

## Studi Pengenalan Ekspresi Wajah Berbasis Convolutional Neural Network

L. M. Rasdi Rere, Sunarto Usna dan Soegijanto  
STMIK Jakarta STI&K  
Jl. BRI No.17, Radio Dalam, Kebayoran Baru, Jakarta Selatan 12140  
{rasdirere267, soegijantostik}@gmail.com, sunartousna@yahoo.com,

### ABSTRAK

*Pengenalan ekspresi wajah adalah topik penelitian yang aktif dilakukan dalam beberapa dekade terakhir. Aplikasinya dapat ditemukan dalam biometric security, robotika, kesehatan, neuromarketing sampai animasi avatar. Mengenal ekspresi pada wajah bukanlah merupakan masalah yang mudah dalam pembelajaran mesin, karena setiap orang dapat memiliki cara yang berbeda untuk memperlihatkan ekspresinya. Bahkan citra pada orang yang sama dalam satu ekspresi, dapat menghasilkan fitur yang bervariasi, tergantung pencahayaan, posisi maupun latarbelakancitra. Karena itu pengenalan ekspresi wajah masih merupakan masalah yang menantang dalam bidang computer vision. Pendekatan tradisional untuk pengenalan ekspresi wajah sangat bergantung pada hand-crafted features seperti SIFT, HOG maupun LBP, yang kemudian dilanjutkan dengan training data pada citra. Dalam beberapa tahun terakhir, metode deep learning mulai dipergunakan dalam pengenalan ekspresi wajah dengan hasil yang secara umum lebih baik. Penelitian ini menggunakan salah satu metode deep learning yang umum dipakai pada citra, yaitu Convolutional Neural Network untuk pengenalan ekspresi wajah, dengan sejumlah variasi parameter menggunakan dataset JAFFE. Hasil terbaik yang diperoleh adalah nilai akurasi sebesar 87,5% untuk arsitektur hanya dengan 3 lapisan training pada jaringan.*

**Kata Kunci :** *Klasifikasi ekspresi wajah dengan deep learning*

### PENDAHULUAN

Emosi adalah bagian yang tidak terhindarkan dari setiap pribadi seseorang pada saat berkomunikasi. Bentuknya dapat diekspresikan dengan berbagai macam cara, yang terkadang tidak dapat diketahui secara langsung. Meskipun demikian, dengan suatu sistem peralatan yang tepat, hal seperti ini sangat mungkin untuk dapat dideteksi dan dikenali.

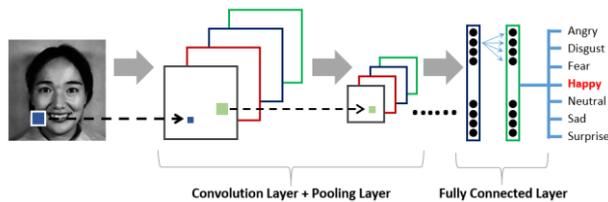
Sistem kecerdasan buatan untuk mengenal emosi seseorang telah menarik banyak minat penelitian, karena potensi aplikasi sistem semacam ini sangat berlimpah, mencakup bidang pemasaran untuk lebih memperhatikan minat pelanggan, pemantauan kesehatan, sistem keamanan, sampai dengan antarmuka robot yang cerdas secara emosional.

Pengenalan emosi dapat dilakukan menggunakan fitur yang berbeda, seperti ekspresi wajah, ucapan, teks dan gerakan tubuh. Di antara beragam fitur ini, pengenalan emosi dengan ekspresi wajah adalah salah satu yang paling populer dengan sejumlah alasan seperti ekspresinya dapat dilihat secara langsung, mengandung

banyak fitur berguna untuk pengenalan emosi, serta tersedia dataset yang cukup berlimpah [1].

Pengenalan ekspresi pada wajah dengan pendekatan tradisional umumnya menggunakan pembelajaran mesin dalam dua langkah. Pertama adalah dengan mengekstraksi sejumlah fitur, yang kemudian dilanjutkan dengan klasifikasi untuk mendeteksi emosi.

