

**Jurnal Teknologi Informasi dan Komunikasi****Vol: 14 No 01 2023****E-ISSN: 2477-3255**

Diterima Redaksi: 03-04-2023 | Revisi: 24-04-2023 | Diterbitkan: 26-05-2023

CNN-RNN Hybrid Model for Diagnosis of COVID-19 on X-Ray Imagery**Novem Berlian Uly¹, Hendry², Ade Iriani³**^{1,2,3}Program Studi Magister Sistem Informasi Fakultas Teknologi Informasi Universitas Kristen Satya Wacana^{1,2,3}Jl. Dr. O. Notohamidjojo No.1-10, Blotongan, Kec. Sidorejo, Kota Salatiga, Jawa Tengah, Indonesiae-mail: ¹972019012@student.uksw.edu, ²hendry@uksw.edu, ³ade.iriani@uksw.edu**Abstract**

This research aims to implement deep learning in determining Covid-19 or normal cases using X-Ray imagery. The method used is CNN (ResNet50) and RNN (LSTM). The research phase begins with data collection, data preprocessing, method modeling, method testing and method evaluation. The data was taken from the kagle.com site with the amount of data used 1.000 images where 500 covid data and 500 normal data, the data is divided into 80% training data, 10% validation data and 10% test data. The results of the evaluation by calculating the ResNet50-LSTM confusion matrix have a value of 95% accuracy, 96% precision, 94% recall and 95% F1-score. At the method testing stage, the researcher got the results of the proposed method experiencing overfitting seen by the comparison of the loss values in the validation data which were not as good as the loss values of the training data. From the results of evaluation and method testing, research can be used as a recommendation in cases of Covid-19 or normal.

Keywords: Deep Learning, ResNet50, LSTM, Image Classification, Covid-19.**Hybrid Model CNN-RNN untuk Diagnosis COVID-19 pada Citra X-Ray****Abstrak**

Penelitian ini bertujuan mengimplementasikan deep learning dalam menentukan kasus Covid-19 atau normal dengan menggunakan citra X-Ray. Metode yang digunakan adalah CNN (ResNet50) dan RNN (LSTM). Tahapan penelitian dimulai dengan pengumpulan data, preprocessing data, pemodelan metode, pengujian metode dan evaluasi metode. Data diambil dari situs kagle.com dengan jumlah data yang dipakai 1.000 citra dimana 500 data covid dan 500 data normal, data dibagi menjadi 80% data latih, 10% data validasi dan 10% data tes. Hasil evaluasi dengan perhitungan confusion matrix ResNet50-LSTM memiliki nilai akurasi 95%, presisi 96%, recall 94% dan F1-score sebesar 95%. Pada tahap pengujian metode, peneliti mendapat hasil metode yang diusulkan mengalami overfitting dilihat dengan perbandingan nilai loss pada data validasi tidak sebesar nilai loss data latih. Dari hasil evaluasi dan pengujian metode, penelitian dapat dijadikan sebagai rekomendasi dalam kasus covid-19 atau normal.

Kata kunci: Deep Learning, ResNet50, LSTM, Klasifikasi Citra, Covid-19

1. Pendahuluan

COVID-19 adalah virus yang menyebar dengan cepat dan mempengaruhi sistem pernapasan manusia. Ini pertama kali muncul di Wuhan, Tiongkok pada akhir tahun 2019 dan secara resmi diumumkan sebagai pandemi oleh WHO pada tahun 2020 [1]. Untuk menekan penyebaran COVID-19, skrining dan reaksi klinis yang cepat pada pasien yang terkontaminasi adalah kebutuhan yang mendesak. Strategi skrining tingkat tertinggi yang digunakan untuk pengujian pasien COVID-19 adalah tes *Reverse Transcription Polymerase Chain Response (RT-PCR)* pada spesimen pernapasan. Prosedur ini merupakan strategi yang paling umum digunakan untuk pengujian identifikasi COVID-19 namun prosedur ini merupakan proses manual dan memakan waktu dengan tingkat positif hanya 63% [2]. Diagnosis COVID-19 lainnya berupa pemeriksaan gejala klinis, riwayat epidemiologi dan citra radiografik positif (*Computed Tomography (CT)/Chest Radiografi (CXR)*) serta uji patogenik positif. Pada interpretasi citra radiografi oleh ahli radiologi terbatas karena kapasitas manusia dalam mendeteksi fitur visual halus pada gambar [3]. Untuk mengatasi hal tersebut para peneliti mengaplikasikan *Deep Learning* untuk menemukan pola pada *rontgen* dada yang biasanya tidak dikenali ahli radiologi dan membantu dokter dalam memutuskan skala penanganan pasien [4]. Oleh karena itu, desain dan penerapan *Deep Learning* untuk klasifikasi citra COVID-19 dalam waktu singkat menjadi kebutuhan mendesak.

