dari website Kaggle yang bernama fer2013. Selain data citra, terdapat juga data bobot untuk algoritma YOLO dan metode CNN. Format bobot dari algoritma YOLO berformat *.weights* dan model CNN berformat *.h5*.

BAB 4

Hasil dan Pembahasan

4.1 Metode Pengujian

Proses pengujian untuk rancangan sistem aplikasi pendeteksian wajah menggunakan algoritma YOLO dan pengenalan citra ekspresi wajah manusia menggunakan metode CNN dibagi menjadi dua bagian, yaitu pengujian tampilan antarmuka dan pengujian sistem program dengan algoritma YOLO dan metode CNN. Pengujian tampilan antarmuka dilakukan dengan metode Black Box Testing. Pengujian sistem program algoritma YOLO dan metode CNN dilakukan dengan cara mendeteksi lokasi wajah manusia pada citra dan mengenali ekspresi wajah yang ada pada citra wajah manusia tersebut. Selain itu pengujian program juga akan dilakukan dengan menjalankan aplikasi menggunakan beberapa data citra dan menganalisa hasil yang diberikan oleh system.

4.1.1 Pengujian *Black Box Testing*

Black Box Testing atau yang sering dikenal dengan sebutan pengujian fungsional merupakan metode pengujian perangkat lunak yang digunakan untuk menguji perangkat lunak tanpa mengetahui struktur internal kode atau program. Black Box Testing berisi keluaran pada program yang dihasilkan dari sistem yang sesuai dengan apa yang pengguna inginkan. Pengujian menggunakan Black Box Testing dilakukan dari proses awal sistem dijalankan hingga akhir dan melihat keluaran yang dihasilkan.

Pada modul halaman utama dilakukan pengujian dengan melihat apakah keluaran sudah sesuai dengan apa yang diharapkan. Pada modul Register dilakukan pengujian dengan membuat akun pengguna baru. Pada modul Login dilakukan pengujian dengan beberapa kondisi seperti memasukkan data pengguna dengan benar dan data pengguna salah.

Pada modul Upload Video dilakukan pengujian dengan membuat konten baru oleh pengguna terdaftar. Pada modul Edit Video dilakukan pengujian dengan mengubah informasi konten pengguna. Pada modul My Video dilakukan pengujian dengan melihat apakah keluaran konten sudah sesuai dengan apa yang pengguna buat.

Pada modul Search dilakukan pengujian dengan melihat apakah keluaran konten sudah sesuai dengan apa yang pengguna cari. Pada modul Konten Modal PopUp dilakukan pengujian dengan menggunakan masukkan citra wajah menggunakan webcam kepada salah satu konten yang tersedia dan melihat teks prediksi pengenalan ekspresi wajah yang dihasilkan dan tampilan jumlah ekspresi penonton dalam konten tersebut. Pada modul Setting dilakukan

pengujian dengan mengubah informasi pengguna. Hasil pengujian black box testing dapat dilihat pada Lampiran 1.

4.1.2 Pengujian Algoritma YOLO

Pengujian untuk algoritma YOLO yang dilatih dengan jumlah epoch sebanyak 150 epoch dan menggunakan 1200 data citra berupa wajah manusia dan anotasinya. Pengujian dilakukan kepada 28 citra wajah. Citra wajah tersebut diambil dari 4 orang berbeda secara manual menggunakan webcam. Pada proses pengujian dibutuhkan file .h5, .txt, dan data citra agar proses pengujian dapat dilakukan dengan baik. Pengujian dilakukan dengan menjalankan kode program yang telah dibuat menggunakan library *Keras*.

Keras merupakan interface library yang dibangun untuk mensesederhanakan implementasi algoritma-algoritma *Deep Learning* di atas *Tensorflow*. *Tensorflow* sendiri merupakan *platform high performance computing* berbasis alur *graph*. *Library keras* dapat digunakan untuk merancang algoritma YOLO. Hasil keluaran dari algoritma YOLO adalah berupa lokasi wajah pada citra yang dideteksi dan tingkat keyakinannya.

4.1.3 Pengujian Algoritma CNN

Pengujian untuk model CNN dilakukan terhadap model yang sudah dilatih dengan jumlah epoch sebanyak 200 epoch dan menggunakan 35.887 data citra berupa ekspresi wajah manusia yang didapatkan dari *website* Kaggle dengan nama *fer2013* menggunakan arsitektur ResNet50. Detil jumlah data pada setiap kelas untuk proses pelatihan dapat dilihat pada **Tabel 1**.

