

Klasifikasi Penyakit *Tuberculosis* (TB) Organ Paru Manusia Berdasarkan Citra Rontgen Thorax Menggunakan Metode *Convolutional Neural Network* (CNN)

Aulia Rasyid¹, Lukman Heryawan²

^{1,2}Departemen Ilmu Komputer dan Elektronika, FMIPA UGM

E-mail: ¹aulia.rasyid@mail.ugm.ac.id, ²lukmanh@ugm.ac.id

Abstract

Tuberculosis (TB) is an infectious disease caused by Mycobacterium tuberculosis and can attack various organs in the human body, especially the lungs. Diagnostic examination of pulmonary TB disease can be done by looking at the patient's clinical complaints or symptoms, examination of culture, microscopic examination, radiology or tuberculin test. So far, paramedics diagnose TB patients only based on X-ray images manually. Therefore, it is necessary to have Artificial Intelligence (AI) to assist and facilitate doctors in classifying pulmonary TB disease based on chest X-ray images. This study designed an x-ray image classification of tuberculosis based on Convolutional Neural Network (CNN) using the MobileNet architecture. The dataset is a lung X-ray image that will be used as input for image processing. The first stage is input in the form of an X-ray image, then the image preprocessing process (resizing, grayscale, contrast), followed by a segmentation process (thresholding), followed by a data augmentation process, finally the image will be classified into three classes, namely normal, TBR (tuberculosis) right side, TBRL (right and left tuberculosis). Tuberculosis left side is not included in the classification because tuberculosis is more likely to be affected on the right or in both parts. In this study, we tested the effect of image (without image processing, preprocessing and segmentation), the effect of batch size, the effect of epoch variations, the effect of augmentation and comparison on classification performance. The final result of this study shows that the best optimizer is Adam using preprocessing CLAHE on epoch 50 batch size 32 and produces validation accuracy values of 96.837% accuracy, 95% precision, 93% recall, 93% F-1 score, and 0.210 loss.

Keywords : *Tuberculosis, Convolutional Neural Network (CNN), X-ray, MobileNet.*

Abstrak

Tuberculosis (TB) adalah suatu penyakit infeksi menular yang disebabkan oleh Mycobacterium Tuberculosis dan dapat menyerang berbagai organ dalam tubuh manusia terutama paru-paru. Pemeriksaan diagnosis pada penyakit TB-paru dapat dilakukan dengan melihat keluhan atau gejala klinis pasien, pemeriksaan biakan, pemeriksaan mikroskopis, radiologi ataupun tuberculin test. Selama ini, paramedis mendiagnosis penderita TB tersebut hanya berdasarkan citra rontgen secara manual. Oleh karena itu, diperlukan adanya Artificial Intelligence (AI) untuk membantu dan mempermudah dokter dalam mengklasifikasi penyakit TB paru berdasarkan citra rontgen thorax. Penelitian ini merancang klasifikasi citra X-ray penyakit tuberculosis berbasis Convolutional Neural Network (CNN) menggunakan arsitektur MobileNet. Dataset berupa citra X-ray paru yang akan digunakan sebagai masukan untuk proses pengolahan citra atau image processing. Tahapan pertama yaitu input berupa citra X-ray, selanjutnya dilakukan proses preprocessing citra (resizing, grayscale, contrast), dilanjutkan dengan proses segmentation (thresholding), dilanjutkan dengan proses augmentasi data, terakhir citra akan diklasifikasikan menjadi tiga kelas yaitu normal, TBR (tuberculosis bagian kanan), TBRL (tuberculosis bagian kanan dan kiri). Tuberculosis bagian kiri tidak masuk ke dalam klasifikasi dikarenakan tuberculosis lebih cenderung terkena pada bagian kanan atau di kedua bagian. Dalam penelitian ini dilakukan pengujian terhadap pengaruh citra (tanpa image processing, preprocessing dan segmentation), pengaruh batch size, pengaruh variasi epoch, pengaruh augmentasi dan perbandingan terhadap performa klasifikasi. Hasil akhir dari penelitian ini menunjukkan optimizer terbaik yaitu Adam menggunakan preprocessing CLAHE pada epoch 50 batch size 32 dan menghasilkan nilai akurasi validasi sebesar accuracy sebesar 96.837%, precision 95%, recall 93%, F-1 score sebesar 93%, dan loss sebesar 0.210.

Kata Kunci: *Tuberculosis, Convolutional Neural Network (CNN), X-ray, MobileNet.*

PENDAHULUAN

Paru-paru adalah organ utama pada sistem pernapasan pada manusia (respirasi), berhubungan dengan sistem peredaran darah (sirkulasi) berfungsi untuk menukar oksigen dari udara dengan karbon dioksida dari darah disebut bernafas. Kebutuhan manusia untuk bernafas sama eratnya dengan darah yang memompa melalui darah. Jumlah setiap manusia bernafas dalam sehari bergantung pada usia dan tingkat aktivitasnya misalnya pada bayi 30-60 kali per menit, anak remaja dan dewasa biasanya mencapai 17.000-30.000 nafas per hari atau lebih. Sehingga jika terjadi kerusakan pada organ paru-paru maka akan mengancam kehidupan penderita yang diakibatkan karena kesulitan bernafas.

