

# Deep Learning Course

## Final Project Report



KHAERURROZIKIN(H071221069)  
Andi Ahmad Salwan (H071221048)  
Muhammad Aditya Permana (H071221063)

December 5, 2024

## Contents

<b>1</b>	<b>INTRODUCTION</b>	<b>2</b>
1.1	Latar Belakang . . . . .	2
1.2	Masalah . . . . .	2
1.3	Tujuan . . . . .	2
<b>2</b>	<b>RELATED WORKS</b>	<b>2</b>
<b>3</b>	<b>DATASET MATERIAL</b>	<b>3</b>
3.1	Dataset . . . . .	3
3.1.1	Sumber Dataset . . . . .	3
<b>4</b>	<b>Result and Discussion</b>	<b>4</b>
4.1	Generalisasi . . . . .	8
4.2	Fine Tuning . . . . .	8
4.3	Confusion Matrix . . . . .	12
<b>5</b>	<b>Conclusion</b>	<b>15</b>

## 1 INTRODUCTION

### 1.1 Latar Belakang

Perkembangan seni rupa modern telah menghasilkan berbagai aliran seni dengan karakteristik unik yang sering kali sulit dibedakan oleh orang awam. Beberapa aliran seni populer seperti abstrak, kubisme, minimalisme, dan surrealisme memiliki elemen visual yang khas, tetapi perbedaan di antara mereka dapat menjadi tantangan untuk dikenali, terutama bagi individu yang tidak memiliki latar belakang seni. Dalam era digital saat ini, teknologi kecerdasan buatan (AI) membuka peluang untuk membantu mengatasi tantangan ini. Dengan pendekatan berbasis Computer Vision dan Deep Learning, pengenalan aliran seni dapat dilakukan secara otomatis melalui analisis gambar.

### 1.2 Masalah

Kesulitan dalam mengenali aliran seni seperti abstrak, kubisme, minimalisme, dan surrealisme tidak hanya dialami oleh masyarakat umum yang mungkin tidak memiliki latar belakang seni, tetapi juga dirasakan oleh berbagai pihak yang terlibat dalam industri seni dan kreatif. Para kolektor seni, misalnya, sering menghadapi tantangan dalam menentukan keaslian dan klasifikasi sebuah karya, terutama ketika mereka harus membuat keputusan pembelian berdasarkan ciri khas aliran seni tertentu. Kurator pameran seni juga dihadapkan pada kesulitan serupa ketika mereka perlu menyusun koleksi yang merepresentasikan gaya atau tema tertentu, tetapi tidak memiliki alat yang memadai untuk memastikan keakuratan klasifikasi tersebut.

### 1.3 Tujuan

Tujuan dari pembuatan dataset ini adalah untuk mendukung pengembangan sistem berbasis kecerdasan buatan yang mampu mengenali dan mengklasifikasikan empat aliran seni rupa utama, yaitu abstrak, kubisme, minimalisme, dan surrealisme. Dataset ini dirancang untuk menjadi sumber data yang representatif dalam melatih dan menguji model Convolutional Neural Network (CNN), sehingga dapat mengenali pola visual seperti bentuk, warna, dan tekstur yang khas dari setiap aliran seni. Dengan dataset ini, diharapkan sistem yang dihasilkan mampu membantu masyarakat umum, pelajar, dan pelaku industri seni dalam memahami dan membedakan karakteristik visual dari keempat aliran seni tersebut.

## 2 RELATED WORKS

Berbagai algoritma deep learning telah digunakan dalam pengenalan pola visual, khususnya dalam klasifikasi gambar seperti aliran seni rupa. Algoritma seperti Convolutional Neural Network (CNN), Inception, dan ResNet telah menjadi pilihan utama karena kemampuannya dalam mengekstraksi fitur visual yang kompleks. Berikut adalah beberapa penelitian terkait yang mengaplikasikan algoritma tersebut.

CNN merupakan algoritma deep learning yang paling umum digunakan dalam klasifikasi gambar karena kemampuannya untuk mengenali pola-pola lokal

dalam data visual. Penelitian oleh Krizhevsky et al. (2012) dengan AlexNet menunjukkan efektivitas CNN dalam menangani dataset gambar berskala besar. Studi ini menjadi dasar untuk berbagai aplikasi, termasuk pengenalan aliran seni rupa. CNN bekerja dengan mengidentifikasi fitur-fitur visual seperti tekstur, warna, dan bentuk, yang sangat relevan untuk membedakan aliran seni seperti abstrak dan kubisme.

CNN telah lama menjadi standar untuk tugas klasifikasi citra karena kemampuannya dalam mengekstraksi fitur hierarkis yang relevan dari gambar. Penelitian oleh Mahmoud et al. (2023) menunjukkan bagaimana CNN terus digunakan dalam berbagai aplikasi pengenalan objek, meskipun Vision Transformers (ViT) mulai menunjukkan potensi di beberapa area. CNN tetap menjadi pilihan utama pada dataset yang lebih kecil atau masalah klasifikasi yang sederhana.

Zou et al. (2018) ResNet, yang dikembangkan oleh He et al. (2016), memperkenalkan konsep residual learning untuk mengatasi masalah degradasi performa pada model deep learning yang sangat dalam. ResNet menggunakan shortcut connections yang memungkinkan aliran informasi lebih efisien, sehingga meningkatkan akurasi pada klasifikasi gambar. Dalam penelitian seni rupa, ResNet terbukti efektif dalam mengenali elemen-elemen kompleks seperti kombinasi warna dan tekstur dalam karya seni surrealisme dan kubisme. Arsitektur ResNet juga sering digunakan dalam transfer learning untuk meningkatkan performa pada dataset spesifik, seperti dataset seni rupa yang sedang dikembangkan.

Sebuah tinjauan terbaru oleh Bernardino et al. (2023) membahas penerapan teknik-teknik baru, termasuk grup konvolusi dalam ResNeXt, yang membantu dalam pengolahan dataset besar dengan lebih efisien. Hal ini juga menekankan pentingnya perhatian dalam jaringan Inception untuk meningkatkan kualitas prediksi pada berbagai masalah klasifikasi citra.

