

Perbandingan Klasifikasi Citra CT-Scan Kanker Paru-Paru Menggunakan Contrast Stretching Pada CNN dengan EfficientNet-B0

Alfitra Salam, Febi Yanto*, Surya Agustian, Siti Ramadhani

Fakultas Saints dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹11950115004@students.uin-suska.ac.id, ^{2,*}febiyanto@uin-suska.ac.id, ³surya.agustian@uin-suska.ac.id,

⁴siti.ramadhani@uin-suska.ac.id

Email Penulis Korespondensi: febiyanto@uin-suska.ac.id

Abstrak Data dari Organisasi Kesehatan Dunia (WHO) menunjukkan bahwa pada tahun 2020, sekitar 10 juta orang meninggal akibat kanker. Merokok diidentifikasi sebagai faktor utama yang menyebabkan kanker paru-paru, karena rokok mengandung lebih dari 60 zat beracun yang dapat memicu perkembangan penyakit kanker. Tingkat kanker paru-paru telah meningkat dengan cepat akibat konsumsi rokok yang berlebihan. Mendeteksi nodul pada paru-paru biasanya memerlukan waktu sekitar 10-30 menit. Dalam penelitian ini, digunakan algoritma CNN dengan arsitektur EfficientNet-B0 untuk mengklasifikasikan kanker paru-paru. Proses preprocessing yaitu *contrast stretching*, dan berbagai teknik optimasi *hyperparameter* seperti Adam, Adagrad, dan SGD digunakan untuk meningkatkan kinerja CNN dan menggunakan *Average pooling* dengan *out dense* 64, 32, 16, 1. Analisis kinerja dilakukan dengan menggunakan matriks klasifikasi (*confusion matrix*). Hasil klasifikasi tertinggi dicapai menggunakan optimizer ADAM dan learning rate 0.01, di mana mencapai 72.48%, presisi 71.52%, recall 64.2%, dan skor F1 64.76%. Sementara itu, hasil yang diperoleh dari dataset asli menunjukkan perbedaan. Hasil klasifikasi tertinggi diperoleh optimizer ADAM dan learning rate 0.01, dengan akurasi mencapai 64.22%, presisi 52.69%, recall 50.52%, dan skor F1 43.51%. Hasil ini menunjukkan bahwa penggunaan *contrast stretching* dalam *preprocessing* klasifikasi kanker paru-paru sangat efektif dalam meningkatkan akurasi.

Kata Kunci: Kanker paru; CNN; EfficientNet-B0; Klasifikasi; Contrast Stretching

Abstract—Data from the World Health Organization (WHO) indicates that in 2020, approximately 10 million people died from cancer. Smoking has been identified as a primary factor causing lung cancer, as cigarettes contain over 60 toxic substances that can trigger the development of the disease. The rate of lung cancer has rapidly increased due to excessive cigarette consumption. Detecting nodules in the lungs typically takes about 10-30 minutes. In this study, a Convolutional Neural Network (CNN) algorithm with EfficientNet-B0 architecture is employed to classify lung cancer. The preprocessing process involves contrast stretching, and various hyperparameter optimization techniques such as Adam, Adagrad, and SGD are used to enhance the CNN's performance. Average pooling with output dense layers of 64, 32, 16, 1 is utilized. Performance analysis is conducted using a confusion matrix. The highest classification results are achieved using the ADAM optimizer with a learning rate of 0.01, where accuracy reaches 72.48%, precision is 71.52%, recall is 64.2%, and the F1 score is 64.76%. Meanwhile, results obtained from the original dataset show differences. The highest classification result is obtained using the ADAM optimizer with a learning rate of 0.01, achieving an accuracy of 64.22%, precision of 52.69%, recall of 50.52%, and an F1 score of 43.51%. These results indicate that the use of contrast stretching in lung cancer classification preprocessing is highly effective in improving accuracy.

Keywords: Lung Cancer; CNN; EfficientNet-B0; Classification; Contrast Stretching

1. PENDAHULUAN

Kanker menjadi permasalahan kesehatan yang banyak terjadi di seluruh dunia. Data dari Organisasi Kesehatan Dunia (WHO) menunjukkan bahwa jumlah kematian akibat kanker mencapai 10 juta pada tahun 2020[1]. Indonesia menempati peringkat ke-8 di Asia Tenggara dan ke-23 di Asia dalam hal jumlah kasus kanker, dengan persentase 136,2% per 100.000 penduduk. Pria Indonesia memiliki tingkat kemungkinan tertinggi untuk mengidap kanker paru-paru, yakni sebanyak 19,4% per 100.000 penduduk, dan angka kematian mencapai 10,9% per 100.000 penduduk. Kanker hati juga menjadi masalah serius dengan tingkat kasus sebanyak 12,4% per 100.000 penduduk dan tingkat kematian sebesar 7,6% per 100.000 penduduk. Sementara itu, wanita Indonesia memiliki tingkat kemungkinan tertinggi untuk mengidap kanker payudara, yakni 42,1% per 100.000 penduduk, dengan angka kematian mencapai 17% per 100.000 penduduk. Kanker serviks juga menjadi perhatian dengan tingkat kasus sebanyak 23,4% per 100.000 penduduk dan tingkat kematian sebesar 13,9% per 100.000 penduduk[2]. Melalui data ini, diketahui bahwa kanker merupakan masalah serius dalam konteks kesehatan di Indonesia, dengan tingkat kemungkinan tertinggi terjadinya beberapa jenis kanker yang memerlukan perhatian khusus.

Kematian akibat kanker terus meningkat setiap tahunnya. Menurut penelitian yang dilakukan oleh Dr. Chin Tan Min, terjadi peningkatan jumlah kematian akibat kanker paru-paru pada tahun 2020, dengan mencatat sebanyak 34.783 kasus, yang menunjukkan peningkatan angka kematian sebesar 18% dibandingkan dengan tahun 2018[3]. Merokok merupakan faktor utama yang menyebabkan kanker paru-paru. Rokok mengandung lebih dari 60 zat beracun yang dapat memicu perkembangan penyakit kanker ini[4]. Kanker paru meningkat dengan sangat cepat karena konsumsi rokok berlebihan[5]. Selain itu, faktor-faktor lain yang dapat menjadi penyebab kanker paru-paru meliputi paparan asbes, radiasi, serta berbagai zat kimia seperti arsenik, krom, nikel, kloromil eter, gas mustard, dan partikel-partikel berbahaya di lingkungan kerja[6]. upaya pencegah kanker paru-paru dengan meningkatkan kesadaran tentang bahaya merokok serta pemahaman tentang risiko lainnya seperti paparan asbes, radiasi, dan zat kimia di lingkungan kerja.

Citra medis merujuk pada representasi visual dua dimensi dari bagian dalam tubuh manusia yang umumnya oleh profesional kesehatan untuk mendeteksi dan menganalisis penyakit pada pasien. Hasil citra yang dihasilkan tidak sesuai dengan harapan karena beberapa faktor potensial, seperti masalah dalam kontras, keaburan, noise atau gangguan visual, goresan, serta ketidakjelasan dalam menggambarkan sebagian detail, yang semuanya bisa terjadi selama proses pengambilan citra[7]. Pada citra medis yang menunjukkan adanya kemungkinan kanker paru-paru, dapat terlihat adanya bercak putih atau nodul yang muncul pada gambaran paru-paru yang terkena penyakit kanker [8]. Dilansir dari halodoc, waktu yang dibutuhkan untuk memindai nodul pada paru-paru membutuhkan waktu sekitar 10-30 menit[9]. Dengan waktu yang cukup lama akan melelahkan saat mendeteksi citra medis yang jumlahnya banyak. Berdasarkan kasus yang terjadi, para peneliti banyak melakukan penelitian terkait dengan mendeteksi citra medis pada kasus kanker.

