

PENGENALAN EKSPRESI WAJAH BERBASIS *CONVOLUTIONAL NEURAL NETWORK* MENGGUNAKAN ARSITEKTUR *RESIDUAL NETWORK-50*

FACIAL EXPRESSION RECOGNITION BASED ON CONVOLUTIONAL NEURAL NETWORK USING RESIDUAL NETWORK 50 ARCHITECTURE

Sena Pahlevi Ristiawanto¹, Budhi Irawan², Casi Setianingsih³

^{1,2,3} Universitas Telkom, Bandung

senapahlevi@student.telkomuniversity.ac.id¹, budhiirawan@telkomuniversity.ac.id²,
setiacasie@telkomuniversity.ac.id³

Abstrak

Emosi adalah suatu kondisi dimana ada faktor faktor yang dampak pada perilaku manusia selain itu emosi bisa menggambarkan ekspresi dari manusia itu sendiri. Ekspresi wajah manusia juga berkaitan dengan emosi karena dikatakan sebagai bentuk komunikasi nonverbal. Banyak penelitian terkait ekspresi wajah diantara nya milik Tang et al dan Gupta et al terkait pengenalan ekspresi wajah menggunakan dataset dari website Kaggle yakni FER2013 menggunakan *Residual Network* (Resnet-50). Sistem yang dirancang menggunakan arsitektur CNN seperti dari penelitian sebelumnya yakni menggunakan *Residual Network* (Resnet-50 *state of the art*) dan *Residual Network* (Resnet-50) modifikasi dengan menggunakan pengujian berbeda menggunakan baik dari pengujian augmentasi dan nonaugmentasi. Dari dua arsitektur tersebut tanpa *transfer learning*. Kemudian langkah selanjutnya membandingkan pengujian hasil dari kedua arsitektur hasil yang jauh lebih baik dengan penelitian milik Tang et al dan Gupta et al. Model yang telah dirancang didapatkan bahwa menggunakan augmentasi dan Resnet-50 modifikasi terlihat jauh lebih baik dengan akurasi pelatihan sebesar 64.28%, tes sebesar 60.71%, validasi sebesar 60.49%. Hasil model penulis jauh lebih bagus sekitar 8% sebesar 64.28% dibandingkan 56.21 milik Gupta et al dan 12 % lebih bagus dibandingkan milik Tang et al 52.08%.

Abstract

Emotion is a condition where there are factors that have an impact on human behavior other than that emotion can describe the expression of the human itself. Human facial expressions are also related to emotions because they are said to be a form of nonverbal communication. Many studies related to facial expressions, including Tang et al and Gupta et al, related to facial expression recognition using a dataset from Kaggle, namely FER2013 using Residual Network (Resnet-50). The system is designed using the CNN architecture as from previous studies, namely using Residual Network (Resnet-50 state of the art) and Residual Network (Resnet-50) modifications by using different tests using both augmentation and nonaugmentation (parameters) tests. Of the two architectures without transfer learning. Then the next step is to compare the results of the two architectures with significantly better results with those of Tang et al and Gupta et al. The designed model, it was found that using augmentation and modified Resnet-50 looks much better with training accuracy of 64.28%, test of 60.71%, validation of 60.49%. The results of the author's model are much better about 8% by 64.28% compared to Gupta et al's 56.21% and 12% better than Tang et al's 52.08%.

Keywords: CNN, FER2013, *Residual Network*.

1. Pendahuluan

Emosi adalah suatu kondisi dimana ada faktor faktor yang dampak pada perilaku manusia. Ekspresi wajah adalah suatu komunikasi nonverbal berupa wajah atau mimik seseorang dalam

mengungkapkan sesuatu emosi positif maupun emosi negatif [1], [2]. Terkait dengan *facial expressions recognition* penelitian sebelumnya yang dilakukan oleh Tang et al dan Gupta et al menggunakan arsitektur *Convolutional Neural network* (CNN). Pada kedua penelitian tersebut menggunakan model yang dipakai adalah Resnet-50 (*Residual Network*) kemudian menggunakan *dataset* FER2013 sebagai ujicoba nya dalam menguji akurasi. Hasil akurasi yang didapatkan oleh Tang et al memakai model Resnet 50 sebesar 52.08% [3]. Kemudian pada penelitian yang dilakukan oleh Gupta et al menggunakan Resnet 50 dengan pretrained model yang berbeda yakni mendapatkan hasil akurasi sebesar 33.46% menggunakan pretrained ImageNet dan 56.21% menggunakan pretrained VGGface [4]. Alasan banyak penelitian terkait pengenalan ekspresi wajah menggunakan metode berbasis Resnet adalah karena Resnet cocok untuk melakukan klasifikasi dan recognition pada suatu objek gambar selain itu juga Resnet merupakan salah satu arsitektur CNN yang mampu meningkatkan performansi kompleks pada model dengan *layer* yang semakin dalam. Tugas akhir ini bertujuan untuk membandingkan hasil akurasi sebelumnya pada paper dari Tang et al dan Gupta et al mengenai hasil akurasi pada arsitektur Resnet 50 dengan fokus obyek penelitian nya adalah menggunakan *dataset* FER2013 ditinjau dari epoch, *learning rate*, *optimizer* serta nonaugmentasi dan augmentasi kemudian dilihat hasil akurasi dan analisa emosi wajah. Rancangan sistem yang dibuat ekstraksi fitur menggunakan konvolusi dan *Pooling layer* kemudian untuk klasifikasi gambar emosi wajah menggunakan softmax. Diharapkan hasil dari penelitian ini bisa memberikan peningkatan akurasi pelatihan pada *dataset* FER2013 menggunakan arsitektur Resnet-50 yang cukup baik dibandingkan dari penelitian sebelumnya.

2. Dasar Teori

2.1 Facial Expression Recognition

Facial expression recognition merupakan salah satu penerapan pada *deep learning* di bidang pengolahan citra yang berguna untuk mendeteksi emosi manusia menggunakan algoritma *neural network*. Dalam penerapan *facial expression recognition* secara umum untuk mendeteksi emosi langkah yang digunakan adalah dengan cara melakukan dari nilai piksel tiap emosi dan secara urutan untuk mendapatkan informasi dalam *frame dataset* [5], [6].