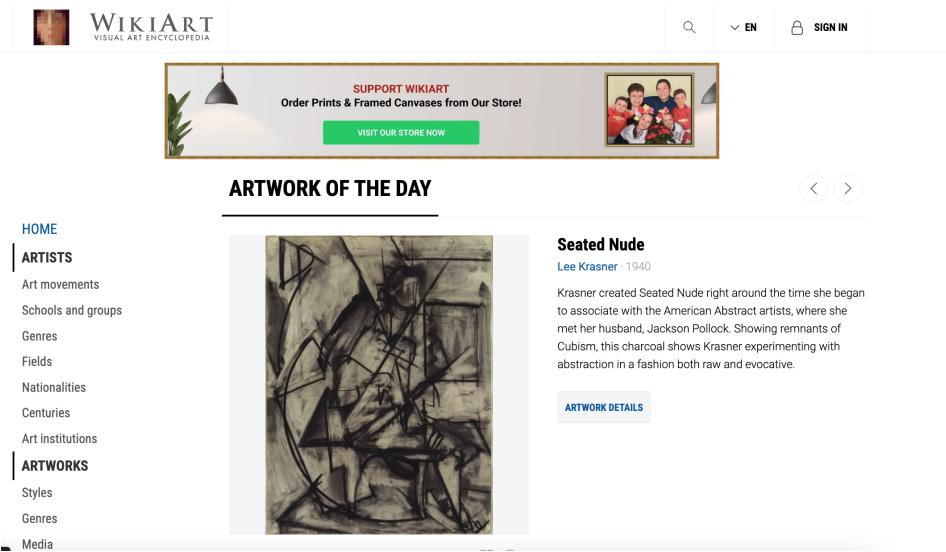
### **3 DATASET MATERIAL**

#### **3.1 Dataset**

Dalam Proyek ini, kami membuat dataset khusus untuk mendeteksi 4 class gambar yaitu abstrak, kubisme, surrealisme dan minimalisme

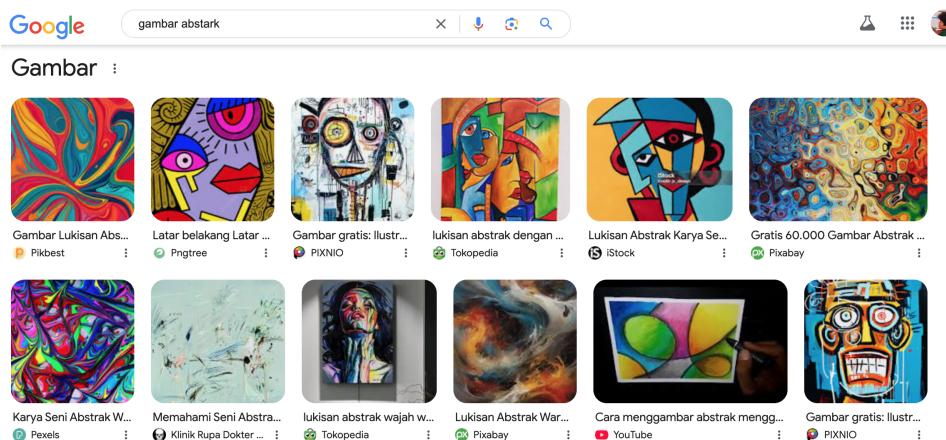
##### **3.1.1 Sumber Dataset**

- WebScraping : Wikiart menyediakan berbagai jenis lukisan mulai dari pelukisnya dan tahun berapa lukisan tersebut dilukis



Gambar 1. WIKIART penyedia berbagai jenis gambar lukisan

- Google: Gambar dari berbagai jenis lukisan didapatkan melalui google images.

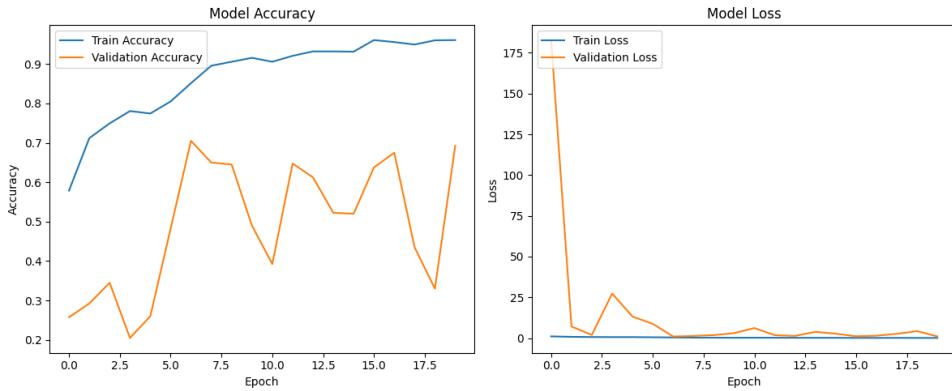


Gambar 2. Pengambilan Gambar Di Google Images

#### 4 Result and Discussion

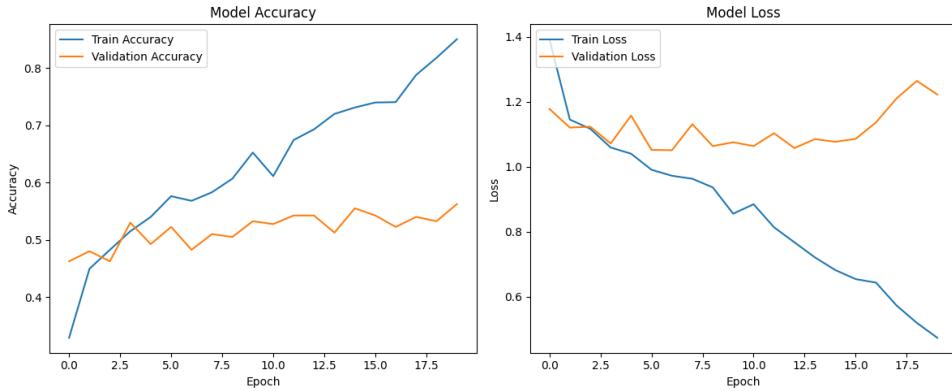
Hasil penelitian ini menunjukkan bahwa penggunaan model Inception dalam klasifikasi memiliki akurasi prediksi yang memuaskan dengan angka 74% dan indikasi overfitting kecil. Setiap model ditrain dengan 20 epoch, masing-masing memiliki 50 step per epoch. Pada awalnya, setiap model ditrain tanpa adanya optimisasi yang perlu. Model Inception mencapai 50% akurasi pada epoch pertama, lebih tinggi dibandingkan ResNet yang mencapai 19% dan model CNN biasa yang memiliki akurasi 32%. Akurasi akhir dari Inception, CNN, dan ResNet secara berturut-turut adalah 96%, 85% dan 42%. Perbandingan akurasi dan loss dari ketiga model adalah sebagai berikut:

- Inception



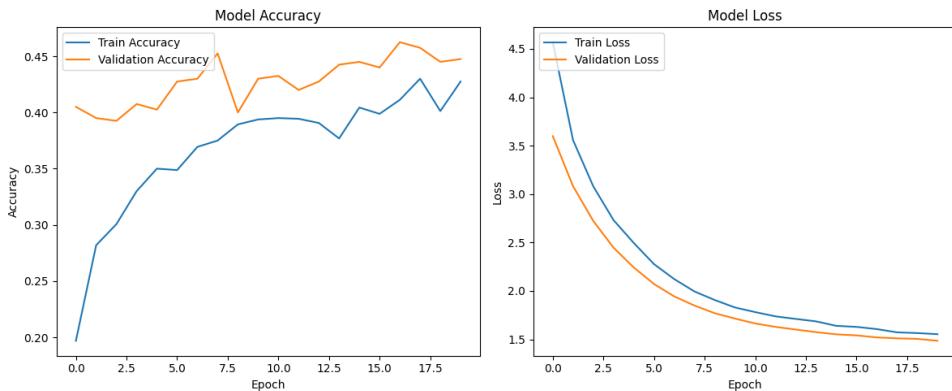
Gambar 3. Akurasi dan loss awal pada Inception

- CNN



Gambar 4. Akurasi dan loss awal pada CNN

- ResNet



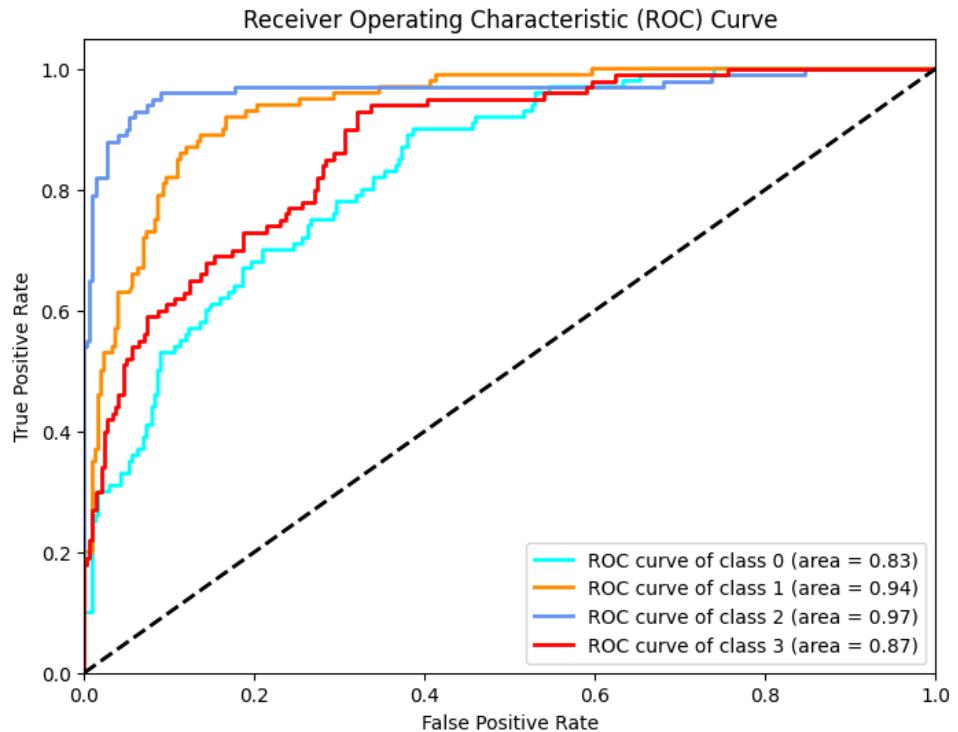
Gambar 5. Akurasi dan loss awal pada ResNet

Model Inception memiliki perbedaan signifikan pada akurasi train dan validasi, mengindikasikan adanya overfitting. Disisi lain terdapat kesenjangan pada akurasi dan loss pada train

dan validasi di CNN yang juga mengindikasikan adanya overfitting, dikarenakan model mampu untuk mengklasifikasikan data train, namun gagal untuk mengklasifikasikan data validasi secara akurat. Sebaliknya model ResNet mengalami underfitting, hal ini diindikasikan dari akurasi yang cukup rendah dan model loss yang bernilai lebih dari satu.

