

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Di era digital saat ini, teknologi berkembang pesat dan merambah berbagai sektor, termasuk seni. Salah satu inovasi yang dilakukan adalah terciptanya teknologi kecerdasan buatan (AI), yang dapat melakukan pengenalan citra dan deteksi objek. Salah satu teknologi yang menarik perhatian adalah Vision Transformer (ViT), yakni *deep learning* yang dirancang untuk tugas pengenalan visual dalam mengidentifikasi gambar. Dalam dunia seni, Vision Transformer berpotensi membantu mengotomatisasi proses pengklasifikasian karya seni. Penelitian ini berfokus pada penerapan Vision Transformer untuk otomatisasi kurasi aliran seni pada karya seni 2D yang diunggah ke situs galeri digital, dengan tujuan untuk mempermudah proses kurasi seni dalam konteks digital.

Proses kurasi karya seni cukup memakan waktu, para kurator bertugas untuk menyatukan seniman, pasar, publik, dan karya. Proses kurasi pameran seni secara umum meliputi tahap riset, penentuan tema atau konsep, pemilihan karya seni, penataan ruang pameran, dan publikasi (Permana, 2021). Dengan semakin banyaknya karya seni yang dihasilkan secara digital, proses kurasi juga harus menggunakan pendekatan secara digital. Oleh karena itu, diperlukan alat untuk membantu mengatasi masalah ini. Salah satu solusi adalah dengan memanfaatkan teknologi kecerdasan buatan, salah satunya Vision Transformer, untuk mengotomatisasi proses kurasi karya seni 2D berdasarkan karakteristik visualnya.

Seiring dengan meningkatnya jumlah karya seni digital, galeri seni juga perlu beradaptasi dengan pendekatan kurasi yang lebih modern dan efisien. Salah satu tantangan utama dalam galeri digital adalah mengklasifikasikan karya seni berdasarkan aliran atau gaya seninya secara cepat dan akurat. Vision Transformer (ViT) hadir sebagai solusi yang menjanjikan dengan kemampuannya dalam menganalisis dan mengenali

pola visual secara otomatis. Dengan memanfaatkan arsitektur berbasis transformer yang unggul dalam pemrosesan citra, ViT mampu melakukan kurasi karya seni secara handal, sehingga mendukung efektivitas pengelolaan galeri digital di era modern.

Vision Transformer (ViT) adalah salah satu model deep learning yang dikembangkan berdasarkan arsitektur transformer, yang awalnya digunakan dalam pemrosesan bahasa alami (Leonardi & Chandra, 2024). Vision Transformer (ViT) menawarkan pendekatan baru dalam pengenalan gambar dengan memecah gambar menjadi *patch* kecil (Iliadis et al., 2022), mirip dengan cara kata-kata dalam kalimat diproses oleh transformer dalam pemahaman teks. Pendekatan ini memungkinkan Vision Transformer (ViT) untuk mengenali pola dan fitur visual dengan lebih akurat dibandingkan model konvolusi tradisional (Schaerf et al., 2024). Dalam konteks kurasi seni, Vision Transformer (ViT) memberikan landasan teori yang kuat karena kemampuannya untuk mempelajari hubungan kompleks antara elemen visual pada gambar. Dengan menerapkan Vision Transformer (ViT), penelitian ini memanfaatkan model dalam mendekripsi dan mengklasifikasikan karya seni 2D berdasarkan karakteristik visual, sehingga dapat lebih efektif dalam otomatisasi proses kurasi seni. Vision Transformer (ViT) meningkatkan akurasi dalam mengenali aliran seni, dan mempercepat proses kurasi dalam skala besar, hal ini akan menjadikan alat yang relevan untuk dunia seni digital.

Penelitian ini bertujuan untuk membuat sebuah galeri digital dengan otomatisasi kurasi seni yang dapat mengklasifikasikan aliran seni pada karya seni 2D secara efisien menggunakan Vision Transformer (ViT). Dengan menerapkan Vision Transformer (ViT) dalam kurasi seni, sistem ini diharapkan dapat meningkatkan kecepatan dan konsistensi proses klasifikasi. Tujuan utama penelitian ini adalah menghasilkan model yang mampu mengenali karakteristik visual khas dari berbagai aliran seni yang umum diketahui, seperti realisme, kubisme, romantisme, impresionisme, atau ekspresionisme, dan secara otomatis mengkategorikannya berdasarkan aliran tersebut. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi penting dalam dunia seni digital, terutama

dalam galeri digital untuk mengelola karya seni dengan lebih efektif dan tepat.