Beberapa metode ekstraksi fitur yang populer dipergunakan untuk pengenalan ekspresi wajah adalah *scale invariant feature transform* (SIFT), *histogram of oriented gradients* (HOG), *local binary patterns* (LBP), *Gabor wavelets* dan *Haar features*. Untuk teknik klasifikasi biasanya menggunakan metode *support vector machine* (SVM), *neural network* atau *random forest*. Pendekatan tradisional pada dasarnya dapat bekerja cukup baik untuk dataset sederhana, akan tetapi akan mengalami masalah untuk dataset yang lebih kompleks dan bervariasi [1] [2].



**Gambar 1.** Arsitektur umum CNN pengenalan ekspresi wajah

Dalam beberapa tahun terakhir, metode *deep learning* telah memberikan sukses besar untuk klasifikasi citra dan masalah *computer vision* lainnya [3][4][5][6][7], tidak terkecuali untuk penelitian yang berhubungan dengan pengenalan ekspresi wajah. Dengan menggunakan *Boosted Deep Belief Network* (BDBN) misalnya, Liu dkk [4] mengkombinasikan ekstraksi fitur dan klasifikasi dalam sebuah *single looped network* sehingga dapat menghasilkan *state-of-the-art* pada dataset pengenalan ekspresi wajah CK+ dan JAFFE. Han dkk [8] mengusulkan *incremental boosting convolutional neural network* (IB-CNN) untuk meningkatkan pengenalan ekspresi wajah secara spontan. Sedangkan Zang dkk [2] menggunakan konsep *inter-identity CNN* (I<sup>2</sup>CNN), melalui deteksi pada bagian-bagian wajah, sehingga mendapatkan hasil yang lebih baik pada dataset CK+ dan JAFFE.

Pada dasarnya sebagian besar metode *deep learning* yang dipergunakan pada pengenalan ekspresi wajah, menggunakan CNN dengan berbagai macam usulan, untuk mendapatkan hasil optimal [9][10][11][12]. Karena itu pada penelitian ini juga dipergunakan CNN pada dataset JAFFE, dengan beberapa eksperimen yang dilakukan berdasarkan variasi parameter yang ada.

## METODE CONVOLUTIONAL NEURAL NETWORK

Salah satu metode yang paling penting dalam *deep learning*, terutama pada citra adalah *Convolutional Neural Network* (CNN). Kelebihan utama metode CNN ini, adalah kemampuannya dalam mereduksi dimensi data, mengekstraksi fitur secara sekuensial serta mengklasifikasi jaringan dalam struktur tunggal [13]. Karenanya,

metode ini paling sering digunakan untuk berbagai aplikasi dalam *computer vision* dan *machine learning*.

Sejarah perkembangan CNN dimulai pada tahun 1962, ketika Hubel dan Wiesel mengusulkan model arsitektur jaringan yang terinspirasi dari korteks visual pada kucing. Komputasi pertama dari model ini dibuat oleh Fukushima di tahun 1980, dan kemudian berdasarkan hal ini, Yan LeCun pada tahun 1989, melakukan training jaringan konvolusi menggunakan *error gradient*, yang menghasilkan *state-of-the-art* untuk kinerja pada pengenalan pola [14].

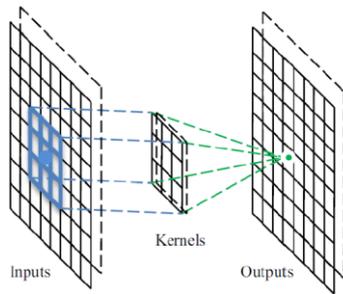
Arsitektur awal dari CNN pada dasarnya merupakan perluasan dari metode MLP yang didasarkan pada tiga ide utama yaitu *local receive field*, *weights sharing* dan *spatial sub-sampling*. Ketiga ide utama ini dalam CNN dinyatakan dalam dua jenis lapisan (*layer*), yaitu *convolution layer* dan *pooling layer*, di mana ke dua *layer* disusun dalam *planes* yang disebut dengan peta fitur (*fitur map*). Bersama sama dengan *fully connected layers*, peta fitur merupakan *main neural layers* dalam arsitektur umum dari CNN [15], seperti diperlihatkan pada Gambar 1, untuk penelitian pengenalan ekspresi wajah yang dilakukan.