Convolutional Neural Network (CNN) merupakan teknik *deep learning* yang terkenal dalam pengenalan pola pada citra. CNN memiliki kemampuan yang sangat baik dalam mengekstraksi fitur-fitur yang rumit dengan otomatis dan efisien, sehingga cocok digunakan untuk mengklasifikasikan citra dalam skala yang besar[10]. Secara dasar, *Convolutional Neural Network (CNN)* dapat dipisahkan menjadi dua tahap utama, yaitu tahap pembelajaran fitur (*feature learning*) dan tahap klasifikasi. Pada tahap pembelajaran fitur, terdapat komponen seperti lapisan *Convolutional, Rectified Linear Unit (ReLU)*, dan lapisan *Pooling*. Sedangkan pada tahap klasifikasi, fokusnya adalah mengklasifikasikan neuron-neuron ke dalam kelas-kelas yang sesuai[11]. *Convolutional Neural Network (CNN)* telah digunakan dalam berbagai kasus pengklasifikasian, salah satunya adalah dalam pengklasifikasian citra X-Ray paru-paru menjadi kategori *pneumonia* dan *non-pneumonia* [12], klasifikasi daging sapi dan daging babi[13][14][15], klasifikasi citra penyakit daun[16], klasifikasi citra stroke[17], klasifikasi citra *kolposkopi*[18], klasifikasi kanker kulit[19], klasifikasi menggunakan citra *computed tomography*[20], Deteksi dan klasifikasi kanker paru otomatis (ALCDC)[21]. Dengan demikian, CNN sangat baik digunakan dalam pemrosesan pada pola citra.

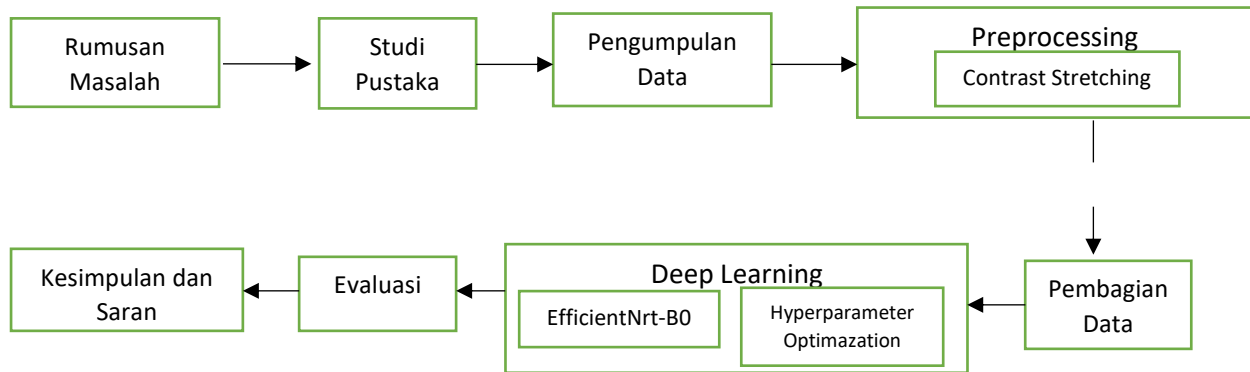
EfficientNet salah satu komponen dalam metode *Convolutional Neural Network (CNN)*. Mencakup serangkaian model yang berbeda, dimulai dari *EfficientNet-B0* hingga *EfficientNet-B7*, yang semula merupakan turunan dari model *EfficientNet-B0*[11]. Kelebihan utama dari *EfficientNet* adalah kemampuannya mencapai akurasi yang tinggi dalam tugas-tugas pengenalan gambar, sementara juga berhasil mengurangi jumlah parameter dan *Floating Point Operations Per Second (FLOPS)* yang dibutuhkan oleh model. Hal ini membuat *EfficientNet* menjadi solusi yang efisien dan efektif dalam mengoptimalkan kinerja model [10], arsitektur *efficientNet-B0* memiliki inputan 224x224[11]. *EfficientNet* mengungguli hampir semua arsitektur CNN, arsitektur lainnya seperti *alexNet*, *ResNet*, *GoogleNet*[22]. Terkait *efficientNet-B0* beberapa penelitian telah dilakukan, seperti klasifikasi kanker kulit menunjukkan nilai akurasi keseluruhan 87%, akurasi top-2 96%, dan akurasi top-3 99%[19]. Selain itu, dalam klasifikasi patologi makula pada retina berdasarkan citra retinal OCT, penelitian tersebut mencapai tingkat akurasi pengujian sebesar 90,60% dengan nilai loss pengujian sebesar 0,27[23]. *EfficientNet* menjadi arsitektur CNN yang efektif dalam mengoptimalkan kinerja model.

Metode *Contrast Stretching* meningkatkan kontras dalam citra dengan memperluas rentang intensitas warna. Tujuannya adalah membuat citra lebih tajam dan membedakan area yang berbeda dengan lebih jelas dengan memperluas rentang dinamis level grayscale pada gambar menjadi 0 hingga 255.[7]. *Contrast Stretching* adalah salah satu teknik peningkatan citra (*image enhancement*). Teknik ini dikenal memiliki hasil yang lebih baik dibandingkan dengan teknik lainnya dalam mengelola citra. *Contrast Stretching* bertujuan untuk meningkatkan kontras dalam citra sehingga citra tersebut memiliki detail yang lebih tajam dan perbedaan antara objek yang berbeda menjadi lebih jelas[24]. Teknik ini sering digunakan untuk memperbaiki citra yang memiliki kontras yang rendah atau terlalu gelap sehingga membuat informasi dalam citra sulit terlihat [25]. Penerapan teknik *enhancement* citra, khususnya pada citra MRI otak, telah diuji. Hasil menunjukkan bahwa penggunaan *contrast stretching* lebih baik dari pada *Histogram Equalization* dan *Adaptive Histogram Equalization* dalam meningkatkan kontras citra MRI otak(7). *Contrast stretching* efektif dalam memperbaiki visualisasi dan meningkatkan detail dalam citra.

Penelitian ini menggunakan *Convolutional Neural Network (CNN)* dengan arsitektur *EfficientNet-B0* untuk mengklasifikasikan kanker paru-paru. pada *preprocessing* dilakukan proses *image enhancement*, Untuk meningkatkan performa CNN, penelitian menggunakan optimisasi hyperparameter menggunakan tiga jenis algoritma yaitu adam, adagrad, dan SGD. Analisis dilakukan untuk mengevaluasi model yang dihasilkan dengan menggunakan matriks klasifikasi. [26]. Harapannya, hasil penelitian ini akan memberikan solusi yang dapat mempercepat dan menyederhanakan proses klasifikasi citra medis sehingga hasil yang diperoleh menjadi lebih akurat dan efisien. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi positif dalam dunia medis, membantu dokter dan profesional kesehatan dalam mendiagnosis penyakit serta meningkatkan pengobatan pasien melalui penggunaan teknologi klasifikasi citra yang lebih baik.

2. METODOLOGI PENELITIAN

Pada bagian ini, akan diuraikan metode yang telah peneliti terapkan dalam menjalankan penelitian ini. Gambar 1 akan menampilkan serangkaian langkah-langkah yang telah dilaksanakan selama penelitian ini dilakukan.