Pengujian dilakukan kepada 28 citra wajah. Citra wajah tersebut diambil dari 4 orang berbeda secara manual menggunakan *webcam*. Masing-masing wajah yang digunakan akan dibagi menjadi 6 kelas ekspresi wajah yang berbeda-beda. Proses pengujian model CNN menghasilkan tingkat keyakinan rata-rata sebesar 90% dalam mengenali ekspresi wajah pada masing-masing citra. Detil jumlah data pada setiap kelas untuk proses pengujian dapat dilihat pada **Tabel 2**.

Tabel 1 Detil Jumlah Data Untuk Proses Pelatihan

Nama Kelas	Jumlah Data	Nama Kelas
0	4.953	Angry
1	547	Disgust
2	5.121	Fear
3	8.989	Happy
4	6.077	Sad
5	4.002	Shocked
6	6.198	Neutral

Tabel 2 Detil Jumlah Data Untuk Proses Pengujian

Nama Kelas	Jumlah Data	Nama Kelas
0	4	Angry
1	4	Disgust
2	4	Fear
3	4	Happy
4	4	Sad
5	4	Shocked
6	4	Neutral

4.1.4 Pengujian Keseluruhan Sistem

Pengujian keseluruhan sistem dilakukan kepada 4 orang dengan mencobakan aplikasi yang sudah dikembangkan. Masing-masing orang akan menonton 6 video yang telah disediakan dan sistem akan mengambil citra wajah masing-masing pengguna. Citra wajah yang diambil adalah yang berhasil terdeteksi oleh algoritma YOLO untuk menghasilkan lokasi dari wajah pengguna. Setelah itu, citra wajah akan dipotong berdasarkan lokasi wajah hasil perhitungan algoritma Yolo.

Setelah proses pemotongan selesai, citra baru dari hasil pemotongan akan dilakukan *pre-processing* sehingga menjadi *grayscale*. Citra yang sudah diubah menjadi *grayscale* lalu diubah ukurannya menjadi 48x48 piksel sesuai dengan masukkan model CNN. Setelah dilakukan normalisasi dilakukan pengenalan ekspresi wajah dengan model CNN menggunakan arsitektur ResNet50 yang sudah dilatih sebelumnya.

Hasil dari pengenalan menggunakan model CNN lalu dimasukkan ke dalam basis data. Setiap ekspresi yang dikenali akan menambah 1 poin pada masing-masing kelas ekspresi dalam basis data.

4.2 Hasil dan Pembahasan Pengujian

4.2.1 Pengujian Algoritma Yolo

Pada proses pengujian terhadap algoritma Yolo, dari 20 citra yang digunakan, sistem berhasil mendeteksi bagian wajah manusia secara tepat untuk seluruh citra. Citra yang digunakan adalah citra yang diperoleh dari hasil pengambilan secara langsung menggunakan *webcam*. Tingkat keyakinan dari hasil yang didapatkan pun sangat tinggi, yakni dengan rata-rata 85%. Hasil pengujian secara detail dapat dilihat pada **Lampiran 2**. Hasil ini sudah sangat baik, akan tetapi tingkat keyakinan ini bisa saja diperbesar jika jumlah data latih diperbanyak.

4.2.2 Hasil Pengujian Algoritma CNN

Hasil pengujian untuk mengenali ekspresi wajah manusia menggunakan CNN, mendapatkan hasil yang kurang baik. Model belum dapat mengenali beberapa ekspresi wajah dengan benar. Terdapat 28 tes yang dilakukan, dari tes tersebut didapatkan hasil hanya sebagian saja yang berhasil. Hasil pengujian menunjukkan rata-rata nilai akurasi sebesar 46,42%. Hasil pengujian secara detail dapat dilihat pada **Lampiran 3**.

Nilai akurasi yang kecil ini mungkin terjadi karena tidak seimbangnya jumlah data latih dan data uji untuk masing-masing kelas. Selain itu, algoritma CNN ini juga membutuhkan jumlah data yang besar, sehingga seharusnya nilai akurasi ini bisa diperbaiki apabila jumlah data diperbesar dan merata untuk seluruh kelas.

4.2.3 Hasil Pengujian Keseluruhan Sistem

Setelah dilakukan pengujian keseluruhan sistem, hasil yang didapatkan masih kurang baik karena data citra dengan kelas jijik, kaget, dan takut tidak dapat dikenali dengan sempurna. Hal itu dikarenakan oleh ketidakseimbangannya jumlah data latih yang dimiliki pada masing-masing kelas ekspresi wajah. Hasil pengujian keseluruhan sistem dapat dilihat pada Tabel 3.