2.2 Convolutional Neural network (CNN)

Convolutional Neural network merupakan salah satu contoh algoritma *deep learning* dimana arsitekturnya mirip dengan jaringan saraf pada manusia karena memiliki neuron satu sama lain saling terhubung salah satu penerapan CNN ini adalah pada pengolahan citra digital dua dimensi. Secara garis besar arsitektur dari CNN ini adalah mempunyai 3 *layer* secara umum konvolusi *layer*, *Pooling layer*, *Fully connected layer* [7]

2.2.1 Konvolusi layer (CNN)

Konvolusi *layer* merupakan bagian *layer* pada CNN sebagai inti dari sistem untuk komputasi atau perhitungan perkalian pada lebar, tinggi, kedalaman suatu gambar dan kernel. Fungsi dari Konvolusi *layer* ini melakukan *filter* dan feature map pada masukan gambar [8].

2.2.2 Subsampling layer

Subsampling adalah suatu metode yang digunakan pada lapisan konvolusional *layer* pada arsitektur CNN untuk mengurangi ukuran dari spasial gambar dari konvolusional. *Subsampling* biasanya disebut sebagai *Pooling layer* [9].

2.2.3 Fully Connected layer (FC)

Fully connected layer bertujuan menerima masukan berupa flatten data dari keluaran kemudian biasanya format nilai pada neuron akan diubah ke bentuk vektor kolom kemudian data yang telah diolah di *fully connected* data tersebut akhirnya bisa diklasifikasi pada kasus prediksi gambar [10].

2.2.4 Softmax

Softmax merupakan salah satu klasifikasi yang populer digunakan oleh CNN dimana kelas yang memiliki jumlah lebih dari satu dapat diberi nilai probabilitas yang sesuai jika dijumlahkan keseluruhan menghasilkan nilai satu [11].

2.3 SGD optimizer

Algoritma yang digunakan pada *deep learning* untuk *update* parameter pada saat melakukan pelatihan dengan secara terus menerus mengurangi kesalahan model dalam memprediksi sedemikian rupa dan meningkatkan akurasi sehingga bisa terus turun ke gradien paling minimum selain itu SGD sendiri memiliki gambaran umum model dalam *update* bobot yang sangat cocok diterapkan bagi pemula [12].

2.4 ADAM optimizer

Adam *optimizer* merupakan algoritma yang digunakan mengoptimalkan untuk model pada deep learning. Kemampuan Adam *optimizer* adalah membutuhkan memori yang jauh lebih kecil saat melakukan proses komputasi [12].

2.5 Aktivasi fungsi

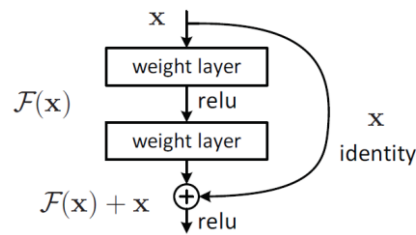
Aktivasi fungsi merupakan salah satu bagian penting pada *neural network* yang berguna untuk menambahkan fungsi nonlinear pada algoritma sehingga model dapat menyelesaikan permasalahan yang kompleks [13].

2.6 Transfer learning

Transfer learning merupakan salah satu proses pada CNN model menggunakan pre-trained atau sebelumnya sudah dilakukan proses pelatihan oleh orang lain atau peneliti dengan menggunakan *dataset* yang sangat besar salah satunya menggunakan *dataset* imagenet [8], [4].

2.7 Residual Network (Resnet)

Residual Network merupakan network buatan pada CNN yang melakukan tugas untuk mengatasi vanish gradient pada arsitektur *Residual Network*. Resnet menggunakan sistem skip sehingga melewati bagian komputasi perkalian pada model sehingga sampai ke bagian *layer* pertama sehingga bisa menghindari kehilangan informasi penting dari gambar. *Residual Network* merupakan network buatan pada CNN yang melakukan tugas untuk mengatasi vanish gradient pada arsitektur *Residual Network*. Resnet menggunakan sistem skip sehingga melewati bagian komputasi perkalian pada model sehingga sampai ke bagian *layer* pertama sehingga bisa menghindari kehilangan informasi penting dari gambar .

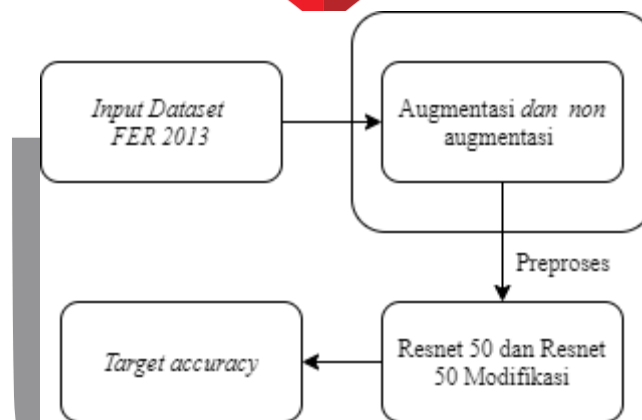


Gambar 2.3. Arsitektur Resnet [14]

3. Perancangan Sistem

3.1 Desain Sistem

Sistem yang dibangun pada pengenalan ekspresi wajah menggunakan model arsitektur yang berbasis *Convolutional Neural network* (CNN) yakni Resnet-50 yang telah dimodifikasi tanpa menggunakan *transfer learning* dan menggunakan Resnet-50 *state of the art* tanpa menggunakan *transfer learning*.