Penelitian ini merujuk pada beberapa penelitian yang menggunakan teknik *Deep Learning* untuk mendeteksi Covid-19 melalui gambar medis seperti CT scan, rontgen dada, dan gambar klinis. Penelitian sebelumnya, *Deep Learning* seperti CNN telah digunakan untuk memperoleh fitur-fitur penting dari gambar medis dan mengklasifikasikannya sebagai positif atau negatif Covid-19 dengan tingkat akurasi yang baik. Pada penelitian awal, dilakukan evaluasi terhadap kemungkinan penggunaan gambar CT dari organ dada sebagai alternatif untuk mendeteksi infeksi Covid-19 [5]. Untuk tujuan ini, sebuah model *Convolutional Neural Network (CNN)* dengan arsitektur VGG16 dilatih dengan menggunakan spesifisitas dua *optimizer* yaitu SGD dan Adam. Setelah melalui tahap pelatihan, hasil evaluasi menunjukkan bahwa model yang dibangun mampu mendeteksi infeksi Covid-19 dengan tingkat sebesar 92,86% dan tingkat sensitivitas sebesar 85,71%. Hal ini menunjukkan bahwa penggunaan gambar CT dari organ dada dapat menjadi pilihan yang potensial dalam mendeteksi infeksi Covid-19 menggunakan teknik *Deep Learning* seperti CNN. Penelitian berikutnya dilakukan perbandingan antara performa dan waktu eksekusi antara algoritma *Convolutional Neural Network (CNN)* dan metode *Machine Learning* klasik seperti *Support Vector Machine (SVM)* dan *K-Nearest Neighbors (KNN)* dalam mendeteksi Covid-19 berdasarkan citra X-Ray dada [6]. Setelah melalui uji coba, hasil evaluasi menunjukkan bahwa CNN memberikan performa yang lebih baik dibandingkan dengan SVM dan KNN. Oleh karena itu, peneliti merekomendasikan penggunaan CNN untuk klasifikasi pada tugas mendeteksi Covid-19 menggunakan citra X-Ray dada. Selain itu, waktu eksekusi CNN juga lebih cepat dibandingkan dengan metode SVM dan KNN. Penelitian lain, dilakukan pengukuran performa model *Convolutional Neural Network (CNN)* dalam melakukan klasifikasi citra rontgen untuk menentukan apakah seseorang terpapar COVID-19 atau tidak [7]. Hasil evaluasi menunjukkan bahwa model CNN tersebut mencapai tingkat akurasi sebesar 98,69%, sensitivitas sebesar 97,71%, dan spesifisitas sebesar 98,90%. Hal ini menunjukkan bahwa CNN memiliki performa yang sangat baik dalam melakukan klasifikasi citra rontgen untuk tujuan tersebut. Dengan demikian, penelitian tersebut dapat menyimpulkan bahwa CNN dapat menjadi pilihan yang baik untuk mendukung diagnosis COVID-19 berdasarkan citra *rontgen*. Penelitian selanjutnya, peneliti mengusulkan penggunaan berbagai arsitektur *transfer learning*, seperti VGG19, MobileNetV2, InceptionResNetV2, dan ResNet (ResNet101V2, ResNet152V2, dan ResNet50V2), untuk menganalisis performa model dalam mendeteksi COVID-19 melalui citra medis [8]. Berdasarkan hasil pengujian, model CNN ResNet50V2 memiliki tingkat akurasi yang lebih baik dibandingkan dengan model lainnya, dengan tingkat akurasi sekitar 0,95 (95%). Dengan demikian, model CNN ResNet50V2 dapat digunakan sebagai identifikasi awal untuk mengklasifikasikan pasien sebagai terinfeksi COVID

atau sehat dengan tingkat akurasi yang tinggi. Dalam penelitian selanjutnya, peneliti mengusulkan *Convolutional Neural Network (CNN)* dan *Recurrent Neural Network (RNN)* [9]. Model CNN dengan prapelatihan VGG19, ResNet152V2, dan DenseNet121 masing-masing digabungkan dengan model RNN *Long Short-Term Memory (LSTM)* dan *Gated Recurrent Unit (GRU)*. Metode yang diusulkan dievaluasi menggunakan dataset berjumlah 16.210 gambar X-ray dan CT (5252 gambar COVID-19, 6154 gambar pneumonia, dan 4804 gambar normal) yang diambil dari beberapa *database* dengan ukuran gambar yang berbeda, tingkat kecerahan, dan tampilan sudut. ResNet152V2-LSTM dengan model GRU menjadi arsitektur terbaik dengan akurasi 93,37%, skor F1 93,54%, presisi 93,73%, dan recall 93,47%. Dari hasil percobaan, metode yang diusulkan sangat efektif dalam membedakan penyakit paru-paru. Selain itu, baik gambar CT maupun X-ray dapat digunakan sebagai masukan untuk klasifikasi, memungkinkan deteksi COVID-19 dengan cepat dan mudah.

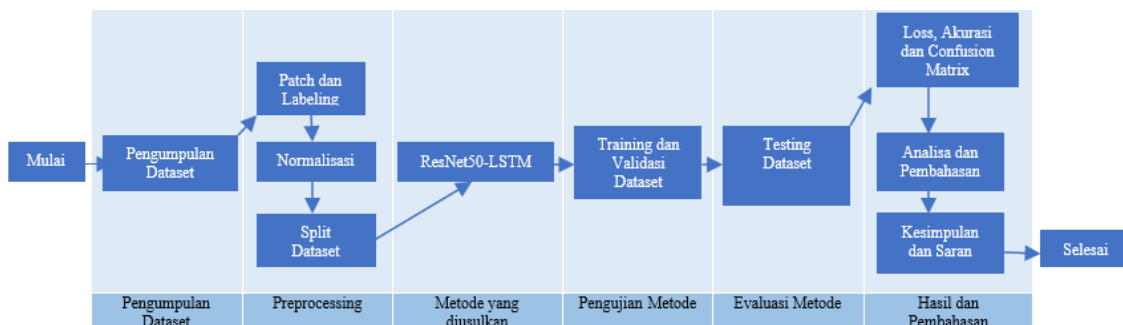
Berdasarkan beberapa penelitian sebelumnya menunjukkan bahwa model *Deep Learning* (CNN dan RNN) yang digunakan dalam dunia medis telah terbukti dapat mengklasifikasi dengan baik apakah seseorang terdampak Covid-19 atau tidak. Hal ini dapat dilihat pada penelitian awal, bahwa penggunaan gambar X-Ray dari organ dada dapat menjadi pilihan yang potensial dalam mendeteksi infeksi Covid-19 dengan mencapai tingkat akurasi lebih dari 90% [5][6][7][8][9] jika dibandingkan dengan proses manual yaitu 63% [2]. Dan jika dilihat dari waktu yang dibutuhkan untuk mendeteksi dengan tingkat akurasi yang tinggi, CNN mampu mendeteksi Covid-19 hanya dalam waktu 3.102 detik [6].

