

# Klasifikasi Ketertarikan Anak PAUD Melalui Ekspresi Wajah Menggunakan Metode CNN

Ajeng Restu Kusumastuti  
Fakultas Sains dan Teknologi  
Institut Sains dan Teknologi Terpadu  
Surabaya institusi  
Surabaya, Indonesia  
reztoeadjenk@gmail.com

Yosi Kristian  
Fakultas Sains dan Teknologi  
Institut Sains dan Teknologi Terpadu  
Surabaya institusi  
Surabaya, Indonesia  
yosi@stts.edu

Endang Setyati  
Fakultas Sains dan Teknologi  
Institut Sains dan Teknologi Terpadu  
Surabaya institusi  
Surabaya, Indonesia  
endang@stts.edu

**Abstract**—The character of emotions in children is different from that of adults, where the characteristics of the emotions in children include, (1) Briefly and ends suddenly, (2) Seems greater or stronger, (3) Temporary or superficial, (4) Frequent, (5) Can be known clearly from behaviour, and (6) Reaction reflects individuality. Emotions that are felt can be expressed through faces, this is continuous with how interested the child is to the material presented in front of him. Measuring the interest classification in PAUD children in this study using CNN. In the process of training the level of interest in PAUD children, the accuracy value of the four models always increases from epoch 25 until 100 with the highest value being the Raj Mehrotra architecture. But during the data testing process, the architecture in this study increased slightly and the highest peak reached an accuracy value of 81.66%. It is 3.33% better than the result obtained with the Raj Mehrotra architecture, and 1.66% better than R. Cui architecture.

**Keywords**— Expression, Faces, CNN.

**Abstrak**— Emosi adalah perasaan atau afeksi yang timbul, ketika seseorang berada dalam suatu keadaan yang dianggap penting oleh individu tersebut. Karakteristik emosi pada anak berbeda dengan karakteristik yang terjadi pada orang dewasa, dimana karakteristik emosi pada anak itu antara lain, (1) Berlangsung singkat dan berakhir tiba-tiba, (2) Terlihat lebih hebat atau kuat, (3) Bersifat sementara atau dangkal, (4) Lebih sering terjadi, (5) Dapat diketahui dengan jelas dari tingkah lakunya, dan (6) Reaksi mencerminkan individualitas. Emosi yang dirasakan dapat diekspresikan melalui wajah, hal ini berkesinambungan dengan seberapa tertariknya anak terhadap tayangan materi yang disajikan dihadapannya. Klasifikasi ketertarikan anak PAUD pada penelitian ini diukur menggunakan metode CNN. Dalam proses training klasifikasi ketertarikan anak PAUD, nilai akurasi keempat model selalu mengalami peningkatan mulai dari epoch 25 hingga 100 dengan nilai tertinggi adalah arsitektur Raj Mehrotra. Tetapi saat proses testing data, arsitektur pada penelitian ini mengalami peningkatan secara perlahan dan puncak tertinggi mencapai nilai akurasi 81,66%. Hal tersebut jauh lebih baik 3,33% dibandingkan hasil yang diperoleh dengan arsitektur Raj Mehrotra, dan 1,66% lebih baik dibandingkan arsitektur R. Cui.

**Keywords**— Ekspresi, Wajah, CNN.

## PENDAHULUAN

Anak prasekolah adalah mereka yang berusia antara 3-6 tahun. Mereka biasanya mengikuti program prasekolah. Sedangkan di Indonesia, umumnya mereka mengikuti program tempat penitipan anak (3 bulan – 5 tahun) dan kelompok bermain (usia 3 tahun), sedangkan pada usia 4-6

tahun biasanya mereka mengikuti program taman kanak-kanak. [1]

Pada dasarnya, pendidikan prasekolah (preschool) adalah pendidikan untuk membantu pertumbuhan dan perkembangan jasmani rohani anak didik di luar lingkungan keluarga sebelum memasuki pendidikan dasar. Taman kanak-kanak atau TK sebagai salah satu bentuk pendidikan prasekolah yang ada di jalur pendidikan sekolah merupakan usaha untuk mengembangkan seluruh segi kepribadian anak didik dalam rangka menjembatani pendidikan dalam keluarga ke pendidikan sekolah. [2] Pendidikan prasekolah kemudian dikenal dengan istilah PAUD atau singkatan dari Pendidikan Anak Usia Dini.

