

Klasifikasi Penyakit Paru-Paru Berbasis Pengolahan Citra X Ray Menggunakan Convolutional Neural Network (Classification Of The Lung Diseases Based On X Ray Image Processing Using Convolutional Neural Network)

1st Razief Moch Diar
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
raziefmo@student.telkomuniversit
y.ac.id

2nd R. Yunendah Nur Fu'Adah
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
yunendah@telkomuniversity.ac.id

3rd Koredianto Usman
Fakultas Teknik Elektro
Universitas Telkom
Bandung, Indonesia
korediantousman@telkomuniversit
y.ac.id

Abstrak

Penyakit pada paru-paru merupakan gangguan yang cukup serius dimana dapat menyerang sistem pernapasan manusia dan bisa berakibat fatal jika tidak ditangani dengan serius. Pada saat ini deteksi penyakit pada paru-paru masih dilakukan secara manual oleh para dokter ahli, namun proses secara manual memakan waktu lama. Oleh karena itu, dalam tugas akhir ini dibuat sistem yang dapat mendeteksi dan mengklasifikasi penyakit paru-paru dengan otomatis. Pada Tugas Akhir ini merancang sistem otomatis untuk mengklasifikasi kondisi paru-paru berdasarkan citra x-ray paru-paru berbasis *Convolutional Neural Network* (CNN) menggunakan arsitektur *MobileNet*. Perancangan pada sistem dibagi menjadi beberapa tahapan dimulai dari menginput data citra x-ray paru-paru, tahap selanjutnya *preprocessing*, pada penelitian ini menggunakan dua jenis *preprocessing*, yaitu *CLAHE*, dan *Gaussian filter*, lalu dari hasil *preprocessing* dilakukan tahap pelatihan dengan dua jenis *optimizer* yang berbeda, yaitu *Stochastic Gradient Descent* (SGD), dan *Adaptive moment* (Adam). Tahap terakhir mengklasifikasi data citra menjadi empat kelas, yaitu *Viral Pneumonia*, *Coronavirus Disease-19* (Covid-19), *Normal* dan *Tuberculosis*. Hasil akhir dari penelitian ini menunjukkan *optimizer* terbaik yaitu Adam menggunakan *preprocessing* *CLAHE* pada epoch 50

dan menghasilkan nilai akurasi sebesar 94,687 dan *loss* sebesar 0,148. Selain itu juga diperoleh hasil dari performansi sistem berupa presisi 95%, *recall* 93%, dan *F-1 score* sebesar 94%.

Kata Kunci : CNN, *MobileNet*, citra x-ray paru-paru, *Viral Pneumonia*, *Coronavirus Disease-19* (Covid-19), *Normal*, *Tuberculosis*.

Abstract

Diseases of the lungs are quite serious disorders which can attack the human respiratory system and can be fatal if not treated seriously. At this time the detection of disease in the lungs is still check manually by expert doctors, but manual process takes a long time. Therefore, in this final project, a system is made that can detect and classify lung diseases automatically. using MobileNet architecture. The design of the system is divided into several stages starting from inputting lung x-ray image data, the next stage is preprocessing, in this study using two types of preprocessing, namely CLAHE, and Gaussian filters, then from the results of preprocessing, the training phase is carried out with two types of optimizers that different, namely Stochastic Gradient Descent (SGD), and Adaptive moment (Adam). The last stage is to classify image data into four classes, namely Viral Pneumonia, Coronavirus Disease-19 (Covid-19), Normal and Tuberculosis. The final result of this study shows that the best optimizer is Adam using CLAHE preprocessing on epoch 50 and produces an accuracy value of 94,687 and a loss of 0.148. In addition, the results of the system performance are 95% precision, 93% recall, and an F-1 score of 94%.

Keywords: CNN, *MobileNet*, lung x-ray images, *Viral Pneumonia*, *Coronavirus Disease-19* (Covid-19), *Normal*, *Tuberculosis*

I. PENDAHULUAN

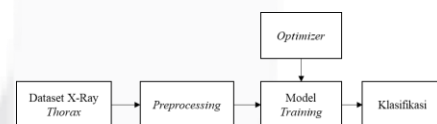
Paru-paru adalah salah satu organ dari sistem pernapasan yang berfungsi sebagai tempat