Convolutional Neural Networks (CNN) adalah salah satu teknik *Deep Learning* yang umum digunakan untuk memproses data citra [10]. Metode pengenalan citra menggunakan CNN dapat menggantikan mata manusia karena akurasi tingkat ketajaman dan kontras pada citra mendapatkan hasil yang sangat jelas [11]. CNN bekerja dengan cara meniru jaringan syaraf otak manusia [12]. CNN terdiri dari beberapa jenis lapisan yang bekerja secara berurutan. Pertama, terdapat lapisan konvolusi yang digunakan untuk mengekstraksi fitur-fitur pada citra. Kemudian, lapisan aktivasi ReLU digunakan untuk menghilangkan nilai negatif pada fitur yang diekstraksi. Setelah itu, lapisan *Pooling* digunakan untuk mereduksi ukuran citra dan mempertajam fitur-fitur yang penting. Akhirnya, terdapat lapisan terhubung secara penuh (*Fully Connected Layer*) yang menghubungkan hasil dari lapisan sebelumnya ke *output* [13]. Dengan struktur ini, CNN dapat melakukan klasifikasi citra dengan akurasi yang tinggi [14]. Berbagai macam CNN telah dikembangkan, salah satu CNN yang paling umum digunakan dalam ekstraksi fitur citra adalah ResNet50. Lapisan pada ResNet50 berjumlah 50 lapisan dengan menggunakan *Skip Connection* (pintasan). Alasan utama menggunakan *Skip Connection* adalah untuk menangani masalah CNN yang sebelumnya ketika kedalaman jaringan bertambah maka akurasi akan menurun [11] dengan cara meneruskan aktivasi dari lapisan sebelumnya ke lapisan berikutnya untuk menghafal parameter dengan lebih baik, yang menyebabkan berkurangnya kemungkinan hilangnya gradien. Model *Deep Learning* yang mengadopsi arsitektur jaringan saraf tiruan lainnya adalah *RNN (Recurrent Neural Network)* yang menggunakan teknik pemrosesan ulang untuk mengolah input data sekuensial. Salah satu varian RNN adalah *Long Short-Term Memory (LSTM)* yang merupakan jaringan berulang yang lebih efektif dalam mempelajari pola data dengan memodelkan isi bobot jaringan untuk mempermudah proses klasifikasi.

Peneliti bermaksud memanfaatkan kemampuan ResNet50 dalam mengekstraksi fitur lokal dan LSTM untuk mempelajari ketergantungan jangka panjang. Pertama, ResNet50 digunakan untuk memproses masukan dan mengekstraksi fitur lokal pada citra kemudian keluaran dari ResNet50 yaitu peta fitur menjadi masukan bagi LSTM untuk dipelajari ketergantungan jangka panjang dan mengklasifikasikannya. Dengan memanfaatkan kemampuan kedua model tersebut, maka peneliti bertujuan untuk mengklasifikasikan Covid-19 pada citra X-Ray menggunakan metode gabungan ResNet50 dan LSTM. Hasil dari klasifikasi diharapkan dapat digunakan sebagai acuan dalam pendektasian awal Covid-19 sehingga dapat mengurangi angka kematian karena lambatnya penanganan pasien yang terpapar Covid-19.

2. Metode Penelitian

Penelitian ini dilakukan melalui beberapa tahapan yaitu pengumpulan data, pra-pemrosesan, pelatihan menggunakan metode yang diusulkan, pengujian metode, evaluasi metode, dan tahapan terakhir pembahasan hasil. Untuk lebih jelasnya bisa dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan Data

Penelitian dimulai dengan pengumpulan dataset dengan nama “COVID-19 Radiography Database” yang diambil dari situs kaggle (www.kaggle.com) berupa citra grayscale X-Ray Covid-19 dan Normal dengan ukuran citra 299x299 piksel. Dataset yang dipakai dalam penelitian ini berjumlah total 1.000 citra dimana 500 citra dipakai untuk dataset Covid-19 dan 500 citra dipakai untuk dataset normal. Dataset tersebut disimpan pada *Google Drive* untuk selanjutnya digunakan pada proses klasifikasi citra.

2.2. Pengolahan Data

Pada proses pengolahan data awal dilakukan sebelum pelatihan dan pengujian metode. Peneliti menggunakan *Google Colab* sebagai alat bantu pada penelitian ini. Proses pengolahan data awal ini dimulai dengan mengimport library yang dibutuhkan selama proses klasifikasi citra pada *Google Colab* kemudian dihubungkan dengan *Google Drive* untuk membaca dataset citra Covid-19 dan normal, dataset tersebut dibuatkan *Path* dan *Label* dalam *Data Frame*. Langkah selanjutnya dilakukan eksplorasi data untuk melihat ukuran citra yang akan dilatih, hasil eksplorasi dataset menunjukkan ukuran citra adalah 299x299 piksel. Setelah mengetahui ukuran citra maka peneliti menentukan parameter untuk data latih, dimana parameter tersebut adalah *num_classes* digunakan untuk menentukan jumlah kategori pada citra yaitu dua kategori untuk Covid dan Normal, parameter berikutnya *Batch_Size* dengan jumlah 32 sampel yang akan digunakan untuk memperbaharui parameter model neural network, kemudian *Epochs* berjumlah 15 yang digunakan untuk berapa kali seluruh dataset akan dimasukkan ke model yang dibangun untuk dilatih dan parameter yang terakhir adalah *IMG_Size* 224x224 piksel yang akan dijadikan sebagai patokan ukuran keseluruhan dimensi citra yang akan diklasifikasikan.

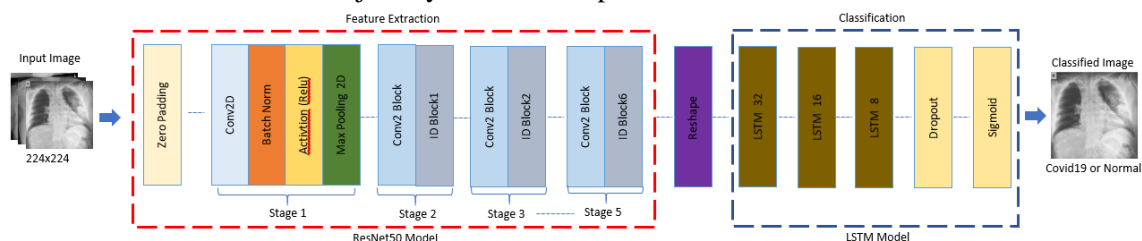
Tahap selanjutnya dilakukan pra-pemrosesan data yaitu dataset perlu di normalisasi dan memiliki ukuran yang sama, gambar yang masih berbentuk *bytes* dalam data frame yang dibuat diawal dibaca kemudian dirubah ke tensor untuk kemudian gambar di normalisasi dengan membagi *img array* dengan 255, kemudian gambar dirubah ukurannya menggunakan parameter *IMG_Size* yang sudah tentukan yaitu 224x224 piksel, langkah terakhir pada pra-pemrosesan gambar yaitu gambar yang masih dalam bentuk grayscale dikonversi ke RGB dengan hasil akhir citra berukuran 224x224 piksel dan memiliki 3 chanel. Setelah pra-pemrosesan gambar selesai selanjutnya dilakukan pra-pemrosesan pada label dengan mengubah nilai label ke integer pada data frame dimana label normal dirubah dengan nilai 0 dan label Covid dengan nilai 1.

