

Contents list available at [www.jurnal.unimed.ac.id](http://www.jurnal.unimed.ac.id)

## CESS (Journal of Computing Engineering, System and Science)

journal homepage: <https://jurnal.unimed.ac.id/2012/index.php/cess>



### Deteksi SARS-CoV-2 Dari Citra CT-Scan Paru Menggunakan Konsep Fine Tuning Pada Model Deep-CNN Terlatih

#### *SARS-CoV-2 Detection from Lung CT-Scan Images Using Fine Tuning Concept on Deep-CNN Pretrained Model*

Simeon Yuda Prasetyo

<sup>1</sup>School of Computer Science, Universitas Bina Nusantara

Jakarta Barat, DKI Jakarta, Indonesia 11480

email: [1simeon.prasetyo@binus.ac.id](mailto:1simeon.prasetyo@binus.ac.id)

#### ABSTRAK

Permasalahan COVID-19 (SARS-CoV-2) yang menyebar dengan cepat dan mendunia. Deteksi dan diagnosis penyakit ini sangat penting untuk memastikan terapi yang tepat sehingga perlu dilaksanakan melalui berbagai pendekatan yang praktis. Pada penelitian sebelumnya, pendeteksian penyakit melalui pencitraan medis sudah mulai banyak bermunculan dan mendapatkan akurasi yang cukup baik sekitar 80 hingga 90 persen menggunakan machine learning. Pada era *deep learning*, beberapa eksperimen mendapatkan akurasi yang lebih baik mencapai rata-rata akurasi 95% menggunakan pendekatan *deep learning* biasa. Kini, *deep learning* telah berkembang lebih pesat, secara khusus untuk klasifikasi citra. Dengan demikian perlu eksperimen dengan pendekatan model terlatih terhadap citra medis. Selain itu, pendekatan *fine tuning* juga akan menjadi aspek pendekatan yang akan dilakukan dalam eksperimen ini untuk dibandingkan dan untuk mengetahui efeknya, secara khusus pada citra CT-Scan paru untuk pendeteksian COVID 19. Hasil dari eksperimen ini mendapati bahwa pendekatan *Pretrained* model bisa mendapatkan akurasi yang cukup tinggi, bahkan akurasi testing terendah dalam eksperimen ini sudah mencapai 94,78% dari arsitektur Xception tanpa *fine tuning*, hal ini tentunya sudah mengalahkan pendekatan *machine learning* pada eksperimen terdahulu yang tidak mencapai akurasi 90%. Pada akurasi terbaik mendapatkan akurasi sebesar 97,59% menggunakan arsitektur VGG 16 dengan menerapkan *fine tuning*. Hasil eksperimen ini sekaligus menunjukkan bahwa tahapan *fine tuning* (pada 10 layer paling atas) bisa meningkatkan akurasi model yang dibangun.

**Kata Kunci:** deteksi SARS-CoV-2, citra CT-Scan paru, model deep-CNN terlatih, fine tuning

## ABSTRACT

The problem of the spread of COVID-19 (SARS-CoV-2) is spreading fleetly and worldwide. Beforehand discovery and opinion of complaint is veritably important to ensure the right remedy so that it needs to be enforced through various practical approaches. In former studies, complaint discovery through medical imaging has started to appear and get a good delicacy of around 80 to 90 percent using machine learning. In the deep learning era, some trials get better accuracy of 95 percent using the traditional deep learning approach. Now, deep learning has developed more fleetly, especially for image classification. therefore, it's necessary to experiment with a pretrained model approach to medical images. In addition, the fine tuning approach will also be an aspect of the approach that will be carried out in this trial to be compared and to find out its effect, specifically on CT-Scan images of the lungs for the discovery of COVID 19. The results of this experiment showed that the pretrained model approach can get high accuracy. Relatively high accuracy, the smallest testing accuracy in this trial reached 94.78 percent of the Xception without fine tuning phase, this result has beaten the machine learning approach which is didn't reach 90 percent of accuracy. The best experiment testing accuracy get 97.59 percent on the VGG 16 by applying fine tuning. The results of this trial also show that the fine tuning stage (for the top 10th layers) can increase the accuracy of the model.

**Keywords:** *SARS-CoV-2 detection, lung CT-Scan imaging, deep-CNN pretrained model, fine tuning*

---

## 1. PENDAHULUAN

Pada 31 Desember 2019, Kota Wuhan, Provinsi Hubei, China melaporkan 27 kasus pneumonia yang tidak diketahui etiologinya. Dengan populasi lebih dari 11 juta, Wuhan adalah kota terpadat di Cina tengah. Gejala klinis yang paling menonjol dari pasien ini adalah batuk kering, dispnea (sesak napas), demam, dan infiltrat paru bilateral pada pencitraan. Semua insiden itu terkait dengan Pasar Grosir Makanan Laut Huanan di Wuhan, yang menjual ikan serta sejumlah hewan hidup seperti ayam, kelelawar, marmut, dan ular. Pada 7 Januari 2020, Pusat Pengendalian dan Pencegahan Penyakit China (CCDC) mengidentifikasi pelakunya dari sampel swap tenggorokan, yang diberi nama *Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)*. WHO (*World Health Organization*) menetapkan penyakit itu sebagai COVID-19 [1].