terjadinya pertukaran karbondioksida dan oksigen dalam darah. Permasalahan yang sering terjadi pada sistem pernapasan adalah kualitas udara yang telah tercemar, sehingga udara yang dihirup mengandung banyak bakteri yang dapat menyerang sistem pernapasan terutama paru-paru [1]. Gangguan pada paru-paru merupakan gangguan yang cukup serius dimana dapat menyerang sistem pernapasan manusia dan bisa berakibat fatal jika tidak ditangani dengan serius. Hal ini dapat menyebabkan penderitaanya sulit bernafas, sulit beraktivitas, dan kekurangan oksigen bahkan apabila tidak cepat ditangani dapat menyebabkan kematian [2]. Adapun gangguan pada paru-paru yang umum kita jumpai, diantaranya *Viral Pneumonia*, *Coronavirus Disease-19* (Covid-19) dan *Tuberculosis*. Dalam dunia medis, salah satu cara untuk mengidentifikasi kelainan-kelainan yang ada pada paru-paru adalah melihat citra dari paru-paru tersebut yang diperoleh dari foto rontgen (X-Ray). Selain melakukan pemeriksaan X-Ray, citra paru-paru juga bisa didapatkan menggunakan CT-scan (*Computerized Tomography Scan*) dan MRI (*Magnetic Resonance Imaging*) akan tetapi pada umumnya digunakan citra X-Ray untuk mengidentifikasi penyakit pada paru-paru. Disini peran dari seorang dokter atau tenaga medis sangatlah dibutuhkan, akan tetapi semakin sering seorang dokter atau tenaga medis membaca citra X-Ray maka hal itu akan membuat mereka kelelahan terlebih lagi dalam jumlah yang banyak. Hal ini menyebabkan seorang tenaga medis membutuhkan waktu yang lama dalam membaca citra dan bisa saja melakukan kesalahan dalam membaca citra X-Ray tersebut karena kurang istirahat [1]. Oleh sebab itu, dibutuhkan sebuah sistem yang dapat membantu tenaga medis untuk mendeteksi penyakit dan mengklasifikasikan citra X-ray secara otomatis. Telah banyak dilakukan penelitian tentang klasifikasi citra berupa X-Ray paru-paru secara otomatis dengan berbagai model, seperti menggunakan *Pre-trained* model *Convolutional Neural Network* yang berfokus pada penyakit menular *Tuberculosis* dan Normal. Dengan menggunakan beberapa model (*ResNet50*, *DenseNet121*, *MobileNet*, *Xception*, *InceptionV3*, *InceptionResNetV2*) dan didapatkan hasil akurasi tertinggi sebesar 91,57% menggunakan model *DenseNet121*. Akan tetapi, model ini hanya menggunakan dua kelas saja dan jumlah dataset yang sedikit [3]. Penelitian lainnya juga telah dilakukan menggunakan model *Convolutional Neural Network* yang menghasilkan tingkat akurasi

sebesar 89,58%, tetapi model ini hanya menggunakan dua kelas saja, yaitu *Normal* dan *Viral Pneumonia* [4]. Namun pada penelitian ini, data tidak disegmentasi terlebih dahulu sehingga tidak diperoleh hasil yang optimal. Ada juga penelitian terhadap X-Ray paru-paru dengan ekstraksi fitur *Local Binary Pattern* menggunakan *Support Vector Machine* dengan tingkat akurasi 65,63 %. Akan tetapi, penelitian ini hanya menggunakan dua kelas saja [5]. Pada Tugas Akhir ini, akan digunakan *Convolutional Neural Network* dengan menggunakan arsitektur *MobileNet*. Perbedaan *MobileNet* dengan arsitektur lainnya adalah penggunaan konvolusi *layer* dengan ketebalan *filter* yang sesuai dengan ketebalan *image* yang diinput. Konvolusi pada *MobileNet* terbagi menjadi *depthwise convolution* dan *pointwise convolution* [6]. Sistem diharapkan dapat mendeteksi penyakit dan mengklasifikasi paru-paru yang terdiri dari empat kelas yaitu *Viral Pneumonia*, *Coronavirus Disease-19* (Covid-19) dan *Tuberculosis*.

II. KAJIAN TEORI

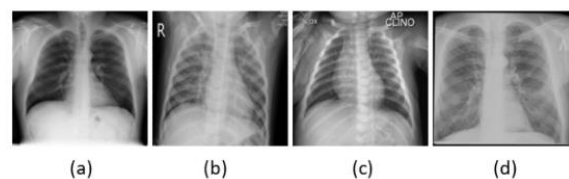
Penelitian ini dibagi menjadi 5 tahap, yaitu tahap pengambilan data, tahap *preprocessing* yaitu proses peningkatan kualitas pada citra. Tahap selanjutnya klasifikasi menggunakan CNN dengan arsitektur *mobilenet*. Kemudian tahap pengujian sistem menggunakan dua skenario pengujian. Gambar 1 merupakan blok diagram tahapan penelitian



Gambar 1. Metode Penelitian

a. DATASET CITRA X-RAY PARU-PARU

Pada tugas akhir ini digunakan data set dari oleh Tawsifur Rahman yang dapat diakses melalui *Kaggle* secara *online*. Disini digunakan data set dengan kategori gambar paru-paru normal, covid-19, *Viral Pneumonia* dan *tuberculosis* dengan pembagian sebagai Gambar berikut.