Gambar 1. Metodologi Penelitian

2.1 Rumusan Masalah

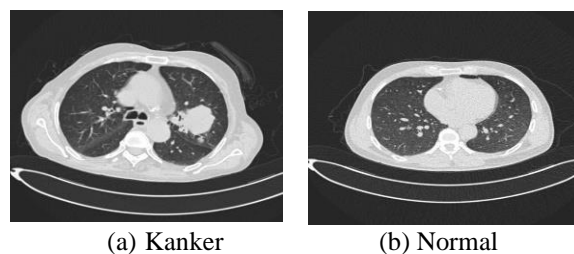
Penelitian ini difokuskan pada merumuskan permasalahan yang berkaitan dengan penerapan teknik *contrast stretching* pada *Convolutional Neural Network* (CNN) dengan menggunakan model *EfficientNet-B0* dalam konteks klasifikasi kanker paru-paru. Dalam kerangka penelitian ini, CNN digunakan untuk membedakan citra CT-Scan paru-paru, mengidentifikasi apakah citra tersebut menunjukkan indikasi kanker atau dalam kondisi normal.

2.2 Studi Pustaka

Mengidentifikasi permasalahan yang memerlukan penyelesaian dan melakukan tinjauan pustaka serta penelitan yang relevan untuk mencari solusi. mengkaji berbagai teori terkait antara lain teori tentang gambar CT-Scan, klasifikasi kanker, teknik pengolahan gambar seperti *contrast stretching*, dan konsep terkait metode *Convolutional Neural Network* (CNN).

2.3 Pengumpulan Data

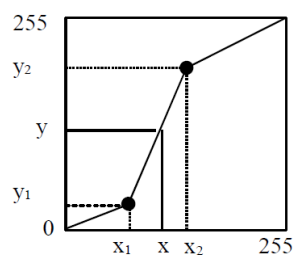
Data bersumber dari dataset pada www.kaggle.com berisi data CT-Scan paru-paru. Total terdapat 1097 gambar dalam format JPG. Data ini dibagi menjadi dua kelas, yaitu kelas "kanker" dengan 681 gambar dan kelas "normal" dengan 416 gambar. Gambar 2 menunjukkan contoh dataset CT-Scan paru-paru yang terdiri dari citra-citra kanker dan normal.



Gambar 2. Dataset Kanker Dan Normal Paru

2.4 Preprocessing

Langkah selanjutnya adalah *preprocessing* data citra yang terdapat pada dataset. *Preprocessing* yang dilakukan adalah *image enhancement*. Teknik yang digunakan pada tahap *image enhancement* adalah *contrast stretching* untuk meningkatkan kontras. *Contrast stretching* teknik yang meningkatkan kontras citra dengan memperluas rentang level *grayscale* dari 0 hingga 255[24]. Gambar 3 merupakan transformasi *contrast stretching*.



Gambar 3. Transformasi Contrast Stretching

Dalam Gambar 3, x mengacu pada nilai intensitas atau *pixel* pada citra sebelum mengalami proses, sementara y mengacu pada nilai intensitas atau *pixel* pada citra setelah melalui proses tersebut. Posisi titik (x_1, y_1) dan (x_2, y_2) akan memengaruhi bagaimana bentuk fungsi transformasi dan penyebaran tingkat keabuan pada citra yang dihasilkan.

Mengasumsikan bahwa citra ini memiliki tingkat keabuan dalam rentang 0 hingga 255. Untuk menghitung nilai transformasi tersebut, dapat menentukannya dengan mencari nilai yang sesuai untuk nilai-nilai yang sudah diketahui:

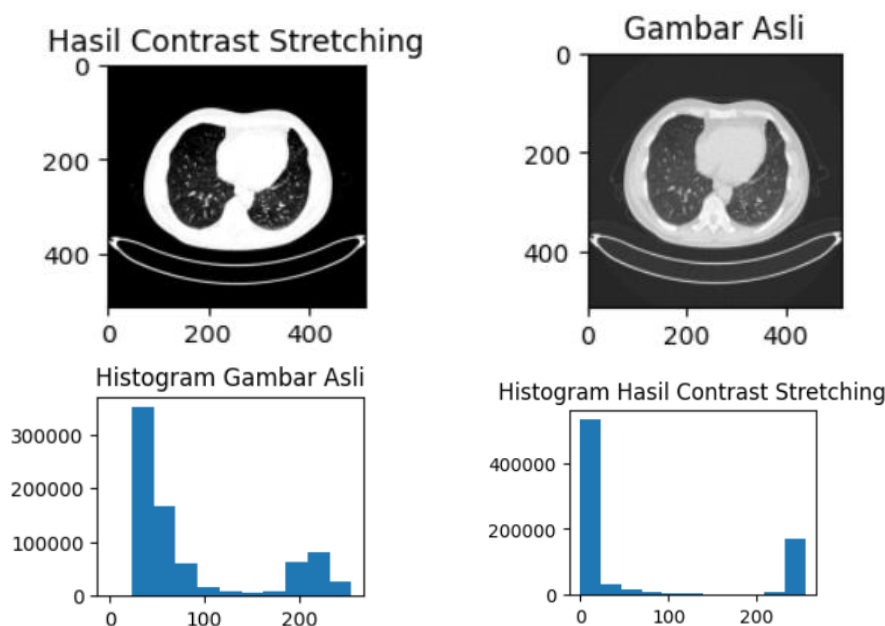
$$0 \leq x < x_1, \text{ maka } y = y_1 \times \left(\frac{x}{x_1}\right) \tag{1}$$

$$x_1 \leq x < x_2, \text{ maka } y = y_1 + (y_2 - y_1) \times \frac{(x-x_1)}{(x_2-x_1)} \tag{2}$$

$$x_2 \leq x < 225, \text{ maka } y = y_2 + (225 - y_2) \times \frac{(x-x_2)}{(225-x_2)} \tag{3}$$

Dengan x_1, x_2, y_1, y_2 memiliki nilai besaran dari 0 sampai 225, ini dilakukan untuk

menghindari menghasilkan nilai negatif dari operasi $(x_2 - x_1)$ atau $(y_2 - y_1)$, sehingga mendapatkan nilai intensitas atau pixel baru yang selalu berada dalam rentang 0 hingga 255. Gambar 4 melihatkan dataset citra asli dan hasil *contrast stretching*.



Gambar 4. Citra Asli Dan Hasil Contrast Stretching

2.5 Pembagian Data

Setelah proses *preprocessing*, langkah selanjutnya adalah membagi dataset menjadi dua kelompok utama: data latih (training data) dan data uji (testing data). Pembagian data train, test, dan validasi dilakukan secara manual dengan membedakan setiap kasus yang terdapat pada dataset. Dengan demikian, tidak ada data yang sama disetiap pembagian data. Rasio pembagian data latih dan uji yang diterapkan adalah 90:10. Sebagian dari data latih juga akan digunakan untuk pembentukan data validasi, di mana 10% dari data latih akan dialokasikan untuk percobaan dengan perbandingan 90:10.

Tabel 1. Dataset Original dan Contrast Stretching Rasio Pembagian 90:10

Dataset Original			Dataset Contrast Stretching		
Train (90%)	Validation (10%)	Test	Train (90%)	Validation (10%)	Test
891	97	109	891	97	109

Tabel 1 juga menampilkan dataset dalam dua bentuk, data asli dan data yang ditingkatkan kontrasnya. Perbedaan signifikan terletak pada rasio pembagian dataset, yaitu 90:10 antara data pelatihan dan data uji, serta 90:10 antara data pelatihan dan data validasi. Tabel ini memberikan informasi tentang jumlah data di setiap bagian dataset, baik yang masih dalam bentuk asli maupun yang telah ditingkatkan kontrasnya. Pembagian dataset ini memegang peranan penting dalam pembelajaran mesin.