Gambar 3.1 Desain Sistem

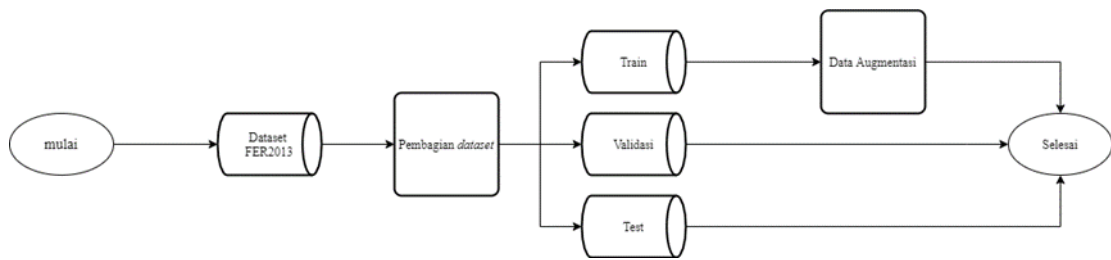
3.2 Alur Desain Sistem

3.2.1 Data akusisi (dataset FER2013)

Dataset FER 2013 seluruh gambar awal mulanya dari file FER2013.CSV. Isi dari format file .csv tersebut berupa kolom emosi, usage, dan piksel masing-masing kelas emosi dan piksel berupa nilai nilai yang kemudian akan dikonversi menjadi nilai string (normalisasi) yakni nilai $(-1,1)$ dari nilai piksel ke rentang $(-1,1)$ karena solusi masalah penyebaran gradien sehingga akan mudah terbaca dengan format tersebut oleh model Resnet-50 *state of the art* dan Resnet-50 modifikasi. Proses pembagian *dataset* FER2013 terbagi tiga yakni: *train* (pelatihan), *public* (validasi), dan *private* (tes)..

3.2.2 Preproses

Diagram preproses ini secara umum dilakukan proses berkaitan dengan data yakni persiapan supaya bisa dilakukan proses pelatihan. Contoh augmentasi untuk solusi *overfitting* dikarenakan penambahan jumlah citra jauh lebih banyak jika tidak memakai FER2013.



Gambar 3.2 Preproses

3.2.3 Data akuisisi (dataset FER2013)

Ukuran *dataset* semua emosi adalah 48x48 piksel dan *dataset* semuanya berwarna grayscale (abu-abu). *Dataset* FER2013 mempunyai file dengan format CSV. *dataset* FER2013 berisi label emosi berisi nilai-nilai piksel dari masing-masing citra tersebut pada masing masing kolom emotion, pixels, usage. Kolom emosi di file CSV terdapat range nilai dari 0 sampai 6 yakni menunjukkan bahwa nilai pada kolom sebagai label dari seluruh kelas emosi berikut (0=marah,1=jijik,2=takut,3=senang,4=sedih,5=kaget,6=netral). Sedangkan citra pada *dataset* FER2013 pada kolom piksel terdiri dari matriks berukuran 48x48 yang mempunyai nilai warna sangat bervariasi mulai dari 0-255. Sedangkan kolom usage untuk kegunaan label citra *dataset train* (data latih), *public test* (validasi), *private* (data tes).



Gambar 3.3 Kelas emosi FER2013

Tabel 3.1 Jumlah gambar tiap kelas *dataset* FER2013 (data pelatihan)

FER2013 data pelatihan	
Kelas emosi wajah	Jumlah gambar
Marah	3995
Jijik	436
Takut	4097
Senang	7215
Sedih	4830
Kaget	3171

FER2013 data pelatihan	
Kelas emosi wajah	Jumlah gambar
Netral	4965
Total	28709

Tabel Jumlah gambar tiap kelas *dataset* FER2013 (data validasi).

FER2013 data validasi	
Kelas emosi wajah	Jumlah gambar
Marah	467
Jijik	56
Takut	496
Senang	895
Sedih	653
Kaget	415
Netral	607
Total	3589

Tabel Jumlah gambar tiap kelas *dataset* FER2013 (data tes).

FER2013 data Tes	
Kelas emosi wajah	Jumlah gambar
Marah	491
Jijik	55
Takut	528

FER2013 data Tes	
Kelas emosi wajah	Jumlah gambar
Senang	879
Sedih	594
Kaget	416
Netral	626
Total	3589

3.2.4 Pembagian Dataset

Tiga jenis kategori dari isi file FER2013.csv yakni: *train* (pelatihan), *public* (validasi), *private* (tes) kemudian diimplementasikan menggunakan pustaka scikit-learn. Rasio yang digunakan pada tugas akhir ini adalah 80% untuk pelatihan, 10% untuk validasi dan 10% untuk tes

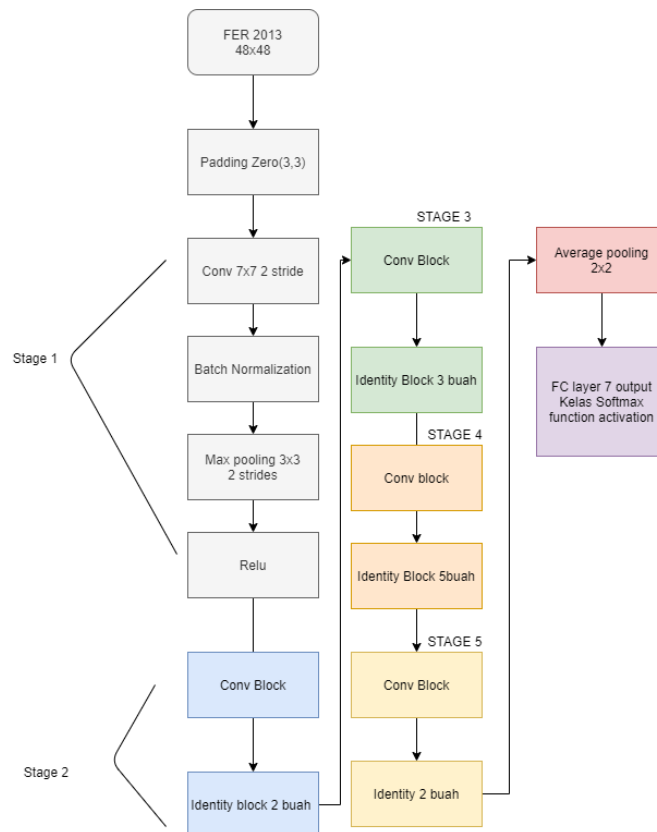
3.2.5 Augmentasi Dataset

Neural network terdapat jumlah *dataset* yang belum cukup untuk meningkatkan performansi saat klasifikasi adanya augmentasi sehingga *dataset* jumlah data citra semakin banyak karena dilakukan beragam proses. Berikut adalah Augmentasi yang digunakan rotasi dengan dilakukan berputar sebesar 10, shear sebesar 10, width shift dengan pergeseran secara horizontal sebesar 0.1, height shift dengan pergeseran secara vertikal sebesar 0.1, zoom atau perbesaran gambar sebesar 0.1, horizontal flip atau dengan cara dicerminkan secara horizontal