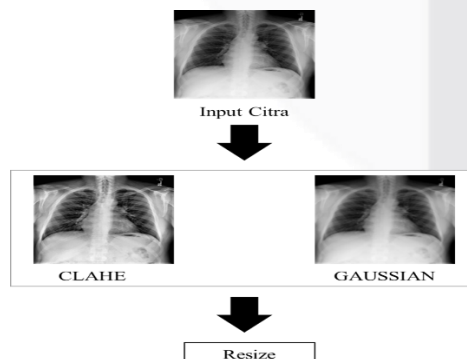
Gambar 2. Dataset X-ray paru-paru (a) Normal, (b) *Viral Pneumonia*, (c) Covid-19, (d) *Tuberculosis*

Dataset yang digunakan sejumlah 15.853 citra X-Ray *Paru-paru* dengan 3.616 citra X-Ray *Paru-paru* Covid, 10.192 citra X-Ray *Paru-paru* Normal, 700 citra X-Ray *Paru-paru* *Tuberculosis*, dan 1.345 citra X-Ray *Paru-paru* *Viral Pneumonia*.

1. *Viral Pneumonia* merupakan infeksi pada saluran pernapasan yang disebabkan oleh bakteri, virus, dan jamur. Infeksi terjadi pada salah satu maupun kedua paru-paru yang dapat menyebabkan kantung udara pada paru-paru sulit untuk menyerap oksigen [10].
2. *Coronavirus* (Covid-19) adalah penyakit yang menyebabkan infeksi pada saluran pernapasan yang disebabkan oleh virus. Penyakit ini dapat menyebabkan demam tinggi, batuk, dan kesulitan bernapas. Parahnya lagi, infeksi ini dapat menyebabkan *Viral Pneumonia*, kegagalan pada ginjal, dan kematian [11].
3. *Tuberculosis* (TB) adalah penyakit menular yang menyerang paru-paru disebabkan karena adanya *Mycobacteria Tuberculosis* pada paru-paru. Penularannya penyakit ini dapat melalui udara lalu masuk ke sistem pernapasan. Penyakit ini harus segera ditangani agar tidak memperburuk kondisi paru-paru dan lebih parahnya dapat menyebabkan kematian [3].

b. PREPROCESSING

Preprocessing dilakukan untuk mendapatkan kualitas citra yang lebih baik dari sebelumnya, agar mempermudah untuk mendapatkan ciri pada citra. Pada Gambar 3, berikut tahapan dan penjelasan pada proses *preprocessing*.



Gambar 3. Alur kerja tahap *preprocessing*

Langkah pertama pada proses *preprocessing* yaitu melakukan input citra X-Ray *Paru-paru* dan citra tersebut akan diproses

dengan dua jenis *preprocessing* yaitu *Gaussian* dan CLAHE. Pada CLAHE akan dilakukan peningkatan kontras pada data citra. CLAHE bekerja dengan menetapkan nilai batas atau *clip limit* pada histogram. Pada *Gaussian filter* sebagai metode untuk menghaluskan, mengurangi *noise*, dan meningkatkan kualitas *detail* gambar. Keluaran dari proses *preprocessing* ini didapatkan dua model gambar berbeda. *Gaussian Filter* adalah filter yang *respons impuls* nya berupa fungsi *Gaussian*. *Gaussian Filter* merupakan salah metode yang biasa digunakan dalam pengolahan citra untuk menghaluskan gambar, mengurangi *noise* pada gambar, dan meningkatkan kualitas *detail* gambar [22]. *Gaussian* adalah model *noise* yang terdistribusi standar dengan rata-rata nol dan standar deviasi 1. Persamaan *Gaussian Filter* sebagai berikut.

$$G(X, Y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (1)$$

Dimana σ adalah standar deviasi. Semakin besar σ , jarak antar piksel lemah atau jauh, tapi untuk deteksi (tepi, *noise*, dsb) semakin bagus. Semakin kecil σ , jarak antar piksel bagus, akan tetapi deteksi (tepi, *noise*, dsb) lemah. (x, y) merupakan elemen matriks kernel *gauss*, π adalah konstanta dengan nilai 3,14, e adalah konstanta bilangan dengan nilai 2,718281828.