Penyakit *Tuberculosis* (TB) merupakan suatu penyakit infeksi kronis atau menahun dan menular langsung yang disebabkan oleh bakteri *mycobacterium tuberculosis*. Biasanya penyakit ini menyerang paru-paru tanpa memandang usia dan jenis kelamin (Purnamasari, 2013). Bakteri ini merupakan bakteri basil yang sangat kuat sehingga memerlukan waktu lama untuk mengobatinya. Bakteri ini lebih sering menginfeksi organ paru-paru dibandingkan bagian lain dari tubuh manusia, sehingga selama ini kasus tuberkulosis yang sering terjadi di Indonesia adalah kasus tuberkulosis paru (Indriani, 2005).

Pada data WHO (*World Health Organization*) region jumlah kasus penyakit *tuberculosis* di Asia Tenggara pada tahun 2015 mencapai 2.656.560 kasus. Sedangkan di Indonesia mencapai 6.375.585 kasus. (Region WHO, 2016). Pemeriksaan diagnosis pada penyakit TB-paru dapat dilakukan dengan melihat keluhan/gejala klinis, pemeriksaan biakan, pemeriksaan mikroskopis, radiologik dan tuberculin test. Deteksi dini terhadap TB meningkatkan hasil pengobatan, mengurangi penularan TB dan memungkinkan pengendalian resistensi obat dan TB yang resisten terhadap obat (Dirvi et al, 2017). Salah satu teknik diagnosis yang paling sering digunakan adalah melalui pemeriksaan X-ray paru. Setiap penyakit paru memiliki ciri dan pola yang berbeda beda, begitu pula dengan *tuberculosis*. Selama ini, diagnosis yang dilakukan masih sangat tergantung pada subjek yang melakukan diagnosis, sehingga terkesan subjektif.

Penanganan penyakit ini dilakukan selama 6-9 bulan dengan paling sedikit 3 macam obat. Masyarakat

masih menganggap batuk selama berbulan-bulan merupakan batuk biasa, apabila dicermati batuk dalam jangka waktu yang panjang merupakan salah satu gejala yang ditimbulkan oleh penyakit TBC. Umumnya masyarakat baru memeriksakan penyakitnya ketika sudah parah.

Perkembangan Teknologi *Artificial Intelligence* (AI) merupakan salah satu teknologi yang sedang populer saat ini. Berbagai bidang industri sudah memanfaatkan teknologi tersebut, mulai dari kesehatan, keuangan, dan lain-lain. *Artificial intelligence* atau yang sering disebut dengan kecerdasan buatan merupakan ilmu yang memiliki berbagai cabang salah satu cabangnya yaitu *Machine Learning* (ML). *Machine Learning* dapat didefinisikan sebagai aplikasi komputer dan algoritma matematika yang diadopsi dengan cara pembelajaran yang berasal dari data dan menghasilkan prediksi di masa yang akan datang (Goldberg & Holland, 1988). Saat ini sudah sangat berkembang dan telah dimanfaatkan di berbagai bidang termasuk salah satunya di bidang kesehatan.

Tidak hanya sampai disana, beberapa peneliti lain mengembangkan *Machine Learning* dan membuat cabang baru yaitu *Deep Learning* (DL), dimana data yang digunakan bersifat lebih kompleks dan berjumlah besar. *Deep Learning* adalah bagian dari *Neural Network* yang memiliki memiliki arsitektur lebih kompleks dan lebih banyak jumlah layer yang digunakan, sehingga diharapkan mampu menangani permasalahan yang lebih rumit dengan lebih banyak data (Widayat, 2018).

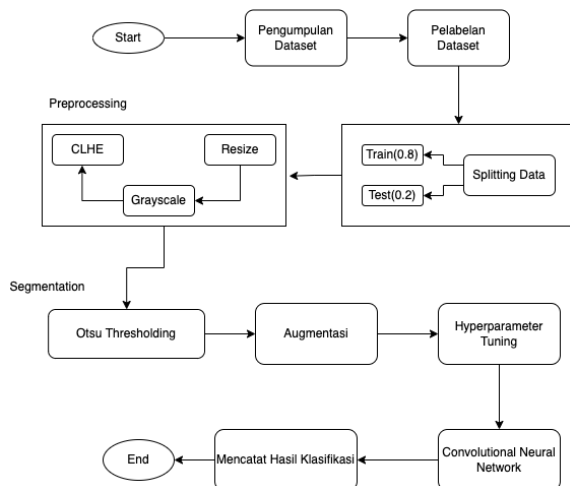
Kedua cabang tersebut sering digunakan dalam berbagai penelitian untuk mendeteksi dan mengklasifikasikan beberapa penyakit seperti : pneumothorax (Chan, Zeng, Wu, Wu, & Sun, 2018), kanker payudara (Kaushal, Bhat, Koundal, & Singla, 2019), tumor otak (Rajan & Sundar, 2019), batu ginjal (Nithya, Appathurai, Venkatadri, Ramji, & Anna, 2020), pneumonia (Biswas, Ghosh, Bhattacharyya, Platos, Snasel, & Chakrabarti, 2020) dan kanker kulit (Savera, Suryawan, & Setiawan, 2020).

Dalam penelitian ini, peneliti akan melakukan penelitian dengan judul “ Klasifikasi Penyakit *Tuberculosis* (TB) Organ Paru Manusia Berdasarkan Citra Rontgen Thorax Menggunakan Metode *Convolutional Neural Network* (CNN). *Convolutional Neural Network* yang akan

digunakan yaitu 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* (Howard et al., 2017). Tujuan penelitian ini adalah untuk dapat mengklasifikasi penyakit *tuberculosis* (TB) pada organ paru-paru berdasarkan citra rontgen thorax atau X-ray menggunakan metode *Convolutional Neural Network* (CNN).