Untuk mempelajari etiologi COVID-19, para peneliti dari beberapa bidang berkolaborasi dengan profesional kesehatan masyarakat. Mereka juga segera membangun metode dengan pembuat kebijakan untuk menghentikan penyebaran penyakit baru ini. Penelitian terbaru tentang pola pencitraan pada computed tomography (CT) untuk pasien COVID-19 telah mencatat kekeruhan paru bilateral pada 40 dari 41 (98 persen) CT dada pada pasien yang terinfeksi di Wuhan, dengan fokus konsolidasi lobular dan subsegmental menjadi hasil yang paling umum. Pasien yang diduga terinfeksi COVID-19 sering dievaluasi menggunakan radiografi toraks. Deteksi dan diagnosis penyakit dini sangat penting untuk memastikan terapi yang tepat. Dari sudut pandang kesehatan masyarakat, isolasi pasien segera sangat penting untuk mengendalikan penyakit menular ini dan memanfaatkan sumber daya yang ada, yang dengan cepat habis dan kewalahan oleh jumlah pasien yang meningkat secara geometris dan waktu perawatan yang berlarut-larut [2].

Pada penelitian sebelumnya terkait pendeteksian penyakit melalui pencitraan paru, salah satunya pengembangan model *machine learning* untuk pendeteksian penyakit paru melalui pencitraan. Sejumlah penelitian, *machine learning* telah memiliki kemampuan yang baik dalam proses klasifikasi, termasuk klasifikasi citra. Tej Bahadur Chandra dan Keshri Verma telah membuktikan kemampuan model *machine learning* untuk mendeteksi gambar rontgen dada dengan membandingkan beberapa model klasifikasi *machine learning* seperti Regresi Logistik, *Multilayer Perceptron*, *Random Forest*, dan *Sequential Minimal Optimization*. Penelitian ini juga menggunakan ekstraksi ciri sebelum klasifikasi. Hasil optimal dari penelitian ini adalah menggunakan Regresi Logistik dengan akurasi 95,53%. Meski memiliki akurasi yang baik, model *machine learning* memiliki kekurangan dalam mengelola data dalam jumlah besar. Hasil optimal yang diperoleh dari penelitian ini menggunakan data dalam jumlah terbatas dengan 412 data citra hasil rontgen dada [3].

Pada penelitian lain yang dilakukan escobar,et.al melakukan deteksi Analisis citra ultrasound (USG) paru dengan bantuan komputer menunjukkan potensi besar untuk mendiagnosis kondisi paru, menjadi alternatif yang layak untuk skrining dan mendiagnosis COVID-19. Dengan mengadaptasi berbagai arsitektur *deep learning* yang telah dilatih sebelumnya (pretrained model), termasuk VGG19, InceptionV3, Xception, dan ResNet50. Pada penelitian ini, arsitektur InceptionV3 mencapai akurasi rata-rata terbaik sebesar 89,1%, dengan ROC (97,1%) untuk deteksi COVID-19 [4].

Penelitian lainnya memanfaatkan teknik *deep learning* yang didasarkan pada arsitektur convolutional neural networks (CNN). Naskah pada penelitian ini berfokus pada membedakan gambar CT scan COVID-19 dan non-COVID 19 CT menggunakan teknik *deep learning* yang berbeda. Model yang dikembangkan sendiri dinamakan CTnet-10 dirancang untuk diagnosis COVID-19, dengan akurasi 82,1%. Selain itu, model lain yang diuji dalam penelitian ini adalah DenseNet-169, VGG-16, ResNet-50, InceptionV3, dan VGG-19. VGG-19 terbukti lebih unggul dengan akurasi 94,52% dibandingkan dengan semua model pembelajaran mendalam lainnya. Diagnosis otomatis COVID-19 dari gambar CT scan dapat digunakan oleh dokter sebagai metode cepat dan efisien untuk skrining COVID-19 [5].

Pada penelitian lainnya, membahas tentang peran teknik *machine learning* untuk mendapatkan hal penting seperti apakah pemindaian computed tomography (CT) paru-paru harus menjadi tes skrining/alternatif pertama untuk reaksi berantai transkriptase-polimerase balik waktu nyata (RT-PCR), adalah COVID-19 pneumonia berbeda dari pneumonia virus lainnya dan jika ya bagaimana membedakannya menggunakan gambar CT scan paru dari data CT scan paru pasien yang terinfeksi COVID-19 dari berbagai rumah sakit di Italia, Cina, Moskow, dan India yang dipilih. Untuk pelatihan dan pengujian sistem yang diusulkan, dalam penelitian ini menggunakan perangkat *computer vision* khusus yakni Microsoft Azure. Akurasi terbaik dari keseluruhan eksperimen mencapai hampir 91% untuk klasifikasi COVID-19 menggunakan metode yang diusulkan [6].