3.2.6 Target Akurasi

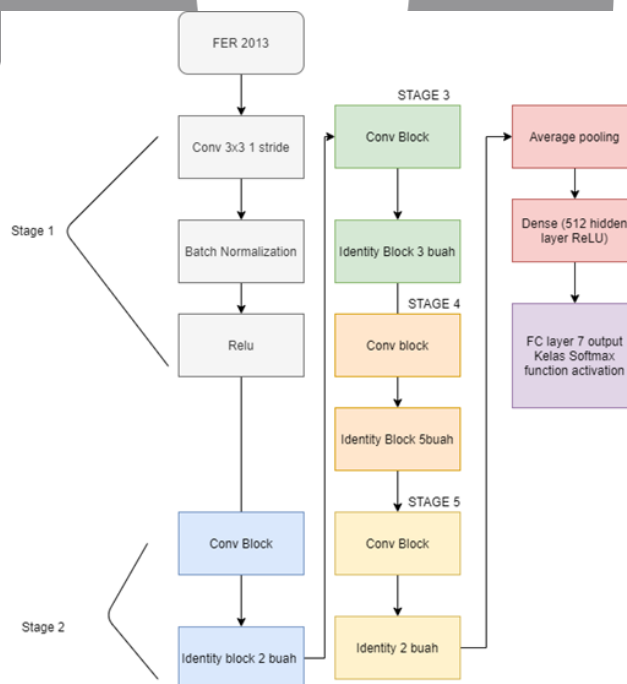
Model yang telah dibuat akan dilakukan pengujian kemudian dari hasil pengujian akan dilakukan analisis emosi. Setelah itu model yang terbaik dari penulis lalu dibandingkan dengan penelitian sebelumnya yang menggunakan arsitektur dan *dataset* yang sama yakni Resnet-50 dan *dataset* FER2013.

3.3 Blok Arsitektur *Residual Network* (Resnet-50 *State of the art*)

Gambar 3.4 Arsitektur Resnet-50 *State of the art*

Arsitektur model Resnet-50 menggunakan semua beberapa parameter baik dari *filter*, *kernel size*, *stride* kecuali pada bagian masukan dan keluaran dikarenakan menyesuaikan pada *dataset* FER2013 yakni memiliki ukuran 48x48 piksel dan kelas sebanyak 7 emosi.

3.4 Blok Arsitektur *Residual Network* (Resnet-50 modifikasi)

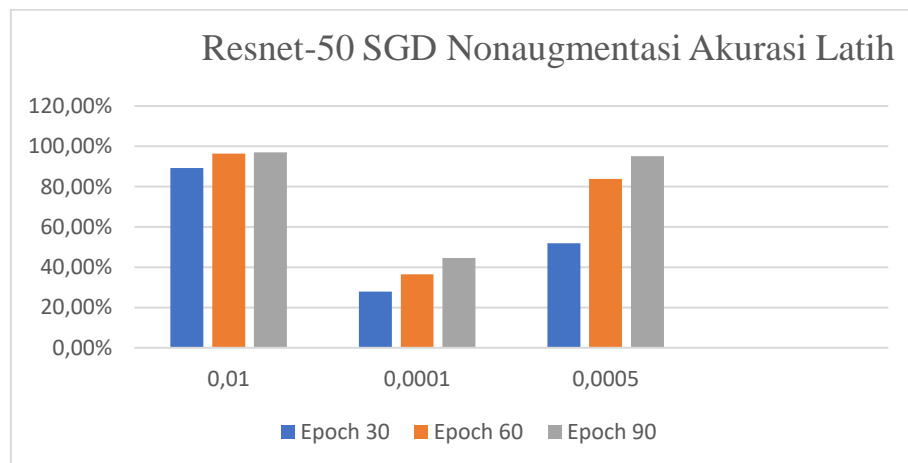


Gambar 3.5 Arsitektur Resnet-50 modifikasi

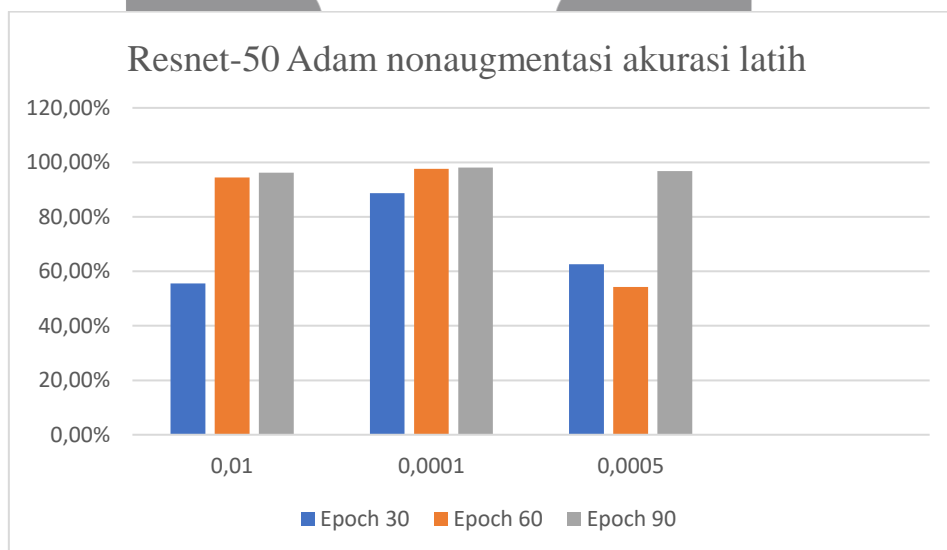
Arsitektur model Resnet-50 modifikasi menggunakan semua parameter baik dari *filter*, *kernel size*, *stride* kecuali pada bagian masukan dan keluaran dikarenakan menyesuaikan ukuran pada *dataset* FER2013 yakni memiliki ukuran 48x48 dan kelas sebanyak 7 emosi.