METODE

Data yang telah diperoleh sebanyak gambar 2564 dari berbagai sumber seperti *google images*, *duckduckgo images*, dan *kaggle*, terbagi atas 2 kategori besar. Gambar X-ray *tuberculosis* dan gambar X-ray paru-paru normal. Gambar-gambar yang diperoleh dibuat seragam melalui tahap *preprocessing* yaitu: *Resizing*, *Grayscale* dan *Contrast*. Setelah tahap *preprocessing* akan dilakukannya tahap *segmentation* yaitu: *Otsu Thresholding*. Kemudian dilanjutkan dengan tahap augmentasi data. Setelah itu kita menggunakan *Hyperparameter tuning*. Tahap terakhir yaitu klasifikasi *Convolutional Neural Network* (CNN) dengan menggunakan arsitektur *MobileNet*.



Gambar 1. Skema Penelitian

Deskripsi pada gambar 1 menjelaskan secara garis besar bagaimana jalannya proses penelitian dari pengumpulan dataset sampai dengan hasil klasifikasi berupa tiga kelas yaitu: normal, TB kanan, dan TB kanan kiri.

Data yang dikumpulkan terdiri atas dua kategori yaitu gambar X-ray *tuberculosis* dan gambar X-ray paru-paru normal. Data citra yang digunakan berasal dari kaggle diperoleh sebanyak 564 citra X-ray *tuberculosis* dan 2000 citra X-ray paru-paru normal, dataset bisa diakses melalui link : (<https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>). (<https://www.kaggle.com/tawsifurrahman/Covid19-radiography-database>)

Data *tuberculosis* yang telah diunduh lalu diseleksi dan dibagi menjadi tiga folder, yaitu *tuberculosis* bagian kanan, dan *tuberculosis* di kedua bagian. Pelabelan file dilakukan dengan mengikuti format "d-I" dimana i adalah angka yang dimulai dari 0 hingga total jumlah gambar yang tersedia. Data yang telah dibagi didapatkan sebanyak 182 gambar untuk *tuberculosis* bagian kanan, dan 382 gambar *tuberculosis* di kedua bagian dan untuk paru-paru normal terdapat 2000 gambar.

Splitting data

Data yang telah diproses, akan dibagi menjadi 2 bagian yaitu *training*, *testing* dan *validation*. Pembagian atas gambar tersebut dilakukan secara manual yang mana akan melakukan pembagian terhadap dataset yang berada dalam satu folder sesuai dengan rasio yang diinginkan yaitu 80% *training* : 20% *testing* dan untuk *validation* diambil 20% dari data *training*.

Preprocessing

Citra yang sudah dilabel selanjutnya akan diolah untuk menghasilkan citra yang lebih baik. Tahapan *preprocessing* terdiri dari *resizing*, *grayscale* dan *contrast*.

Resizing: gambar-gambar diubah menjadi ukuran 224x224 pixel

Grayscale: bertujuan untuk penyeragaman warna keabuan pada citra yang akan diproses.

Contrast : bertujuan untuk meningkatkan kontras citra X-ray menjadi lebih jelas.

Segmentation

Tahapan ini adalah tahap pemisahan objek yang satu dengan objek yang lain dalam suatu gambar (citra) menjadi objek-objek berdasarkan karakteristik khusus yang dimiliki suatu citra Tahapan *segmentation* terdiri dari *thresholding otsu*.

Thresholding Otsu: Citra yang telah di preprocessing dilanjutkan menggunakan thresholding metode Otsu bertujuan untuk memisahkan objek dari background. Proses *thresholding* citra akan diubah menjadi citra biner untuk memberikan informasi tentang objek dan background. Dimana citra akan berwarna hitam dan putih saja.

Augmentasi

Pada tahap ini, augmentasi dilakukan untuk meningkatkan jumlah data dengan memodifikasi data sehingga program mengenali data tersebut adalah data yang berbeda. Proses ini dapat membantu penelitian dengan data berjumlah sedikit untuk mencegah *overfitting* maupun *underfitting*. Selain itu, proses augmentasi dapat meningkatkan kinerja model dalam proses klasifikasi gambar (Loey, dkk, 2020).

Hyperparameter

Menerapkan parameter-parameter sebelum dilakukannya pelatihan (*training*), hal ini bertujuan agar meningkatkan hasil klasifikasi CNN. Pada penelitian ini, dilakukan metode hyperparameter tuning untuk menentukan parameter *optimizer*, *learning rate*, dan *dense layer*. Contoh pada tabel 1 berikut:

Tabel 1. Parameter Pemanding

	Parameter
Optimizer	Adam
Learning Rate	0.001
Dense	1024,512

Pelatihan model

Dengan menggunakan library dari bahasa pemrograman python yaitu Tensorflow dan Scikit-learn, model-model dari machine learning dan *deep learning* dapat dibuat hanya dengan mengubah-ubah parameter yang sesuai dengan keinginan. Beberapa parameter yang akan diubah seperti iterasi, *optimizer*, dataset, dan lainnya. Pelatihan model adalah proses pelatihan mengenali objek dan mengklasifikasikannya sesuai dengan kelasnya. Sistem pelatihan model pada penelitian ini menggunakan model *MobileNet*. Pada Gambar 2 merupakan proses pelatihan model *MobileNet*.