Dalam penelitian sejenis lainnya, peneliti melakukan melakukan identifikasi pada model Convolutional Neural Network (CNN) yang sesuai melalui studi komparatif. Pemilihan model awal dari beberapa model CNN populer. Kemudian pada penelitian ini dilakukan mengoptimalkan model VGG19 yang dipilih untuk modalitas gambar untuk menunjukkan bagaimana model dapat digunakan untuk kumpulan data COVID-19 yang sangat langka dan menantang. Kami juga mengusulkan tahap pra-pemrosesan gambar untuk membuat kumpulan data gambar yang dapat dipercaya untuk mengembangkan dan menguji model pembelajaran mendalam. Pendekatan baru ini bertujuan untuk mengurangi noise yang tidak

diinginkan dari gambar sehingga model *deep learning* dapat fokus pada pendeteksian penyakit dengan fitur spesifik darinya. Hasil kami menunjukkan bahwa gambar USG memberikan akurasi deteksi yang unggul dibandingkan dengan X-Ray dan CT scan. Hasil eksperimen menyoroti bahwa dengan data yang terbatas, sebagian besar jaringan yang lebih dalam berjuang untuk berlatih dengan baik dan memberikan konsistensi yang lebih sedikit pada tiga mode pencitraan yang digunakan. Model VGG19 yang dipilih, yang kemudian disetel secara ekstensif dengan parameter yang sesuai, bekerja dalam tingkat deteksi COVID-19 yang cukup besar terhadap pneumonia atau normal untuk ketiga mode gambar paru-paru dengan presisi hingga 86% untuk X-Ray, 100% untuk Ultrasound dan 84% untuk CT scan [7].

Penelitian yang dilakukan oleh Zebari, et.al., menyajikan skema fusi model berdasarkan sistem machine learning menggunakan tiga fitur tekstur yang signifikan, yaitu, Local Binary Pattern (LBP), Fractal Dimension (FD), dan Gray-Level Co-occurrence matrix (GLCM). Dalam hasil eksperimen, untuk menunjukkan efisiensi skema yang diusulkan, eksperimen ini menggunakan 300 gambar CT scan dari database yang tersedia untuk umum. Hasil eksperimen menunjukkan kinerja LBP, FD, dan GLCM memperoleh akurasi masing-masing sebesar 89,87%, 87,84%, dan 90,98%, sedangkan skema yang diusulkan memberikan hasil yang lebih baik dengan mencapai akurasi 96,91% [8].

Serangkaian algoritma pembelajaran mendalam, dilatih dalam dari 1280 pasien untuk melokalisasi pleura parietal/parenkim paru diikuti dengan klasifikasi pneumonia COVID-19, dapat mencapai akurasi hingga 90,8%, dengan sensitivitas 84% dan spesifisitas 93%, sebagai dievaluasi dalam set tes independen (tidak termasuk dalam pelatihan dan validasi) dari 1337 pasien. Kontrol normal termasuk CT dada dari onkologi, darurat, dan indikasi terkait pneumonia. Tingkat positif palsu pada 140 pasien dengan pneumonia lain yang dikonfirmasi laboratorium (non COVID-19) adalah 10%. Algoritme berbasis AI dapat dengan mudah mengidentifikasi CT scan dengan pneumonia terkait COVID-19, serta membedakan pneumonia non-COVID dengan spesifisitas tinggi pada populasi pasien yang beragam [9].

Berdasarkan sejumlah penelitian yang telah ada sebelumnya cukup banyak yang masih menggunakan metode pengenalan citra konvensional dengan pendekatan *machine learning*. Selain itu, terdapat pula penelitian yang memanfaatkan metode terbaru *deep learning* yang memanfaatkan model yang telah dilatih sebelumnya (*pretrained model*) seperti VGG, ResNet dan Inception, namun pada penelitian tersebut tidak menjelaskan bagaimana peneliti sebelumnya memanfaatkan *pretrained model* baik memanfaatkan *Feature Based Transfer Learning* atau hingga melaksanakan tahapan *Fine Tuning*. Penelitian sebelumnya juga belum menampilkan implementasi dari *callbacks* dalam proses pelatihan model seperti *learning rate scheduler* dan *save best model*. Untuk itu, penelitian ini akan mengakomodasi hal-hal tersebut dengan membandingkan beberapa model *Deep-CNN* seperti VGG16, VGG19 dan Xception dengan harapan setelah mengimplementasikan tahapan *Transfer Learning* dan *Fine Tuning* dan mengimplementasikan *callbacks* bisa mengoptimalkan kinerja model.

## 2. TINJAUAN TEORI

### 2.1. Transfer Learning

Pada tahun 2010, Pan et al mengusulkan konsep belajar pengetahuan yang tidak diketahui melalui pengetahuan yang ada yang disebut *Transfer Learning*. Konsep inti dari pembelajaran ini adalah untuk menemukan kesamaan antara pengetahuan yang ada dan pengetahuan yang belum diketahui. Beberapa domain pengetahuan terlalu abstrak untuk dipelajari,

menghasilkan biaya belajar keseluruhan yang tinggi. Oleh karena itu, menggunakan pengetahuan yang ada untuk membantu pembelajaran menjadi penting [10].