## Lapisan Konvolusi

Lapisan konvolusi (*convolution layer*) adalah blok bangunan inti dari metode CNN, dan merupakan bagian utama dari beban komputasi jaringan. Lapisan ini berkorelasi langsung dengan peta fitur, ukuran kernel dan koneksi dengan lapisan sebelumnya. Pada lapisan ini dilakukan perkalian titik antara dua matriks, dimana satu matriks adalah himpunan parameter yang dapat dipelajari, yang dikenal dengan kernel, dan matriks lainnya adalah bagian terbatas dari bidang reseptif. Secara matematis suatu perkalian konvolusi untuk fungsi diskrit satu dimensi, dapat dinyatakan dalam persamaan berikut:

$$f(x) * g(x) = \sum f(\alpha)g(x - \alpha) \quad (1)$$

dimana  $g(x)$  adalah filter atau kernel konvolusi yang dioperasikan dengan cara menggeser pada sinyal masukan  $f(x)$ .

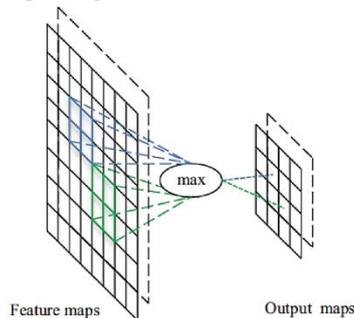


Gambar 2. Proses lapisan Konvolusi [16]

Terdapat tiga keuntungan dalam proses konvolusiseperti diperlihatkan pada Gambar 2. Pertama adalah mekanisme berbagi bobot (*weigh sharing*) pada peta fitur yang sama, sehingga dapat mereduksi sejumlah parameter yang ada. Kedua adalah konektivitas lokal, sehingga dapat dipelajari korelasi antar piksel, serta keuntunganyang ketiga adalah invarian ke lokasi objek tertentu [16].

### Lapisan Pooling

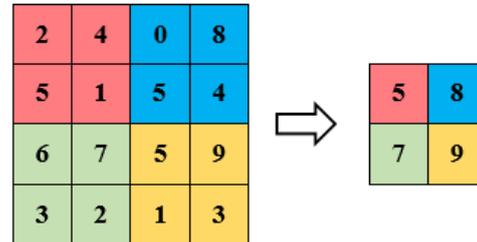
Secara umum lapisan *pooling* dapat dipergunakan untuk mengurangi dimensi peta fitur dan parameter jaringan. Proses pada lapisan ini dilakukan setelah lapisan konvolusi. Ada beberapa strategi yang biasanya dipakai pada lapisan ini, tetapi yang sering dipakai adalah *max-pooling*. Seperti diperlihatkan pada Gambar 3, sebuah peta fitur ukuran  $8 \times 8$  dengan operator *max-pooling* yang memiliki ukuran  $2 \times 2$  dan *stride* 2, maka dimensinya akan berkurang menjadi  $4 \times 4$ .



Gambar 3. Proses lapisan max-pooling [16]

Ilustrasi hasil perhitungan matriks citra dengan menggunakan *max-pooling* diperlihatkan seperti pada Gambar 4, di mana *output* dari lapisan konvolusi dibagi

menjadi beberapa *grid* kecil dan kemudian mengambil nilai yang terbesar dari setiap *grid*, sehingga akan terbentuk matriks citra baru yang telah mengalami reduksi.