Contrast limited Adaptive Histogram Equalization (CLAHE) adalah penggabungan dari metode *Adaptive Histogram Equalization* (AHE) dan *Histogram Equalization* (HE). Metode ini menggunakan teori probabilitas dalam tingkat keabuan dalam meningkatkan dan membatasi masalah kontras pada citra digital dengan memberikan nilai yang diketahui untuk ambang batas atau *clip limit* [23]. Nilai batas atau *clip limit* merupakan batas maksimum suatu *histogram*. Cara menghitung *clip limit* suatu *histogram* dapat dilihat pada persamaan berikut [24].

$$\beta = \frac{M}{N} \left(1 + \frac{\sigma}{100} (S_{max} - 1) \right) \quad (2)$$

Variabel M menyatakan luas region, N menyatakan nilai *grayscale* dan σ merupakan *clip factor* menyatakan penambahan batas *limit histogram* yang bernilai antara 0 sampai 100. *Histogram* diatas nilai *clip limit* dianggap kelebihan (*excess*) piksel yang akan didistribusikan kepada area sekitar dibawah *clip limit* sehingga *histogram* merata [24].

c. FUNGSI OPTIMIZER

Optimizer adalah algoritma optimisasi berfungsi untuk mendapatkan bobot optimal, meminimalkan kesalahan dan memaksimalkan akurasi. Pada proses pelatihan, parameter (bobot) dapat diubah untuk mencoba dan meminimalkan fungsi kerugian, agar mampu memprediksi seakurat mungkin Adapun algoritma optimisasi yang akan digunakan pada penelitian ini antara lain *adaptive moment optimization* (Adam), dan *Stochastic Gradient Descent* (SGD). *Adaptive moment optimization* (Adam) adalah salah satu adaptif *learning rate* optimasi yang mengkombinasikan *rmsprop* dan *momentum*. *Adaptive moment optimization* (Adam) merupakan algoritma optimasi berbasis gradien orde pertama dimana efisien secara komputasi membutuhkan jumlah memori yang sedikit [25]. Adapun persamaan matematis dari Adam sebagai berikut.

$$\theta_{t+1} = \theta_t - \alpha \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t \quad (3)$$

Dengan θ adalah parameter yang akan diperbaiki (bobot), α adalah *step size* atau *learning rate*, $\hat{m}_t$ adalah estimasi momentum dengan bias yang dikoreksi pada saat itu, $\hat{v}_t$ adalah estimasi *subgradien* adaptif dengan bias yang dikoreksi pada saat itu.

Stochastic Gradient Descent (SGD) adalah *optimizer* yang populer dalam melakukan optimasi pada *Artificial Neural Network* dan *optimizer* yang paling sederhana tetapi membutuhkan waktu yang cukup lama untuk mendekati konvergen. Algoritma ini digunakan untuk mengupdate sebuah parameter seperti bobot (*weight*) dan bias. Keunggulan dari SGD adalah meminimalisir memori yang digunakan pada saat mencari *weight* baru [26]

$$\theta = \theta - \alpha \times \nabla_{\theta} J(\theta; x^{(i)}; y^{(i)}) \quad (4)$$

Dengan θ adalah bobot atau bias, $\nabla_{\theta} J(\theta; x^{(i)}; y^{(i)})$ merupakan *gradient* terhadap bobot atau bias, x dan y , x adalah input, y adalah label, α *learning rate*.

d. PARAMETER UJI

1. Akurasi

Akurasi merupakan rasio prediksi benar (positif dan negatif) terhadap keseluruhan data. Nilai akurasi dapat diperoleh dengan rumus sebagai berikut [27].

$$\text{Akurasi} = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (5)$$

Dengan TP (*True Positive*) merupakan kasus dimana diprediksi (*Positive*), hasil *positive*, TN (*True Negative*) merupakan kasus dimana diprediksi (*Negative*), hasil *positive*, FP (*False Positive*) merupakan kasus dimana yang diprediksi *positive*, ternyata tidak. Hasil salah (*Negative*), FN (*False Negative*) merupakan kasus dimana yang diprediksi *positive*, ternyata tidak. Hasilnya salah (*negative*).

2. Presisi

Presisi adalah rasio prediksi benar positif terhadap hasil prediksi positif secara keseluruhan. Nilai Presisi dapat diperoleh dengan rumus sebagai berikut [27].

$$\text{Presisi} = \frac{(TP)}{(TP+FP)} \quad (6)$$

Dengan TP (*True Positive*) merupakan kasus dimana diprediksi (*Positive*), hasil *positive*, FP (*False Positive*) merupakan kasus dimana yang diprediksi *positive*, ternyata tidak. Hasil salah (*Negative*).

3. Recall

Recall adalah rasio prediksi positif benar terhadap keseluruhan data positif benar. Nilai *recall* dapat diperoleh dengan rumus berikut sebagai berikut [27].

$$\text{Recall} = \frac{(TP)}{(TP+FN)} \quad (7)$$

Dengan TP (*True Positive*) merupakan kasus dimana diprediksi (*Positive*), hasil *positive*. FN (*False Negative*) merupakan kasus dimana yang diprediksi *positive*, ternyata tidak. Hasilnya salah (*negative*).

4. F1- Score

F1-Score adalah kalkulasi evaluasi dalam pencarian informasi yang menggabungkan hasil dari *recall* dan *presisi*. Pada situasi tertentu, nilai *recall* dan *presisi* dapat memiliki bobot yang berbeda. Nilai F-1 score dapat diperoleh dengan persamaan sebagai berikut [28].