Konsep inti dari *transfer learning* adalah bagaimana menemukan relevansi antara yang diketahui dan yang tidak diketahui dan mempelajari pengetahuan baru. Dalam *transfer learning*, pengetahuan yang ada biasanya disebut Domain Sumber, dan pengetahuan yang tidak diketahui disebut Domain Target. Pembelajaran ini terutama mempelajari bagaimana memigrasikan pengetahuan dari Domain Sumber ke Domain Target. Di bidang pembelajaran mesin, *transfer learning* berfokus pada penerapan pengetahuan yang ada ke yang tidak diketahui melalui model yang ditetapkan.

## 2.2. Deep Convolutional Neural Network

CNN adalah jenis jaringan saraf tiruan yang memiliki banyak lapisan dan ahli dalam memproses volume data yang tinggi dengan akurasi yang lebih tinggi dan biaya komputasi yang lebih sedikit. Struktur dasar CNN terdiri dari *convolution*, *pooling*, *flattening*, dan *fully connected layer* [11]. Setelah langkah pre-processing dasar, langkah selanjutnya adalah mengekstrak fitur sesuai kebutuhan. Dalam usulan karya (CNN) ini digunakan pendekatan metode pembelajaran model *transfer learning* yang dilanjutkan *fine tuning* pada *Deep Convolutional Neural Network* (DCNN) terlatih. CNN memiliki banyak keunggulan. Pertama, ekstraksi fitur dan klasifikasi menjadi perhatian CNN itu sendiri dengan struktur tunggal. Kedua, jaringan ini mengekstrak fitur 2-D yang lebih dalam. Ketiga, sepenuhnya adaptif dan invarian terhadap perubahan lokal dan geometris pada citra. Tiga jenis layer utama yang ada di CNN: *Convolution layer*, *Pooling layer (Sub-sampling)*, dan *Output layer*.

## 2.3. Xception

Xception (*extreme inception*) adalah jaringan saraf convolutional dengan 71 lapisan dalam. Arsitektur model xception merupakan konvolusi yang dapat dipisahkan secara mendalam dengan memperbaiki arsitektur model Inception. Dalam model Xception, arsitektur jaringan saraf konvolusi dalam yang revolusioner diusulkan. Hal ini menyebabkan model Xception mengungguli inception V3 dalam hal klasifikasi citra dengan data yang lebih besar, terdiri dari 350 juta citra dan 17.000 kelas. Penggunaan parameter model yang lebih efektif daripada kapasitas tambahan menghasilkan peningkatan kinerja pada model Xception. Parameter yang digunakan pada model Xception juga sama dengan yang ada pada model InceptionV3 [12]

## 2.4. VGG

VGG merupakan model arsitektur yang terdiri dari 16-layer terlatih yang memiliki bobot. Arsitektur model VGG berisi arsitektur model CNN yang menggunakan spesifikasi lapisan konvolusi dari filter konvolusi 3 x 3 kecil. Karena ukuran *convolutional filter* yang kecil, *neural network* dapat memiliki lebih banyak *convolutional layer*. Model ini dibuat pada tahun 2015 dan dilatih pada satu juta gambar dari ImageNet [13].

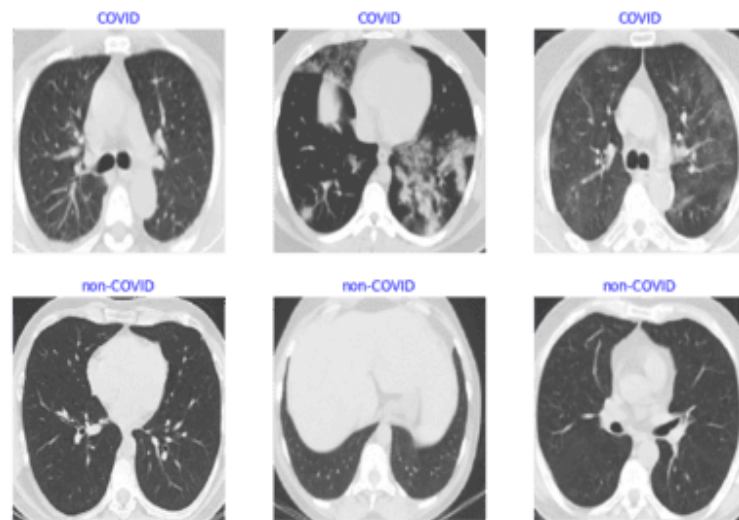
## 3. TAHAPAN PENELITIAN

Tahap ini berisi penjelasan tentang tahapan metode yang digunakan dalam penelitian ini. Pendeteksian COVID-19 menggunakan citra CT-Scan dilakukan dengan arsitektur model Deep-CNN. Metode penelitian ini dilakukan melalui beberapa tahap yakni:

- a. Pra pemrosesan data  
Pada dataset yang sudah didapatkan dilakukan tahapan pra-pemrosesan. Pada penelitian ini tahapan pra-pemrosesan yang digunakan adalah melakukan *resize* citra menjadi yang semula berukuran 348 x 256 menjadi berukuran 224 x 224. Tahapan *resize* citra ini bertujuan agar gambar yang akan digunakan dalam pelatihan berukuran sama dengan ukuran *default* dari model Deep-CNN.
- b. Perancangan Eksperimen  
Perancangan Eksperimen ini memanfaatkan 3 buah model Deep-CNN (DCNN) terlatih yakni VGG16, VGG19 dan Xception. Pada proses pelatihan juga dilakukan penerapan *callbacks* berupa *learning rate scheduler* yang bermanfaat agar dalam proses pelatihan terjadi pengurangan learning rate secara berkala yang tujuannya agar tidak terjebak pada local minima melainkan mencapai global minima loss pada gradient descent. Selain itu juga dilakukan implementasi *save best model* untuk menyimpan bobot pada akurasi validasi terbaik, hal ini bertujuan apabila suatu saat model yang dilatih mencapai *overfit* maka bobot yang mengalami *overfit* tersebut tidak akan digunakan, hal ini tentunya juga akan mendukung model mencapai *loss* validasi yang terendah.
- c. Training Model transfer learning dan fine tuning  
Pada tahapan ini dilakukan proses *transfer learning* dan *fine tuning* terhadap model DCNN seperti VGG16, VGG19 dan Xception.
- d. Evaluasi model  
Pada tahapan ini dapat dilihat metrik evaluasi berupa akurasi baik pada data training validasi ataupun data testing. Kemudian memilih akurasi terbaik dari komparasi tersebut dan menyajikan Presisi, *Recall* dan *F1 Score*.

### 3.1. SARS-CoV-2 CT-scan Dataset

Dataset yang digunakan pada eksperimen ini adalah SARS-CoV-2 CT-scan Dataset. Dataset ini terdiri dari 2 label yakni Covid dan non-Covid. Dataset ini berisi 1251 CT scan yang positif untuk infeksi SARS-CoV-2 (COVID-19) dan 1230 CT scan untuk pasien yang tidak terinfeksi SARS-CoV-2, dengan total citra 2481 CT scan secara keseluruhan. Data ini dikumpulkan dari pasien nyata di rumah sakit dari Sao Paulo, Brasil. Tujuan dari kumpulan data ini adalah untuk mendorong penelitian dan pengembangan metode kecerdasan buatan yang mampu mengidentifikasi apakah seseorang terinfeksi SARS-CoV-2 melalui analisis CT scan-nya. Dataset ini bisa didapatkan melalui situs *kaggle* ataupun *github* [14]. Contoh citra dari dataset ini dapat dilihat pada Gambar 1.

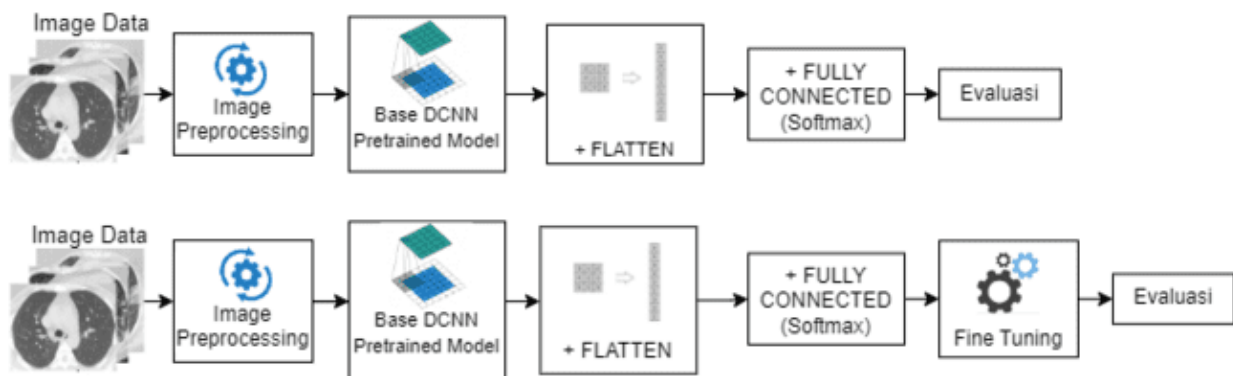


**Gambar 1.** Contoh Citra dari SARS-CoV-2 CT-scan dataset

Dalam eksperimen ini, dilakukan pembagian data menjadi 3 set untuk fase *training*, fase validasi dan fase *testing*. Setelah melakukan split data maka jumlah data pada *training*, *validasi* dan *testing* secara berurutan menjadi 2008, 224 dan 249 atau dengan proporsi 81%: 9% : 10%.

### 3.2. Rancangan Eksperimen

Dalam menjawab masalah dalam penelitian ini, dibutuhkan pembentukan suatu model untuk klasifikasi dari citra CT-Scan. Tahapan atau langkah eksperimen yang dilakukan dalam penelitian ini dibedakan menjadi dua skema. Pada skema pertama menggunakan konsep transfer learning biasa, sedangkan pada skema kedua menggunakan konsep transfer learning dan dilanjutkan pada tahapan fine tuning. Alur langkah eksperimen disajikan pada Gambar 2 seperti dibawah ini.