Type / Stride	Filter Shape	Input Size
Conv / s2	3 × 3 × 3 × 32	224 × 224 × 3
Conv dw / s1	3 × 3 × 32 dw	112 × 112 × 32
Conv / s1	1 × 1 × 32 × 64	112 × 112 × 32
Conv dw / s2	3 × 3 × 64 dw	112 × 112 × 64
Conv / s1	1 × 1 × 64 × 128	56 × 56 × 64
Conv dw / s1	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 128	56 × 56 × 128
Conv dw / s2	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 256	28 × 28 × 128
Conv dw / s1	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 256	28 × 28 × 256
Conv dw / s2	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 512	14 × 14 × 256
5× Conv dw / s1	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw / s2	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 1024	7 × 7 × 512
Conv dw / s2	3 × 3 × 1024 dw	7 × 7 × 1024
Conv / s1	1 × 1 × 1024 × 1024	7 × 7 × 1024
Avg Pool / s1	Pool 7 × 7	7 × 7 × 1024
FC / s1	1024 × 1000	1 × 1 × 1024
Softmax / s1	Classifier	1 × 1 × 1000

Gambar 2. Arsitektur MobileNet (Howard et al., 2017)

Pada gambar 2 menunjukkan bahwa pelatihan model *MobileNet* memiliki layer sebanyak 27 *convolutional layer*, satu *pooling layer*, satu *Fully connected layer* yang diikuti oleh lapisan softmax. *MobileNet* sendiri membagi dua *layer* konvolusi yaitu menjadi *depthwise convolution* dan *pointwise convolution*. Semua citra akan mengalami normalisasi menjadi 224×224 piksel yang berguna untuk mempercepat proses pelatihan. Kemudian dilakukan proses flatten yaitu proses mengubah feature map hasil pooling layer ke dalam format vector. Proses ini biasanya disebut dengan 19 langkah dari *fully Connected layer*. Setelah citra melalui fully connected layer, hasilnya adalah input fungsi aktivasi *softmax* yang menghitung probabilitas dari data hasil train terhadap objek yang terdiri dari tiga kelas.

Mencatat hasil dan membuat kesimpulan

Setelah pelatihan model dilakukan hasil akan berupa akurasi, *recall*, *precision*, dan *f-1 score*. Hal ini dapat diperoleh menggunakan confusion matrix ada empat istilah yang merupakan representasi dari hasil klasifikasi pada confusion matrix yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, dan *False Negative (FN)*.

HASIL DAN PEMBAHASAN

Pada bagian hasil dan pembahasan akan berisi tentang hasil dan analisis dari proses penelitian

yang telah dilakukan terhadap klasifikasi penyakit pada paru-paru menggunakan metode *CNN* dengan arsitektur *MobileNet*.

Skenario Pengujian

Skenario yang dilakukan pada penelitian ini terdiri dari tiga bentuk skenario. Hasil terbaik dari setiap skenario akan diuji kembali pada skenario berikutnya hal ini bertujuan agar mendapatkan model terbaik. Berikut skenario-skenario yang digunakan :

1. Skenario 1

Pada skenario pertama akan dilakukan pengujian terhadap citra sebelum proses klasifikasi. Citra yang akan diuji yaitu citra yang tidak diubah, citra *preprocessing* *CLAHE*, dan citra *segmentation thresholding-otsu*. Tujuan dilakukan pengujian ini untuk mendapatkan hasil image processing terbaik, sehingga dapat meningkatkan nilai akurasi pada sistem.

2. Skenario 2

Pada skenario kedua akan dilakukan pengujian terhadap pengaruh *batch size* pada akurasi sistem. Pada penelitian ini dilakukan tiga nilai *batch size* yaitu 16, 32, dan 64. Pengujian ini bertujuan untuk melihat apakah semakin besar jumlah *batch size* akan semakin baik akurasi sistem atau sebaliknya.

3. Skenario 3

Pada skenario ketiga akan dilakukan pengujian terhadap pengaruh jumlah *epoch* pada akurasi sistem. Pada penelitian ini dilakukan tiga nilai maksimal *epoch* yaitu 50, 100, dan 150. Pengujian ini bertujuan untuk melihat apakah semakin besar jumlah *epoch* akan semakin baik akurasi sistem atau sebaliknya.

4. Skenario 4

Pada skenario keempat akan dilakukan pengujian terhadap pengaruh augmentasi pada akurasi sistem. Skenario ini dilakukan dengan menggunakan model terbaik dari skenario sebelumnya. Pengujian ini bertujuan untuk melihat apakah menggunakan augmentasi akan memperbaiki akurasi atau sebaliknya.

5. Skenario 5

Pada skenario kelima akan dilakukan pengujian terhadap performa klasifikasi. Dalam penelitian

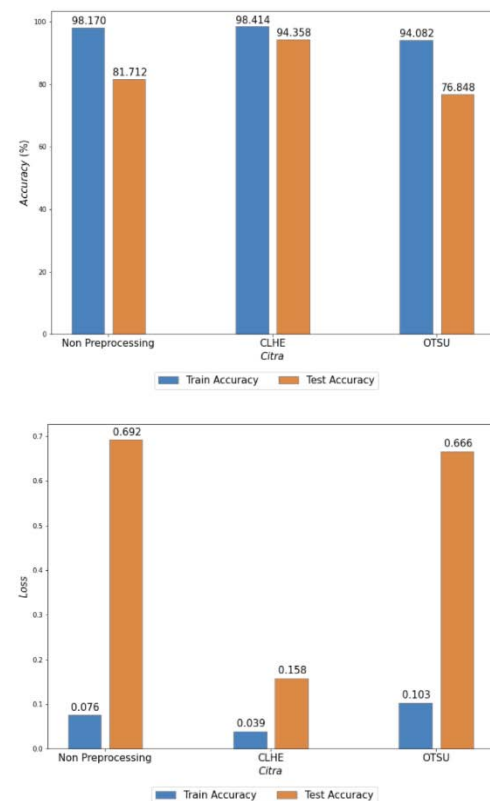
ini terdapat empat parameter performa yang akan dibandingkan yaitu *accuracy*, *recall*, *presisi*, dan *f1-score*. Skenario terbaik dari pengujian sebelumnya akan digunakan untuk pengujian skenario keempat.