Emosi sebagai perasaan atau afeksi yang timbul ketika seseorang berada dalam suatu keadaan yang dianggap penting oleh individu tersebut. Emosi diwakilkan oleh perilaku yang mengekspresikan kenyamanan atau ketidaknyamanan terhadap keadaan atau interaksi yang sedang dialami. Emosi dapat berbentuk rasa senang, takut, marah, dan sebagainya. Karakteristik emosi pada anak berbeda dengan karakteristik yang terjadi pada orang dewasa, dimana karakteristik emosi pada anak itu antara lain, (1) Berlangsung singkat dan berakhir tiba-tiba, (2) Terlihat lebih hebat atau kuat, (3) Bersifat sementara atau dangkal, (4) Lebih sering terjadi, (5) Dapat diketahui dengan jelas dari tingkah lakunya, dan (6) Reaksi mencerminkan individualitas. [3]

Emosi yang dirasakan dapat diekspresikan melalui wajah, seperti yang dijelaskan oleh K Amda dan R Fitriyani melalui bukunya “Membaca Ekspresi Wajah”, dimana ekspresi wajah seseorang menyimpan banyak informasi. Salah satu contoh sederhana adalah ketika anak prasekolah menunjukkan ketertarikan terhadap tayangan yang disajikan dihadapannya melalui ekspresi gembira atau terkejut. Anak yang tertarik akan fokus melihat keseluruhan tayangan, sedangkan jika tidak tertarik maka tatapan mata hingga arah wajah tidak akan fokus melihat tayangan bahkan berpaling seluruhnya dan tidak menampilkan ekspresi gembira (happy). Secara berkesinambungan kondisi pembelajaran yang menyenangkan akan berdampak pada ekspresi kegembiraan anak. Sedangkan ekspresi emosi yang sering terjadi pada anak menurut Hurlock (2001) seperti yang dikutip oleh Dwi Nidika yaitu gembira, sedih, marah, cemburu, dan takut. [4]

Klasifikasi ketertarikan anak PAUD dapat dilakukan dengan mengenali ekspresi wajah. Proses pengenalan ekspresi wajah yang telah banyak diteliti, menggunakan dataset dengan subject orang dewasa seperti FER-2013 dan

CK+. Namun belum adanya penelitian yang melibatkan subject anak PAUD, dimana subject ini memiliki karakteristik yang berbeda dalam menampilkan emosinya. [3] Penelitian ini dilakukan untuk mengklasifikasikan ketertarikan anak PAUD, dengan subject anak-anak berusia 4-5 tahun, yang diambil secara langsung pada salah satu Taman Pendidikan Anak. Pada penelitian ini, algoritma pengenalan ekspresi wajah dilakukan menggunakan CNN (Convolutional Neural Network). Ekspresi yang berhasil dikenali, diharapkan dapat mengklasifikasikan ketertarikan anak PAUD dalam menerima sebuah informasi yang disajikan dihadapannya. Ekspresi yang ditampilkan anak sesuai dengan tampilan tayangan dan merupakan ekspresi dasar positif. Dalam penelitian ini, CNN digunakan untuk melakukan proses training data, sehingga dapat menghasilkan klasifikasi ketertarikan anak sesuai labelling pengajar. Penelitian ini membandingkan hasil pengukuran dari beberapa model CNN untuk mendapatkan nilai akurasi tertinggi dengan error yang kecil. Penelitian ini dapat membantu para pengajar dalam menentukan durasi dan metode pembelajaran online yang tepat digunakan selama masa pandemi. Supaya materi yang disampaikan lebih tepat sasaran dan sesuai dengan keterbatasan interaksi belajar.