Sebelum dataset dilatih, peneliti menentukan fitur dan target dimana fitur merupakan karakteristik dari data citra X-Ray sedangkan target merupakan label atau kelas yang akan diprediksi. Setelah itu, dataset dibagi menjadi tiga bagian, yaitu 80% data latih, 10% data validasi, dan 10% data uji. Untuk keperluan pelatihan data, perlu dibuatkan aliran data atau data pipeline berdasarkan *Batch Size* supaya data dikeluarkan berurutan karena komputer tidak

mampu menampung seluruh data sekaligus apabila data terlalu besar terlebih lagi data berupa gambar.

2.3. Metode yang diusulkan

Pada tahap ini, peneliti menggunakan metode ResNet50 untuk ekstraksi fitur dari citra x-ray dalam dataset yang telah dikumpulkan dan dipreproses sebelumnya. Fitur-fitur ini kemudian digunakan sebagai masukan bagi model LSTM yang bertujuan untuk mengklasifikasikan apakah citra x-ray tersebut menunjukkan kasus Covid-19 atau normal. Dalam hal ini, ResNet50 bertindak sebagai pendeteksi fitur yang handal dan LSTM bertanggung jawab untuk melakukan klasifikasi akhir. Untuk lebih jelasnya bisa dilihat pada Gambar 2.



Gambar 2. Metode yang diusulkan

Pada langkah pertama, setelah tahap *preprocessing* data, inputan citra dengan ukuran 224x224 piksel, model menggunakan teknik *Zero Padding* untuk mengubah ukuran citra input menjadi ukuran matriks input yang sama. Hal ini bertujuan untuk memudahkan proses konvolusi pada layer selanjutnya dan menjaga informasi pada tepian citra. Setelah itu, citra input dikonvolusikan pada *Convolutional Layer* dengan ukuran filter 7x7 dan Stride 2. Konvolusi ini berfungsi untuk mengekstraksi fitur-fitur penting dari citra, seperti tepian, tekstur, dan bentuk. Ukuran filter 7x7 dipilih karena ukuran tersebut dapat menangkap fitur-fitur yang lebih kompleks dalam citra. Selanjutnya, hasil konvolusi dinormalisasi menggunakan *Batch Normalization*. *Batch Normalization* bertujuan untuk menjaga distribusi data pada setiap layer dan mencegah terjadinya *vanishing atau exploding gradient*, sehingga proses *training* menjadi lebih stabil. Hasil normalisasi kemudian masuk ke lapisan aktivasi yang berisi fungsi ReLU. Fungsi ReLU digunakan untuk membuat hasil ekstraksi fitur menjadi non-linear dan mempertahankan nilai positif dari setiap piksel citra. Hal ini membantu meningkatkan kemampuan model dalam mengekstraksi fitur yang lebih kompleks dan beragam dari citra.

Setelah melewati lapisan konvolusi pertama dan normalisasi *batch*, nilai *output* dari fungsi ReLU akan direduksi pada lapisan *Max Pooling* dengan ukuran matriks 3x3. Proses ini berguna untuk mengurangi dimensi dari *feature map* dan juga membantu mengurangi *overfitting* pada model. Selanjutnya, hasil dari *max pooling* akan diteruskan ke tahap konvolusi kedua. Pada tahap konvolusi kedua sampai kelima, proses ekstraksi fitur dilakukan dengan kombinasi *Convolutional Block* dan *Identity Block*. *Convolutional Block* terdiri dari tiga lapisan konvolusi, sedangkan *Identity Block* terdiri dari dua lapisan konvolusi dan satu *shortcut connection* untuk menghindari *vanishing gradient problem*. *Shortcut connection* menghubungkan layer-layer pada kedua blok secara langsung, sehingga memungkinkan informasi dari layer awal lebih mudah mengalir ke layer akhir dan mencegah masalah gradien menghilang saat melatih model. Setelah proses ekstraksi fitur selesai, kemudian *Feature Map* di *Reshape* ukuran matriks menjadi 2x2 dengan menyesuaikan nilai layer sebelumnya yaitu 2048 untuk menjadi masukan bagi LSTM. Pada tahap klasifikasi citra peneliti menggunakan lima gerbang layer yang dimana tiga macam LSTM digunakan yaitu 32, 16 dan 8 neuron hidden layer, gerbang berikutnya layer *Dropout* digunakan untuk mencegah model dari *overfitting* serta *Danse* layer untuk menghubungkan setiap masukan dengan setiap keluaran dengan menggunakan aktivasi sigmoid.

2.4. Pengujian Metode

Pada tahap ini, data yang digunakan yaitu data *training* dan data validasi. Data *training* digunakan untuk melatih model dan data validasi digunakan untuk mengevaluasi performa model, dengan kata lain, model akan diberikan data validasi untuk diprediksi, dan kemudian

hasil prediksi akan dibandingkan dengan label aslinya. Dari perbandingan ini, dapat dihitung berbagai metrik evaluasi performa model seperti akurasi, presisi, *recall*, dan *F1-score*. Metrik-metrik ini digunakan untuk mengevaluasi seberapa baik model dalam mengambil keputusan.