Analisis Hasil Pengujian

Setelah dilakukan pengujian berdasarkan beberapa skenario diatas, didapatkan hasil pengujian sebagai berikut.

Hasil Pengujian Terhadap Citra

Pada skenario pertama dilakukan pengujian terhadap citra. Pada pengujian ini akan menggunakan tiga buah citra, pertama adalah citra yang tidak di *preprocessing* maupun di *segmentation*, kedua adalah citra *preprocessing* *CLAHE*, dan ketiga adalah citra *segmentation otsu-thresholding*. Pada pengujian skenario pertama akan menggunakan parameter yang sama dengan *optimizer adam*, *learning rate* 0.001, *epoch* 50, dan *batch size* 32. Hasil pengujian terhadap *preprocessing* dan *segmentation* dapat dilihat pada gambar berikut:



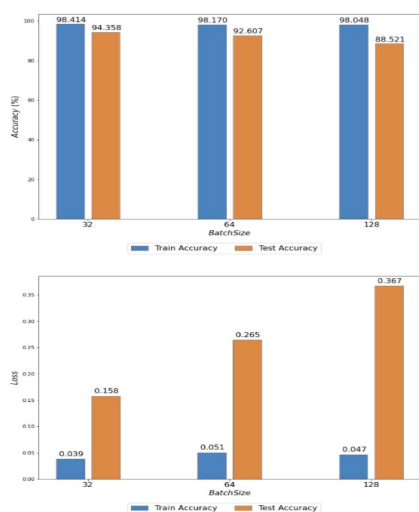
Gambar 3. Hasil Pengujian Terhadap Citra

Berdasarkan Gambar 3, dapat dilihat dimana hasil akurasi terbaik diperoleh pada *preprocessing*

menggunakan *CLAHE* dengan akurasi latih 98,414% dan akurasi uji 94,358%. Pada pengujian terhadap *loss*, dapat kita lihat pada gambar 3 dimana *loss* terendah diperoleh pada preprocessing menggunakan *CLAHE* dimana nilai *train loss* sebesar 0,039 dan nilai *test loss* sebesar 0,158. Dari hasil pengujian yang telah dilakukan dapat ditentukan bahwa melakukan *preprocessing* citra tanpa *segmentation* dapat mengoptimalkan akurasi dan memperkecil nilai *loss*. Proses *preprocessing* menggunakan *CLAHE* merupakan *preprocessing* terbaik dalam melakukan pengujian terhadap klasifikasi *tuberculosis* menggunakan citra X-ray menggunakan model *CNN* dengan arsitektur *MobileNet*. Preprocessing menggunakan *CLAHE* dianggap terbaik dikarenakan 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 *tuberculosis* pada citra X-ray yang awalnya tak terlihat dengan jelas akan menjadi terlihat jelas dikarenakan peningkatan kontras dari citra tersebut.

Hasil Pengujian Pengaruh Batch Size

Pada skenario kedua akan dilakukan pengujian terhadap pengaruh *batch size* terhadap performa klasifikasi. Pada pengujian ini menggunakan tiga varian *batch size* yaitu 32, 64, dan 128. Pengujian ini berdasarkan model terbaik dari hasil pengujian sebelumnya yaitu menggunakan *optimizer Adam* dan *preprocessing CLAHE* dengan *learning rate* 0.001 *epoch* 50. Hasil pengujian terhadap pengaruh *batch size* dapat dilihat pada gambar berikut.

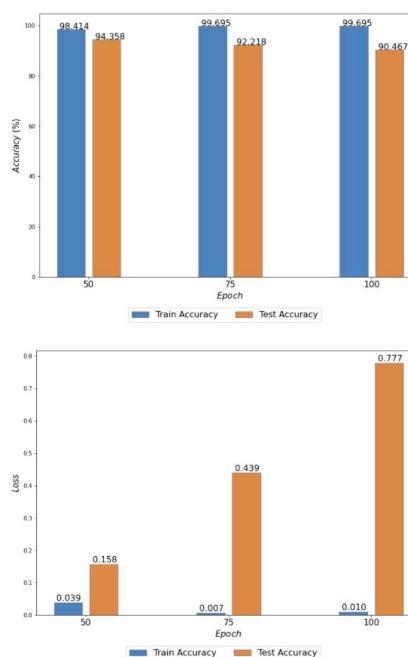


Gambar 4. Hasil Pengujian *Batch size*

Dari gambar diatas dapat dilihat bahwa semakin besar *batch size* yang digunakan, angka *loss* yang didapatkan semakin besar juga. Demikian juga sebaliknya, *batch size* yang kecil menghasilkan angka *loss* yang semakin kecil. Dari ketiga angka *batch size* yang dibandingkan, hasil yang paling optimal adalah *batch size* 32. Biasanya *batch size* besar digunakan karena memungkinkan percepatan komputasi. Jika menggunakan ukuran *batch size* kecil akan menghabiskan waktu yang sangat lama. Akan tetapi, dibalik percepatan komputasi akan membuat tingkat *loss* menjadi besar. Ukuran *batch size* yang terlalu besar akan memberikan hasil yang kurang optimal. Semakin besar ukuran *batch size* semakin tidak teliti hasilnya. Dari hasil pengujian yang telah dilakukan dapat ditentukan bahwa *batch size* yang kecil akan membuahkan hasil yang optimal dibandingkan menggunakan *batch size* yang lebih besar.