#### MODEL CNN

Model yang pertama merupakan penelitian Yiliang Xie Et Al [5] yang berjudul "Improving The Lenet With Batch Normalization And Online Hard Example Mining For Digit Recognition". Dataset yang digunakan peneliti adalah MNIST untuk mengelompokkan digit recognition. Arsitektur CNN yang digunakan oleh peneliti adalah 2 conv layer dengan pembagian layer pertama diterapkan filter 3x3 sebanyak 32 menggunakan ReLU. Kemudian dilanjutkan layer kedua dengan filter dan activation function yang sama sebanyak 64. Setelah itu diikuti max pooling berukuran 2x2 sebelum memasuki proses regularisasi dropout sebesar 0,25. Setelah melalui proses regularisasi, maka hasil dimasukkan kedalam fully connected layer dengan 128 kelas, untuk kemudian diregularisasi kembali dengan dropout sebesar 0,5. Proses terakhir adalah memprosesnya kedalam fully connected layer dengan 3 kelas.

Model CNN kedua, diambil melalui penelitian Yassine Ghouzam [6] dalam menjawab tantangan Kaggle terkait permasalahan deep learning berjudul "Introduction to CNN Keras". Dataset dalam penelitian ini, sama dengan yang digunakan oleh penelitian [5]. Arsitektur CNN yang digunakan oleh peneliti adalah 4 conv layer. Layer pertama diterapkan filter 3x3 sebanyak 32 dengan activation function ReLU, hal ini diulang pada layer kedua. Setelah itu diikuti max pooling berukuran 2x2 sebelum memasuki proses regularisasi dropout sebesar 0,25. Kemudian pada layer ketiga dan keempat, masing-masing menggunakan filter 3x3 sebanyak 64 dengan activation function ReLU. Setelah melalui layer tersebut, maka diterapkan max pooling berukuran 2x2 sebelum memasuki regularisasi dropout sebesar 0,25. Setelah melalui proses regularisasi, maka hasil dimasukkan kedalam fully connected layer dengan 256 kelas, untuk kemudian diregularisasi kembali dengan dropout sebesar 0,5. Proses terakhir adalah memprosesnya kedalam fully connected layer dengan 3 kelas.

Model CNN ketiga, merupakan penelitian Raj Mehrotra [7] yang berjudul "Flower Recognition CNN Keras". Dataset

dalam penelitian ini, sama dengan yang digunakan oleh penelitian [5][6]. Arsitektur CNN yang digunakan adalah 4 conv layer. Layer pertama diterapkan filter 5x5 sebanyak 32 dengan activation function ReLU, dan diikuti max pooling berukuran 2x2. Hal yang sama diterapkan pada layer kedua hingga keempat dengan filter 3x3 sebanyak filter yang berbeda. Layer kedua sebanyak 64, sedangkan layer ketiga dan keempat sebanyak 96. Hasil terakhir dimasukkan kedalam fully connected layer dengan 3 kelas.

Model CNN keempat, merupakan penelitian R. Cui Et Al [8] yang berjudul, "Facial Expression Recognition Based on Ensemble of Multiple CNNs". Dataset yang digunakan peneliti adalah CK+. Adapun arsitektur yang digunakan pada penelitian ini adalah 2 conv layer dimana layer pertama diterapkan filter 7x7 sebanyak 32 dengan activation function ReLU dan diikuti max pooling berukuran 2x2. Pada layer kedua diterapkan filter 5x5 sebanyak 32 dengan activation function dan max pooling berukuran yang sama. Setelah melalui proses regularisasi, maka hasil dimasukkan kedalam fully connected layer dengan 128 kelas, untuk kemudian diproses kedalam fully connected layer dengan 3 kelas.