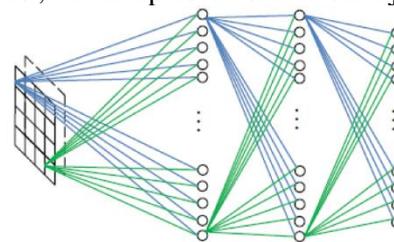


Gambar 4. Perhitungan max-pooling

Lapisan *pooling* sama halnya dengan lapisan konvolusi juga *translation invariant*. Hal ini karena pada proses komputasinya mengambil piksel ketetanggaan dalam perhitungan, dengan mengurangi resolusi peta fitur.

### Lapisan fully connected

Setelah proses dengan lapisan konvolusi dan *pooling*, struktur terakhir dari arsitektur CNN adalah lapisan *fully connected*. Lapisan ini bekerja sebagai masukan untuk serangkaian lapisan terhubung sepenuhnya yang melaksanakan tugas klasifikasi. Seperti diperlihatkan pada Gambar 5, setelah proses pada lapisan *pooling* pada jaringan terakhir, ada beberapa lapisan *fully connected* yang mengubah peta fitur dua dimensi, menjadi hanya vektor satu dimensi, untuk representasi fitur selanjutnya.



Gambar 5. Proses lapisan fully connected [16]

Kekurangan lapisan *fully connected* adalah memiliki banyak parameter, sehingga menghasilkan waktu komputasi yang cukup lama untuk melatihnya. Hal ini disebabkan kerja pada lapisan *fully connected* seperti halnya jaringan syaraf tradisional, yang dalam metode CNN dapat mengandung sekitar 90% parameter [16].

**DESAIN EKSPERIMEN**

Secara umum arsitektur CNN yang dipergunakan dalam penelitian ini menggunakan tiga lapisan *fully connected layer*, seperti diperlihatkan pada Tabel 1. Pada ketiga lapisan, dipergunakan *convolution layer* 2D, dengan fungsi aktivasi ReLu dan *max pooling layer* dengan *region(2, 2)*. Setelah proses *flatten*, pada *fully connected layer*, digunakan *dense\_1* dengan fungsi aktivasi ReLu dan *dense\_2* dengan fungsi aktivasi *softmax*.

Dalam penelitian ini, ada 8 skenario eksperimen yang dilakukan seperti diperlihatkan pada Tabel 2. Eksperimen pertama dan kedua menggunakan citra ukuran 64x64, dengan jaringan *dropout*, serta optimizer adam dan *stochastic gradient descent* (SGD). Eksperimen tiga dan empat dengan ukuran citra 128x128, tanpa jaringan *dropout*, menggunakan optimizer adam dan SGD.

Eksperimen lima dan enam sama dengan eksperimen tiga dan empat, akan tetapi menggunakan *dropout*. Sedangkan eksperimen tujuh dan delapan sama dengan eksperimen lima dan enam, akan tetapi perbandingan data training dan testingnya menjadi 70% berbanding 30%. Pada eksperimen sebelumnya adalah 85% data training dan 15% data testing.

**Tabel 1. Arsitektur CNN**

Jenis layer	Bentuk Output	
Conv2d_1	64, 64, 6	128, 128, 6
Activation_1	64, 64, 6	128, 128, 6
MaxPooling_1	32, 32, 6	64, 64, 6
Conv2d_2	32, 32, 16	64, 64, 16
Activation_2	32, 32, 16	64, 64, 16
MaxPooling_2	16, 16, 16	32, 32, 16
Conv2d_3	12, 12, 120	28, 28, 120
Activation_3	12, 12, 120	28, 28, 120
Dropout_1	12, 12, 120	28, 28, 120
Flatten_1	17280	94080
Dense_1	84	84
Activation_4	84	84
Dropout_2	84	84
Dense_2	7	7
Activation_5	7	7

Pada CNN dengan *dropout* jaringan, nilai dari *Dropout\_1* adalah 0,25, sedangkan

*Dropout\_2* bernilai 0,50. Semua eksperimen dijalankan dengan 50 *epoch*, menggunakan Keras dengan *Tensorflow backend*. Untuk mengevaluasi eksperimen yang dilakukan, dengan arsitektur CNN yang telah ditentukan, digunakan dataset *Japanese Female Facial Expression* (JAFPE).