2.6 Deep Learning

Dataset yang telah diproses pada tahap sebelumnya akan digunakan untuk melatih arsitektur *EfficientNet-B0*. Pada dasarnya, *Convolutional Neural Network* (CNN) memiliki dua tahap utama, yaitu tahap *feature learning* dan tahap klasifikasi. Dalam penelitian ini, dilakukan optimasi *hyperparameter*, yang merupakan parameter-parameter yang memiliki pengaruh signifikan pada proses pembelajaran model dan penyesuaian bobot selama proses optimasi. Beberapa algoritma optimasi yang digunakan dalam penelitian ini adalah SGD, Adam, dan Adagrad, yang berfungsi untuk

mengoptimalkan parameter-parameter dalam model. Selain itu, *learning rate* digunakan untuk mengendalikan sejauh mana perubahan bobot model terjadi selama proses pelatihan. Pengaturan yang optimal dari *learning rate* dapat memengaruhi kinerja model secara keseluruhan

2.7 Evaluasi

Setelah mendapatkan hasil dari serangkaian pengujian, langkah selanjutnya adalah tahap evaluasi. Evaluasi dilakukan untuk mengukur kinerja dari model yang telah dibuat. Pengukuran performa dalam penelitian ini menggunakan matriks klasifikasi, yang dikenal sebagai *confusion matrix*. *Confusion matrix* adalah alat pengukuran performa yang umum digunakan dalam masalah klasifikasi dalam *machine learning*, terutama saat output dapat berupa dua kelas atau lebih. Matriks konfusi memiliki empat sel utama, yang mewakili empat kemungkinan hasil prediksi dalam masalah klasifikasi biner:

- True Positive (TP): kasus yang benar-benar positif yang diklasifikasi dengan benar oleh model.
- True Negative (TN): kasus yang benar-benar negatif yang diklasifikasi dengan benar oleh model.
- False Negative (FN): kasus yang benar-benar negatif, tetapi diklasifikasi positif oleh model.
- False Positive (FP): kasus yang benar-benar positif, tetapi diklasifikasi negatif oleh model.

Berikut adalah rumus-rumus yang digunakan untuk menghitung beberapa metrik evaluasi kinerja klasifikasi berdasarkan matriks konfusi:

- Akurasi: Akurasi mengukur sejauh mana model benar-benar memprediksi dengan benar kelas-kelas target. Rumusnya adalah:

$$akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (4)$$

- Presisi: Presisi mengukur sejauh mana prediksi positif yang dibuat oleh model adalah benar. Ini memberikan gambaran tentang berapa banyak dari prediksi positif yang sebenarnya benar. Rumusnya adalah:

$$presisi = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

- Recall: Sensitivitas mengukur sejauh mana model mampu mendeteksi semua kasus positif yang sebenarnya. Ini memberikan gambaran tentang seberapa baik model mengidentifikasi kasus positif. Rumusnya adalah:

$$recall = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

- F1 Score: F1 score mencerminkan sejauh mana model mencapai akurasi pada seluruh dataset. Rumusnya adalah:

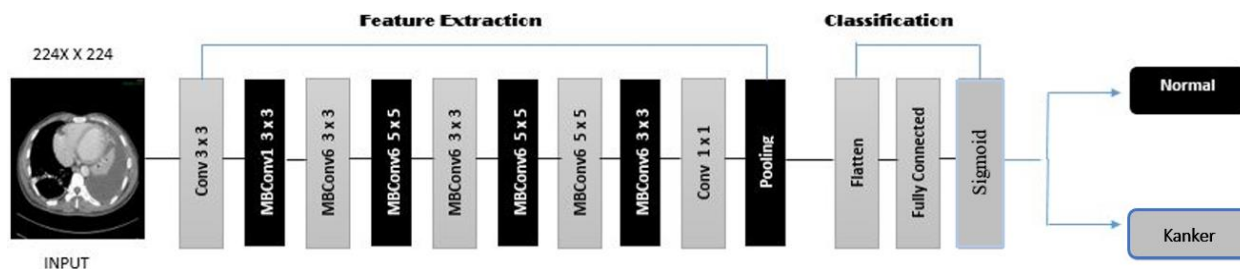
$$f1\ score = 2 \times \frac{presisi \times recall}{presisi + recall} \quad (7)$$

2.8 Kesimpulan dan Saran

Tahap akhir, penulis menyajikan ringkasan kesimpulan akhir penelitian dan memberikan saran untuk penelitian selanjutnya.

3. HASIL DAN PEMBAHASAN

Dalam penelitian ini, digunakan algoritma *Convolutional Neural Network* (CNN) dengan menggunakan arsitektur *EfficientNet-B0*. Gambar 4 menggambarkan rancangan arsitektur *EfficientNet-B0* yang digunakan dalam penelitian ini. *EfficientNet-B0* merupakan salah satu model yang efisien dan efektif dalam pengolahan citra, dan digunakan untuk tugas klasifikasi citra kanker paru-paru dalam penelitian ini.



Gambar 5. Arsitektur EfficientNet-B0

Gambar 5 diatas mengilustrasikan serangkaian langkah-langkah yang terkandung dalam arsitektur EfficientNet-B0 digunakan dalam penelitian bertujuan mengklasifikasikan gambar-gambar CT-Scan paru-paru. Rangkaian ini mencakup berbagai komponen dalam arsitektur EfficientNet-B0 yang digunakan dalam tugas klasifikasi:

- Convolutional Blocks: Terdiri dari beberapa blok konvolusi yang menggabungkan konvolusi depthwise dan pointwise untuk mengekstraksi fitur dari input gambar.

- b. Input Layer: Layer awal pada jaringan saraf menerima gambar sebagai masukan dengan dimensi umumnya adalah (224, 224, 3), yang menandakan gambar berwarna dengan ukuran 224x224 piksel
- c. Width Scaling, Depth Scaling, dan Resolution Scaling: Konsep penskalaan dalam lebar, kedalaman, dan resolusi digunakan untuk menciptakan variasi model yang efisien untuk berbagai tugas.
- d. Squeeze-and-Excitation (SE) Blocks: Beberapa versi EfficientNet menggunakan blok SE untuk meningkatkan kemampuan model dalam memfokuskan fitur-fitur yang signifikan.
- e. Global Average Pooling (GAP): Menggantikan fully connected layers di akhir model dengan GAP untuk mengurangi jumlah parameter dan mencegah overfitting.
- f. Output Layer: Lapisan output yang menghasilkan prediksi berdasarkan fitur-fitur yang diekstraksi dari blok-blok sebelumnya.
- g. Batch Normalization: Batch Normalization sering digunakan untuk mempercepat pelatihan dan meningkatkan stabilitas model.
- h. Depthwise Separable Convolution: EfficientNet-B0 menggunakan konvolusi yang terpisah secara mendalam (depthwise separable convolution), yang memisahkan operasi konvolusi menjadi dua tahap: konvolusi depthwise dan konvolusi pointwise. Hal ini membantu mengurangi jumlah parameter dan operasi komputasi.
- i. Pooling Layers: Lapisan pooling, seperti Average Pooling atau Max Pooling, dimanfaatkan untuk mengurangi dimensi spasial dari fitur yang telah diekstraksi.