Model CNN kelima, merupakan penelitian Kuang Liu Et Al [9] yang berjudul, "Facial Expression Recognition with CNN Ensemble". Dataset yang digunakan peneliti adalah FER-2013. Adapun arsitektur yang digunakan pada penelitian ini adalah 3 conv layer dimana layer pertama diterapkan filter 5x5 sebanyak 64 dengan activation function ReLU dan diikuti max pooling berukuran 2x2. Pada layer kedua diterapkan filter 5x5 sebanyak 128 dengan activation function dan max pooling berukuran yang sama. Kemudian pada layer ketiga diterapkan filter 5x5 sebanyak 256 yang diikuti activation function juga maxpooling berukuran sama. Setelah melalui proses regularisasi, maka hasil dimasukkan kedalam fully connected layer dengan 4096 kelas, untuk kemudian diregularisasi kembali dengan dropout sebesar 0,5. Proses terakhir adalah memprosesnya kedalam fully connected layer dengan 3 kelas. Arsitektur yang diujicobakan pada penelitian ini sesuai dengan subnet1.

Model CNN keenam, merupakan penelitian Yangho Ji Et Al [10] yang berjudul, "Sign Language Learning System with Image Sampling and Convolutional Neural Network". Dataset yang digunakan peneliti sesuai dengan penelitian [5][6]. Adapun arsitektur yang digunakan pada penelitian ini adalah 3 conv layer dimana layer pertama diterapkan filter 3x3 sebanyak 32 dengan activation function ReLU dan diikuti max pooling berukuran 4x4. Pada layer kedua diterapkan filter 3x3 sebanyak 64 dengan activation function ReLU diikuti max pooling berukuran 2x2. Kemudian pada layer ketiga diterapkan filter 3x3 sebanyak 128 yang diikuti activation function ReLU diikuti maxpooling berukuran 2x2. Setelah melalui proses regularisasi, maka hasil dimasukkan kedalam fully connected layer dengan 128 kelas, diikuti fully connected layer dengan 3 kelas.

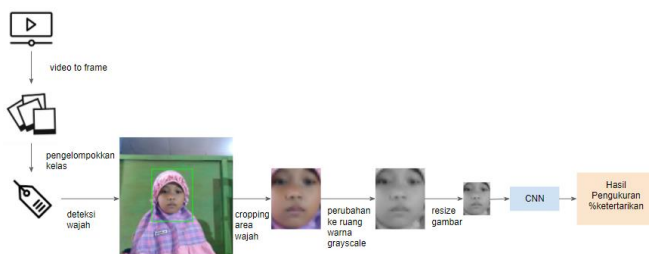
Model CNN ketujuh, merupakan penelitian Su Yang Et Al [11] yang berjudul, "Video-Based Chinese Sign Language Recognition Using Neural Network". Dataset yang digunakan peneliti sesuai dengan penelitian [5][6][10]. Adapun arsitektur yang digunakan pada penelitian ini adalah 2 conv layer dimana layer pertama diterapkan filter 5x5 sebanyak 16 dengan activation function ReLU dan diikuti max pooling berukuran 5x5. Pada layer kedua diterapkan filter 5x5 sebanyak 32 dengan activation function ReLU diikuti max pooling berukuran 3x3. Setelah melalui proses

regularisasi, maka diterapkan regularisasi dengan dropout sebesar 0,2. Setelah itu hasil dimasukkan kedalam fully connected layer dengan 128 kelas, untuk kemudian diregularisasi kembali dengan dropout sebesar 0,5. Proses terakhir adalah memprosesnya kedalam fully connected layer dengan 3 kelas.

Seluruh kompilasi dalam arsitektur model CNN yang diujicobakan pada penelitian ini, menggunakan adam optimizer. Penelitian [5][6][10][11] diujicobakan penerapan metode CNN sebagai pengenalan ekspresi wajah. Fully connected layer terakhir pada setiap penelitian telah disesuaikan dengan kebutuhan pengklasifikasian yaitu 3 kelas yang terdiri dari Cukup Tertarik, Tertarik, dan Tidak Tertarik.