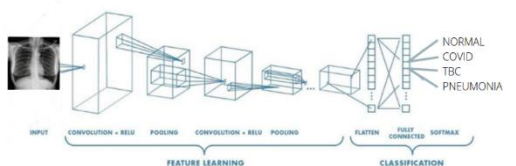
$$F1 \text{ Score} = 2 \times \frac{(\text{Recall} \times \text{Presisi})}{(\text{Recall} + \text{Presisi})} \times 100 \quad (8)$$

5. Confusion Matrix

Confusion Matrix adalah metode yang biasa digunakan untuk mengukur suatu classifier dalam melakukan prediksi pada kategori yang berbeda. Ada empat istilah yang merupakan representasi hasil proses klasifikasi pada *confusion matrix* yaitu *Total True Positif* (TTP), *Total True Negatif* (TTN), *Total False Positif* (TFP), dan *Total False Negatif* (TFN).

III. METODE

Convolutional Neural Networks (CNN) merupakan jaringan saraf buatan. Model algoritma ini digunakan untuk melakukan klasifikasi secara langsung melalui gambar, *video*, teks atau *audio*. CNN juga salah satu algoritma yang paling terkenal pada metode *deep learning*. CNN bisa digunakan untuk mendiagnosis data berbasis gambar



(misalnya sinar-X dan OCT) yang digunakan untuk mendiagnosis keadaan medis. CNN mempunyai dua tahapan utama, yaitu *feature learning* dan *classification*. *Feature learning* terdiri menurut *convolution layer*, ReLU (fungsi aktivasi) dan *pooling layer* sedangkan pada *classification* terdiri dari *flatten*, *fully-connected layer*, dan *softmax* [15].

Gambar 4. Convolutional Neural Network

a. Convolutional Layer

Convolution Layer merupakan lapisan yang paling penting berdasarkan metode CNN. Hasil dari *convolution layer* merupakan citra yg berupa citra masukan. Setiap citra masukan, memakai *filter convolution layer* pada prosesnya. *Convolutional layer* terdiri dari neuron yang tersusun membangun *filter* panjang dan tinggi (*pixel*) [16].

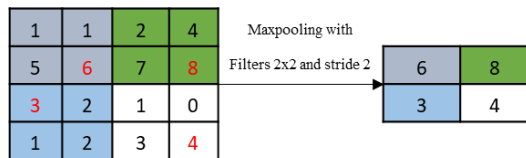
b. ReLU

ReLU adalah fungsi aktivasi yang bertanggung jawab untuk menormalkan nilai yang dihasilkan oleh *convolutional Layer*, misalkan dalam fungsi aktivasi *rectifier* menormalisasikan nilai sehingga tidak terdapat nilai yang berada di bawah 0 menggunakan fungsi $\max(0, x)$. Jika terdapat matriks input x maka nilai ReLU bisa diperoleh pada Persamaan [17].

$$f(x) = \max(0, x) \quad (9)$$

c. Pooling Layer

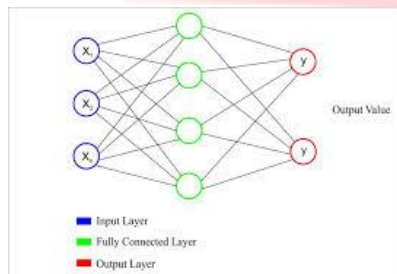
Pooling layer Merupakan lapisan fungsi *Feature Maps* sebagai masukan dan mengolahnya menggunakan berbagai operasi statistik pada nilai piksel terdekat. Lapisan *pooling* yang dimasukkan diantara lapisan *konvolusi* secara berturut-turut pada susunan arsitektur CNN secara progresif bisa mengurangi ukuran dari *volume output* pada *Feature Map* [18]. *Pooling layer* membuat data menjadi lebih kecil, lebih mudah dikelola, dan lebih mudah untuk mengontrol *overfitting* [16]. Pada *layer pooling* dilakukan beberapa proses, yaitu *MaxPooling*, *Average Pooling* dan *Sum Pooling*. Lapisan *pooling* bekerja di setiap susunan *Feature Maps* dan mengurangi ukurannya. Bentuk umum lapisan *pooling* adalah menggunakan filter dengan ukuran 2x2 yang diaplikasikan dengan langkah sebanyak dua dan kemudian beroperasi pada setiap irisan input. *Pooling* yang biasa digunakan adalah *MaxPooling*. *MaxPooling* dilakukan dengan menetapkan nilai maksimum pada suatu *window* [16].