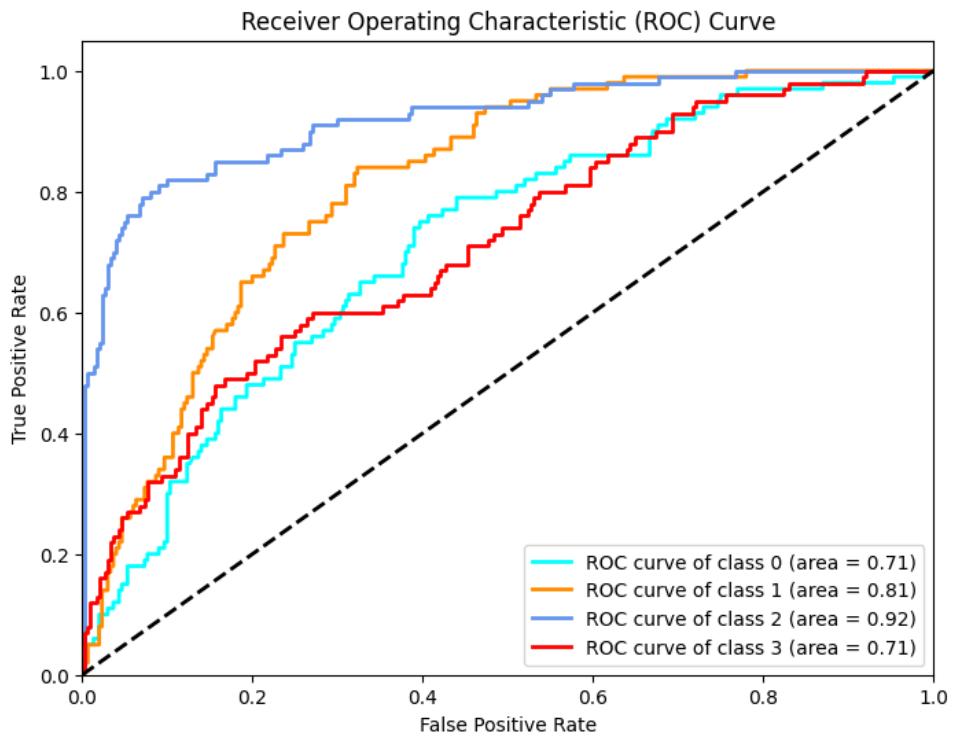
Berikut adalah grafik Receiver Operating Characteristic (ROC) dari ketiga model :

- Inception



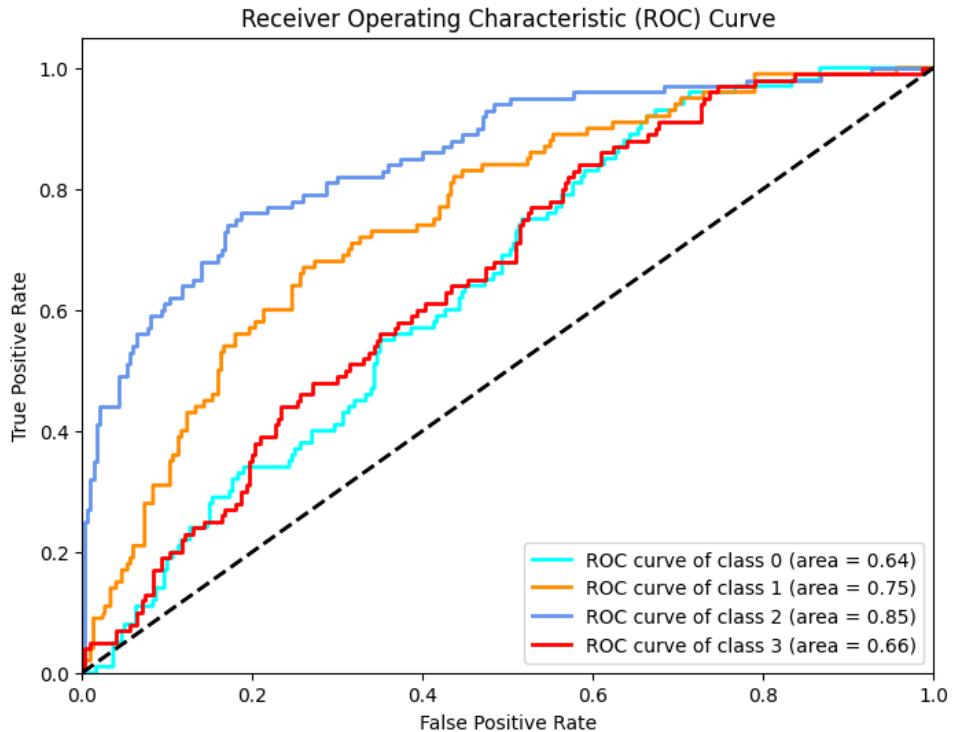
Gambar 6. ROC awal pada Inception

- CNN



Gambar 7. ROC awal pada CNN

- ResNet



Gambar 8. ROC awal pada ResNet

Dari ketiga model diatas, kita dapat menyimpulkan bahwa model inception memiliki performa yang paling baik, namun karena terdapat indikasi overfitting, sehingga grafik

ROC menjadi kurang valid. Dari grafik ROC kelas 2 (minimalisme) memiliki akurasi yang cukup tinggi, hal ini mengimplikasikan bahwa kelas minimalisme memiliki perbedaan fitur yang mampu membedakannya dengan kelas lain.

Dari hasil train diatas, dapat dilakukan beberapa optimisasi untuk menaikkan akurasi model dan meminimalisir overfitting, metode yang dilakukan yakni generalisation dan fine tuning.

## 4.1 Generalisasi

Generalisasi adalah kemampuan suatu model untuk memberikan prediksi akurat terhadap data baru yang tidak digunakan selama proses pelatihan. Karena tujuan utama pengembangan model pembelajaran mesin adalah untuk memaksimalkan akurasi tidak hanya pada data pelatihan tetapi juga pada data yang belum pernah dilihat sebelumnya, generalisasi penting saat mengevaluasi kinerja model.

Laporan ini mengukur generalisasi dengan membandingkan metrik kinerja seperti akurasi, kerugian, dan AUC-ROC antara data pelatihan, validasi, dan pengujian. Suatu model dianggap *overfitted* jika performa model pada data pelatihan jauh lebih baik dibandingkan pada data validasi atau pengujian. Hal ini menunjukkan bahwa model terlalu spesifik terhadap pola dalam data pelatihan dan gagal menangkap pola yang lebih umum.

```
from tensorflow.keras.preprocessing.image import ImageDataGenerator
import tensorflow as tf
# Regularisasi
datagen = ImageDataGenerator(
    rotation_range=20,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    fill_mode='nearest',
)
datagen.fit(X_train)
train_generator = datagen.flow(X_train, y_train, batch_size=32)
```

Gambar 9. Generalisasi model pada data pelatihan, validasi, dan pengujian.

## 4.2 Fine Tuning

Fine-tuning adalah proses penyesuaian ulang parameter model dengan tujuan meningkatkan performa dan kemampuan generalisasi pada dataset tertentu. Proses ini dilakukan terutama pada model yang telah dilatih sebelumnya (*pre-trained model*), di mana bobot awal model digunakan sebagai titik awal, kemudian disesuaikan dengan data spesifik yang sedang ditangani.