### METODE PENELITIAN

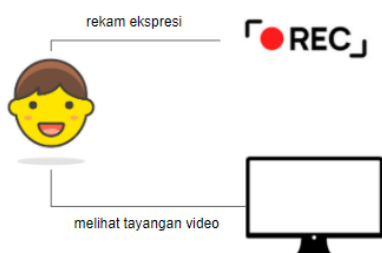
Penelitian ini digunakan untuk mengklasifikasikan ketertarikan anak PAUD melalui ekspresi wajah. Tahapan penelitian yang dilakukan yaitu dengan mengambil hasil rekaman ekspresi wajah anak, kemudian dilakukan pengolahan menggunakan metode CNN.



Gambar 1. Tahapan Penelitian

### A. Dataset

Dataset yang digunakan dalam penelitian ini melibatkan 30 subject dengan rentang usia 4-5 tahun, yang kemudian dikenal dengan anak TK A. Setiap anak dilakukan proses perekaman video menggunakan 5 unit notebook dengan perincian: 1 video kamera internal, 2 video kamera eksternal, dan 2 kamera SLR dalam membantu proses perekaman wajah.



Gambar 2. Proses pengambilan data

Setiap anak melihat tayangan video berdurasi maksimal 5 menit melalui notebook yang disediakan, seperti yang dijelaskan pada gambar 1.

Setelah proses perekaman selesai dilakukan, maka tahapan penelitian dilanjutkan dengan melakukan konversi video ke gambar. Gambar yang didapatkan kemudian dilakukan pengelompokan oleh validator yaitu pengajar.

Hasil pengelompokan tadi kemudian dilakukan proses deteksi wajah, cropping area wajah, perubahan ke ruang warna grayscale. Hal tersebut dijelaskan pada tahapan preprocessing, sebelum diolah pada CNN dan menghasilkan prosentase ketertarikan.

### B. Daftar Kelas

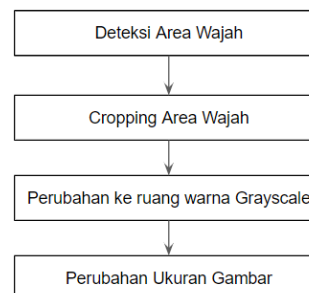
Hasil dari proses perekaman wajah, kemudian diolah menjadi gambar dengan memanfaatkan konverter *video to image* pada OpenCV. Total dataset yang digunakan sebanyak 243 gambar, terbagi kedalam 3 kelas yang dikelompokkan secara langsung oleh pengajar yaitu Tertarik (T), Cukup Tertarik (CT), dan Tidak tertarik (TT). Contoh gambar siswa yang mewakili setiap kelas dijelaskan pada tabel 2.

TABEL II. DAFTAR KELAS KETERTARIKAN ANAK PAUD

| Nama Kelas          | Contoh Gambar |
|---------------------|---------------|
| Tertarik (T)        |               |
| Cukup Tertarik (CT) |               |
| Tidak Tertarik (TT) |               |

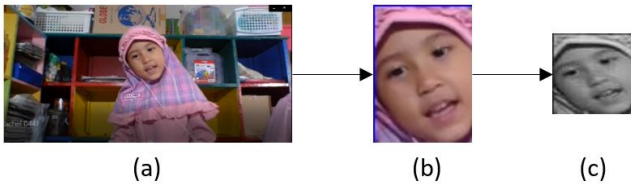
### C. Preprocessing

Tahap preprocessing dilakukan untuk memudahkan proses training, sehingga waktu yang dibutuhkan pada setiap *epoch* lebih singkat. Tahapan preprocessing dilakukan dengan cara mendeteksi area wajah dengan memanfaatkan algoritma *viola jones*. [8] Setelah area wajah dideteksi, maka selanjutnya dilakukan pemotongan area tersebut yang dilanjutkan dengan perubahan ruang warna dari RGB ke grayscale dan perubahan ukuran menjadi 48x48 pixel. Adapun alur preprocessing dapat dilihat pada gambar 1.