Tabel 3. Hasil Pengujian Keseluruhan Sistem

Video	Happy	Sad	Shocked	Disgust	Angry	Fear	Neutral	Kelas Dominan Yang Diharapkan
1	44	226	0	0	9	14	395	<i>Happy</i>
2	1	153	0	0	36	8	502	<i>Sad</i>
3	34	168	0	1	39	53	320	<i>Shocked</i>
4	24	114	0	1	20	85	456	<i>Angry</i>
5	16	219	0	0	41	12	468	<i>Fear</i>
6	12	118	1	0	83	37	333	<i>Disgust</i>

Seluruh fungsi pada sistem dari hasil pengujian keseluruhan sistem berjalan dengan lancar. Proses pengambilan citra menggunakan *webcam* berjalan dengan lancar. Proses pemotongan citra wajah yang sudah dideteksi lokasinya oleh algoritma YOLO berjalan dengan lancar. Proses mengubah citra menjadi *greyscale* dan mengubah ukuran menjadi 48x48 piksel pada citra wajah yang sudah dipotong berjalan dengan lancar sehingga proses pengenalan menggunakan metode CNN dapat berjalan dengan baik. Memasukkan poin ekspresi ke basis data berjalan dengan baik.

BAB 5

Kesimpulan dan Saran

5.1 Kesimpulan

Dari serangkaian penelitian yang telah dilaksanakan, didapatkan kesimpulan sebagai berikut:

1. Model YOLO yang dilatih menggunakan jumlah epoch sebanyak 150 epoch dengan tingkat keyakinan rata-rata sebesar 85%. Dengan jumlah epoch sebanyak ini, YOLO dapat mendeteksi wajah pada citra dengan baik. Citra wajah yang diambil pun memiliki beragam ekspresi dan sudut.
2. Model CNN berarsitektur ResNet50 yang digunakan dalam penelitian ini memiliki performa yang kurang baik. Hal ini dibuktikan dengan akurasi model yang hanya sebesar 46,42%. Faktor yang menyebabkan model tidak dapat mengenali karakter dengan benar yaitu ketidakseimbangan jumlah data latih pada kelas data.
3. Secara keseluruhan sistem hanya 3 konten yang mendapatkan hasil yang sesuai dengan kelas ekspresinya. Semua konten mewakili 6 kelas ekspresi yaitu ekspresi marah, jijik, takut, senang, sedih, dan terkejut.

5.2 Saran

Saran yang dapat diberikan untuk pengembangan serta perbaikan sistem aplikasi secara keseluruhan adalah sebagai berikut:

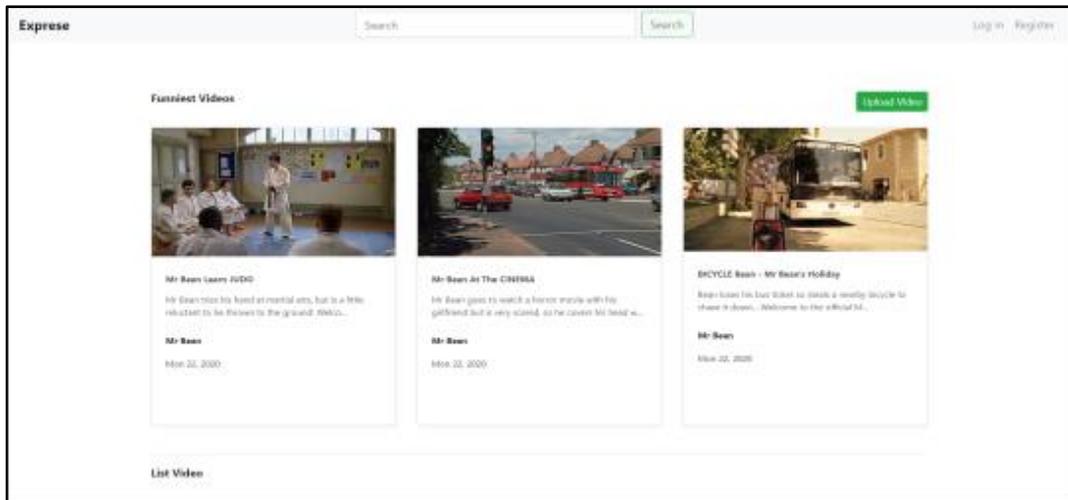
1. Untuk proses deteksi wajah dapat menggunakan algoritma lain yang lebih ringan dikarenakan algoritma YOLO membutuhkan spesifikasi komputer yang cukup tinggi.
2. Untuk pengembangan model CNN dapat menambah jumlah data latih untuk kelas tertentu yang masih sedikit.