Hasil Pengujian Pengaruh Jumlah Epoch

Pada skenario ketiga dilakukan pengujian terhadap pengaruh jumlah *epoch* terhadap performa klasifikasi. Pada pengujian ini menggunakan tiga perbandingan *epoch* yaitu 50, 75, dan 100. Pengujian ini berdasarkan model terbaik dari hasil pengujian sebelumnya yaitu menggunakan *optimizer Adam* dan *preprocessing CLAHE* dengan *learning rate* 0.001 dan *batch size* 32. Hasil pengujian terhadap pengaruh *epoch* dapat dilihat pada gambar berikut.

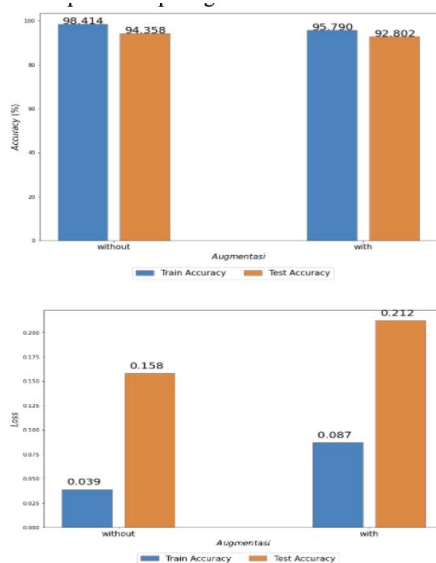


Gambar 5. Hasil pengujian *epoch*

Dari gambar diatas , dapat dilihat dimana kenaikan jumlah *epoch* pada penelitian ini mempengaruhi kenaikan nilai akurasi. Semakin besar nilai *epoch* yang kita gunakan belum tentu memperoleh akurasi yang lebih baik. Jadi kenaikan *epoch* tidak berbanding lurus dengan kenaikan dari akurasi sistem. Pada *epoch* 50 diperoleh nilai akurasi latih sebesar 98,414% dan akurasi uji sebesar 94,358%, pada *epoch* 75 diperoleh nilai akurasi latih sebesar 99,645% dan akurasi uji sebesar 92,218%, pada *epoch* 100 mendapatkan nilai akurasi latih sebesar 99,695% dan akurasi uji sebesar 90,467%. Bisa dilihat pada *epoch* 75 dan *epoch* 100 untuk akurasi uji mengalami penurunan , hal ini juga terjadi pada nilai loss yang diperoleh. Berdasarkan data diatas dapat disimpulkan bahwa peningkatan jumlah *epoch* sedikit mempengaruhi akurasi, akan tetapi peningkatan jumlah *epoch* belum tentu mendapatkan hasil yang lebih baik dari jumlah *epoch* yang lebih sedikit.

Hasil Pengujian Pengaruh Augmentasi

Pada skenario keempat dilakukan pengujian terhadap augmentasi. Pada pengujian ini menggunakan dataset yang sudah di augmentasi dan belum di augmentasi. Pengujian ini berdasarkan model terbaik dari hasil pengujian sebelumnya yaitu menggunakan optimizer Adam dan *preprocessing* CLAHE dengan *learning rate* 0.001, *batch size* 32 dan *epoch* 50. Hasil pengujian terhadap pengaruh augmentasi dapat dilihat pada gambar berikut.

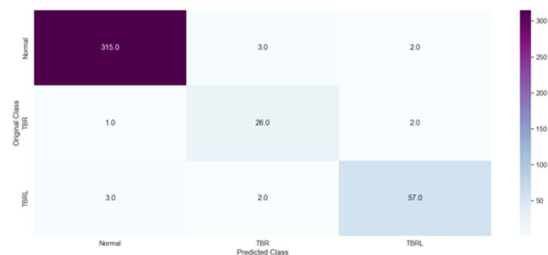


Gambar 6. Hasil pengujian augmentasi

Dari gambar diatas dapat dilihat pemakaian augmentasi mempengaruhi nilai akurasi. Pada pemakaian augmentasi nilai akurasi latih sebesar 95.424% , dan nilai akurasi uji sebesar 92.802%. Untuk yang tidak memakai augmentasi nilai akurasi latih mendapatkan 98.414% dan nilai akurasi uji sebesar 94.358%. Dapat disimpulkan bahwa pemakaian augmentasi belum tentu mendapatkan nilai akurasi yang lebih baik dari pada yang tidak memakai augmentasi.

Hasil Pengujian Performa Klasifikasi

Pada skenario kelima adalah pengujian terkait performa klasifikasi. 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 *batch size* 32 tanpa augmentasi dengan jumlah citra sebanyak 2.564 citra.