Gambar 3. Tahapan Preprocessing

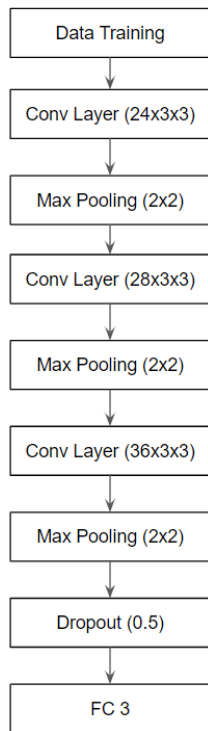
Langkah awal dalam tahapan ini adalah melakukan deteksi area wajah dengan memanfaatkan algoritma viola jones. [8] Setelah area wajah dideteksi, maka selanjutnya dilakukan pemotongan area tersebut yang dilanjutkan dengan perubahan ruang warna dari RGB ke grayscale dan perubahan ukuran menjadi 48x48 pixel.



Gambar. 4. (a) Hasil video to frame (b) Hasil deteksi dan cropping area wajah (c) Hasil perubahan ukuran dan warna

D. Model Penelitian

Arsitektur yang diujicoba oleh peneliti merupakan hasil kombinasi dan perubahan, dengan melihat hasil training dan testing penelitian sebelumnya. Dengan melakukan perubahan pada conv layer dan nilai dropout.



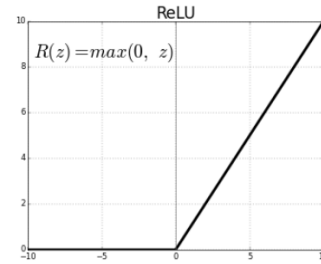
Gambar. 5. Arsitektur CNN

Dalam penelitian ini menggunakan 4 conv layer dimana layer pertama menggunakan filter 3x3 sebanyak 24 kernel, dan diterapkan activation function ReLU lalu diikuti max pooling 2x2. Layer kedua menggunakan filter 3x3 sebanyak 28 kernel, dan diterapkan activation function ReLU dan diikuti max pooling 2x2. Layer ketiga menggunakan filter 3x3 sebanyak 32 kernel, dan diterapkan activation function ReLU dan diikuti max pooling 2x2. Layer kedua menggunakan filter 3x3 sebanyak 36 kernel, dan diterapkan activation function ReLU dan diikuti max pooling 2x2. Optimizer yang digunakan adalah RMSprop dengan lr = 0,0004 dan diregularisasi menggunakan dropout sebesar 0,5. Regularisasi dropout diterapkan sebesar 0,5 untuk

menghindari terjadinya overfitting. Layer terakhir menggunakan fully connected layer dengan hasil keluaran 3 kelas untuk mewakili setiap tingkat ketertarikan.

Setiap conv layer terdapat hidden layer dengan activation function ReLU. Hal tersebut berengaruh pada efisiensi waktu komputasi dan meminimalisir terjadinya overfitting. Adapun fungsi ReLU dijelaskan pada rumus (1), dan visualisasi dijelaskan pada gambar 6. Activation function ini juga diadaptasi oleh peneliti berdasarkan arsitektur penelitian sebelumnya.

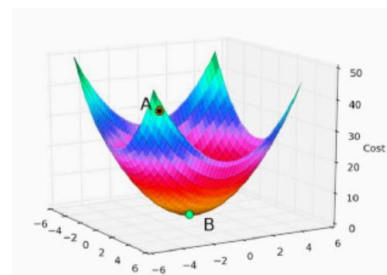
$$f(x) = \max(0, x) \tag{1}$$



Gambar. 6. Aktivasi ReLU

Sedangkan digunakan pooling layer dengan jenis max pooling berukuran 2x2. Pada prosesnya hal tersebut bergeser keseluruhan bagian dari feature map, dengan mengambil nilai tertinggi. Hal ini digunakan untuk mengurangi dimensi gambar sebagai upaya mempercepat proses komputasi dalam meminimalisir overfitting.