Daftar Pustaka

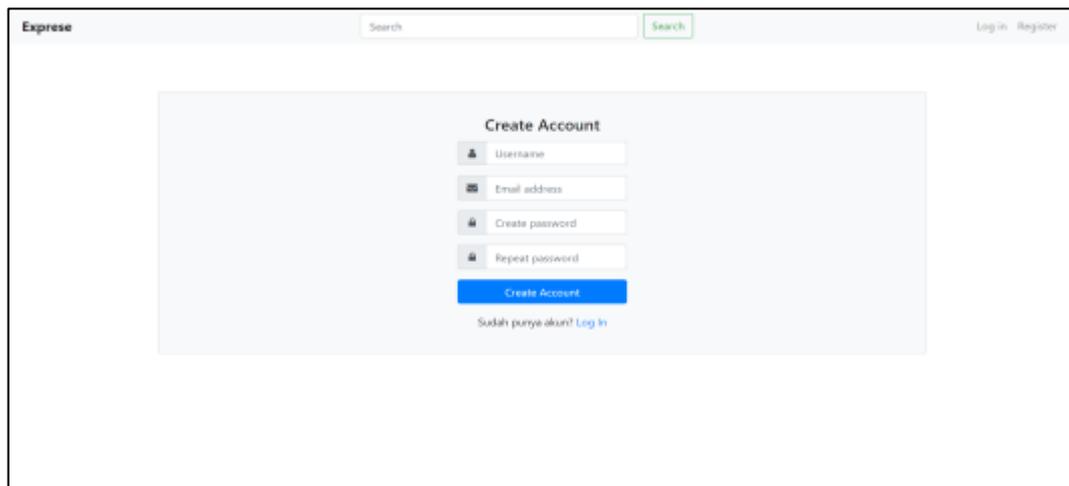
- [1] Albelwi, Saleh; and Mahmood, Ausif. A Framework For Desaigning the Architectures of Deep Convolutional Neural Networks. <https://www.mdpi.com/1099-4300/19/6/242>. 24 Mei 2017.
- [2] Das, Siddharth. CNN Architectures:LeNet, AlexNet, VGG, Google Net, ResNet and more <https://medium.com/@sidereal/cnnsarchitectures-lenet-alexnet-vgg-googlenet-resnet-and-more666091488df5>. 7 September 2019.
- [3] Dataaspirant. Difference Between Softmax Function And Sigmoid Function. <https://dataaspirant.com/2017/03/07/difference-betweensoftmax-function-and-sigmoid-function/>. 18 Januari 2019.
- [4] Hallstrom, Erik. BackPropagation from the beginning. <https://medium.com/@erikhallstrm/backpropagation-from-thebeginning-77356edf427d>. 18 Januari 2019.
- [5] Hollemans, Matthijs. Real-Time Object Detection With YOLO. <http://machinethink.net/blog/object-detection-with-yolo>. 2 Oktober 2019.
- [6] Ji, Qingge. Optimized Deep Convolutional Neural Networks for Identification of Macular Diseases from Optical Coherence Tomography Images. https://www.researchgate.net/figure/Left-ResNet50-architecture-Blocks-with-dotted-line-represents-modules-that-might-be_fig3_331364877. Januari 2019.
- [7] Maj, Michal. What is object detection? Introduction to YOLO algorithm. <https://appsilon.com/object-detection-yolo-algorithm/>. 22 Agustus 2018.
- [8] Rosebrock, Adrian. YOLO object detection with OpenCV. <https://www.pyimagesearch.com/2018/11/12/yolo-object-detection-with-opencv/>. 12 November 2018.
- [9] Saha, Sumit. A Comprehensive Guide To Convolutional Neural Networks – the ELI5 way. <https://towardsdatascience.com/acomprehensive-guide-to-convolutional-neural-networks-the-eli5way-3bd2b1164a53>. 3 September 2019.
- [10] Sena, Samuel. Pengenalan Deep Learning Part 7 : Convolutional Neural Network (CNN). <https://medium.com/@samuelsena/pengenalan-deep-learning-part-7-convolutional-neural-network-cnn-b003b477dc94>. 13 November 2017.
- [11] Sharma. Activation Functions: Neural Networks. <https://towardsdatascience.com/activation-functions-neuralnetworks-1cbd9f8d91d6>. 18 Januari 2019.