Fine-tuning memungkinkan model untuk memanfaatkan pengetahuan umum yang diperoleh dari dataset besar, seperti pola dasar dalam gambar atau teks, sekaligus beradaptasi dengan karakteristik unik dataset target. Dalam proyek ini, *fine-tuning* dilakukan dengan memanipulasi *learning rate* menjadi 0.001 dan membekukan semua *layer* jaringan neural kecuali 50 *layer* terakhir.

```

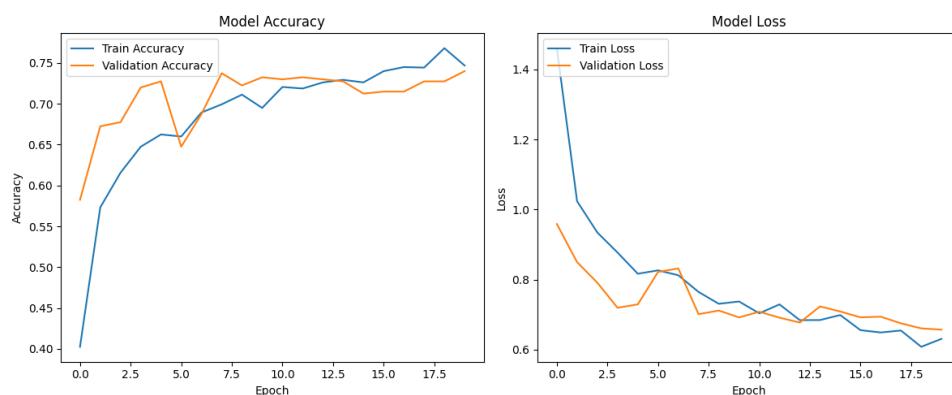
from tensorflow.keras import regularizers
model = models.Sequential([
    resnet_base,
    layers.GlobalAveragePooling2D(),
    layers.Dense(128, activation='relu', kernel_regularizer=regularizers.l2(0.01)),
    layers.Dropout(0.5),
    layers.Dense(10, activation='softmax', kernel_regularizer=regularizers.l2(0.01)) # Output layer untuk klasifikasi
])

```

Gambar 10. Proses fine-tuning model pada dataset spesifik.

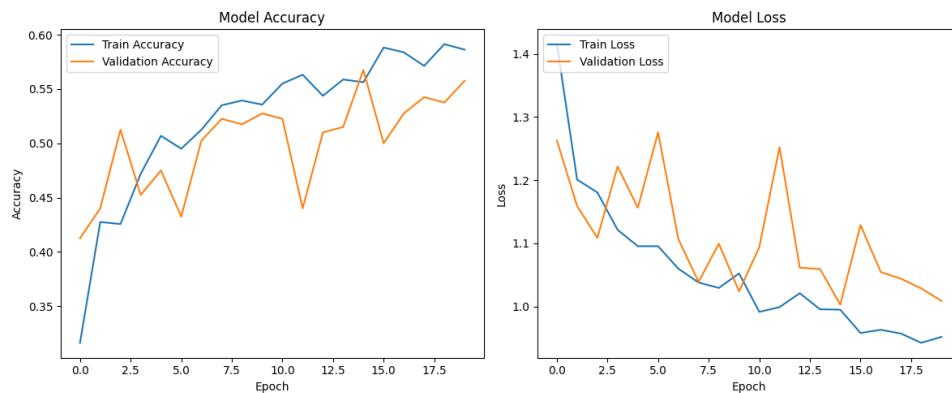
Setelah dilakukan optimisasi, pelatihan kembali dilakukan untuk mengevaluasi model. Hasilnya, model Inception memiliki akurasi 74%, CNN mencapai 55%, dan ResNet mencapai 51%. Grafik perbandingan hasil akurasi model adalah sebagai berikut:

- Inception



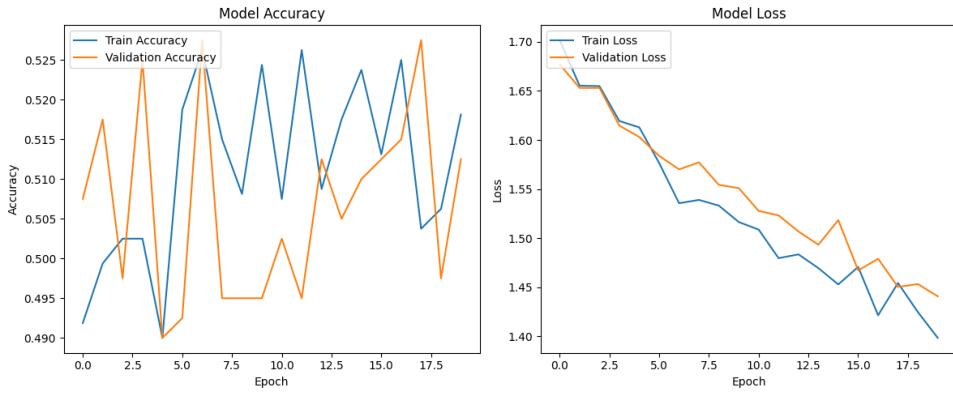
Gambar 11. Akurasi dan loss pada Inception