4. Hasil Pengujian dan Analisis

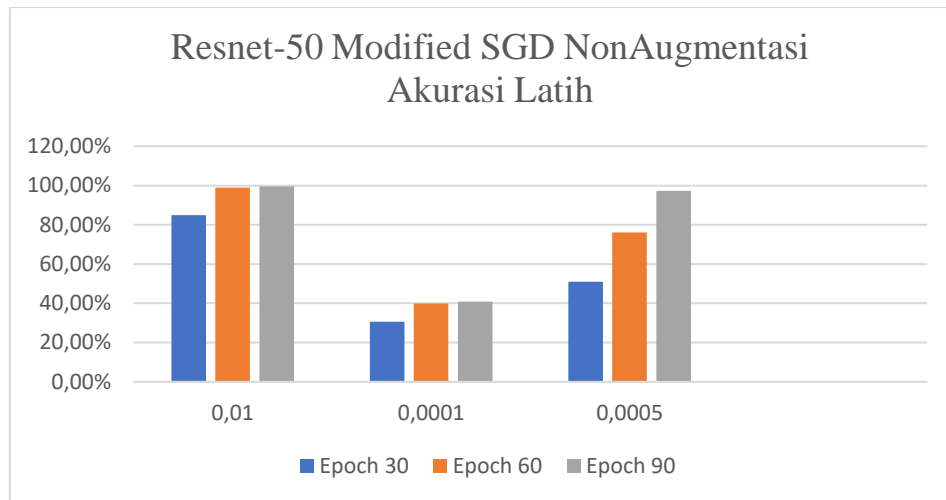
4.1 Pengujian Ke-1 : Nonaugmentasi



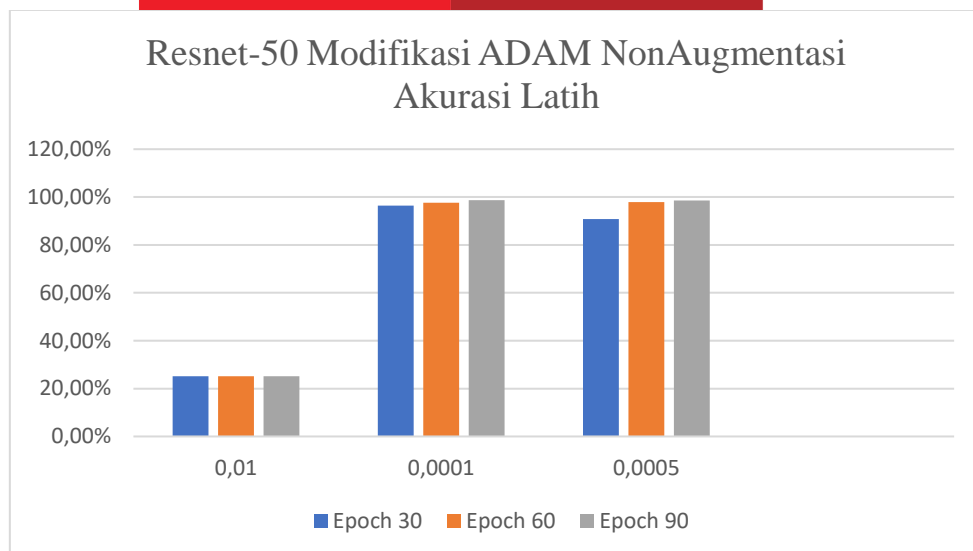
Gambar 4.1 Grafik Resnet-50 *state of the art* SGD akurasi latih nonaugmentasi.



Gambar 4.2 Grafik Resnet-50 *state of the art* Adam akurasi latih nonaugmentasi.



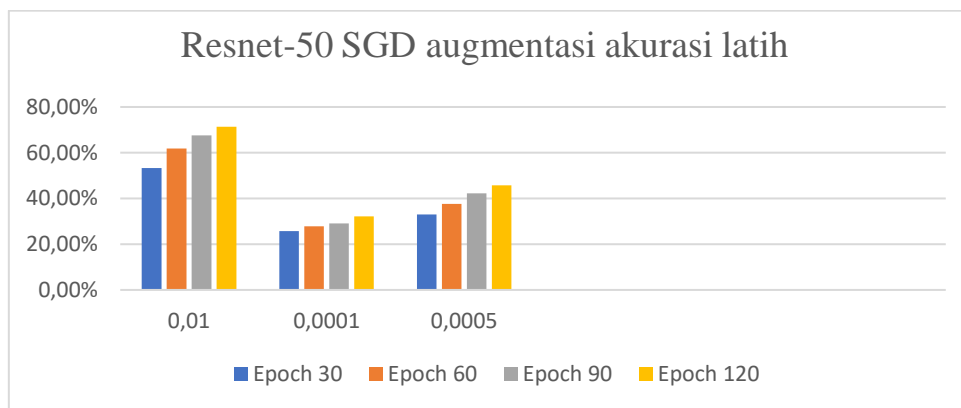
Gambar 4.3 Grafik Resnet-50 modifikasi SGD akurasi latih nonaugmentasi.