Lampiran 1

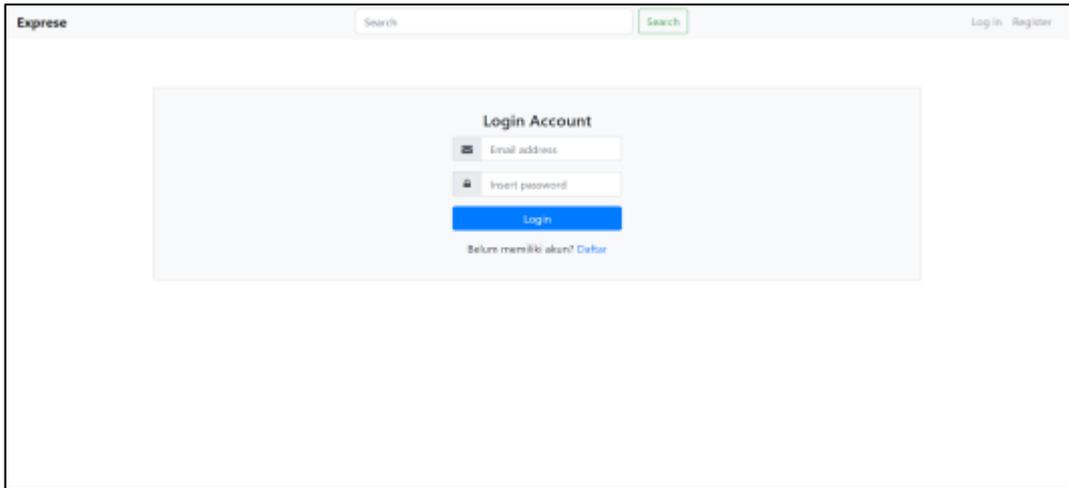
Hasil Black Box Testing



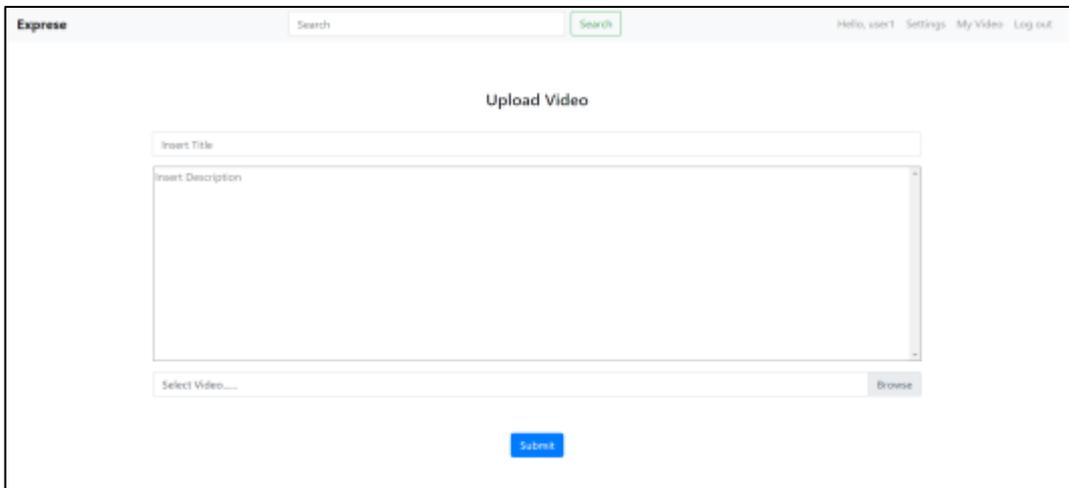
Gambar 10 Tampilan Antarmuka Program Modul Halaman Utama