Gambar 5. Pooling layer

d. Fully Connected Layer

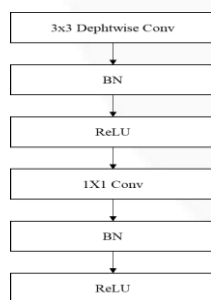
Lapisan *Fully-connected* adalah lapisan yang semua *neuron* aktif lapisan sebelumnya terhubung dengan *neuron* lapisan berikutnya sehingga dapat diimplementasikan oleh jaringan syaraf tiruan. Setiap aktivitas dari lapisan sebelumnya harus diubah menjadi data satu dimensi sebelum dapat ditautkan ke semua *neuron* di lapisan *Fully-Connected* [16].



Gambar 6. Fully Connected Layer

e. MobileNet

MobileNet merupakan salah satu arsitektur dari *Convolutional Neural Network* yang dapat digunakan untuk mengatasi kebutuhan akan *computing resource* berlebih. Konvolusi pada *MobileNet* terbagi menjadi *depthwise convolution* dan *pointwise convolution* [6]. *MobileNet* memiliki ukuran *weight* lebih kecil dan waktu komputasi



Gambar 7. depthwise separable convolutions

lebih cepat pada proses pelatihan, yang dapat dengan mudah diimplementasikan sesuai dengan kebutuhan aplikasi *mobile* dan *embedded applications* [20]. *MobileNet* didasarkan pada konvolusi *depthwise*, yaitu proses konvolusi faktorial, di mana dilakukan perhitungan konvolusi standar menjadi konvolusi *depthwise* [6]. Konvolusi *depthwise* menggunakan *filter* pada setiap saluran input, dan Konvolusi

pointwise menggunakan konvolusi 1×1 untuk menggabungkan output dengan konvolusi *depthwise*. Konvolusi biasanya memfilter dan menggabungkan

input menjadi satu set keluaran yang baru dalam satu langkah. Akan tetapi konvolusi *separable depthwise* membagi ini menjadi dua lapisan, yaitu lapisan penyaringan dan lapisan penggabungan. Hal ini menyebabkan berkurangnya jumlah kalkulasi dan ukuran mode [6].

IV. HASIL DAN PEMBAHASAN

Spesifikasi perangkat yang digunakan pada pengujian sistem ini, yaitu bahasa pemrograman Python 3.8, *software pycharm community*, dan *hardware*, dengan spesifikasi Intel(R) Core (TM) i7-8550U CPU @ 1.80GHz. 2.0 GHz, dan memori 8.00gb RAM.

a. Pengujian preprocessing

Bagian ini merupakan pengujian terhadap pengaruh *preprocessing* pada sebuah citra sebelum proses klasifikasi. *Preprocessing* yang digunakan yaitu CLAHE, dan *gaussian filter*. Pengujian ini membandingkan hasil citra original dengan citra hasil *preprocessing*. *Optimizer* yang digunakan adalah Adam dengan *learning rate* 0.001, *batch size* 32, dan *epoch* 50. Tabel 1 merupakan hasil pengujian *preprocessing* terhadap performa akurasi dan *loss*.

Tabel 1. Pengaruh Preprocessing terhadap Performa

	Akurasi		Loss	
	Train	Test	Train	Test
Normal	98,125	93,984	0,057	0,207
CLAHE	98,711	95,859	0,033	0,152
GAUSSIAN	96,758	93,242	0,095	0,203

Berdasarkan table dapat dilihat dimana hasil akurasi terbaik diperoleh pada *preprocessing* menggunakan CLAHE dengan akurasi latih 98,125% dan akurasi uji 95,859%. Pada pengujian terhadap *loss*, dapat kita lihat pada Gambar 4.4 dimana *loss* terendah diperoleh pada *preprocessing* menggunakan CLAHE dimana nilai *train loss* sebesar 0,033 dan nilai *test loss* sebesar 0,152. Dari

hasil pengujian yang telah dilakukan dapat ditentukan bahwa melakukan *preprocessing* citra dapat mengoptimalkan akurasi dan memperkecil nilai *loss*. Proses *preprocessing* menggunakan CLAHE merupakan *preprocessing* terbaik dalam melakukan pengujian terhadap klasifikasi penyakit pada paru-paru menggunakan citra X-ray menggunakan model CNN dengan arsitektur *MobileNet*. *Preprocessing* menggunakan CLAHE dianggap terbaik karena dapat meningkatkan kualitas dari citra dengan kontras yang rendah. Citra dengan kontras rendah akan mengalami peningkatan intensitas warna, sehingga bagian paru-paru yang terkena penyakit pada citra X-ray yang awalnya tak terlihat dengan jelas akan menjadi terlihat jelas dikarenakan peningkatan kontras dari citra tersebut.