Gambar 7. Confusion Matrix

Dari gambar diatas dapat kita lihat hasil pengujian menggunakan data validation. Terdapat pengujian terhadap tiga kategori citra X-ray paru paru dengan jumlah 400 citra X-Ray paru-paru normal, 29 citra X-Ray paru paru *tuberculosis* bagian kanan, 62 citra X-Ray paru-paru *tuberculosis* bagian kanan dan kiri. Pada gambar kita lihat bahwa sistem bekerja dengan baik dalam melakukan deteksi pada citra X-ray paru-paru normal dimana memperoleh hasil benar 313 citra dan salah 5 citra. Sedangkan sistem sulit mendeteksi pada citra X-ray paru-paru *tuberculosis* dimana memperoleh hasil benar 26 citra dan salah 3 citra. Berdasarkan gambar dari confusion matrix diatas dilakukan perhitungan parameter accuracy, recall, precision, dan F-1 score.

Tabel 2. Performa System Klasifikasi Tuberkulosis

	<i>Precision</i>	<i>Recall</i>	<i>F-1 score</i>
Normal	99%	98%	99%
<i>Tuberculosis</i> kanan	84%	90%	87%
<i>Tuberculosis</i> kanan kiri	93%	92%	93%

Berdasarkan perhitungan diatas, hasil yang diperoleh untuk klasifikasi penyakit pada paru-paru menggunakan CNN dengan arsitektur *MobileNet* memiliki accuracy sebesar 96.837%, precision 95%, recall 93%, dan F-1 score sebesar 93%, hal ini menunjukkan bahwa kinerja dari arsitektur *MobileNet* pada penelitian ini berjalan dengan cukup baik, dimana dapat mengklasifikasi penyakit pada setiap citra dengan optimal.

SIMPULAN

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 3 kelas diantaranya normal, *tuberculosis* kanan, dan *tuberculosis* kanan kiri. Hasil yang diperoleh dari pengujian diatas diantaranya akurasi latih sebesar 98,414% dan akurasi uji sebesar 94,358% dengan nilai *loss* latih 0,039 dan *loss* uji 0,158. Hasil ini didapatkan dengan menggunakan parameter terbaik dari setiap skenario yaitu menggunakan *optimizer Adam*, *preprocessing CLAHE*, *epoch 50*, *learning rate 0,001* dan *batch size 32*. Selain itu juga diperoleh hasil dari performa klasifikasi dengan menggunakan data validation berupa accuracy rata-rata sebesar 96.837% , precision 95%, recall 93%, dan F-1 score sebesar 93 %.

DAFTAR PUSTAKA

Ahmad Hania, Abu. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. Jurnal Teknologi Indonesia.

Ahmad. U, 2005, Pengolahan Citra Digital dan Teknik Pemrogramannya, Graha Ilmu

Anbarjafari, G., Jafari, A., Jahromi, M. N., Ozcinar, C., & Demirel, H. (2015). Image Illumination Enhancement With an Objective no-reference measure of illumination assessment based on Gaussian distribution Mapping. Engineering Science and Technology, an International Journal.

Astuti, Fitri Andri. 2021. Pemanfaatan Teknologi Artificial Intelligence untuk Penguatan Kesehatan dan Pemulihan Ekonomi Nasional. Bandung

B. Khasoggi and E. Samsuryadi, "Efficient mobilenet architecture as image recognition on mobile and embedded devices," Indonesian Journal of Electrical Engineering and Computer Science, vol. 16, no. 1, p. 389~394, 2019

Center of Disease Control and Prevention. Transmission and Pathogenesis of Tuberculosis. 2013

Chitradevi, B., & Srimathi, P. (2014). An Overview Image Processing Techniques. International Journal of Innovative Research in Computer and Communication Engineering. India: ISSN.

Cristianini, N. dan Shawe-Taylor, J., 2006, An introduction to support vector machines: And other kernel-based learning methods, Cambridge: Cambridge University Press

Depkes RI. 2008. Pedoman Nasional Penanggulangan Tuberkulosis, edisi 2, Jakarta: Dirjen P2M&PL.

Eko, Dirvi.J.S., Asih,Setiari. 2017. Identifikasi Bakteri pada Citra Dahak Penderita Tuberculosis (TBC) Menggunakan Metode Watershed. Journal of Electrical Electronic Control and Automotive Engineering (JEECAE). Madiun. Vol.2, No.1.

Evelyn CP, 2009. Anatomi dan Fisiologi untuk Paramedis. Jakarta. Gramedia

Goldberg, D. E., & Holland, J. H. (1988). Genetic algorithms and machine learning. Machine Learning, 3(2), 95–99.

Gonzales, Rafael C. & Richard E. Woods & Steven L. Eddins. 2004. Digital Image Processing. Prentice-Hall