**Gambar 2.** Skema Eksperimen

Pada skema pertama (atas) citra CT-Scan dilakukan pra pemrosesan terhadap citra terlebih dahulu. Yang dilakukan dalam tahapan pra pemrosesan citra ini adalah mengubah ukuran citra menjadi berukuran 224x224 yang merupakan standar dari pelatihan dengan arsitektur Deep CNN. Selanjutnya melakukan penyusunan arsitektur Deep-CNN dan memuat bobot pada arsitektur tersebut dengan bobot pra pelatihan dengan imageNet. Pada tahapan ini pula perlu

melakukan freezing atau pembekuan terhadap seluruh layer. Tahapan freezing ini bertujuan agar bobot yang sudah dilatih sebelumnya tidak rusak oleh update pada fase training.

Setelah selesai melakukan load model, maka model untuk ekstraksi fitur sudah jadi. Selanjutnya, fitur yang sudah ada akan dalam bentuk multi dimensi, karena kebanyakan classifier tidak bisa menangani data yang multi dimensi, maka diperlukan perubahan ke bentuk dimensi yang lebih rendah. Proses perubahan dimensi tersebut bisa dilakukan dengan menggunakan *Flatten Layer*. sampai pada tahap ini, model sudah mendapatkan fitur yang bentuknya sudah sesuai dengan kebutuhan model, namun disini model belum mempunyai *classifier* yang mana bertugas untuk melakukan klasifikasi atau sebagai otak dalam tugas ini. Oleh sebab itu, ditambahkan *fully connected layer* atau yang mana pada konsep *machine learning* dikenal dengan jaringan syaraf tiruan.

Penggunaan jaringan syaraf tiruan pada tahapan ini menggunakan fungsi aktivasi *softmax* dan output dengan 2 buah neuron karena merupakan kasus klasifikasi biner. Pada umumnya, untuk klasifikasi biner menggunakan aktivasi *sigmoid*, namun disini menggunakan *softmax* dan hal tersebut tidak akan menimbulkan perbedaan pada hasil dan akurasi sebab pada *sigmoid* akan mengembalikan nilai pada kelas satu apabila hasil *sigmoid* lebih dari 0.5 dan apabila kurang dari 0.5 akan menghasilkan sebaliknya. Disini, 0.5 berarti digunakan sebagai *threshold*. Sedangkan pada *softmax* akan mengembalikan nilai [O1, O2] dimana O1 merupakan probabilitas dari kelas 1 (kelas negatif) dan O2 merupakan probabilitas pada kelas 2 (kelas positif). Atau apabila dinotasikan perubahan dari *sigmoid* menjadi *softmax* akan menjadi [1-*sig*, *sig*] sehingga apabila ditotalkan pada *softmax* akan menghasilkan nilai 1.

### 3.3. Metriks Evaluasi

Metriks evaluasi bertujuan untuk meringkas dan mengevaluasi suatu sistem. Biasanya, tugas klasifikasi menggunakan metriks berupa akurasi, *precision*, *recall* dan *f1-score*. Akurasi merupakan pendefinisian persentase seluruh data yang diklasifikasikan secara benar pada kelas positif ataupun negatif. Sehingga berarti seluruh data prediksi yang bernilai benar dibagi dengan seluruh data yang ada. Akurasi dapat dirumuskan menjadi seperti berikut ini

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

Seluruh value pada persamaan tersebut biasanya bisa didapatkan dari *confusion matrix* [15]. *Precision* adalah metrik yang membahas seberapa presisi/akurat sebuah model. Hal ini dapat diukur dari hasil prediksi positif (berapa banyak yang benar-benar positif). Sehingga presisi dapat dirumuskan sebagai:

$$Precision = \frac{True\ Positive}{Total\ Predicted\ Positive} \quad (2)$$

Jadi dari persamaan awal, bagian total yang diprediksi positif dapat diubah menjadi TP + FP, sehingga rumusnya menjadi:

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

Recall merupakan metrik yang menghitung berapa banyak Positif Aktual yang dirujuk oleh model dengan melabelinya sebagai Positif (Benar-Benar Positif). Diberikan akan menjadi



metrik model yang digunakan untuk memilih model terbaik ketika ada biaya tinggi yang terkait dengan Negatif Palsu. Jadi penarikan kembali dapat dirumuskan sebagai:

$$Recall = \frac{True\ Positive}{Total\ Actual\ Positive} \quad (3)$$

Sehingga dari persamaan awal tersebut, bagian pada *total actual positive* bisa diubah menjadi TP + FN, sehingga perumusannya menjadi:

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

Skor F1 adalah metrik untuk mengukur dengan mempertimbangkan keseimbangan antara presisi dan recall sehingga skor f1 adalah nilai negatif yang sebenarnya. Maka skor f1 dapat dirumuskan:

$$F1 - Score = \frac{(2 \times Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

#### 4. HASIL DAN PEMBAHASAN

Seluruh eksperimen dilakukan menggunakan bahasa pemrograman *python* dan menggunakan *coding environment* berupa google colab pro. Pada eksperimen ini dijalankan menggunakan akselerator perangkat berupa GPU (K80 atau P100 atau T4) dan RAM sebesar 25GB. Seluruh model arsitektur dan bobot pelatihan dari *imageNet* sudah disediakan melalui library Keras dan Tensorflow. Seluruh eksperimen pada tahapan ini menggunakan *crossentropy loss* dan *epoch* sejumlah 10. Pada fase *fine tuning*, dilakukan *unfreeze layer* model sejumlah 10-layer yang dimulai dari *top layer*.