Tabel 3. Hyperparameter

Optimizer	Learnig Rate
SGD	0.1
ADAM	0.01
ADAGRAD	0.001
	0.0001

Tabel 3 menyajikan informasi mengenai *hyperparameter* yang akan digunakan dalam proses pelatihan model. *Hyperparameter* adalah parameter-parameter yang telah ditentukan sebelum dimulainya proses pelatihan dan berperan penting dalam pengaruh metode pembelajaran serta penyesuaian bobot selama proses optimasi. Tabel ini juga mencakup berbagai algoritma pengoptimal seperti SGD, Adam, dan Adagrad yang berfungsi untuk mengoptimalkan parameter-parameter dalam model. Selain itu, *learning rate* digunakan untuk mengendalikan seberapa cepat atau lambat bobot model akan berubah selama proses pelatihan.

3.1 Proses Pemilihan Model

Dalam tahap ini dilakukan pemilihan model terhadap dataset Original dan dataset Contrast Stretching. Berbagai skenario pooling seperti Global Average Pooling dan Global Max Pooling, serta lapisan output Dense dengan ukuran 64, 32, 16, 1 dan 32, 16, 8, 1 dieksplorasi. Proses ini bertujuan untuk mengidentifikasi model optimal yang akan digunakan dalam pelatihan. Evaluasi dilakukan berdasarkan nilai F1 score tertinggi yang diperoleh dari setiap skenario yang diuji.

Tabel 4. Hasil Pemilihan Model Data Original

No	Pooling	Optimizer	Dense Out Layers	Learning Rate	F1 Score	Akurasi
1	Average	SGD	64, 32, 16, 1	0.1	54.47	54.55
2	Max	SGD	64, 32, 16, 1	0.1	49.24	49.49
3	Average	SGD	32, 16, 8, 1	0.1	40.22	46.46
4	Max	Adam	64, 32, 16, 1	0.1	53.04	59.60
5	Max	Adam	32, 16, 8, 1	0.1	41.27	43.43
6	Max	Adagrad	64, 32, 16, 1	0.1	49.99	50.51
7	Average	Adagrad	32, 16, 8, 1	0.1	41.42	41.41

Tabel 5. Hasil Pemilihan Model Data Contrast Stretching

No	Pooling	Optimizer	Dense Out Layers	Learning Rate	F1 Score	Akurasi
1	Average	SGD	64, 32, 16, 1	0.1	55.58	57.58
2	Max	SGD	64, 32, 16, 1	0.1	41.95	50.51
3	Average	SGD	32, 16, 8, 1	0.1	49.06	56.57
4	Max	Adam	64, 32, 16, 1	0.1	39.31	46.46
5	Max	Adam	32, 16, 8, 1	0.1	53.15	53.54
6	Max	Adagrad	64, 32, 16, 1	0.1	46.82	47.47
7	Average	Adagrad	32, 16, 8, 1	0.1	49.23	51.52

Tabel 4 dan 5 merupakan hasil yang didapatkan pada saat proses pemilihan model, pemilihan dilakukan dengan berbagai skenario yang bertujuan untuk melihat perbedaan disetiap hasil pemilihan model. Tabel diatas memberikan informasi yaitu tentang jenis pooling, optimizer, dense out layers, learning rate, F1 score, akurasi. Hasil yang didapatkan dari dataset Original dan Contrast Stretching bahwa Average Pooling dengan Dense Out Layers 64, 32, 16, 1 mendapatkan

hasil yang lebih tinggi dibandingkan variasi yang lainnya. Oleh karena itu, diputuskan Average Pooling dengan Dense Out Layers 64, 32, 16, 1 akan digunakan pada proses pelatihan model.

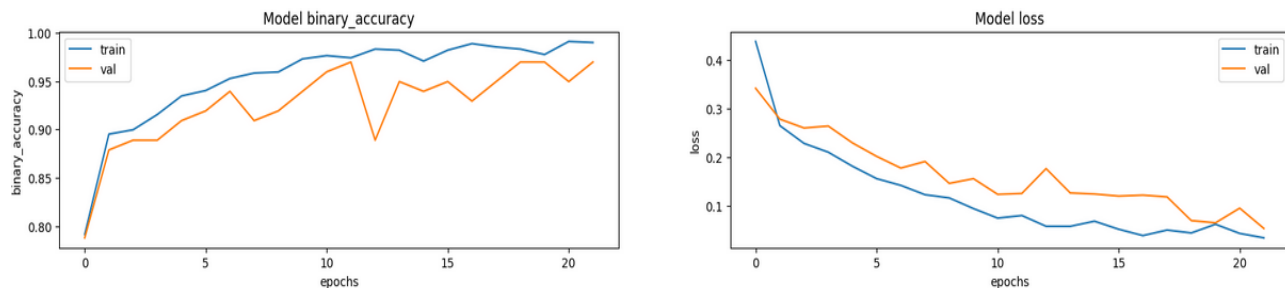
3.2 Proses Pelatihan Model

Pada tahap pelatihan model EfficientNet-B0, berbagai variasi hyperparameter seperti optimizer dan learning rate diterapkan. Proses pelatihan ini dilakukan dengan menggunakan library Tensorflow dan Keras. Untuk menjalankan eksperimen secara komprehensif, pelatihan model diulang sebanyak 24 kali untuk setiap rasio pembagian data yang digunakan. Pelatihan dilakukan selama 50 epoch dengan menerapkan metode early stopping, yang berfungsi untuk memantau nilai validasi loss. Jika terjadi peningkatan validasi loss, maka proses pelatihan akan dihentikan untuk mencegah overfitting. Batch size yang digunakan adalah 64, yang berarti model diperbarui setiap kali menggunakan batch berisi 64 sampel data. Pemilihan batch size ini dapat memengaruhi kecepatan pelatihan dan penggunaan sumber daya komputer. Pelatihan model diterapkan terhadap data Train dan validasi, hasil pelatihan model ditampilkan dalam tabel.

Tabel 6. Hasil Pelatihan Model

No	Dataset	Optimizer	Learning Rate	Training Loss	Training Binary Accuracy	Validasi Loss	Validasi Binary Accuracy
1	Original	SGD	0.1	0.1859	0.9312	0.2878	0.9091
2	Original	SGD	0.01	0.0813	0.9786	0.0691	0.9798
3	Original	SGD	0.001	0.1535	0.9493	0.2024	0.9192
4	Original	SGD	0.0001	0.071	0.9842	0.1058	0.9596
5	Original	Adam	0.1	0.0339	0.9899	0.0532	0.9697
6	Original	Adam	0.01	0.0313	0.9899	0.1516	0.9495
7	Original	Adam	0.001	0.0298	0.9876	0.0605	0.9697
8	Original	Adam	0.0001	0.0344	0.9944	0.0961	0.9798
9	Original	Adagrad	0.1	0.4086	0.8794	0.3935	0.9091
10	Original	Adagrad	0.01	0.3415	0.8726	0.4077	0.8283
11	Original	Adagrad	0.001	0.315	0.8805	0.35	0.8586
12	Original	Adagrad	0.0001	0.2685	0.8884	0.3627	0.8889
13	Contrast Stretching	SGD	0.1	0.0969	0.9696	0.0985	0.9596
14	Contrast Stretching	SGD	0.01	0.0983	0.9808	0.1119	0.9495
15	Contrast Stretching	SGD	0.001	0.1122	0.9707	0.1891	0.9091
16	Contrast Stretching	SGD	0.0001	0.1465	0.9448	0.1648	0.9293
17	Contrast Stretching	Adam	0.1	0.0394	0.9842	0.1031	0.9697
18	Contrast Stretching	Adam	0.01	0.032	0.9899	0.0354	0.9899
19	Contrast Stretching	Adam	0.001	0.0223	0.991	0.0243	0.9899
20	Contrast Stretching	Adam	0.0001	0.0255	0.9921	0.1997	0.9495
21	Contrast Stretching	Adagrad	0.1	0.2471	0.9312	0.2662	0.9192
22	Contrast Stretching	Adagrad	0.01	0.2789	0.9019	0.3189	0.8485
23	Contrast Stretching	Adagrad	0.001	0.3139	0.885	0.3234	0.8687
24	Contrast Stretching	Adagrad	0.0001	0.3123	0.8963	0.3074	0.8889