- Hadiarto, Mangunnegoro. 2015. Diagnosis dan Penatalaksanaan Asma. *Jurnal Respiriologi Indonesia*. 15(3): 113-119.
- Howard, A. G., M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, . M. Andreetto and H. Adam, 2017 . "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision," *Computer Science*.
- Icksan A.G, Luhur R., 2008. *Radiologi Thorax Tuberculosis Paru*. 1st edition, Jakarta: Sagung Seto. Hal 2-3, 7, 16.
- Indriani, D., Adiningsih, S., Mahmudiono, T. (2005). Faktor resiko yang mem-pengaruhi kejadian TB paru pada anak jalanan dengan studi kasus di Yayasan Insani Surabaya. *Jurnal FKM UA*. Surabaya.
- Jayanti, N. 2013. Perbandingan Kapasitas Vital Paru pada Atlet Pria Cabang Olahraga & Lari Cepat Persiapan Olahraga Provinsi 2013 di Bandar Lampung. *Majority Journal*. 2(5): 113-118.
- Jong, J.S, 2005, *Jaringan Saraf Tiruan dan Pemrogramannya Menggunakan Matlab*, C.V Andi Offset, Yogyakarta.
- Karargyris, A., Jaegar, S., Antani, S., & Thoma, G. (2012). Detection Tuberculosis in Radiographs Using Combined Lung Masks. *Annual International Conference*. San Diego, California USA: IEEE EMBS.
- Khan, A. H. (2016). *Tuberculosis Control in Sindh, Pakistan: Critical Analysis of its Implementation*. King Saud Bin Abdulaziz University for Health Sciences. Pakistan: Elsevier Limited.
- Knechel, N. A., RN, MSN, & ACNP. (2009). Tuberculosis; Pathophysiology, Clinical Features, and Diagnosis. *Clinical Article*.
- L. Perez and J. Wang. The effectiveness of data augmentation in image classification using deep learning, 2017.
- Loey, M., Manogaran, G., & Khalifa, N. E. M. (2020). A deep transfer learning model with classical data augmentation and CGAN to detect Covid-19 from chest CT radiography digital images. *Neural Computing and Applications*, 0123456789.
- Maladkar, K., 2020, Overview of convolutional neural network in image classification, *Analytics India Magazine*, Available at: <https://analyticsindiamag.com/convolutional-neural-network-image-classification-overview/> (Accessed October 23, 2021).
- Mitha.N, Wara. 2014. Analisis Citra Sinar-X Tulang Tangan Menggunakan Metode Thresholding Otsu Untuk Identifikasi Osteoporosis. Skripsi. Pontianak: Universitas Tanjungpura.
- Munir, R., 2004, *Pengolahan Citra Digital dengan Pendekatan Algoritmik*, Informatika, Bandung.
- Patil, N. M., & Nemade, M. U. (2017). Music Genre Classification Using MFCC , K-NN and SVM Classifier. 4(2), 43–47.
- Pasi, Nani Sulviana. 2018. Identifikasi Penyakit Tuberculosis (TB) Organ Paru Berdasarkan Citra X-ray Menggunakan Probabilistic Neural Network (PNN), Skripsi. Jurusan Teknologi Informasi Universitas Sumatera Utara Medan.
- Purnamasari, R. W., 2013, *Implementasi jaringan Syaraf Tiruan Backpropagation Sebagai Sistem Deteksi Penyakit TBC*, Skripsi, Jurusan Matematika UNNES Semarang.
- Rohmah, R. N., Susanto, A., & Soesanti, I. (2013). Lung Tuberculosis Identification Based on Statistical Feature of Thoracic X-ray. *IEEE Quality in Research* (pp. 19-20). Indonesia: IEEE.
- Rohman, Feri Fahrur dan Ami Fauzijah. 2008. Rancang bangun aplikasi sistem pakar untuk menentukan jenis gangguan perkembangan pada anak. *Jurnal Media Informatika Vol 6 No 1* ISSN: 0854-4743.
- Santosa, B. 2007. *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Santosa, P. A. (2014). Implementasi Jaringan Syaraf Tiruan dalam Memilih Metode Enhanced Oil Recovery. *proceeding Simposium Nasional IATMII*.
- Suciati, Nanik & Rosdiana Rahmawati. (2006). Dekomposisi Morfologi Bentuk Biner Dua dimensi Menjadi Poligon Konveks dengan Pendekatan Heuristik. *Jurusan Teknik Informatika, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember*.

- Suhendra, Adang. 2004. Catatan Kuliah Pengantar Pengolahan Citra. Universitas Gunadarma.
- Sutoyo, T. 2009. Teori Pengolahan Citra Digital. Yogyakarta: Andi
- Tanaka, M.; Okutomi, M. A novel inference of a restricted boltzmann machine. Pattern Recognition (ICPR), 2014 22nd International Conference on. 2014; pp 1526–1531.
- Tua, Jodiaman. 2017. Klasifikasi Penyakit Stroke Menggunakan Convolutional Neural Network. Skripsi. Program Studi Teknologi Informasi. Fakultas Ilmu Komputer dan Teknologi Informasi. Universitas Sumatera Utara. Medan
- Ucar, F., & Korkmaz, D. (2020). Covidiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (Covid-19) from X-ray images. Medical Hypotheses, 140(April), 109761. <https://doi.org/10.1016/j.mehy.2020.1097>
- Utami, S. R. (2014). Perancangan Aplikasi Perbaikan Citra Hasil Pengambilan Webcam Menerapkan Metode Contrast Stretching. Jurnal Pelita Informatika Budi Darma Volume VII, pp.39-43.
- 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.
- W. Widayat, “Pengaruh Word Embedding Dimension Reduction Terhadap Kinerja LSTM untuk Analisis Sentimen,” 2018.
- Wang, Yu, Chen, Y., Yang, N., Zheng, L., Dey, N., Ashour, A. S., ... Shi, F. (2018). Classification of mice hepatic granuloma microscopic images based on a deep convolutional neural network. Applied Soft Computing Journal. <https://doi.org/10.1016/j.asoc.2018.10.006>
- Wang, S.C., 2003, Artificial Neural Network, Interdisciplinary Computing in Java Programming, 81–100.
- Wayan. I.S.E.P., Arya, Yudhi.W., Rully, Soelaiman. 2016. Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101. Jurnal Teknik ITS. Surabaya. Vol. 5, No. 1.
- WibowoD., Paryana W., 2009. Anatomi Tubuh Manusia. Yogyakarta: Graha Ilmu.
- World Health Organization (WHO Region (2016). Tuberculosis Control In The SouthEast Asia Region. New Delhi, India: Region Office for South-East As