Hasil eksperimen berdasarkan metrik berupa akurasi, baik akurasi pada data training, validasi ataupun testing pada tiap arsitekturnya dan pada tiap fasenya tersaji pada Tabel 1 dibawah ini.

**Tabel 1.** Hasil Eksperimen

Model	Akurasi Transfer Learning			Akurasi Fine Tuning		
	Training	Validation	Testing	Training	Validation	Testing
<b>VGG16</b>	0.9833	0.9821	0.9518	<b>1.0000</b>	<b>0.9866</b>	<b>0.9759</b>
<b>VGG19</b>	0.9739	0.9554	0.9558	0.9965	0.9821	0.9719
<b>Xception</b>	1.0000	0.9464	0.9478	0.9975	0.9866	0.9639

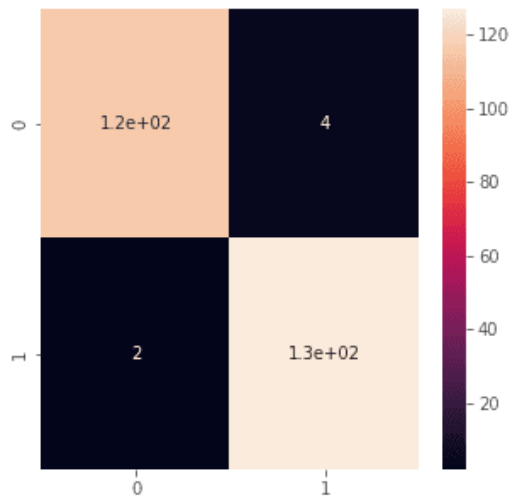
Berdasarkan Tabel 1, Arsitektur VGG 16 mendapatkan akurasi tertinggi pada data pengujian (testing), secara spesifik pada fase fine tuning. Hal ini menunjukkan bahwa kedalaman model memang tidak menjamin memiliki akurasi yang lebih baik (pada kasus ini VGG16 merupakan arsitektur DCNN yang memiliki kedalaman arsitektur paling minimum). Namun hal ini tentu bergantung dengan data yang digunakan pula, memang pada kasus pendeteksian SARS-CoV-2 yang merupakan kasus klasifikasi biner VGG16 menunjukkan performa yang lebih baik dan hal ini tidak berarti pada kasus atau dataset lainnya VGG16 akan lebih unggul. Dalam eksperimen menggunakan DCNN, ada banyak faktor tak tentu yang bisa mempengaruhi baik

buruknya proses training suatu model, misalkan learning rate, jumlah kernel, jumlah layer hingga pemanfaatan callbacks.

Berbagai upaya telah dilakukan untuk menjadikan proses pembelajaran model menjadi lebih optimal, seperti penggunaan *pretrained model*, penerapan *fine tuning*, penggunaan *learning rate scheduler* hingga penggunaan *save best model*. hal ini yang mengakibatkan hasil akurasi menjadi sedikit lebih baik apabila dibandingkan dengan penelitian sebelumnya yang mendapatkan akurasi dengan *machine learning* tidak lebih tinggi dari 90 persen dan mencapai akurasi tertinggi sebesar 96 persen dengan penerapan *deep learning*, yang mana pada penelitian ini mendapatkan akurasi terbaik sebesar 97,6 persen dengan arsitektur VGG16.

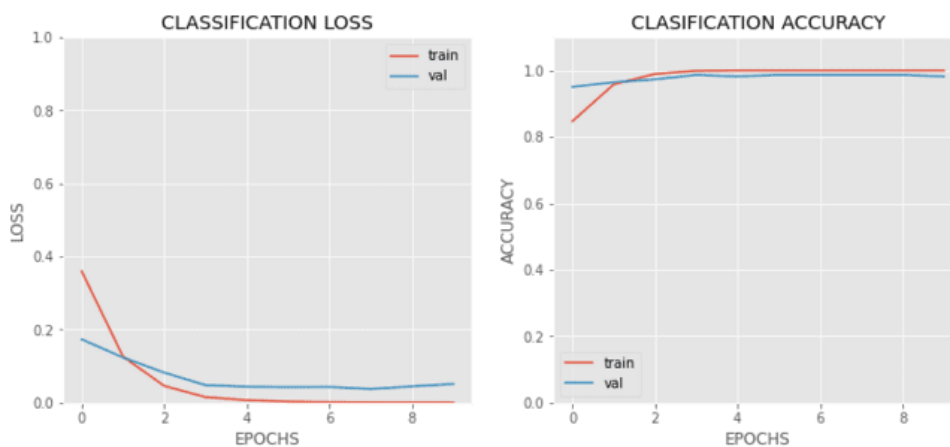
Apabila dilihat secara menyeluruh pada tahapan *transfer learning* biasa, akurasi testing berkisar antara 94 hingga 95 persen dan pada tahapan fine tuning berada pada akurasi testing mencapai 97,6% dan tidak ada akurasi yang kurang dari 96%. Hal ini menunjukkan bahwa tahapan fine tuning berhasil meningkatkan performa dari model DCNN.