- CNN



Gambar 12. Akurasi dan loss pada CNN

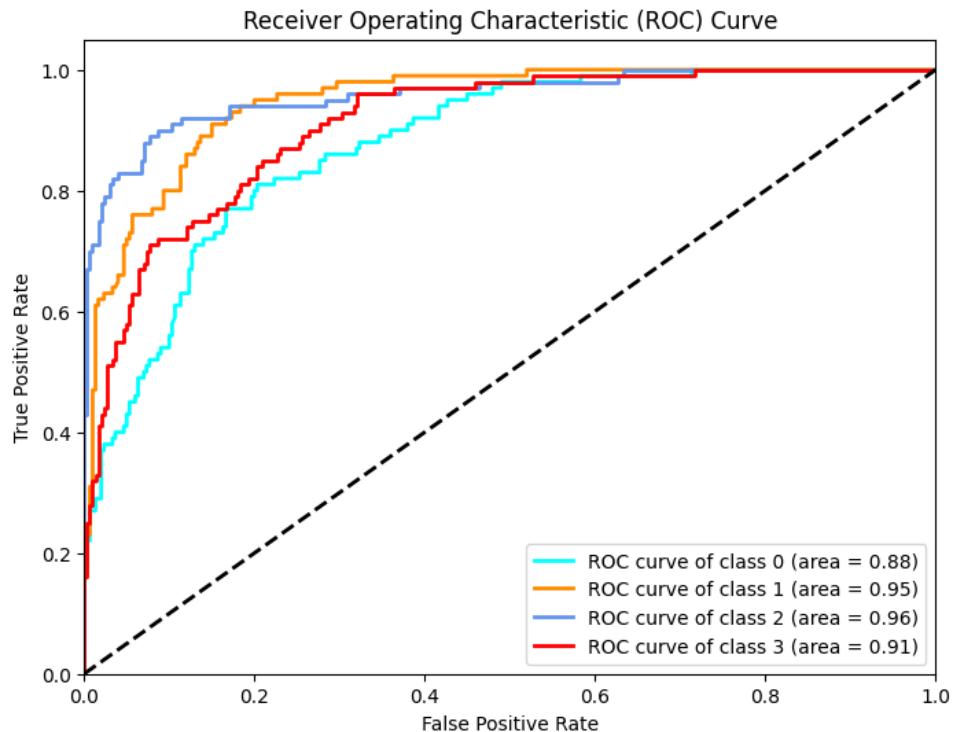
- ResNet



Gambar 13. Akurasi dan loss pada ResNet

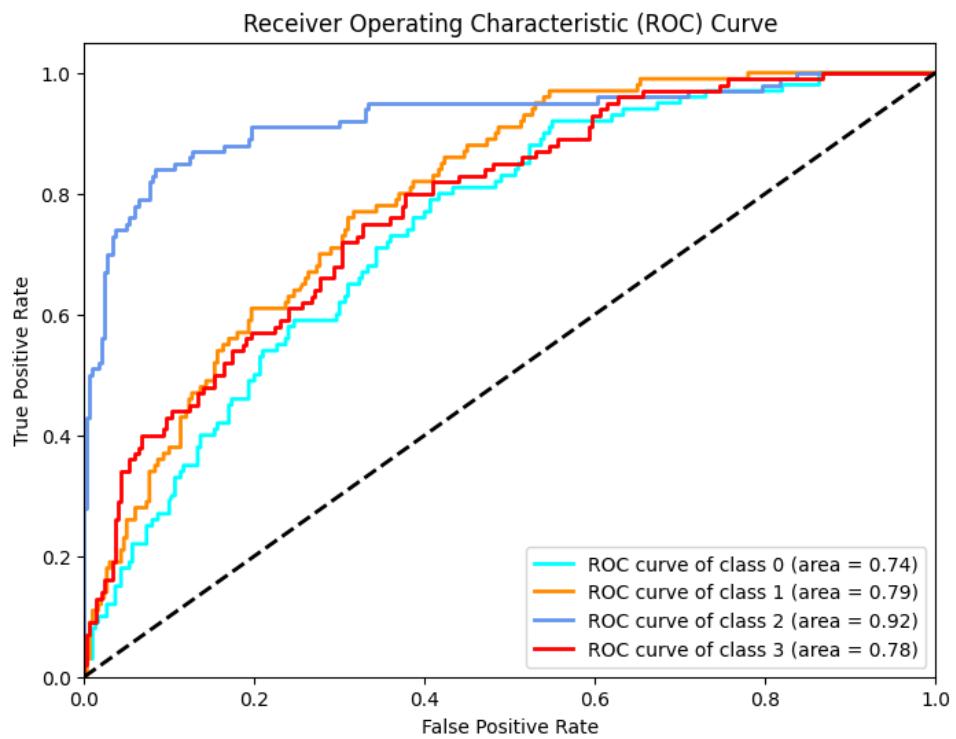
Dari ketiga grafik diatas, model Inception memiliki performa paling baik dengan akurasi dan loss yang tidak begitu fluktuatif. Model CNN mengalami fluktuasi pada validasi yang mengindikasikan model mengalami kendala dalam generalisasi data validasi. ResNet mengalami fluktuasi pada akurasi, hal ini bisa terjadi karena model yang terlalu kompleks atau learning rate yang terlalu tinggi. Kedua hal ini mengindikasikan overfitting pada model CNN dan ResNet. Grafik ROC seperti berikut:

- Inception



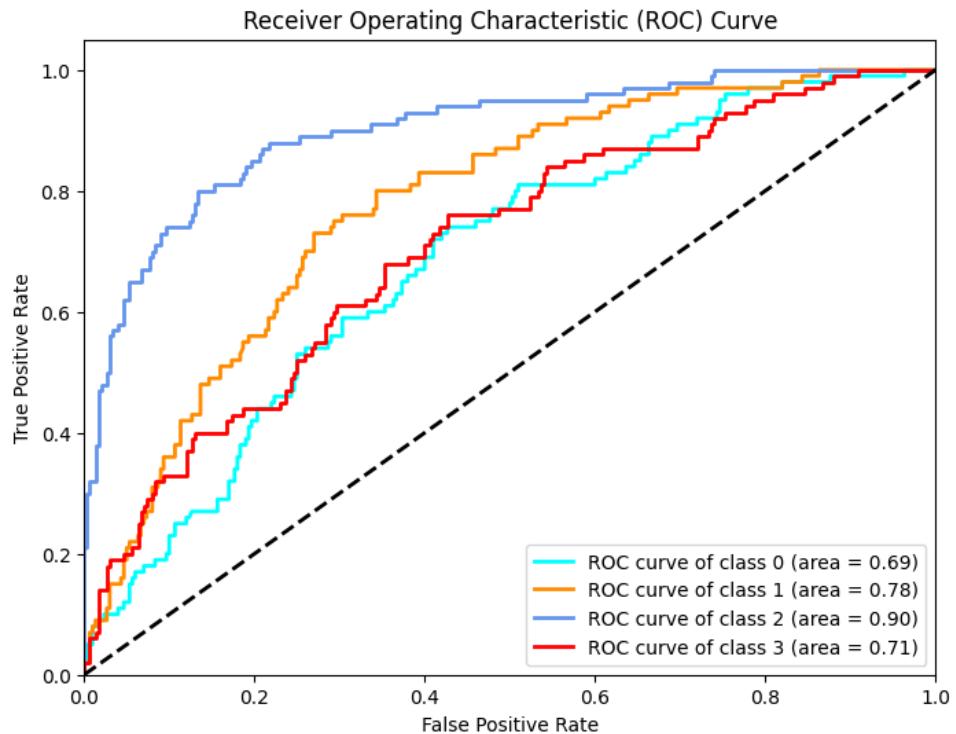
Gambar 14 . ROC pada Inception

- CNN



Gambar 15 . ROC pada CNN

- ResNet



Gambar 16 . ROC pada ResNet

Hasilnya, model Inception memiliki akurasi yang paling tinggi dan stabil dibandingkan model yang lain. Namun perlu diperhatikan bahwa model Inception masih belum optimal

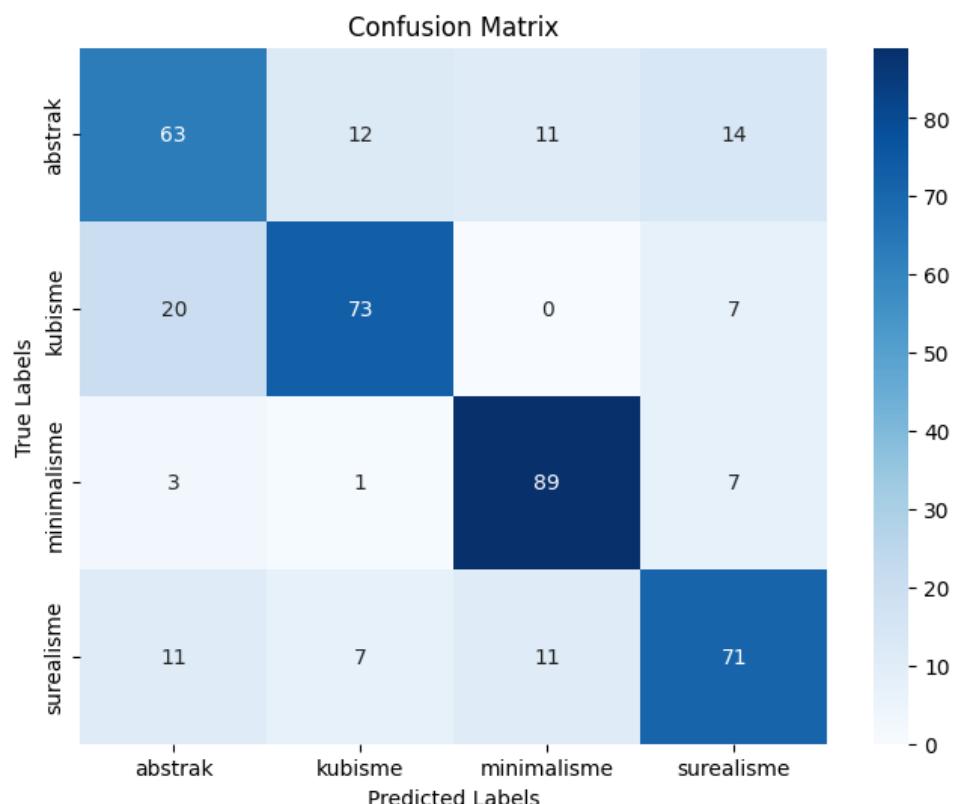
akibat akurasinya yang masih dibawah 80%, hal ini bisa disebabkan oleh kurangnya dataset yang diberikan pada model.

### 4.3 Confusion Matrix

Confusion matrix adalah alat evaluasi kinerja model klasifikasi yang memberikan gambaran detail tentang prediksi model dibandingkan dengan label yang sebenarnya. Matriks ini terdiri dari empat elemen utama untuk setiap kelas dalam dataset:

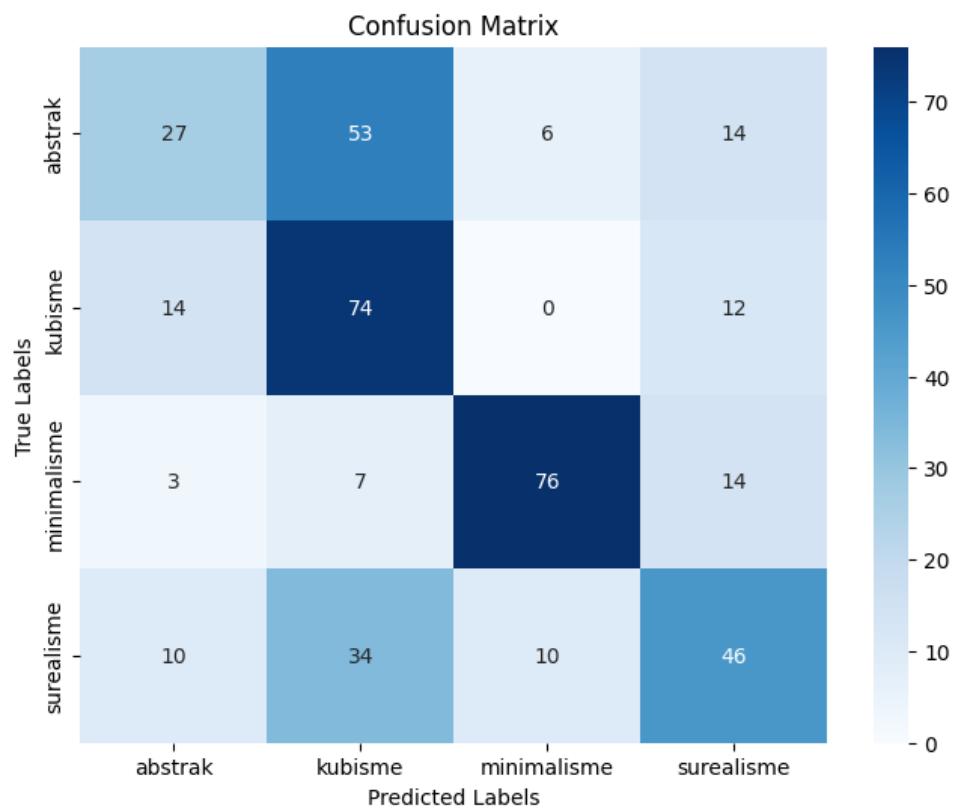
True Positive (TP): Jumlah sampel yang benar-benar termasuk dalam kelas tertentu dan berhasil diprediksi dengan benar oleh model. False Positive (FP): Jumlah sampel yang sebenarnya bukan dari kelas tertentu tetapi diprediksi sebagai kelas tersebut oleh model. False Negative (FN): Jumlah sampel yang sebenarnya dari kelas tertentu tetapi tidak diprediksi sebagai kelas tersebut oleh model. True Negative (TN): Jumlah sampel yang bukan dari kelas tertentu dan tidak diprediksi sebagai kelas tersebut oleh model.