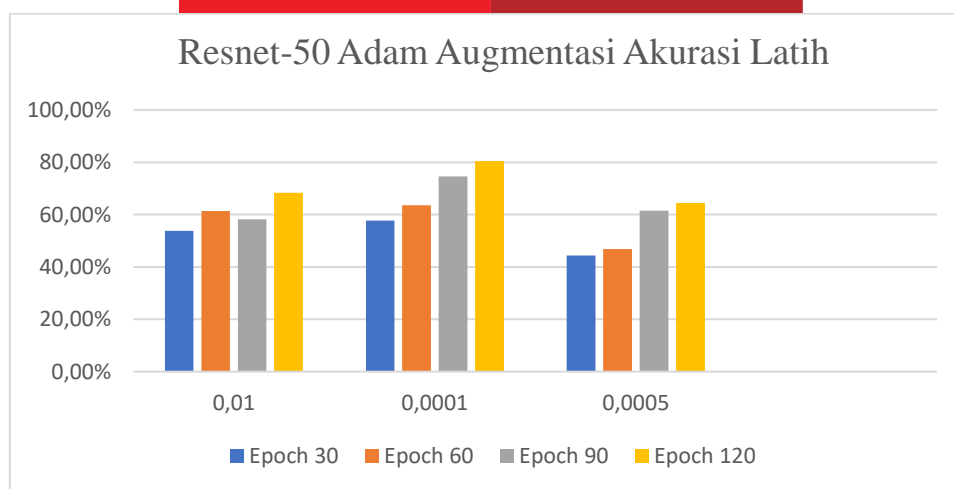
Gambar 4.4 Grafik Resnet-50 modifikasi Adam akurasi latih nonaugmentasi.

Pada gambar 4.1 sampai gambar 4.4 menggunakan pengujian nonaugmentasi pada Resnet-50 *state of the art* dan Resnet-50 modifikasi hasil yang didapatkan cukup bervariasi pada gambar 4.1 sampai gambar 4.4 mengalami *overfitting* karena pada lampiran A pengujian ke-1 hampir dikatakan tidak begitu optimal karena menggunakan nonaugmentasi.

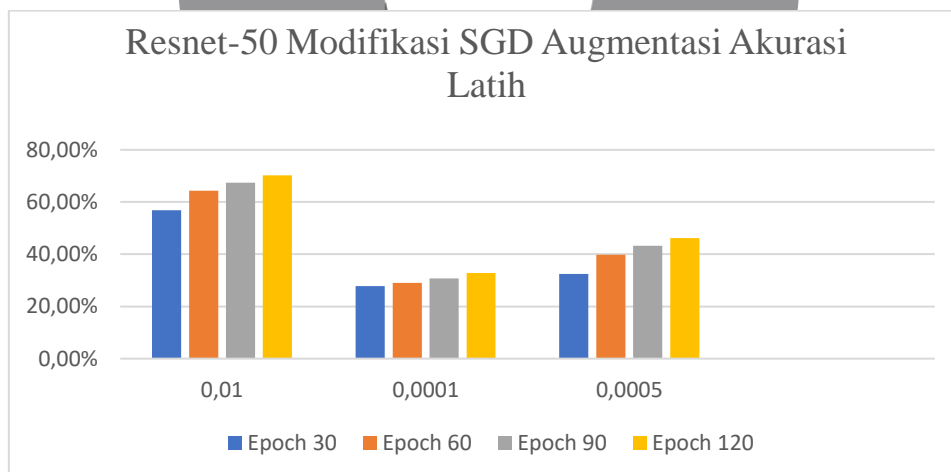
4.2 Pengujian Ke-2: Augmentasi



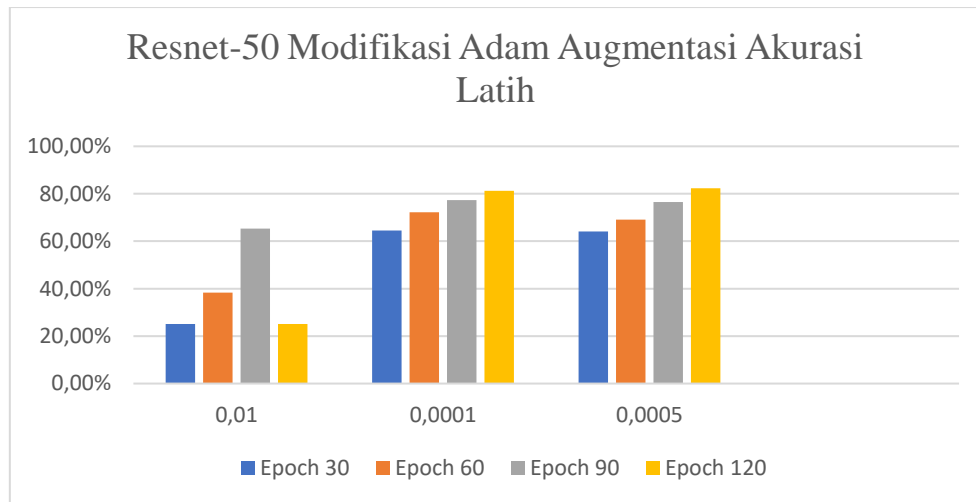
Gambar 4.5 Grafik Resnet-50 *state of the art* SGD akurasi latih augmentasi.



Gambar 4.6 Grafik Resnet-50 *state of the art* Adam akurasi latih augmentasi.



Gambar 4.7 Grafik Resnet-50 modifikasi SGD akurasi latih augmentasi.

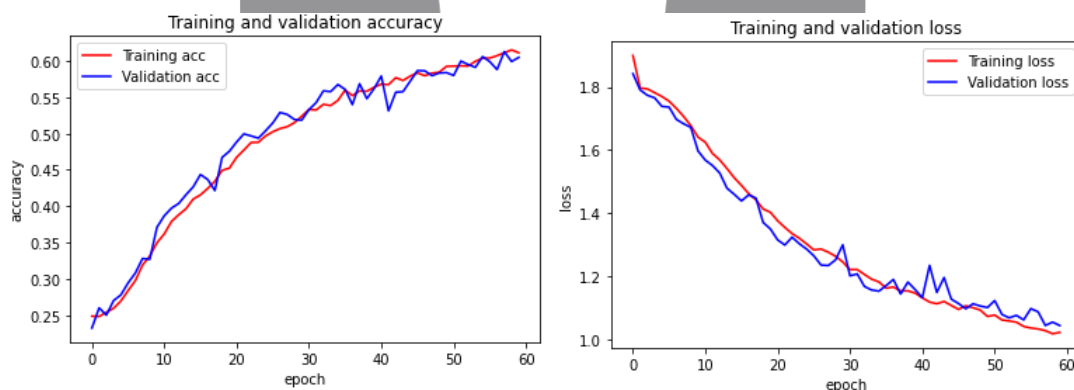


Gambar 4.8 Grafik Resnet-50 modifikasi Adam akurasi latih augmentasi.