Berdasarkan konfigurasi arsitektur model DCNN terbaik (VGG16) yang dibentuk tersebut mendapatkan confusion matrix yang tersaji pada Gambar 3.



Gambar 3. Confusion Matrix

Hasil eksperimen menggunakan arsitektur dasar dari VGG 16 + fine tuning ini mendapatkan learning curve yang tersaji pada Gambar 4.



Gambar 4. Learning Curve

Dengan metrik eksperimen pada data testing dengan VGG 16 dapat dilihat pada Tabel 2 dibawah ini

**Tabel 2.** Metriks Eskperimen

	<b>Precision</b>	<b>Recall</b>	<b>F1- Score</b>
<b>Non Covid</b>	0.98	0.97	0.97
<b>Covid</b>	0.97	0.98	0.98
Akurasi			0.9759
Macro - Average	0.98	0.98	0.98
Micro - Average	0.97	0.98	0.98

## 5. KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Berdasarkan eksperimen yang telah dilakukan pada kasus deteksi COVID-19 menggunakan citra CT-Scan paru didapati bahwa ketiga arsitektur pretrained berhasil mendapatkan akurasi testing yang cukup tinggi, bahkan mencapai akurasi terendah pada akurasi testing sebesar 94.78% pada arsitektur Xception tanpa dilakukan *fine tuning*. Sedangkan pada akurasi testing terbaiknya yakni pada arsitektur VGG 16 dengan mengaplikasikan fine tuning mendapatkan akurasi testing sebesar 97,59%. Selain mendapatkan arsitektur model yang terbaik, pada eksperimen ini juga mendapati bahwa fase *fine tuning* berhasil meningkatkan akurasi model secara cukup signifikan yakni peningkatan sekitar 2 persen. Selain itu, dengan adanya implementasi *save best model* dan *learning rate scheduler* juga bisa meningkatkan performa model yang dibangun yang mana hal ini ditunjukkan dari penelitian sebelumnya yang tidak mengimplementasikan kedua teknik ini walaupun dengan model yang sama namun mendapatkan hasil yang lebih rendah.

### 5.2. Saran

Untuk eksperimen kedepannya, peneliti bisa mencoba beberapa arsitektur lainnya seperti efficientNet, Inception dan ResNet. Selain itu, peneliti juga bisa mengkombinasikan model seperti pada konsep ensemble model seperti Ensemble Voting yang mana bisa terdiri dari soft voting dan Hard Voting.

## REFERENSI

- [1] C. Sohrabi, Z. Alsafi, N. O. Neill, M. Khan, and A. Kerwan, "World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)," no. January, 2020.
- [2] T. Ai *et al.*, "Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases."
- [3] T. B. Chandra and K. Verma, "Pneumonia Detection on Chest X-Ray Using Machine Learning Paradigm," *Proceedings of 3rd international conference on computer vision and image processing*, 2020.

- [4] J. Diaz-Escobar *et al.*, "Deep-learning based detection of COVID-19 using lung ultrasound imagery," *PLoS One*, vol. 16, no. 8 August, Dec. 2021, doi: 10.1371/journal.pone.0255886.
- [5] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, and N. Mehendale, "Diagnosis of COVID-19 using CT scan images and deep learning techniques," *Emerg Radiol*, vol. 28, no. 3, pp. 497–505, 2021, doi: 10.1007/s10140-020-01886-y.
- [6] S. Sharma, "Drawing insights from COVID-19-infected patients using CT scan images and machine learning techniques: a study on 200 patients," *Environmental Science and Pollution Research*, vol. 27, no. 29, pp. 37155–37163, 2020, doi: 10.1007/s11356-020-10133-3.
- [7] M. J. Horry *et al.*, "COVID-19 Detection through Transfer Learning Using Multimodal Imaging Data," *IEEE Access*, vol. 8, pp. 149808–149824, 2020, doi: 10.1109/ACCESS.2020.3016780.
- [8] D. A. Zebari, A. M. Abdulazeez, D. Q. Zeebaree, and M. S. Salih, "A Fusion Scheme of Texture Features for COVID-19 Detection of CT Scan Images," *3rd International Conference on Advanced Science and Engineering, ICOASE 2020*, pp. 12–17, 2020, doi: 10.1109/ICOASE51841.2020.9436538.
- [9] S. A. Harmon *et al.*, "Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets," *Nat Commun*, vol. 11, no. 1, pp. 1–7, 2020, doi: 10.1038/s41467-020-17971-2.
- [10] J. C. Hung, K. C. Lin, and N. X. Lai, "Recognizing learning emotion based on convolutional neural networks and transfer learning," *Applied Soft Computing Journal*, vol. 84, p. 105724, 2019, doi: 10.1016/j.asoc.2019.105724.
- [11] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, and V. Singh, "Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet," *Chaos Solitons Fractals*, vol. 138, p. 109944, 2020, doi: 10.1016/j.chaos.2020.109944.
- [12] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 1800–1807, 2017, doi: 10.1109/CVPR.2017.195.
- [13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–14, 2015.
- [14] E. Soares and P. Angelov, "A large dataset of real patients CT scans for COVID-19 identification," *Harv. Dataverse*, vol. 1, pp. 1–8, 2020.
- [15] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," no. January 2008, 2020.