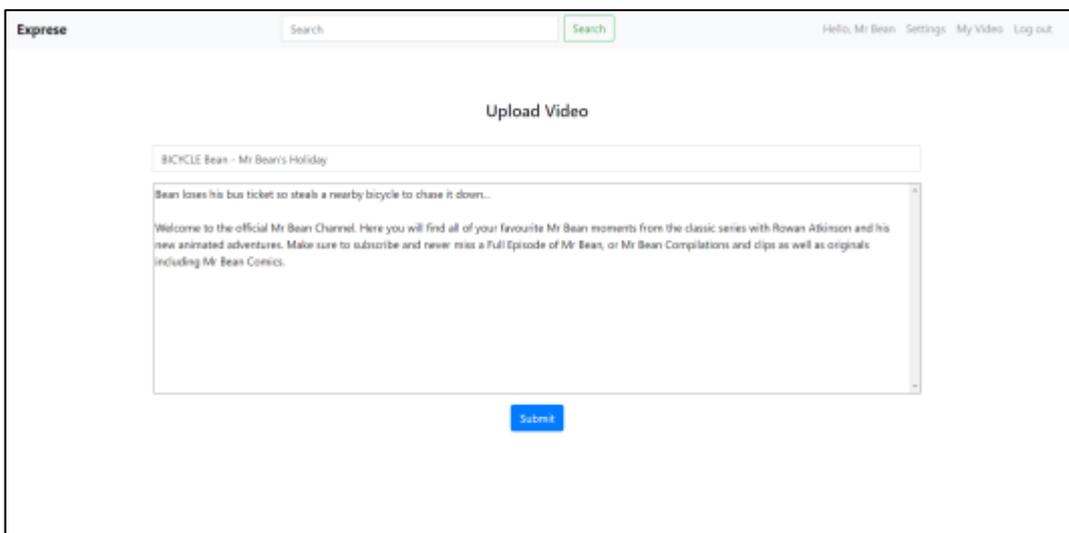
Gambar 11 Tampilan Antarmuka Program Modul Register



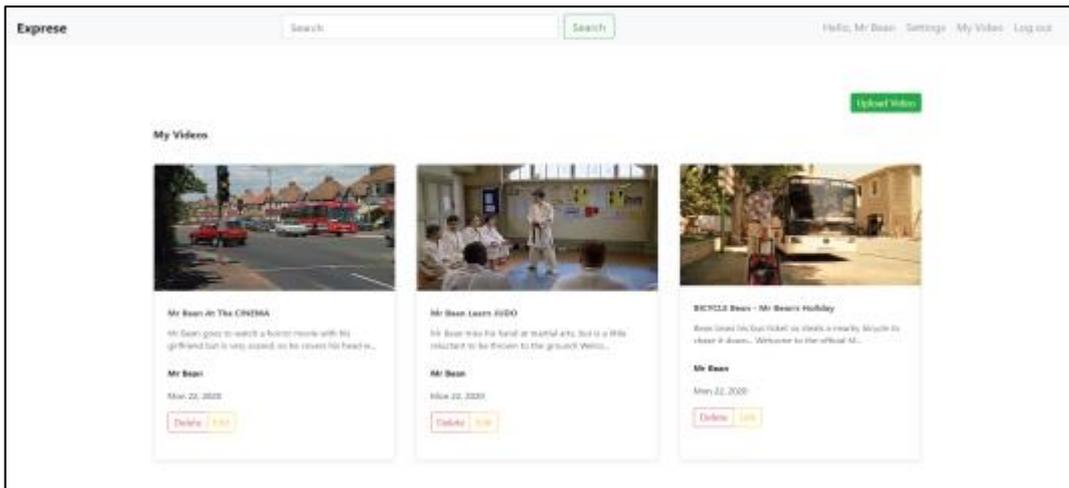
Gambar 12 Tampilan Antarmuka Program Modul Login



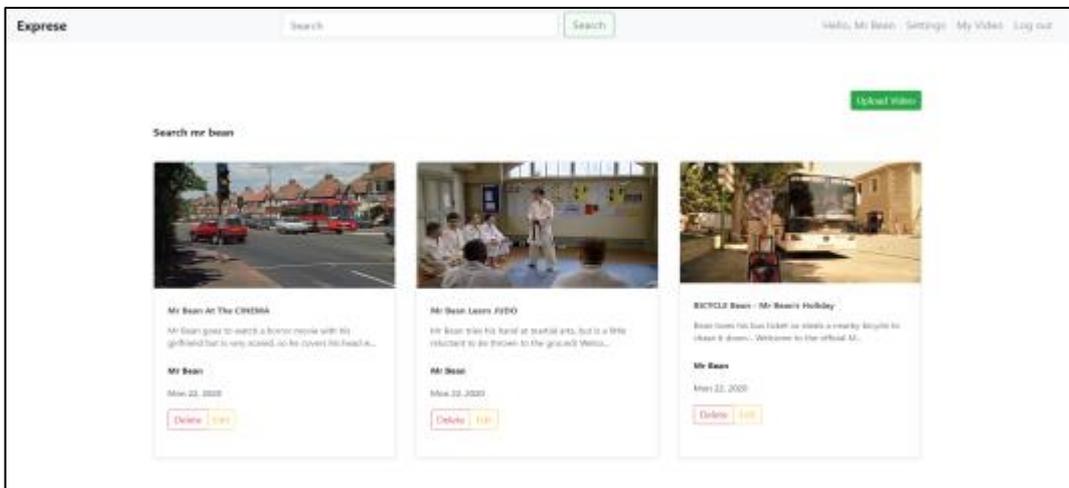
Gambar 13 Tampilan Antarmuka Program Modul Upload Video