Pada gambar 4.5 sampai gambar 4.8 pengujian menggunakan augmentasi didapatkan model optimal milik penulis yaitu Resnet-50 modifikasi SGD epoch 60 model benar benar optimal seperti pada subbab 4.5 grafik menunjukkan tidak *overfitting*serta sudah optimal dari akurasi latih dan validasi. Sisa model mayoritas belum mencapai optimal dan *overfitting*.

4.2 Analisis pengujian

Pada pengujian augmentasi menggunakan Resnet 50 modifikasi didapatkan hasil terbaik dibandingkan menggunakan Resnet-50 *state of the art* lebih jelas nya sebagai berikut



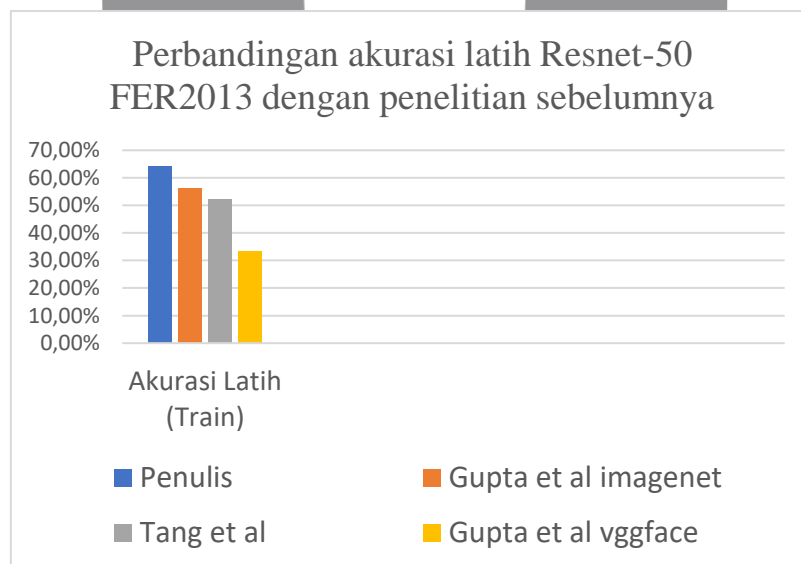
Gambar 4.9. Grafik pelatihan dan validasi dari Resnet-50 modifikasi.

Pada hasil pengujian dipilih yang terbaik/optimal dari yang telah diuji coba baik dari pengujian parameter dan augmentasi. Melihat hasil lebih lengkap nya ada di lampiran A. Hasil yang optimum/terbaik adalah model Resnet-50 modifikasi dengan augmentasi dengan parameter epoch 60, *learning rate* 0.01, dan menggunakan *optimizer* SGD. Pada gambar 4.9 grafik pelatihan dan validasi sangat optimal karena tidak mengalami *overfitting* maupun *underfitting* yakni jarak akurasi pelatihan dan tes tidak terlalu jauh. Pada hasil evaluasi didapatkan hasil pelatihan sebesar 64.28%, validasi sebesar 60.49%, tes sebesar 60.71%.

Angry	252	2	57	21	67	9	72
Disgust	29	9	11	0	9	1	1
Fear	59	1	208	19	129	38	61
Happy	34	0	18	747	26	15	43
Sad	60	0	61	29	319	5	123
Surprise	14	0	78	38	6	247	14
Neutral	42	0	31	74	109	4	397
	Angry	Disgust	Fear	Happy	Sad	Suprise	Neutral

Gambar 4.10 *confusion* matriks dari Resnet-50 modifikasi.

Pada gambar 4.10 *confusion* matriks menunjukkan label kelas emosi diwakili oleh nilai menunjukkan bahwa untuk kelas emosi senang memiliki nilai yang cukup tinggi yakni model mampu mendeteksi emosi senang dengan baik sedangkan terendah pada emosi jijik ini disebabkan karena jumlah *dataset* yang kurang. Terlihat masing-masing semua ekspresi wajah memiliki akurasi precision, recall, dan akurasi pada *confusion* matriks untuk melihat seberapa bagus model memprediksi dan mempelajari tiap emosi yang berbeda bisa dilihat bahwa emosi jijik memiliki nilai TP yang rendah karena kelas jijik tidak memiliki *dataset* yang banyak ketimbang kelas emosi lain sedangkan nilai TP tertinggi oleh kelas senang artinya model bisa mempelajari ekspresi senang dengan baik yakni 747 karena *dataset* kelas senang sangat banyak dibanding kelas lainnya.



Gambar 4.11 Grafik akurasi latih Resnet-50 dengan penelitian sebelumnya menggunakan *dataset* FER2013.

Hasil gambar 4.11 model milik Gupta et al yakni menggunakan beberapa parameter menggunakan Resnet-50 *transfer learning* baik ImageNet dan VGG-FACE menggunakan parameter Epoch 30 dan *learning rate* 0.1. Metode pengujian milik Tang et al menggunakan parameter *learning rate*

0.01. Model penulis menggunakan parameter yang digunakan pada subbab 4.3 model penulis mendapatkan akurasi jauh lebih baik yakni 64.28% meningkat sekitar 8% dari penelitian penggunaan imagenet milik Gupta et al sebesar 56.21%, meningkat 31% dari penelitian menggunakan vgg-face milik Gupta et al dan meningkat sekitar 12% dari penelitian milik Tang et al 52.08%.