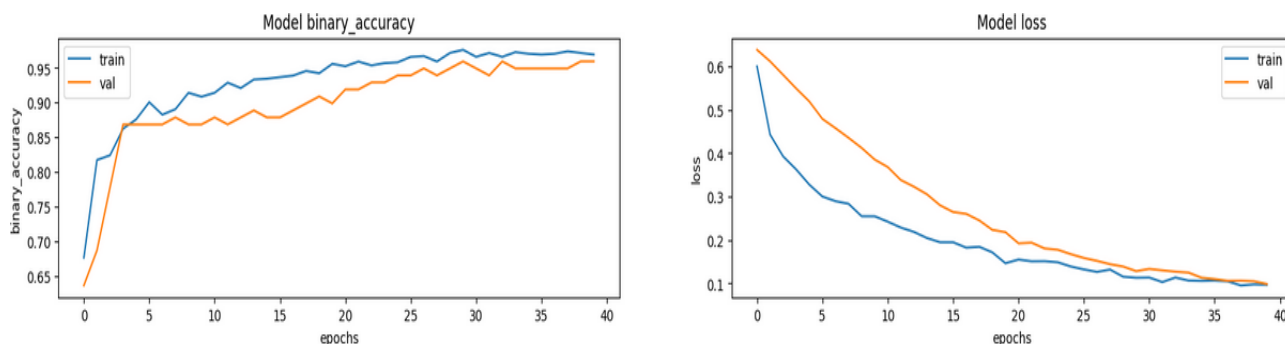
Tabel 6 merupakan hasil yang didapatkan pada saat proses pelatihan, pelatihan dilakukan terhadap data train dan validasi dengan berbagai skenario yang bertujuan untuk melihat perbedaan disetiap hasil pelatihan model. Tabel diatas memberikan informasi yaitu tentang jenis data, optimizer, learning rate, training loss, training binary accuracy, validation loss. Validation binary accuracy. Training loss dan validation loss bertujuan untuk memahami seberapa dekat prediksi model dengan label asli. Selain itu, training binary accuracy dan validation binary accuracy mengukur sejauh mana model mampu mengklasifikasi data dengan benar. Berikut beberapa grafik dari pelatihan model.



Gambar 6. Grafik Pelatihan Model Eksperimen Ke-1 Dataset Original

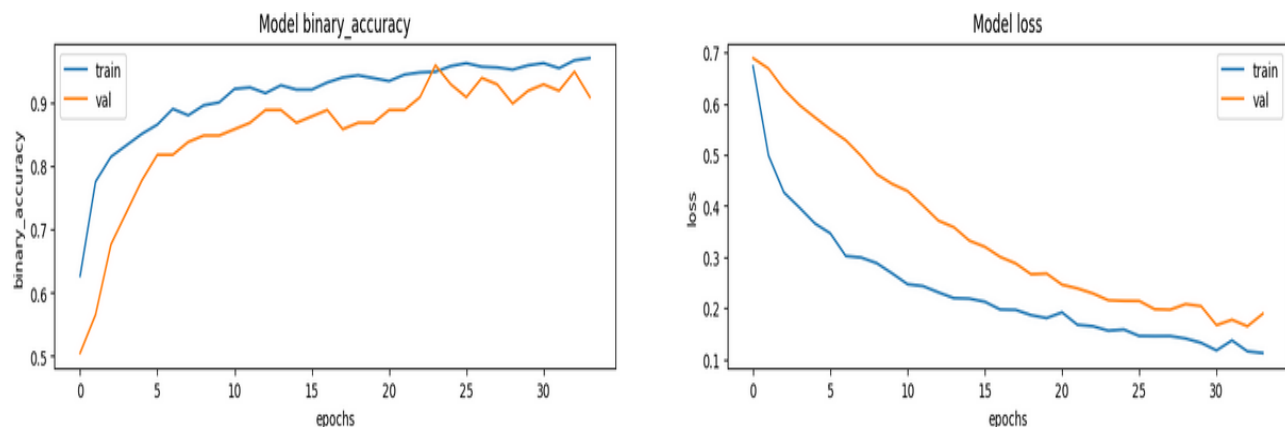
Grafik dalam gambar 6 menggambarkan proses pelatihan dari eksperimen ke-1 dataset original. Pada eksperimen ini, model dilatih menggunakan data original dengan pengaturan SGD dan learning rate 0.1. Hasil pelatihan menunjukkan

bahwa akurasi validasi mencapai 96.97% yang tidak memiliki perbedaan yang signifikan dengan akurasi pelatihan sebesar 98.99%.



Gambar 7. Grafik Pelatihan Model Eksperimen Ke-1 Dataset Contrast Stretching

Grafik dalam gambar 7 menggambarkan proses pelatihan dari eksperimen ke-1 dataset contrasts stretching. Pada eksperimen ini, model dilatih menggunakan data original dengan pengaturan SGD dan *learning rate* 0.1. Hasil pelatihan menunjukkan bahwa akurasi validasi mencapai 95.96% yang tidak memiliki perbedaan yang signifikan dengan akurasi pelatihan sebesar 96.96%.



Gambar 8. Grafik Pelatihan Model Eksperimen Ke-3 Dataset Contrast Stretching

Grafik dalam gambar 7 menggambarkan proses pelatihan dari eksperimen ke-3 dataset contrasts stretching. Pada eksperimen ini, model dilatih menggunakan data original dengan pengaturan SGD dan *learning rate* 0.001. Hasil pelatihan menunjukkan bahwa akurasi validasi mencapai 90.91% yang tidak memiliki perbedaan yang signifikan dengan akurasi pelatihan sebesar 97.07%.

3.3 Proses Pengujian Model

Setelah selesai tahap pelatihan model, langkah selanjutnya adalah melakukan pengujian dengan menggunakan data validasi yang telah disiapkan sebelumnya. Pengujian ini akan dilakukan untuk setiap skenario yang telah diatur. Setelah itu, hasil pengujian dari data asli dan data yang telah ditingkatkan kontrasnya akan dibandingkan. Perbandingan performa model akan dievaluasi menggunakan matriks kebingungan untuk mengukur akurasi dan F1 score. Berikut ini adalah tabel hasil dari proses pengujian.