Nilai dropout diletakkan pada akhir sebelum diterapkannya fully connected layer, sebesar 0,5. Dropout merupakan sebuah teknik regularisasi dimana beberapa neuron dipilih secara acak dan tidak akan digunakan dalam proses training. Pada arsitektur ini, algoritma optimasi yang digunakan adalah RMSprop. RMSprop atau Root Mean Square Propagation digunakan untuk mendapatkan hasil yang lebih baik dengan memperbaharui weight dan bias. Sehingga didapatkan nilai minimum global dimana cost function mencapai nilai terkecil, seperti yang dijelaskan pada gambar 7.

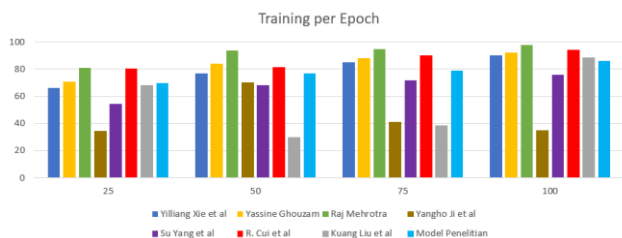


Gambar. 7. Visualisasi 3D RMSprop

HASIL DAN PEMBAHASAN

Proses training data menggunakan laptop dengan spesifikasi Core i5, memory 8 GB, dengan CPU 2,49 GHz. Library yang digunakan pada penelitian ini adalah Tensorflow dan Keras.

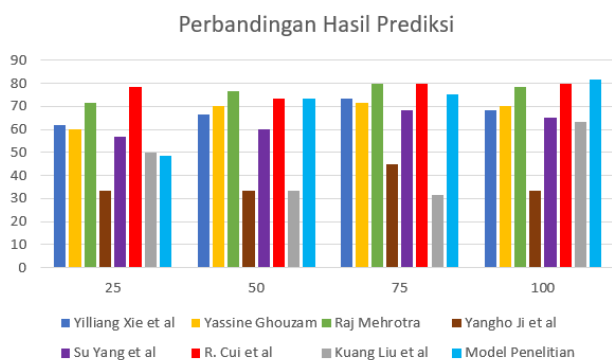
Dataset yang dijelaskan pada poin B bagian II, hanya mencakup keperluan training data yaitu sebanyak 243 data. Sedangkan untuk testing, menggunakan 60 data. Proses training data dibagi kedalam 4 tahapan yaitu 25 *epoch*, 50 *epoch*, 75 *epoch*, dan 100 *epoch*. Sehingga dapat dilihat perbandingan nilai akurasi yang didapatkan dari setiap tahapan.



Gambar 8. Perbandingan Hasil Training

Hasil training per *epoch* yang digambarkan pada gambar 8, menjelaskan uji coba model penelitian menghasilkan nilai yang meningkat secara perlahan yaitu 70, 77, 79, dan 86. Sedangkan nilai tertinggi pada *epoch* tertinggi yaitu 100, berada pada penelitian Raj Mehrotra. Nilai yang didapatkan secara berturut yaitu 81, 94, 95, dan 98. Hasil uji coba dengan arsitektur R. Cui et al merupakan salah satu hasil training tertinggi dengan nilai 80.36, 81.52, 90.18, dan 94.31.

Berdasarkan hasil nilai prediksi, model penelitian yang diuji coba diadaptasi dengan menggunakan 4 conv layer seperti penelitian [6], dengan melakukan perubahan pada filter, mengalami peningkatan. Nilai yang didapatkan yaitu 48.33, 73.33, 75, dan 81.66. Hal tersebut lebih tinggi 1.66% dibandingkan percobaan dengan arsitektur [8] dan 3.33% lebih tinggi dibandingkan percobaan dengan arsitektur [7]. Adapun perbandingan hasil prediksi dijelaskan melalui gambar 9.