2.5. Evaluasi Metode

Confusion Matrix adalah salah satu metode evaluasi pengukuran kinerja klasifikasi yang umum digunakan. *Confusion Matrix* menggambarkan hasil klasifikasi berdasarkan empat kemungkinan hasil yaitu *True Positive (TP)* dimana jumlah data positif yang diklasifikasikan dengan benar, *True Negative (TN)* dimana jumlah data negatif yang diklasifikasikan dengan benar, *False Positive (FP)* dimana jumlah data negatif yang diklasifikasikan salah sebagai positif, dan *False Negative (FN)* dimana jumlah data positif yang diklasifikasikan salah sebagai negatif. Untuk lebih jelasnya bisa dilihat pada Tabel 1.

Tabel 1. *Confusion Matrix*

	Prediksi Positif	Prediksi Negatif
Aktual Positif	TP	FN
Aktual Negatif	FP	TN

Confusion Matrix sangat berguna untuk mengevaluasi kinerja klasifikasi karena memberikan gambaran yang jelas tentang jumlah data yang diklasifikasikan dengan benar dan salah. Dari *Confusion Matrix*, beberapa metrik dapat dihitung seperti Akurasi (*Accuracy*), Presisi (*Precision*), *Recall* (Sensitivitas), dan *F1-Score*. Kinerja model dapat dibandingkan dengan metode lainnya berdasarkan metrik-metrik ini [15].

Akurasi (*Accuracy*) dapat dihitung dengan membagi jumlah prediksi yang benar (*True Positive* dan *True Negative*) dengan total jumlah prediksi (*True Positive*, *True Negative*, *False Positive*, dan *False Negative*). Hal ini memberikan proporsi prediksi yang benar dalam klasifikasi [16].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

Precision adalah metrik evaluasi yang menghitung seberapa akurat model dalam memprediksi jumlah positif yang benar dari semua prediksi positif yang dilakukan. Dengan kata lain, *precision* mengukur seberapa baik model dalam membatasi jumlah *False Positive* (kasus yang sebenarnya negatif namun diprediksi sebagai positif) dalam prediksinya. Secara matematis, *precision* dihitung dengan membagi jumlah *True Positive* (TP) dengan jumlah *True Positive* dan *False Positive* (FP). Rumus untuk menghitung *Precision* adalah sebagai berikut:

$$precision = \frac{TP}{TP + FP} \quad (2)$$

Recall merupakan metrik evaluasi yang mengukur kemampuan model dalam mengklasifikasikan semua data positif secara benar. Dalam kata lain, *recall* menghitung proporsi data positif yang benar diklasifikasikan oleh model dari seluruh data positif yang sebenarnya. Semakin tinggi nilai *recall*, semakin baik kemampuan model dalam mengidentifikasi semua data positif. *Recall* dapat dihitung dengan rumus:

$$recall = \frac{TP}{TP + FN} \quad (3)$$

F1-Score adalah gabungan antara *Precision* dan *Recall* yang digunakan untuk mengukur kinerja model klasifikasi. *F1-Score* dapat memberikan informasi yang lebih lengkap daripada akurasi karena dapat memberikan hasil yang baik bahkan pada dataset yang tidak seimbang. *F1-score* mengambil rata-rata harmonis antara *precision* dan *recall*, sehingga mempertimbangkan kekurangan dalam kinerja klasifikasi yang disebabkan oleh ketidakseimbangan antara *True Positive* dan *False Negative*. *F1-score* dapat dihitung dengan rumus:

$$F1 - score = 2 \frac{(recall * precision)}{(recall + precision)} \quad (4)$$

3. Hasil dan Pembahasan

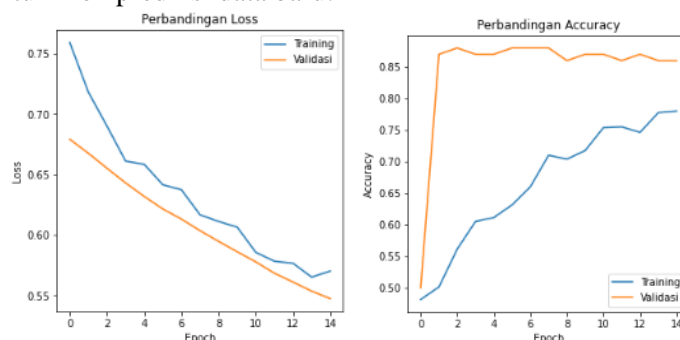
Pada bagian ini peneliti melakukan beberapa skenario untuk membandingkan hasil eksperimen dengan metode lain yang diterapkan pada dataset yang sama untuk mengevaluasi seberapa baik model yang peneliti usulkan. Skenario pertama menggunakan metode ResNet50, kedua LSTM dan ketiga menggunakan metode gabungan ResNet50-LSTM. Peneliti menggunakan nilai *Accuracy* tertinggi dan *Loss* terendah pada data validasi untuk mengevaluasi model yang dibangun.

3.1. Pengaturan Model

Peneliti menggunakan model ResNet50-LSTM sebagai model yang diusulkan dengan menggunakan *Google Colab* sebagai alat untuk melakukan eksperimen pada penelitian ini. Pada semua model peneliti menggunakan rasio pembagian data 80:10:10 dimana 80% data training, 10% data validasi, dan 10% data *test*, *optimizer* menggunakan Adam dengan *Learning Rate* adalah 0,001, aktivasi yang digunakan adalah *Sigmoid* dengan jumlah 15 *Epochs* dan *Batch Size* 32.