b. Pengujian Optimizer



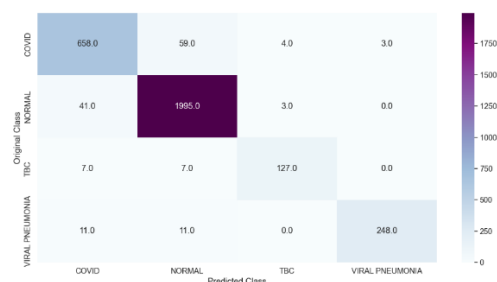
Gambar 8. Performa Gaussian Filter terhadap akurasi dan *loss*

Berdasarkan Gambar 4.1, dapat dilihat hasil akurasi terbaik terdapat pada *optimizer* Adam

dengan akurasi latih 98,125% dan akurasi uji sebesar 93,984%. Untuk hasil uji *optimizer* terhadap *loss* dapat dilihat pada Gambar 4.2, dapat dilihat bahwa *optimizer* Adam menghasilkan nilai *train loss* sebesar 0.05 dan *test loss* sebesar 0.166. Sedangkan untuk *optimizer* SGD menghasilkan nilai *train loss* sebesar 0,092 dan *test loss* sebesar 0,231. Dapat disimpulkan bahwa *optimizer* Adam menghasilkan nilai *loss* yang lebih rendah dibandingkan dengan *optimizer* SGD. Dari hasil pengujian yang telah dilakukan, maka *optimizer* Adam dianggap paling baik dalam melakukan klasifikasi penyakit pada paru-paru menggunakan citra X-ray menggunakan model CNN dengan arsitektur *MobileNet*. Hal ini dikarenakan *optimizer* Adam merupakan extension dari *optimizer* SGD berdasarkan pada estimasi adaptif yang pertama dan kedua. *Optimizer* Adam juga melakukan penyimpanan rata-rata *gradient* proses sebelumnya secara eksponensial sehingga menghasilkan hasil yang optimal. Pada digunakan *learning rate* 0,001 dan *batch size* 32.

c. Pengujian Performa Sistem

Pada skenario 4 adalah pengujian terkait performansi dari sistem. Parameter digunakan diantaranya akurasi, *recall*, presisi, dan *F-1 score*. Pengujian ini menggunakan hasil terbaik pada pengujian sebelumnya yaitu dengan menggunakan *optimizer* Adam *preprocessing* CLAHE, *epoch* 50 *learning rate* 0,001 dan *batch size* 32 dengan jumlah citra sebanyak 15.853 citra.



Gambar 9. Confusion matrix

Dari gambar diatas dapat kita lihat hasil pengujian yang dilakukan oleh sistem dengan hasil sebenarnya. Terdapat pengujian terhadap 4 kategori citra X-ray paru-paru dengan jumlah 2.039 citra X-Ray paru-paru normal, 724 citra X-Ray paru-paru Covid, 141 citra X-Ray paru-paru *tuberculosis*, dan 270 citra X-Ray paru-paru *Viral Pneumonia*. Pada gambar kita lihat bahwa sistem bekerja dengan baik dalam melakukan deteksi pada citra X-ray paru-paru

normal dimana memperoleh hasil benar 1995 citra dan salah 44 citra. Sedangkan sistem sulit mendeteksi pada citra X-ray paru-paru *tuberculosis* dimana memperoleh hasil benar 127 citra dan salah 14 citra. Berdasarkan gambar dari *confusion matrix* diatas dilakukan perhitungan parameter performansi akurasi, *recall*, presisi, dan *F-1 score*

Tabel 2. Performa System klasifikasi penyakit pada paru-paru

	Akuras i	Presis i	Recal l	F1- score
Covid	91%	92%	91%	91%
Normal	98%	96%	98%	97%
Tuberculosi s	90%	95%	90%	92%
Viral Pneumonia	92%	99%	92%	95%

Berdasarkan perhitungan performansi sistem diatas, hasil yang diperoleh untuk klasifikasi penyakit pada paru-paru menggunakan CNN dengan arsitektur *MobileNet* memiliki akurasi rata-rata sebesar 94%, presisi 95%, *recall* 93%, dan *F-1 score* sebesar 94%, hal ini menunjukkan bahwa kinerja dari arsitektur *MobileNet* pada penelitian ini berjalan dengan cukup baik, dimana dapat mengklasifikasi penyakit pada setiap citra dengan optimal.

V. KESIMPULAN

Pada tugas akhir ini telah dirancang sebuah sistem untuk mengklasifikasi penyakit pada paru-paru menggunakan model CNN dengan arsitektur *MobileNet*. Hasil dari tugas akhir ini mendapatkan bahwa kinerja dari penggunaan model CNN dengan arsitektur *MobileNet* sudah cukup baik dalam melakukan klasifikasi penyakit pada paru-paru berdasarkan citra X-ray dari paru-paru dengan 4 kelas diantaranya normal, *covid*, *tuberculosis* dan *viral pneumonia*. Hasil yang diperoleh dari pengujian diatas diantaranya akurasi latih sebesar 98,711% dan akurasi uji sebesar 95,859% dengan nilai *loss* latih 0,033 dan *loss* uji 0,152. Hasil ini didapatkan dengan menggunakan parameter terbaik dari setiap skenario yaitu menggunakan *optimizer* Adam, *prerprocessing* CLAHE, *epoch* 50, *learning rate* 0,001 dan *batch size* 32. Selain itu juga diperoleh hasil dari performansi sistem berupa akurasi rata-rata sebesar 94%, presisi 95%, *recall* 93%, dan *F-1 score* sebesar 94%.