**Tabel 2. Desain eksperimen**

No	Eksperimen	Parameter
1	T1-64A	Citra 64x64, <i>dropout</i> , adam, 85%/15%
2	T2-64SGD	Citra 64x64, <i>dropout</i> , SGD, 85%/15%
3	T3-128A	Citra 128x, <i>dropout</i> , adam, 85%/15%
4	T4-128SGD	Citra 128x, <i>dropout</i> , SGD, 85%/15%
5	T5-128A	Citra 128x, adam, 85%/15%
6	T6-128SGD	Citra 128x, SGD, 85%/15%
7	T7-128A	Citra 128x, adam, 70%/30%
8	T8-128SGD	Citra 128x, SGD, 70%/30%

Dataset JAFPE terdiri dari 213 citra, berasal dari 10 model wanita jepang, dengan 7 ekspresi pada wajah, yaitu *angry*(30 data), *disgust*(29 data), *fear*(32 data), *happy*(31 data), *neutral*(30 data), *sad*(31 data) dan *surprise*(30 data). Beberapa contoh ekspresi wajah dataset JAFPE diperlihatkan seperti pada Gambar 6.



**Gambar 6.** Contoh citra dataset JAFPE

## HASIL DAN PEMBAHASAN

Hasil eksperimen dari penelitian yang dilakukan secara keseluruhan dapat dilihat pada Tabel 3. Pada dasarnya ada 4 parameter yang diujicobakan untuk mengetahui pengaruhnya terhadap arsitektur CNN yang didesain. Parameter pertama adalah ukuran citra, kedua adalah jaringan *dropout*, ketiga adalah perbandingan data training dan validasi, serta keempat adalah jenis optimizer yang dipergunakan.

### Ukuran citra

Citra ekspresi wajah yang dipergunakan pada eksperimen 1 dan 2 berukuran 64×64, dengan optimizer adam dan SGD. Jika dibandingkan dengan eksperimen 3 dan 4, yang menggunakan ukuran citra 128×128 dengan parameter lainnya sama. Maka ukuran citra yang lebih besar menghasilkan nilai yang lebih baik, untuk akurasi training maupun validasi. Dalam hal waktu komputasi, secara umum untuk citra 64×64 memiliki waktu komputasi sekitar 100 detik, sedangkan untuk citra dengan ukuran 128×128, waktu komputasinya adalah sekitar 500 detik.

Dalam eksperimen 1 sampai 4, grafik perbandingan akurasi untuk hasil training dan validasi diperlihatkan seperti pada Gambar 7, 8, 9 dan 10. Secara umum untuk setiap *epoch*, nilai akurasi training lebih baik dari nilai akurasi validasi.

**Tabel 3.** Hasil eksperimen

No	Eksperimen	Akurasi	
		Training	Validasi
1	T1-64A	96,7%	78,1%
2	T2-64SGD	91,2%	62,5%
3	T3-128A	93,4%	87,5%
4	T4-128SGD	88,4%	87,5%
5	T5-128A	99,4%	84,4%
6	T6-128SGD	96,1%	87,5%
7	T7-128A	100%	71,9%
8	T8-128SGD	94,0%	78,1%

### Jaringan dropout

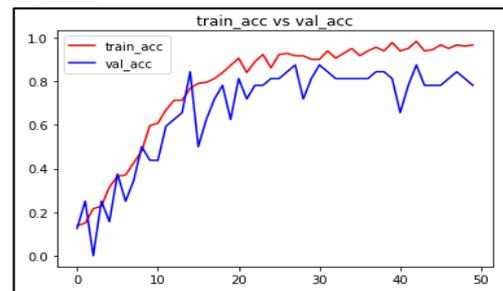
Eksperimen 5 dan 6 yang dilakukan pada dasarnya sama dengan eksperimen 3 dan 4. Perbedaannya adalah bahwa pada eksperimen 3 dan 4 dilakukan *dropout* pada jaringan. Untuk optimizer SGD, tidak ada

kenaikan nilai akurasi pada validasi. Akan tetapi untuk optimizer adam, ada kenaikan nilai akurasi pada validasi dari 84,4% (T5-128A) menjadi 87,5% (T3-128A).