Tabel 7. Hasil Pengujian Model

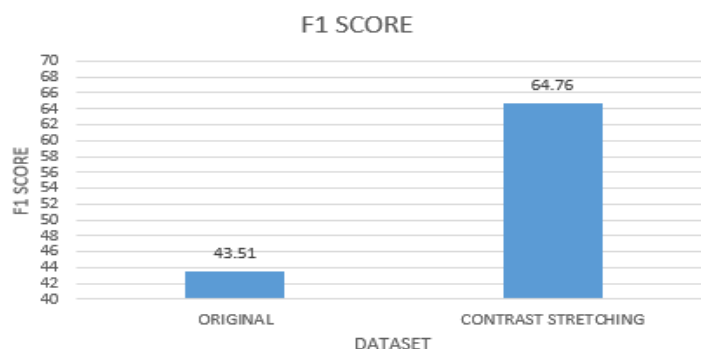
No	Dataset	Optimizer	Learning Rate	F1 Score	Akurasi
1	Original	SGD	0.1	46.2	53.61
2	Original	SGD	0.01	43.39	48.45
3	Original	SGD	0.001	48.7	53.61
4	Original	SGD	0.0001	45.66	50.52
5	Original	Adam	0.1	41.13	46.39
6	Original	Adam	0.01	50.19	54.64
7	Original	Adam	0.001	47.92	52.58
8	Original	Adam	0.0001	45.66	50.52
9	Original	Adagrad	0.1	40.74	48.45
10	Original	Adagrad	0.01	49.33	56.7
11	Original	Adagrad	0.001	46.82	48.45

12	Original	Adagrad	0.0001	40.97	50.52
13	Contrast Stretching	SGD	0.1	48.7	54.64
14	Contrast Stretching	SGD	0.01	45.63	51.55
15	Contrast Stretching	SGD	0.001	50.26	55.67
16	Contrast Stretching	SGD	0.0001	52.57	57.73
17	Contrast Stretching	Adam	0.1	45.66	50.52
18	Contrast Stretching	Adam	0.01	56.98	60.82
19	Contrast Stretching	Adam	0.001	52.45	56.7
20	Contrast Stretching	Adam	0.0001	47.18	52.58
21	Contrast Stretching	Adagrad	0.1	44.04	50.52
22	Contrast Stretching	Adagrad	0.01	52.59	58.76
23	Contrast Stretching	Adagrad	0.001	44.04	50.52
24	Contrast Stretching	Adagrad	0.0001	55.54	59.79

Tabel 7 merupakan hasil yang didapatkan pada saat proses pengujian, pengujian dilakukan terhadap data train dan validasi dengan berbagai skenario yang bertujuan untuk melihat perbedaan disetiap hasil pengujian model. Tabel diatas memberikan informasi yaitu dataset original mendapatkan F1 score dan Akurasi tertinggi terhadap skenario optimizer Adam dan Learning rate 0.01 dengan nilai F1 score 50.19 dan akurasi 54.64, sedangkan dataset contrast stretching mendapatkan F1 score dan akurasi tertinggi terhadap skenario optimizer Adam dan Learning rate 0.01 dengan nilai F1 score 56.98 dan akurasi 60.82.

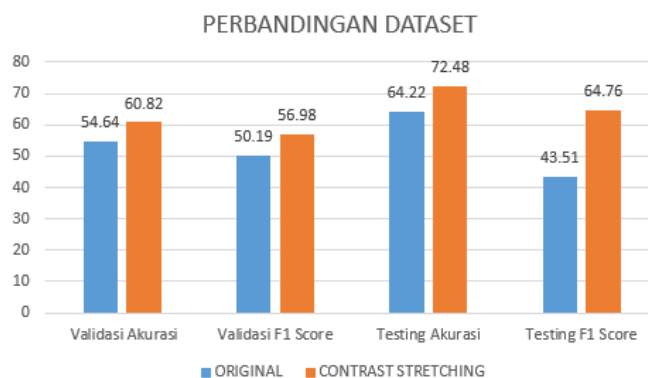
3.4 Proses Pengujian Model Terhadap Data Testing

Langkah selanjutnya adalah melakukan pengujian dengan menggunakan data uji yang telah disiapkan sebelumnya. Pengujian ini akan dilakukan menggunakan hasil skenario tertinggi dari proses pengujian model. Data original menggunakan optimezer SGD dengan learning rate 0.1, average pooling, out dense 64, 32, 16, 1. Sedangkan data contrast stretching menggunakan optimizer Adam dengan learning rate 0.01, average pooling, out dense 64, 32, 16, 1. Setelah itu, hasil pengujian dari data asli dan data yang telah ditingkatkan kontrasnya akan dibandingkan. Perbandingan performa model akan dievaluasi menggunakan matriks kebingungan untuk mengukur akurasi, presisi, recall, dan F1 score. Berikut ini adalah gambar hasil dari proses pengujian.



Gambar 9. Grafik Perbandingan Hasil F1-Score Original dan Contrast Stretching

Gambar 9 memaparkan grafik perbandingan hasil f1 score original dan contrast stretching, dapat dilihat fi score data original dan data contrast stretching menggunakan skenario yang sudah dipersiapkan sebelumnya. Perbedaan kedua data sangat jelas terlihat, dara original f1 score yaitu 43.51%, sedangkan data contrasts stretching f1 score yaitu 64.76%.



Gambar 10. Grafik Perbandingan Dataset Original dan Contrast Stretching

Gambar 10 memaparkan grafik perbandingan akurasi dataset original dan *contrast stretching*, dapat dilihat akurasi dataset original dan dataset *contrast stretching* memiliki perbedaan yang sangat jelas terlihat, dataset original validasi akurasi yaitu 64.22% dan validasi f1 score 43.51%. Sedangkan dataset *contrast stretching* validasi akurasi yaitu 72.48% dan validasi f1 score 64.76%. Sedangkan untuk testing akurasi dataset original yaitu 54.55% dan testing f1 score 54.47%, dataset *contrast stretching* testing akurasi yaitu 60.61% dan testing f1 score 54.22

4. KESIMPULAN

Penelitian ini membandingkan metode klasifikasi kanker paru menggunakan dua pendekatan: pertama, dengan menggunakan teknik *contrast stretching* pada *convolution neural network* (CNN), dan kedua dengan menggunakan data citra asli. Dataset yang digunakan terbagi menjadi dua kelompok, yaitu dataset asli dan dataset yang telah ditingkatkan kontrasnya melalui *contrast stretching*. Totalnya, dataset ini terdiri dari 1097 citra CT-Scan paru. Data dibagi secara manual dengan membedakan setiap kasus, sehingga tidak ada citra yang sama disetiap data. Data dibagi menjadi rasio pembagian 90:10 untuk pelatihan dan validasi model. Untuk mengukur kinerja model, digunakan matriks konfusi (confusion matrix), yang menghasilkan akurasi, presisi, recall, dan skor F1 sebagai metrik evaluasi. Hasil penelitian menunjukkan bahwa penggunaan teknik *contrast stretching* pada citra CT-scan kanker paru memberikan hasil yang lebih baik dibandingkan dengan menggunakan dataset asli. Hasil klasifikasi tertinggi menggunakan optimizer ADAM, dan learning rate 0.01, di mana akurasi mencapai 72.48%, presisi 71.52%, recall 64.2%, dan skor F1 64.76%. Sementara itu, hasil yang diperoleh dari dataset asli menunjukkan perbedaan. Hasil klasifikasi menggunakan optimizer ADAM dan learning rate 0.01, dengan akurasi mencapai 64.22%, presisi 52.69%, recall 50.52%, dan skor F1 43.51%. Hasil ini menunjukkan bahwa penggunaan *contrast stretching* dalam klasifikasi kanker paru-paru sangat efektif dalam meningkatkan akurasi, dan disarankan untuk menggabungkannya dengan teknik pengolahan citra lainnya seperti Directional Wavelet Transform (DWT) serta menggunakan lebih banyak data untuk peningkatan kinerja yang lebih baik.