REFERENSI

- [1] R. Rahmadewi and . R. Kurnia, "KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA RONTGEN," *Jurnal Nasional Teknik Elektro*, Vols. Vol: 5,, no. 1, pp. 2302 - 2949, Maret 2016.
- [2] A. Saputra, "SISTEM PAKAR IDENTIFIKASI PENYAKIT PARU-PARU PADA," *JURNAL TEKNOLOGI DAN INFORMATIKA (TEKNOMATIKA)*, vol. 1, no. 3, p. 203, SEPTEMBER 2011.
- [3] O. Rochmawanti, F. Utaminingrum and F. A. Bachtar, "ANALISIS PERFORMA PRE-TRAINED MODEL CONVOLUTIONAL NEURAL NETWORK DALAM MENDETEKSI PENYAKIT TUBERKULOSIS," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 8, no. 4, pp. 805-814, 2021.
- [4] I. M. D. Maysanjaya, "Klasifikasi Pneumonia pada Citra X-rays Paru-paru dengan," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 9, no. 2, pp. 2301 - 4156, Mei 2020.
- [5] D. Mariana and H. Irsyad, "Klasifikasi Pneumonia pada Chest X-Ray," *Jurnal Ilmiah Betrik*, vol. 12, no. 1, 2021.
- [6] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, . M. Andreetto and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision," *Computer Science*, 2017.
- [7] J. Lin and H. Irsyad, "Klasifikasi Pneumonia Pada Citra X-Rays Paru-Paru," *Jurnal Algoritme*, vol. 1, no. 2, pp. 184-194 , 2021.
- [8] W. Swastika, "STUDI AWAL DETEKSI COVID-19 MENGGUNAKAN CITRA CT BERBASIS DEEP LEARNING," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 7, no. 3, pp. 629-634, 2020.
- [9] A. F. Prasetyo, "Penerapan Gaussian Filter pada Edge Detection," *Informatika ITB*, 2016.
- [10] N. M. S. HADIS and . R. ANGRIAWAN, "NILAI OPTIMAL CLIP LIMIT METODE CLAHE UNTUK MENINGKATKAN

AKURASI PENGENALAN WAJAH PADA VIDEO CCTV," *JURNAL INSTEK*, vol. 5, 2020.

- [11] A. M. REZA, "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement," *Journal of VLSI Signal Processing*, vol. 28, no. 1, pp. 25-44, 2004.
- [12] D. P. Kingma and J. L. Ba, "ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION," *ICLR*, 2015.
- [13] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv*, 2017.
- [14] L. A. Andika, P. A. Nur Azizah and R. , "Analisis Sentimen Masyarakat Terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 Pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier," *Indonesian Journal of Applied Statistics*, vol. 2, p. 2621 – 086, May 2019.
- [15] A. Jamhari, . F. M. Wibowo and . W. . A. Saputra,, "Perancangan Sistem Pengenalan Wajah Secara Real-Time pada CCTV Dengan Metode Eigenface," *Journal of Informatics, Information System, Software Engineering and Applications*, vol. 2, no. 2, pp. 020-032, 2020.
- [16] F. LI, H. CHEN, . Z. LIU, X.-D. ZHANG, M.-S. JIANG, Z.-Z. WU and K.-Q. ZHOU, "Deep learning-based automated detection of retinal diseases using optical coherence tomography images," *Biomedical Optic Express*, vol. 10, no. 12, pp. 6204-6226, 2019.
- [17] A. Peryanto, A. Yudhana and R. Umar, "Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network," *Jurnal Format*, vol. 8, no. 2, pp. 138-147, 2019.
- [18] A. Yusuf, R. C. Wihandika and C. Dewi, "Klasifikasi Emosi Berdasarkan Ciri Wajah Menggunakan Convolutional Network," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 11, pp. 10595-10604, 2019.
- [19] C.-Y. Lee, P. W. Gallagher and Z. Tu, "Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree," *Artificial Intelligence and Statistics*, vol. 51, no. 1, pp. 464-472, 2016.
- [20] W. Wang, Y. Hu, T. Zou, H. Liu, J. Wang and X. Wang, "A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers," *Hindawi*, vol. 1, no. 1, pp. 1-10, 2020.