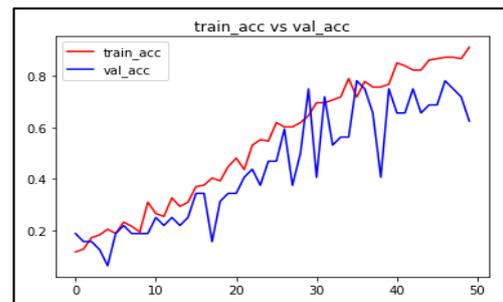
Perbandingan nilai akurasi training dan validasi untuk eksperimen 5 dan 6 diperlihatkan seperti pada Gambar 11 dan 12. Seperti diperlihatkan pada kedua gambar, secara umum selisih nilai akurasi training dan akurasi validasi lebih besar, jika dibandingkan dengan hasil eksperimen 3 dan 4.

### Perbandingan data training dan validasi

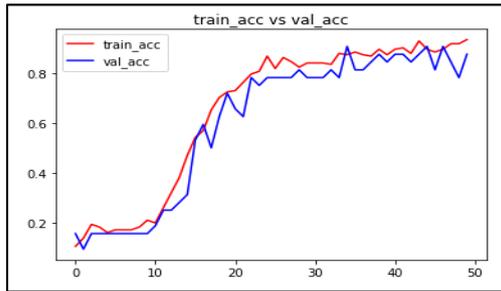
Pada eksperimen 1 sampai dengan 6, perbandingan antara data training dan data validasi adalah 85% berbanding 15%. Dalam eksperimen 7 dan 8 yang dilakukan, perbandingan yang dipergunakan adalah 70% data training dan 30% data validasi.



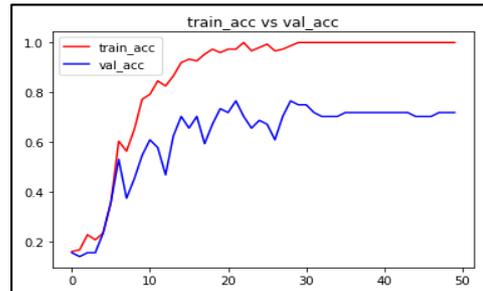
**Gambar 7.** Perbandingan Akurasi training dan validasi untuk eksperimen T1-64A



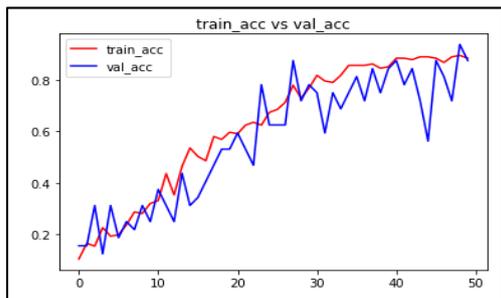
**Gambar 8.** Perbandingan Akurasi training dan validasi untuk eksperimen T2-64SGD



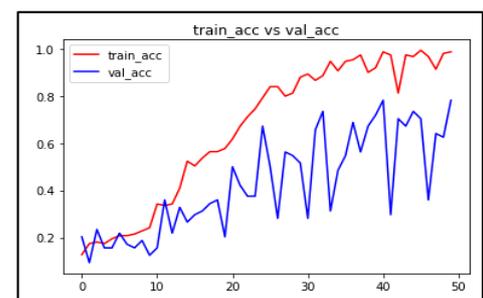
**Gambar 9.** Perbandingan Akurasi training dan validasi untuk eksperimen T3-128A