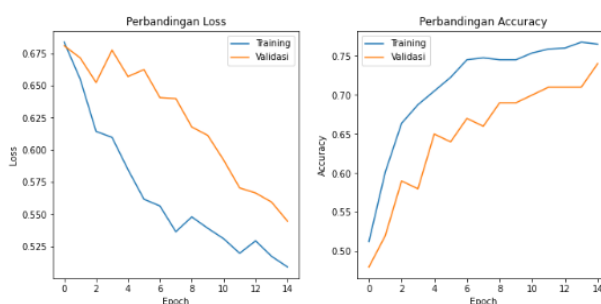
3.2. Pengujian Model

Pada tahapan ini dilakukan pengujian model menggunakan data validasi pada data latih disetiap model yang telah dibuat. Tujuannya adalah untuk memeriksa kinerja model pada data yang belum pernah dilihat sebelumnya dan mengukur seberapa baik model dapat generalisasi pada data baru. Pengujian dilakukan dengan membandingkan nilai *loss* terendah dan akurasi tertinggi pada data latih dan data validasi. Jika nilai *loss* pada data latih sangat rendah tetapi pada data validasi tinggi, maka model mungkin mengalami *overfitting*, yaitu model terlalu fokus pada data latih dan tidak dapat generalisasi pada data baru. Sebaliknya, jika nilai *loss* pada data latih dan data validasi cukup rendah dan akurasi cukup tinggi, maka model dianggap baik dan dapat digunakan untuk memprediksi data baru.



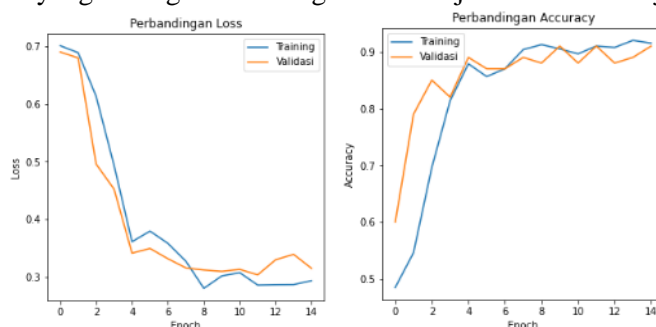
Gambar 3. Perbandingan *Loss* dan *Accuration* ResNet50

Berdasarkan hasil pada Gambar 3. Bisa dilihat bahwa *Training Loss* dan *Validation Loss* mengalami penurunan *loss* tidak mengalami *overfitting* dan pada perbandingan akurasi terlihat bahwa akurasi pada data latih meningkat secara bertahap sedangkan akurasi pada data validasi pada *epoch* kedua sampai *epoch* terakhir prediksi yang dihasilkan sangat bagus dibandingkan dengan hasil akurasi data latih dan grafik menunjukkan tidak mengalami *overfitting*



Gambar 4. Perbandingan *Loss* dan *Accuration* LSTM

Berdasarkan hasil pada Gambar 4. Bisa dilihat bahwa *Training Loss* dan *Validation Loss* keduanya mengalami penurunan *loss* namun grafik menunjukkan model mengalami *overfitting* karena model terlalu focus pada data latih dibandingkan dengan data validasi sehingga data validasi mengalami *loss* yang tinggi dibandingkan dengan data latih, begitu juga pada perbandingan akurasi terlihat bahwa akurasi pada data latih meningkat lebih besar dibandingkan dengan data validasi sehingga mengalami *overfitting*. Hasil kedua perbandingan tersebut model LSTM yang dibangun tidak bagus untuk dijadikan model rujukan



Gambar 5. Perbandingan *Loss* dan *Accuration* ResNet50-LSTM

Berdasarkan hasil pada Gambar 5. Bisa dilihat bahwa *Training Loss* dan *Validation Loss* mengalami penurunan dengan nilai perbandingan *Loss* yang tidak jauh berbeda dan masih pada tahap toleransi dan grafik menunjukkan model mengalami *Overfitting*. Hal ini dikarenakan model terlalu fokus mempelajari data latih dan saat menerima data baru dalam hal ini data validasi maka hasilnya tidak sebagus data latih walaupun perbedaannya sedikit. Sedangkan pada perbandingan akurasi terlihat bahwa akurasi pada data latih dan data validasi meningkat dengan nilai akurasi pada data latih dan data validasi juga tidak berbeda terlalu jauh, hal ini menunjukkan perbandingan akurasi tidak mengalami *Overfitting*. Dengan hasil kedua perbandingan tersebut bisa dikatakan model ResNet50-LSTM yang dibangun sangat bagus untuk dijadikan model rujukan dan tidak mengalami *Overfitting*.

Untuk melihat model dengan akurasi yang paling baik, maka peneliti melihat hasil nilai akurasi validasi tertinggi dengan membandingkan nilai *loss* yang dihasilkan pada data latih dan data validasi, untuk melihat lebih jelas apakah model yang dibangun baik untuk menjadi rujukan bisa dilihat pada Tabel 2 berikut:

Tabel 2. Perbandingan *Loss* dan *Accuration*

Model	Training		Validasi	
	<i>Loss</i>	Acc	<i>Loss</i>	Acc
ResNet50	0.5702	0.7800	0.5473	0.8600
LSTM	0.5090	0.7650	0.5446	0.7400
ResNet50-LSTM	0.2938	0.9150	0.3150	0.9100

Dari hasil pengujian yang dilakukan, model ResNet50 memiliki *loss* dan akurasi yang lebih baik pada data *training* dibandingkan model LSTM, namun pada data validasi, ResNet50 memiliki *loss* yang lebih tinggi namun akurasi yang lebih baik dibandingkan LSTM. Model LSTM memiliki *loss* dan akurasi yang lebih buruk pada data training dibandingkan dengan ResNet50, namun memiliki hasil yang sedikit lebih baik pada data validasi dalam hal *loss*. Namun, model ini memiliki akurasi yang lebih rendah pada data validasi. Model ResNet50 merupakan model yang paling baik karena tidak mengalami *overfitting* dibandingkan dengan model lainnya. Model ResNet50-LSTM merupakan gabungan dari kedua model sebelumnya. Hasilnya menunjukkan bahwa model ini memiliki *loss* dan akurasi yang lebih baik pada data training dan validasi dibandingkan dengan model LSTM dan ResNet50. Meskipun demikian jika dilihat dari tabel 2 maka model ResNet50-LSTM mengalami *overfitting* dengan perbedaan nilai *loss* pada data latih 0.2938 naik menjadi 0.3150 saat memvalidasi data baru, dan untuk nilai akurasi pada data latih dari 0.9150 turun menjadi 0.9100 saat memvalidasi data baru.