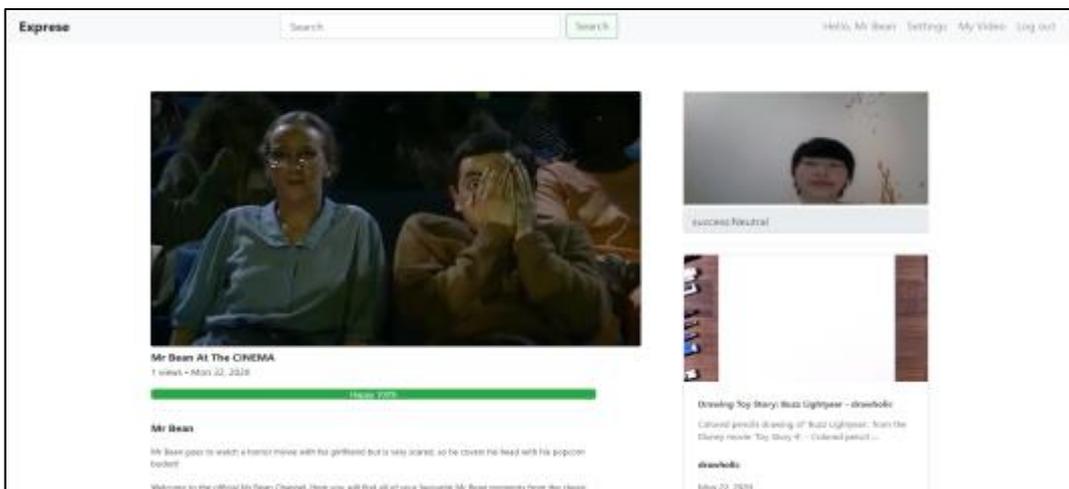
Gambar 14 Tampilan Antarmuka Program Modul Edit Video



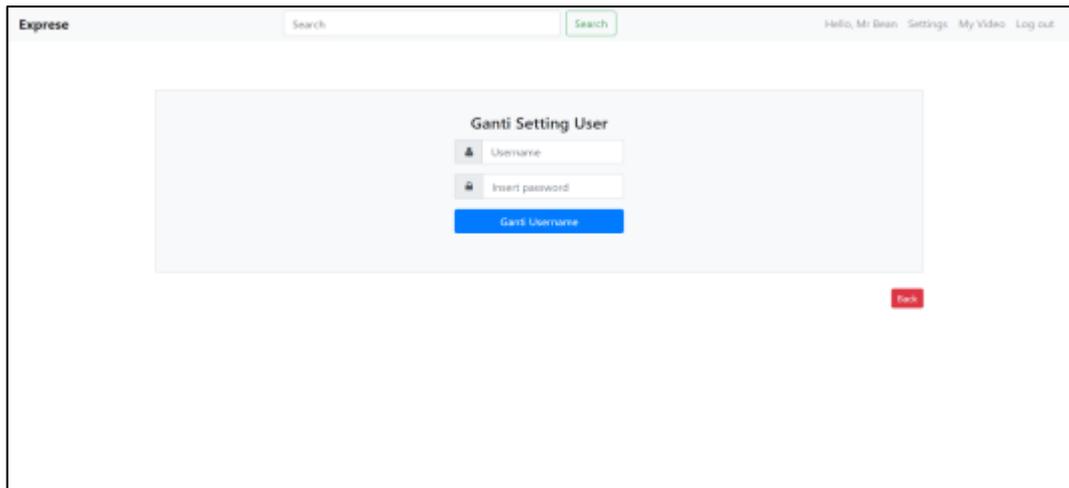
Gambar 15 Tampilan Antarmuka Program Modul My Video