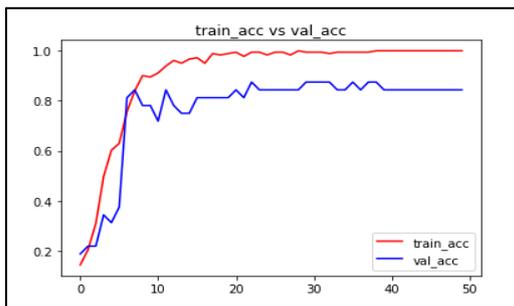
**Gambar 13.** Perbandingan Akurasi training dan validasi untuk eksperimen T7-128A



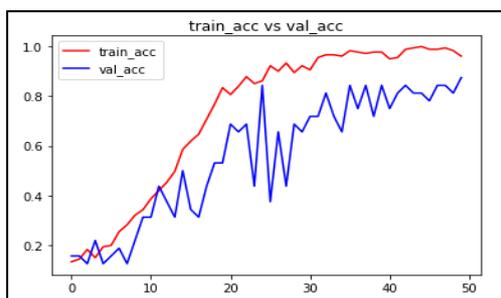
**Gambar 10.** Perbandingan Akurasi training dan validasi untuk eksperimen T4-128SGD



**Gambar 14.** Perbandingan Akurasi training dan validasi untuk eksperimen T8-128SGD



**Gambar 11.** Perbandingan Akurasi training dan validasi untuk eksperimen T5-128A



**Gambar 12.** Perbandingan Akurasi training dan validasi untuk eksperimen T6-128SGD

Kecuali untuk perbandingan data training dan validasi, parameter yang dipergunakan untuk eksperimen 5 dan 6 sama dengan eksperimen 7 dan 8. Seperti dapat dilihat pada Tabel 2, eksperimen 5 dan 6 dengan jumlah data training lebih banyak, menghasilkan nilai akurasi yang lebih baik jika dibandingkan dengan eksperimen 7 dan 8. Berdasarkan grafik perbandingan akurasi training dan validasi, eksperimen 7 dan 8, seperti diperlihatkan Gambar 13 dan 14, terlihat bahwa selisih antara akurasi training dan akurasi validasi, relatif lebih besar, jika dibandingkan dengan eksperimen 5 dan 6 dengan jumlah data training yang lebih besar.

### Jenis optimizer

Dari semua eksperimen yang dilakukan, terlihat bahwa nilai akurasi training untuk optimizer adam lebih baik dari optimizer SGD. Akan tetapi untuk nilai akurasi validasi, optimizer SGD lebih baik dari optimizer adam, kecuali untuk eksperimen T2-64SGD.

Berdasarkan grafik perbandingan akurasi training dan akurasi validasi, seperti diperlihatkan pada Gambar 7 sampai 14, optimizer adam lebih cepat mencapai nilai akurasi dibandingkan optimizer SGD,

terutama untuk training. Pada Gambar 11, eksperimen T5-128A, yang menggunakan optimizer adam misalnya, pada *epoch* ke 10 nilai akurasi trainingnya adalah 89,5% dan akurasi validasinya 78,1%. Sedangkan menggunakan optimizer SGD seperti diperlihatkan Gambar 12, pada eksperimen T6-128SGD, pada epoch ke 10 nilai akurasi trainingnya adalah 34,3% dan akurasi validasinya 31,3%.

## PENUTUP

Berdasarkan penelitian yang telah dilakukan untuk pengenalan ekspresi wajah, ukuran pada citra, penentuan jaringan *dropout*, jumlah data training yang dipergunakan, serta jenis optimizer yang dipakai, sangat mempengaruhi kinerja dari arsitektur CNN yang telah ditentukan. Ukuran citra yang lebih besar dapat lebih meningkatkan nilai akurasi pada sistem, akan tetapi waktu komputasi yang diperlukan menjadi lebih lama. Pemakaian jaringan *dropout*, dan terutama jumlah data training yang lebih banyak, sangat signifikan dalam meningkatkan nilai akurasi. Sedangkan pemilihan optimizer yang tepat, dapat mempercepat nilai akurasi.