3.3. Evaluasi Model

Pada tahap evaluasi model, peneliti menggunakan data *testing* yang terdiri dari 100 data yang akan diprediksi dan diuji, dengan 50 diantaranya aktual terinfeksi Covid-19 dan 50 lainnya aktual tidak terinfeksi (Normal). Peneliti juga menggunakan *Confusion Matrix* untuk melakukan pengukuran performa model untuk memperoleh nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* pada masing-masing model.

Tabel 3. *Confusion Matrix* ResNet50

	Prediksi Covid-19	Prediksi Normal
Aktual Covid-19	47	3
Aktual Normal	4	46

Pada Tabel 3 terlihat bahwa dari 50 data yang sebenarnya terinfeksi Covid-19, model memprediksi 47 data tersebut dengan benar (*True Positive*), sedangkan 3 data terinfeksi Covid-19 salah diprediksi sebagai Normal (*False Negative*), sedangkan dari 50 data sebenarnya Normal, model memprediksi 46 data tersebut dengan benar (*True Negative*), sementara 4 data Normal salah diprediksi sebagai Covid-19 (*False Positive*).

Tabel 4. *Confusion Matrix* LSTM

	Prediksi Covid-19	Prediksi Normal
Aktual Covid-19	42	8
Aktual Normal	8	42

Pada Tabel 4 terlihat bahwa dari 50 data yang sebenarnya terinfeksi Covid-19, model memprediksi 42 data tersebut dengan benar (*True Positive*), sedangkan 8 data terinfeksi Covid-19 salah diprediksi sebagai Normal (*False Negative*), sedangkan dari 50 data sebenarnya Normal, model memprediksi 42 data tersebut dengan benar (*True Negative*), sementara 8 data Normal salah diprediksi sebagai Covid-19 (*False Positive*).

Tabel 5. *Confusion Matrix* ResNet50-LSTM

	Prediksi Covid-19	Prediksi Normal
Aktual Covid-19	47	3
Aktual Normal	2	48

Pada Tabel 5 terlihat bahwa dari 50 data yang sebenarnya terinfeksi Covid-19, model memprediksi 47 data tersebut dengan benar (*True Positive*), sedangkan 3 data terinfeksi Covid-19 salah diprediksi sebagai Normal (*False Negative*), sedangkan dari 50 data sebenarnya Normal, model memprediksi 48 data tersebut dengan benar (*True Negative*), sementara 2 data Normal salah diprediksi sebagai Covid-19 (*False Positive*).

Dari semua hasil *Confusion Matrix* yang menunjukkan hasil prediksi benar paling tinggi dengan tingkat kesalahan prediksi rendah adalah metode gabungan ResNet50-LSTM. Hal ini menunjukkan bahwa model yang diusulkan sangat baik untuk dijadikan acuan. Selanjutnya peneliti melakukan perbandingan performa model berdasarkan *Confusion Matrix* pada masing-masing model untuk melihat *Accuracy*, *Precision*, *Recall* dan *F1-Score*. Berikut hasil perhitungan perbandingan performa model yang dihitung menggunakan formula akurasi (1), presisi (2), *Recall* (3) dan *F1-Score* (4).

Tabel 6. Perbandingan Performa Model

Performa Model	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
ResNet50	0.93	0.92	0.94	0.93
LSTM	0.84	0.84	0.84	0.84
ResNet50-LSTM	0.95	0.96	0.94	0.95

Pada Tabel 6 terlihat bahwa Model ResNet50-LSTM memiliki akurasi tertinggi yaitu 0.95, artinya model tersebut mampu memprediksi dengan benar 95% dari keseluruhan data. Model ResNet50-LSTM juga memiliki nilai *precision* dan *F1-Score* tertinggi yaitu 0.96 dan 0.95 secara berturut-turut. Hal ini menunjukkan bahwa model tersebut memiliki kemampuan yang baik dalam menghindari *False Positive* dan *False Negative*, sedangkan model LSTM memiliki nilai performa yang lebih rendah dibandingkan dengan dua model lainnya, dengan nilai akurasi, *precision*, *recall*, dan *F1-score* 0.84. Dengan hasil perbandingan performa model

maka model ResNet50-LSTM adalah model terbaik karena memiliki nilai performa yang paling tinggi dalam memprediksi kedua kelas dengan baik. Ini menunjukkan bahwa model yang diusulkan pada penelitian ini yaitu ResNet50-LSTM memiliki performa yang baik dan layak dijadikan salah satu acuan dalam mendeteksi Covid-19.

Jika dibandingkan dengan penelitian sebelumnya yang menggunakan metode hampir sama yaitu ResNet152V2+GRU dengan akurasi dari 93,37% naik menjadi 95%, F1-skor 93,54% naik menjadi 95%, presisi 93,73% naik menjadi 96%, dan recall 93,47% naik menjadi 94%. maka model yang diusulkan ResNet50-LSTM pada penelitian ini masih unggul dengan perbedaan yang cukup tinggi, akan tetapi jika dibandingkan dengan penelitian yang lainnya, maka penelitian ini masih harus ditingkatkan. Contohnya model VGG16 mencapai presisi 100% paling tinggi dibandingkan penelitian yang lain walaupun dataset yang digunakan paling sedikit. Sedangkan untuk tingkat akurasi dan *recall* tertinggi dihasilkan dari model CNN dengan nilai 98% dengan jumlah data yang digunakan tidak seimbang antara data Covid-19, Normal dan Viral Pneumonia. Sedangkan untuk nilai F1-Score paling tinggi dihasilkan dari model ResNet50V2 dengan nilai 96%.