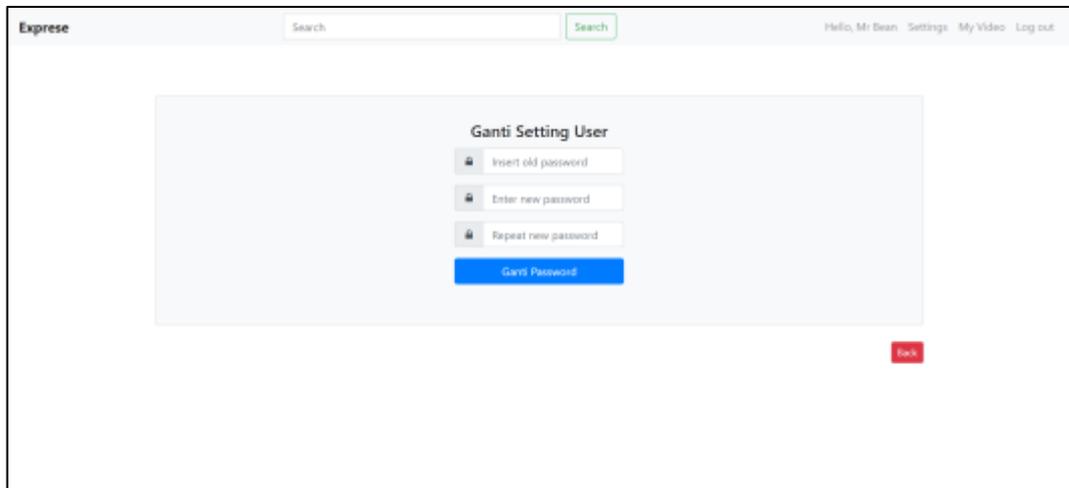
Gambar 16 Tampilan Antarmuka Program Modul Search



Gambar 17 Tampilan Antarmuka Program Modul Konten Modal PopUp



Gambar 18 Tampilan Antarmuka Program Fitur *Setting Username*



Gambar 19 Tampilan Antarmuka Program Fitur *Setting Password*

Lampiran 2
Hasil Pengujian Algoritma YOLO

Tabel 4 Hasil Pengujian YOLO

No	Hasil Pengujian	Tingkat Keyakinan
1	 A close-up photograph of a man's face with a neutral expression, looking directly at the camera.	75,63%
2	 A close-up photograph of the same man's face with a frowning expression, showing his teeth.	87,26%
3	 A close-up photograph of the same man's face with a frowning expression, looking slightly to the side.	82,61%
4	 A close-up photograph of the same man's face with a smiling expression, showing his teeth.	71,74%
5	 A close-up photograph of the same man's face with a neutral expression, looking directly at the camera.	79,93%

Tabel 4 (lanjutan)

6		54,44%
7		91,57%
8		95,59%
9		99,98%
10		94,19%

Tabel 4 (lanjutan)

11		79,30%
12		64,60%
13		98,26%
14		95,84%
15		77,63%

Tabel 4 (lanjutan)

16		77,59%
17		94,35%
18		90,21%
19		94,80%
20		84,36%

Tabel 4 (lanjutan)

21		97,49%
22		83,63%
23		86,81%
24		77,66%
25		92,88%
26		87,45%

Tabel 4 (lanjutan)

27				80,44%
28				84,88%

Lampiran 3
Hasil Pengujian Algoritma CNN

Tabel 5 Hasil Pengujian Model CNN

No	Hasil Pengujian	Hasil Yang Diharapkan	Hasil Pengujian
1		Angry	Angry
2		Disgust	Happy
3		Fear	Happy
4		Happy	Happy
5		Neutral	Neutral

Tabel 5 (Lanjutan)

6		Sad	Sad
7		Shocked	Shocked
8		Angry	Sad
9		Disgust	Happy
10		Fear	Sad

Tabel 5 (Lanjutan)

11		Happy	Happy
12		Neutral	Neutral
13		Sad	Sad
14		Shocked	Shocked
15		Angry	Angry

Tabel 5 (Lanjutan)

16		Disgust	Neutral
17		Fear	Happy
18		Happy	Happy
19		Neutral	Neutral
20		Sad	Happy

Tabel 5 (Lanjutan)

21		Shocked	Happy
22		Angry	Sad
23		Disgust	Anger
24		Fear	Happy
25		Happy	Happy
26		Neutral	Sad

Tabel 5 (Lanjutan)

27			Sad	Sad
28			Shocked	Shocked