Dalam penelitian ini nilai akurasi yang dihasilkan cukup baik yaitu sebesar 87,5%, meskipun hanya 3 lapisan yang dipergunakan pada arsitektur CNN. Untuk penelitian selanjutnya, akan dilakukan eksperimen dengan dataset pengenalan ekspresi wajah lainnya, seperti FER, CK+, face94 dan FERG. Selain itu suatu algoritme metaheuristik akan dicoba untuk ditambahkan, sehingga arsitektur CNN yang dihasilkan menjadi lebih optimal.

## DAFTAR PUSTAKA

- [1]. Minaee, Shervin dan Amirali Abdolrashedi, "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network", arXiv:1902.01019v1 [cs.CV], 2019.
- [2]. Zhang, Chongsheng, Pengyou Wang, Ke Chen, dan Joni-Kristian K., "Identity-aware convolutional neural networks for facial expression recognition", *Journal of Systems engineering and Electronics*, vol. 28, No. 4, pp. 784-792, 2017.
- [3]. Krizhevsky, Alex, Ilya Sutskever, dan Geoffrey E. Hinton. "Imagenetclassification with deep convolutional neural networks." In *Advances in neural information processing systems*, pp. 1097-1105, 2012.
- [4]. P. Liu, S. Han, Z. Meng dan Y. Tong, "Facial expression recognition via a boosted deep belief network", dalam *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1805 - 1812, 2014.
- [5]. Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, dan Yoshua Bengio. "Generative adversarial nets." In *Advances in neural information processing systems*, pp. 2672-2680, 2014.
- [6]. Long, Jonathan, Evan Shelhamer, dan Trevor Darrell. "Fully convolutional networks for semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015.
- [7]. He, K., Gkioxari, G., Dollr, P., dan Girshick, R. "Mask r-cnn" *International Conference on Computer Vision*, pp. 2980-2988, IEEE, 2017.
- [8]. Han, Shizhong, Zibo Meng, Ahmed Shehab Khan, dan Yan Tong. "Incremental boosting convolutional neural network for facial action unit recognition." In *Advances in Neural Information Processing Systems*, pp. 109-117. 2016.
- [9]. Ashlesha Vaidya, Hitee Sachdeva dan R. Brindha, "Facial Expression Recognition using Convolutional Neural Networks", *International Journal of Pure and Applied Mathematics*, vol. 118, No. 22, pp. 245-250, 2018.
- [10]. Masakazu Matsugu, Katsuhito Mori, Yusuke Mitari dan Yuji Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network", *Neural Networks* 16, pp. 555 - 559, 2003.

- [11] . D.V. Sang, N.V. Dat, dan D.P. Thuan, "Facial expression recognition using deep convolutional neural networks", Knowledge and Systems Engineering (KSE) 9th International Conference, 2017.
- [12] . Ke Shan, Junqi Guo, Wenwan You, Di Lu, dan Rongfang Bie, "Automatic Facial Expression Recognition Based on a Deep Convolutional-Neural-Network Structure", SERA 2017, London, UK, IEEE, pp: 123-128, 2017.
- [13] . Bengio, Yoshua. Learning Deep Architecture for AI, volume 2:No.1. Foundation and Trends in Machine Learning, 2009.
- [14] . Yann LeCun, K. Kavukcuoglu, dan C. Farabet. Convolution network and applications in vision. in Proc. IEEE International Symposium on Circuit and Systems, pp. 253–256, 2010.
- [15] . Rere, L.M. Rasdi, M. Ivan Fanany, dan A. Murni Arymurthy, Computational Intelligence and Neuroscience, Volume 2016, Article ID 1537325, 13 pages, 2016.
- [16] . Yanming Guo, Yu Liu, and Song Wu Michael S. Lew Ard Oerlemans, dan Songyang Lao. Deep learning for visual understanding: A review. Neurocomputing, 187:27–48, 2016.