REFERENCES

- [1] W. H. Organization, "Cancer," *World Health Organization*, 2022. <https://www.who.int/news-room/fact-sheets/detail/cancer> (accessed Jan. 06, 2023).
- [2] K. Kesetahan, "Penyakit Kanker di Indonesia Berada Pada Urutan 8 di Asia Tenggara dan Urutan 23 di Asia," *p2p.kemkes.go.id*, 2019. <http://p2p.kemkes.go.id/penyakit-kanker-di-indonesia-berada-pada-urutan-8-di-asia-tenggara-dan-urutan-23-di-asia/> (accessed Jun. 01, 2023).
- [3] E. Pranita, "Situasi Kanker Paru di Indonesia Saat Ini, Prevalensi Kematian Meningkat," *Kompas.com*, 2021. <https://www.kompas.com/sains/read/2021/12/10/183100723/situasi-kanker-paru-di-indonesia-saat-ini-prevalensi-kematian-meningkat?page=all> (accessed Dec. 16, 2022).
- [4] D. Pittara, "Penyebab Kanker Paru-paru Kanker Paru-Paru," *alodokter.com*, 2022. <https://www.alodokter.com/kanker-paru-paru/penyebab> (accessed Jan. 06, 2023).
- [5] N. Aliyah, E. Pranggono, and B. Andriyoko, "Kanker Paru: Sebuah Kajian Singkat," *Indones. J. Chest Emerg. Med.*, vol. 4, no. 1, pp. 28–32, 2016.
- [6] M. Keluarga, "Fakta Seputar Kanker Paru: Penyebab, Gejala, dan Pengobatan," *mitra keluarga*, 2022. <https://www.mitrakeluarga.com/artikel/artikel-kesehatan/kanker-paru> (accessed Dec. 27, 2022).
- [7] I. W. A. Wijaya Kusuma and A. Kusumadewi, "Penerapan Metode Contrast Stretching, Histogram Equalization Dan Adaptive Histogram Equalization Untuk Meningkatkan Kualitas Citra Medis Mri," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 11, no. 1, pp. 1–10, 2020, doi: 10.24176/simet.v11i1.3153.
- [8] N. Jawas and I. W. B. Sentana, "Deteksi Lung Nodule Otomatis dari Citra Medis Computed Tomography (CT) Scan dengan Operasi Morfology dan SVM," *J. Sist. Dan Inform.*, vol. 12, no. 2, pp. 140–148, 2018.
- [9] dr. Rizal Fadli, "Beginilah Cara CT Scan Mendeteksi Kanker Lebih Akurat," *halodoc*, 2019. <https://www.halodoc.com/artikel/cara-ct-scan-mendeteksi-kanker-lebih-akurat> (accessed Jan. 06, 2023).
- [10] R. A. Ramadhani, B. Wahyu, and R. Purbaningtyas, "Klasifikasi Tumor Otak Menggunakan Convolutional Neural Network Dengan Arsitektur EfficientNet-B3," *JUST IT J. Sist. Informasi, Teknol. Informasi, dan Komput.*, vol. 11, no. 3, pp. 55–59, 2021.
- [11] S. Alamgunawan and Y. Kristian, "Klasifikasi Tekstur Serat Kayu pada Citra Mikroskopik Veneer Memanfaatkan Deep Convolutional Neural Network," *J. Intell. Syst. Comput.*, vol. 2, no. 1, pp. 06–11, 2021, doi: 10.52985/insyst.v2i1.152.
- [12] M. Ramadhan, D. Iskandar Mulyana, M. Betty Yel, and S. Tinggi Ilmu Komputer Cipta Karya Informatika Jl Raden, "Optimasi Algoritma Cnn Menggunakan Metode Transfer Learning Untuk Klasifikasi Citra X-Ray Paru-Paru Pneumonia Dan Non-Pneumonia," *J. Tek. Inform. Kaputama*, vol. 6, no. 2, pp. 670–679, 2022.
- [13] I. DLY, J. Jasril, S. Sanjaya, L. Handayani, and F. Yanto, "Klasifikasi Citra Daging Sapi dan Babi Menggunakan CNN Alexnet dan Augmentasi Data," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1176–1185, 2023, doi: 10.47065/josh.v4i4.3702.
- [14] S. Lasniari, J. Jasril, S. Sanjaya, F. Yanto, and M. Affandes, "Klasifikasi Citra Daging Babi dan Daging Sapi Menggunakan Deep Learning Arsitektur ResNet-50 dengan Augmentasi Citra," *J. Sist. Komput. dan Inform.*, vol. 3, no. 4, p. 450, 2022, doi: 10.30865/json.v3i4.4167.
- [15] S. Lasniari, J. Jasril, S. Sanjaya, F. Yanto, and M. Affandes, "Pengaruh Hyperparameter Convolutional Neural Network Arsitektur ResNet-50 Pada Klasifikasi Citra Daging Sapi dan Daging Babi," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 3, pp. 474–481, 2022, doi: 10.32672/jnkti.v5i3.4424.
- [16] D. Putri Ayuni, Jasril, M. Irsyad, F. Yanto, and S. Sanjaya, "Augmentasi Data Pada Implementasi Convolutional Neural Network Arsitektur EfficientNet-B3 Untuk Klasifikasi Penyakit Daun Padi," *Zo. J. Sist. Inf.*, vol. 5, no. 2, pp. 239–249, 2023, doi: 10.31849/zn.v5i2.13874.
- [17] N. H. Putri, Jasril, M. Irsyad, S. Agustian, and F. Yanto, "Klasifikasi Citra Stroke Menggunakan Augmentasi dan Convolutional

- Neural Network EfficientNet-B0,” *J. Media Inform. Budidarma*, vol. 7, no. 2, pp. 650–658, 2023, doi: 10.30865/mib.v7i2.5981.
- [18] B. Y. Phiadelvira, “Klasifikasi kanker serviks berdasarkan citra kolposkopi menggunakan Convolutional Neural Network (CNN) Model Alexnet,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 4, no. 1, pp. 1–103, 2021.
- [19] Y. A. Agus Eko Minarno, “Klasifikasi Citra Kanker Kulit Menggunakan Metode Convolutional Neural Network dengan Arsitektur EfficientNet-B0,” *UMM*, 2022. <https://eprints.umm.ac.id/93049/>
- [20] P. Gupta, S. Garg, and A. Garg, “Convolutional Neural Network For Lung Cancer,” vol. 2, no. 3, pp. 54–59, 2019.
- [21] W. Abdul, “An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network,” *Proc. - Int. Conf. Dev. eSystems Eng. DeSE*, vol. 2020-Decem, pp. 443–446, 2020, doi: 10.1109/DeSE51703.2020.9450778.
- [22] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691–10700, 2019.
- [23] D. A. Anam, L. Novamizanti, and S. Rizal, “Klasifikasi Patologi Makula Pada Retina Berdasarkan Citra Retinal OCT Menggunakan Convolutional Neural Network (Classifying Retinal Pathology Using OCT Retinal Imaging With Convolutional Neural Network),” vol. 8, no. 5, pp. 5064–5071, 2021, [Online]. Available: <https://www.kaggle.com/paultimothymooney/kermany2018>
- [24] S. Supiyanto and T. Suparwati, “Perbaikan Citra Menggunakan Metode Contrast Stretching,” *J. Siger Mat.*, vol. 2, no. 1, pp. 13–18, 2021, doi: 10.23960/jsm.v2i1.2743.
- [25] “Perbandingan Contrast Stretching dan Histogram Equalization Untuk Memperbaiki Kualitas Citra Digital skripsi oleh : Ridho Saragih Program Studi Teknik Informatika STMIK Global Informatika MDP Palembang,” 2020.
- [26] A. Satyo, B. Karno, W. Hastomo, I. Sari, K. Wardhana, and D. Arif, “29 Jenis Penyakit Tanaman Menggunakan Deep Learning EfficientNetB3 Identifikasi,” vol. 2, 2022.