- Inception



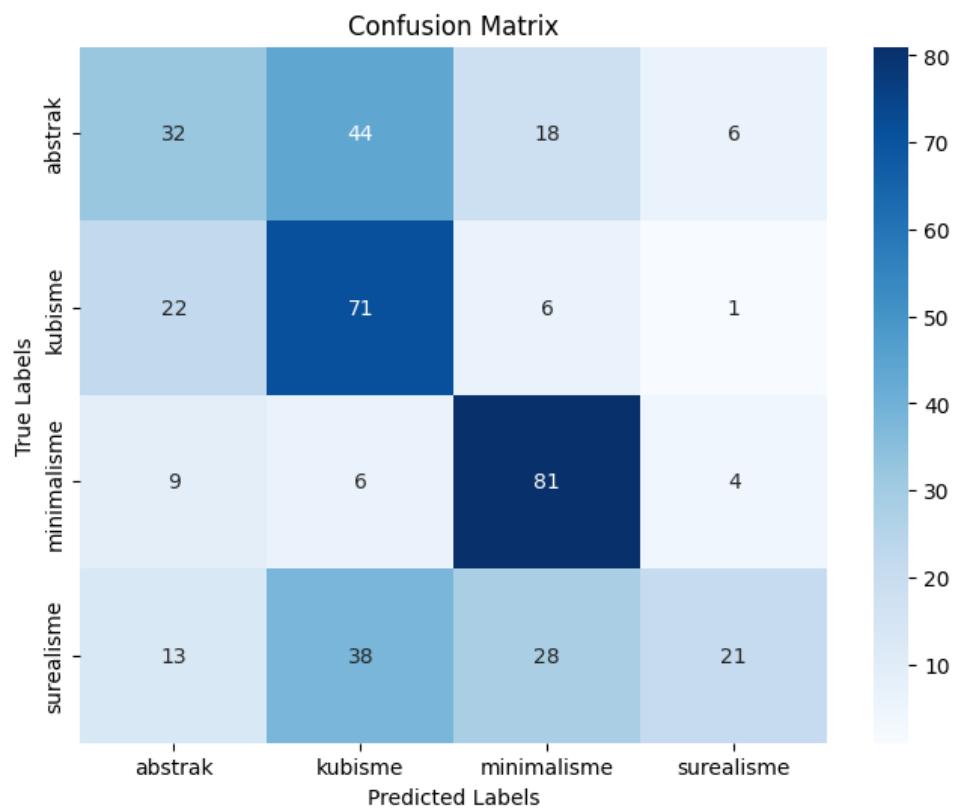
Gambar 17 . Confusion Matrix pada Inception

- CNN



Gambar 18 . Confusion Matrix pada CNN

- ResNet



Gambar 19 . Confusion Matrix pada ResNet

Berdasarkan ketiga grafik confusion matrix yang disajikan, dapat disimpulkan sebagai berikut:

Gambar 17 memperlihatkan hasil model Inception, yang menunjukkan bahwa model ini sudah mampu membedakan kelas 'abstrak', 'kubisme', dan 'surealisme' dengan baik. Namun, model Inception masih mengalami kesulitan dalam membedakan kelas 'minimalism' dari kelas lainnya.

Gambar 18 menampilkan hasil model CNN dengan layer manual, di mana terdapat peningkatan performa dalam membedakan kelas 'abstrak' dan 'kubisme' dibandingkan model ResNet. Akan tetapi, model CNN masih kesulitan dalam membedakan kelas 'minimalism' dan 'surealisme'.

Gambar 19 menunjukkan bahwa model ResNet memiliki kinerja terbaik dalam mengidentifikasi kelas 'surealisme' dengan persentase klasifikasi benar (true positive) sebesar 81%. Namun, model ini masih mengalami kesulitan dalam membedakan kelas 'abstrak' dan 'kubisme', dengan banyak terjadi kesalahan klasifikasi antar kedua kelas tersebut.

Secara keseluruhan, ketiga model menunjukkan performa yang beragam dalam mengklasifikasikan karya seni ke dalam kategori 'abstrak', 'kubisme', 'minimalism', dan 'surealisme'. Performa terbaik terlihat pada kelas 'surealisme', sementara kelas 'minimalism' tampak menjadi yang paling sulit dibedakan oleh semua model. Hasilnya, dapat disimpulkan bahwa model Inception memiliki banyak prediksi yang True Positive atau prediksinya benar

## **5 Conclusion**

Proyek ini bertujuan untuk membandingkan model untuk klasifikasi lukisan pada 4 tipe lukisan dengan menggunakan model CNN, Resnet, dan Inception. Hasil menunjukkan bahwa model Inception memiliki performa yang cukup meyakinkan dengan akurasi akhir 74% dan loss 65% serta rendahnya indikasi overfitting. Hal ini juga diperkuat dengan perbandingan confusion matrix Proyek ini menunjukkan pentingnya pemilihan model yang sesuai dalam membuat suatu perangkat klasifikasi. Model Inception mampu belajar dengan optimal walau ukuran dataset kecil. Sebaliknya model ResNet membutuhkan dataset yang lebih banyak untuk menghasilkan output yang lebih baik. Sehingga mempunyai dataset yang besar untuk melatih model juga mempunyai pengaruh signifikan pada pembuatan algoritma klasifikasi.