Gambar 9. Perbandingan Hasil Prediksi

Melalui grafik dapat dilihat bahwa model penelitian secara stabil mengalami peningkatan hasil prediksi per *epoch*. Sedangkan arsitektur penelitian lain mengalami peningkatan di *epoch* 50-75 dan sedikit menurun atau stabil pada *epoch* terakhir.

## KESIMPULAN

Dengan ditambahkan dropout layer dan RMSprop optimizer pada fully connected layer, mampu meningkatkan performansi arsitektur CNN yang diusung. Hasilnya dapat dilihat melalui training dan testing, dimana model penelitian mampu mendapatkan peningkatan nilai secara perlahan, dan lebih baik dibandingkan uji coba dengan penelitian sebelumnya.

Dalam proses training, hasil tertinggi diperoleh dari hasil uji coba arsitektur Raj Mehrotra dan R. Cui, selisih yang diperoleh yaitu 14 dan 8.6. Sedangkan menurut hasil testing, selisih yang didapatkan antara kedua hasil training tertinggi dengan model penelitian, mengalami penurunan nilai hingga menghasilkan 3.33% dan 1.66% lebih tinggi.

## REFERENSI

- [1] Soemiarti Patmonodewo, "Pendidikan Anak Prasekolah," (Jakarta: PT. Rineka Cipta, 2003), Cet. 2, hlm. 19.
- [2] Lift Anis Ma'shumah, "Pembinaan Kesadaran Beragama Pada Anak, dalam Ismail SM (eds), Paradigma Pendidikan Islam," (Yogyakarta: Pustaka Pelajar, 2001), hlm. 216-217.
- [3] F. Nurmalitasari, P. M. Psikologi, F. Psikologi, and U. Gadjah, "Perkembangan Sosial Emosi pada Anak Usia Prasekolah," vol. 23, No. 2, pp. 103-111, 2015.
- [4] K Amda, R Fitriyani, "Membaca Ekspresi Wajah," Huta Publisher, Depok, 2016, hlm.17.
- [5] Yiliang Xie, Hongyuan Jin, Eric C.C Tsang, "Improving The Lenet With Batch Normalization And Online Hard Example Mining For Digits Recognition," IEEE Int. Conf. on Wavelet and Pattern Recognition. Eng., no. October, pp. 149-153, 2017.
- [6] Yassine Ghouzam, 2017. Introduction to CNN Keras - 0.997 (top 6%). [Online] [https://www.kaggle.com/yassinighouzam/introduction-to-cnn-keras-0-997-top-6#Introduction-to-CNN-Keras---Acc-0.997-\(top-8%\)](https://www.kaggle.com/yassinighouzam/introduction-to-cnn-keras-0-997-top-6#Introduction-to-CNN-Keras---Acc-0.997-(top-8%)) [Diakses 24 Agustus 2020].
- [7] Raj Mehrotra, 2018. Flower Recognition CNN Keras. [Online]. <https://www.kaggle.com/rajmehra03/flower-recognition-cnn-keras> [Diakses 24 Agustus 2020].
- [8] Ruoxuan Cui, Minyi Liu, and Manhua Liu, "Facial Expression recognition Based on Ensemble of Multiple CNNs", Springer International Publishing AG 2016, pp. 511-518, 2016.
- [9] Kuang Liu, Minming Zhang, and Zhigeng Pan, "Facial Expression Recognition with CNN Ensemble", IEEE Int. Conf. on Cyberworlds, pp. 163-166, 2016.
- [10] Yangho Ji, Sunmok kim, and Ki-Baek Lee, "Sign Language Learning System with Image Sampling and Convolutional Neural Network", IEEE Int. Conf. on Robotic Computing. Eng., no. May, pp. 371-375, 2017.
- [11] Su Yang, Qing Zhu, "Video-Based Chinese Sign Language Recognition Using Convolutional Neural Network", IEEE Int. Conf. on Com Software and Network. Eng., no May, pp. 929-934, 2017.