Tabel 7. Perbandingan Model Usulan dengan Penelitian Sebelumnya

<i>Author</i>	<i>Dataset</i>	<i>Method</i>	<i>ACC</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
W. Swastika	140 X-Ray (Binary)	VGG16	93%	100%	86%	-
Naufal, M.F., et.al	3.886 X-Rays (Multi-class)	KNN, SVM, dan CNN	95%	95%	95%	95%
Bambang Pulu Hartato	2.905 X-Rays (Multi-class)	CNN	97%	97%	98%	-
Mawaddah Harahap., et.al	550 X-Ray (Binary)	ResNet50V2, InceptionResNetV2, VGG19, dan MobileNetV2	95%	96%	97%	96%
Kanjanasurat., et.al	9.271 X-rays and 6939 CT (Multi-class)	ResNet152V2 + LSTM + GRU	93%	94%	93%	94%
Proposed method	1.000 X-Ray (Binary)	ResNet50+LSTM	95%	96%	94%	95%

4. Kesimpulan

Dengan memanfaatkan kemampuan ResNet50 dalam mengekstraksi fitur citra dan LSTM dalam mempelajari ketergantungan jangka panjang serta mengklasifikasikan citra, maka metode gabungan ResNet50-LSTM dapat melakukan klasifikasi citra X-Ray sangat baik dengan akurasi yang mencapai 95%. Model yang dibuat diharapkan dapat menjadi salah satu acuan dalam mendeteksi awal kasus Covid-19 sehingga dapat menentukan pasien Covid-19 atau tidak. Namun jika dilihat pada tabel 2 hasil perbandingan *Loss* dan akurasi terjadi *Overfitting* pada model yang dibuat dikarenakan model terlalu fokus pada pembelajaran data latih sehingga pada saat ada data baru, hasilnya tidak sebagus data latih sehingga perbandingan nilai *loss* pada data validasi tidak sebagus nilai *loss* data latih. Perbedaan nilai *loss* pada data latih 0.2938 naik menjadi 0.3150 saat memvalidasi data baru, dan untuk nilai akurasi pada data latih dari 0.9150 turun menjadi 0.9100 saat memvalidasi data baru.

Selain itu, penelitian ini memiliki keterbatasan dimana model perlu dilakukan uji coba lapangan lebih lanjut agar dapat mengetahui akurasi dari model yang dibangun, untuk itu perlu melakukan kerjasama dengan dokter spesialis dan pihak rumah sakit. Penelitian ini juga masih banyak ruang yang bisa dikembangkan mulai dari pengembangan akurasi model, pengembangan pembacaan data citra seperti menambahkan beberapa layer proses ekstraksi citra, mengabungkan dengan metode lainnya.

Daftar Pustaka

- [1] C. R. Dennison Himmelfarb and D. Baptiste, "Coronavirus Disease (COVID-19)," *J. Cardiovasc. Nurs.*, vol. 35, no. 4, pp. 318–321, Jul. 2020.
- [2] D. Wang *et al.*, "Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China," *JAMA - J. Am. Med. Assoc.*, vol. 323, no. 11, pp. 1061–1069, 2020.
- [3] M. Heidari, S. Mirniaharikandehi, A. Z. Khuzani, G. Danala, Y. Qiu, and B. Zheng, "Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms," *Int. J. Med. Inform.*, vol. 144, no. September, p. 104284, 2020.
- [4] H. Y. F. Wong *et al.*, "Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19," *Radiology*, vol. 296, no. 2, pp. E72–E78, Aug. 2020.
- [5] W. Swastika, "Studi Awal Deteksi Covid-19 Menggunakan Citra Ct Berbasis Deep Learning Preliminary Study Of Covid-19 Detection Using Ct Image Based On Deep Learning," vol. 7, no. 3, pp. 629–634, 2020.
- [6] M. F. Naufal *et al.*, "Analisis Perbandingan Algoritma Klasifikasi Citra Chest X-ray Untuk Deteksi Covid-19," *Teknika*, vol. 10, no. 2, pp. 96–103, Jun. 2021.
- [7] Bambang Pilu Hartato, "Penerapan Convolutional Neural Network pada Citra Rontgen Paru-Paru untuk Deteksi SARS-CoV-2," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 747–759, Aug. 2021.
- [8] M. Harahap, Em Manuel Laia, Lilis Suryani Sitanggang, Melda Sinaga, Daniel Franci Sihombing, and Amir Mahmud Husein, "Deteksi Penyakit Covid-19 Pada Citra X-Ray Dengan Pendekatan Convolutional Neural Network (CNN)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 70–77, Feb. 2022.
- [9] I. Kanjanasurat, K. Tenghongsakul, B. Purahong, and A. Lasakul, "CNN–RNN Network Integration for the Diagnosis of COVID-19 Using Chest X-ray and CT Images," *Sensors*, vol. 23, no. 3, pp. 1–12, 2023.
- [10] A. Satyo, B. Karno, W. Hastomo, Y. Efendi, D. Diyah, and R. Irawati, "Arsitektur Alexnet Convolution Neural Network (CNN) Untuk Mendeteksi Covid-19 Image Chest-Xray" Konferensi Nasional Ilmu Komputer (KONIK), vol. 5, no. 1, pp. 482–485, 2021.
- [11] Y. S. HARIYANI, S. HADIYOSO, and T. S. SIADARI, "Deteksi Penyakit Covid-19 Berdasarkan Citra X-Ray Menggunakan Deep Residual Network," *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 8, no. 2, p. 443, May 2020.
- [12] A. B. Sinuhaji, A. Gautama Putrada, and H. H. Nuha, "Klasifikasi Gambar dari Prototipe Camera Trap Menggunakan Model ResNet-50 untuk Mendeteksi Satwa Dilindungi." e-Proceeding of Engineering, Vol.8, No.5 Oktober 2021
- [13] F. Nashrullah, S. A. Wibowo, and D. G. Budiman, "COMPLETE Journal of Computer, Electronic, and Telecommunication Investigasi Parameter Epoch Pada Arsitektur ResNet-50 Untuk Klasifikasi Pornografi," *J. Comput. Electron. Telecommun.*, vol. 1, no. 1, 2020.
- [14] D. Irfansyah *et al.*, "Arsitektur Convolutional Neural Network (CNN) Alexnet Untuk Klasifikasi Hama Pada Citra Daun Tanaman Kopi," *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, Vol.6, No.2, Mei 2021.
- [15] M. T. D. Putra, M. I. Ardiansyah, and D. Aprianti, "Deteksi Konten Pornografi Menggunakan Convolutional Neural Network Untuk Melindungi Anak Dari Bahaya Pornografi," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 4, p. 2401, Oct. 2022.
- [16] L. Aji Andika and P. Amalia Nur Azizah, "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, No. 1, May 2019.