5. Kesimpulan dan Saran

5.1. Kesimpulan

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan pada tugas akhir ini maka didapatkan kesimpulan sebagai berikut:

1. Sistem yang telah dirancang dari dua arsitektur berbeda yakni Resnet-50 *state of the art* dan Resnet-50 modifikasi dari dua pengujian yakni augmentasi dan nonaugmentasi performansi yang terbaik adalah menggunakan model Resnet-50 modifikasi dengan parameter epoch 60, *learning rate* 0.01 dan *optimizer* SGD pada *dataset* FER 2013 dengan jumlah kelas emosi sebanyak 7 terdiri dari emosi. Dengan rasio pembagian jumlah *dataset* FER 2013 *train* 80%, validasi Pada hasil evaluasi didapatkan hasil pelatihan sebesar 64.28% , validasi sebesar 60.49%, tes sebesar 60.71% dibandingkan model lain yang cenderung tidak optimal dan overfitting.
2. Sistem yang telah dirancang oleh penulis mendapatkan akurasi pelatihan sebesar 64.28% lebih unggul 8% dari milik Gupta et al sebesar 56.21% serta lebih unggul 12% milik Tang et al sebesar 52.08%.
3. Sistem yang telah dirancang menggunakan dua jenis model yang berbeda yakni: Resnet-50 *state of the art* dan Resnet-50 modifikasi masing masing tanpa *transfer learning* didapatkan bahwa dengan menggunakan augmentasi performansi model menghasilkan cukup baik dalam melakukan klasifikasi dibandingkan kedua model yakni Resnet-50 *state of the art* dan Resnet-50 modifikasi tanpa menggunakan pengujian augmentasi. Augmentasi sangat cukup baik dalam meningkatkan performansi model sehingga bisa mendapatkan model yang optimal

5.2. Saran

Berdasarkan hasil perancangan dan pengujian pada tugas akhir ini berikut merupakan saran kedepan untuk penelitian terkait ini:

1. Menggunakan metode menyeimbangkan kelas (*Image Balancing Class*) karena jumlah kelas emosi pada *dataset* FER2013 tidak seimbang terutama pada emosi jijik.
2. Melakukan tuning parameter kembali dan melakukan modifikasi kembali supaya hasilnya lebih baik.
3. Melakukan proses preproses dengan meningkatkan kualitas citra pada *dataset* supaya hasilnya lebih baik dibandingkan sebelumnya yakni dengan melakukan proses penajaman kualitas citra dan pencahayaan.

REFERENSI

- [1] J. E. Prawitasllri, "Mengenal Emosi Melalui Komunikasi Nonverbal," *Bul. Psikol.*, vol. 3, no. 1, pp. 27–43, 2016, doi: 10.22146/bpsi.13384.
- [2] N. Samadiani *et al.*, "A review on automatic facial expression recognition systems assisted by multimodal sensor data," *Sensors (Switzerland)*, vol. 19, no. 8, pp. 1–27, 2019, doi: 10.3390/s19081863.

- [3] Y. Tang, X. Zhang, X. Hu, S. Wang, and H. Wang, "Frequency Neural Network," vol. 30, pp. 444–457, 2021.
- [4] A. Gupta, "Transfer:- Deep Inductive Network for Facial Emotion Recognition," *J. Mech. Contin. Math. Sci.*, vol. 15, no. 7, pp. 352–361, 2020, doi: 10.26782/jmcms.2020.07.00029.
- [5] M. H. Siddiqi, R. Ali, A. M. Khan, Y. T. Park, and S. Lee, "Human Facial Expression Recognition Using Stepwise Linear Discriminant Analysis and Hidden Conditional Random Fields," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1386–1398, 2015, doi: 10.1109/TIP.2015.2405346.
- [6] A. Fathallah, L. Abdi, and A. Douik, "Facial expression recognition via deep learning," *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA*, vol. 2017-Octob, pp. 745–750, 2018, doi: 10.1109/AICCSA.2017.124.
- [7] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," *2016 IEEE Winter Conf. Appl. Comput. Vision, WACV 2016*, 2016, doi: 10.1109/WACV.2016.7477450.
- [8] I. Gogul and V. S. Kumar, "Flower species recognition system using convolution neural networks and transfer learning," *2017 4th Int. Conf. Signal Process. Commun. Networking, ICSCN 2017*, no. March, 2017, doi: 10.1109/ICSCN.2017.8085675.
- [9] H. H. Aghdam and E. J. Heravi, *Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification*. 2017.
- [10] G. P. Harvian, S. Al Faraby, and U. N. Wisesty, "Klasifikasi Iris Biometrik Menggunakan Convolutional Neural Network (CNN) Dengan Metode Segmentasi K- Means Iris Biometric Classification Using Convolutional Neural Network (CNN) and K-Means Segmentation Method," *e-Proceeding Eng. Vol.6*, vol. 6, no. 2, pp. 9827–9837, 2017.
- [11] E. Al Hadhrami, M. Al Mufti, B. Taha, and N. Werghi, "Transfer learning with convolutional neural networks for moving target classification with micro-Doppler radar spectrograms," *2018 Int. Conf. Artif. Intell. Big Data, ICAIBD 2018*, no. July, pp. 148–154, 2018, doi: 10.1109/ICAIBD.2018.8396184.
- [12] J. Moolayil, *Learn Keras for Deep Neural Networks: A Fast-Track Approach to Modern Deep Learning with Python*. 2019.
- [13] W. Hao, W. Yizhou, L. Yaqin, and S. Zhili, "The Role of Activation Function in CNN," *Proc. - 2020 2nd Int. Conf. Inf. Technol. Comput. Appl. ITCA 2020*, pp. 429–432, 2020, doi: 10.1109/ITCA52113.2020.00096.
- [14] K. H. Mahmud, Adiwijaya, and S. Al Faraby, "Klasifikasi Citra Multi-Kelas Menggunakan Convolutional Neural Network," *e-Proceeding Eng.*, vol. 6, no. 1, pp. 2127–2136, 2019.