Meskipun teknologi kecerdasan buatan (AI) telah diterapkan dalam berbagai bidang, penerapan Vision Transformer (ViT) dalam kurasi seni masih terbilang terbatas. Banyak teknologi yang serupa dengan Vision Transformer (ViT), seperti Convolutional Neural Networks (CNN), yang masih mendominasi dalam dunia pengenalan citra. Namun, Vision Transformer (ViT) mempunyai pendekatan yang bisa dikatakan lebih canggih dengan kemampuannya untuk memecah gambar menjadi *patch* dan memahami konteks visual secara mendalam. Meskipun sudah ada penelitian awal yang menunjukkan potensi ViT dalam pengenalan pola visual, penggunaannya dalam kurasi otomatis aliran seni 2D belum banyak dijelajahi. Ini menunjukkan adanya peluang untuk penelitian lebih lanjut yang dapat mengeksplorasi dan mengembangkan aplikasi Vision Transformer (ViT) dalam konteks kurasi seni. Dengan memanfaatkan keunggulan ViT, penelitian ini tidak hanya bertujuan untuk memperkenalkan metode baru dalam klasifikasi karya seni, tetapi juga untuk mendorong adopsi teknologi ini dalam dunia seni digital yang semakin berkembang.

## 1.2 Rumusan Masalah

Bagaimana pengembangan dan implementasi ViT dapat membantu untuk mengurasi karya seni digital secara otomatis pada platform galeri digital?

## 1.3 Tujuan

Penelitian ini bertujuan untuk membuat solusi yang inovatif dalam bidang kurasi seni digital dengan memanfaatkan Vision Transformer (ViT). Tujuan utama dari penelitian ini adalah sebagai berikut:

1. Mengembangkan model Vision Transformer (ViT) yang mampu mengenali dan mengklasifikasikan karya seni 2D berdasarkan aliran seni lukis secara otomatis.

2. Mengimplementasikan sistem otomatisasi kurasi seni lukis digital menggunakan Vision Transformer (ViT) pada platform galeri digital.

#### **1.4 Manfaat**

Penelitian ini diharapkan mampu memberikan manfaat dalam pengelolaan dan kurasi karya seni digital, khususnya dalam galeri digital. Beberapa manfaat yang diharapkan adalah sebagai berikut:

1. Mempermudah proses kurasi karya seni lukis 2D secara otomatis, dan mengurangi metode manual yang memakan waktu.
2. Penyelenggara pameran dapat dengan lebih cepat mendapatkan klasifikasi atas karya yang diunggah ke galeri digital.

#### **1.5 Batasan Masalah**

Adapun batasan masalah untuk fokus dan kejelasan penelitian ini dijelaskan sebagai berikut:

1. Penelitian ini akan fokus pada karya seni lukis 2D yang tersedia dalam format digital.
2. Implementasi akan difokuskan pada integrasi Vision Transformer dalam platform galeri digital berbasis web.
3. Penelitian ini dibatasi pada pengenalan karakteristik visual khas dari berbagai aliran seni yang umum dikenal (realisme, kubisme, romantisme, impresionisme, dan ekspresionisme).
4. Aplikasi yang dikembangkan hanya membantu pada tahapan kurasi bagian pemilihan karya seni.

## 1.6 Metodologi Penelitian

Metodologi penelitian yang digunakan dalam pembuatan tugas akhir ini melibatkan beberapa tahapan. Pertama, dilakukan studi literatur terkait Vision Transformer (ViT), aliran seni lukis, dan teknologi berbasis kecerdasan buatan. Setelah itu, dilakukan pengumpulan data gambar karya seni 2D dari berbagai sumber. Data tersebut akan diproses dan dianalisis untuk diintegrasikan ke dalam model ViT

### 1.6.1 Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan di Laboratorium Data Science, STIKI Malang. Laboratorium ini memiliki fasilitas yang digunakan untuk kegiatan penelitian terkait *data science*, yang berlokasi di Kampus STIKI Malang, Kota Malang, Provinsi Jawa Timur.

**Tabel 1.1** Waktu Penelitian

Bulan	Tahun Ajaran 2024/2025
	Kegiatan
Agustus	Mencatat dan menganalisis ide yang akan dibuat
September	Riset mengenai kebutuhan dalam perancangan program dan mengumpulkan jurnal terkait dengan judul
Oktober	Menyusun tugas akhir dan membuat rancangan sistem dan <i>wireframe</i> program
November	Memulai tahap perancangan awal program

### 1.6.2 Bahan dan Alat Penelitian

Penelitian ini dilakukan menggunakan laptop dengan merek Lenovo IdeaPad Gaming 3 15IAH7 dengan sistem operasi Windows 11 64-bit. Laptop ini dilengkapi dengan Prosesor Intel 12th Gen Intel(R) Core(TM) i5, memiliki 16GB RAM, dan menggunakan Solid State Drive berkapasitas 512 TB.

### **1.6.3 Pengumpulan Data dan Informasi**

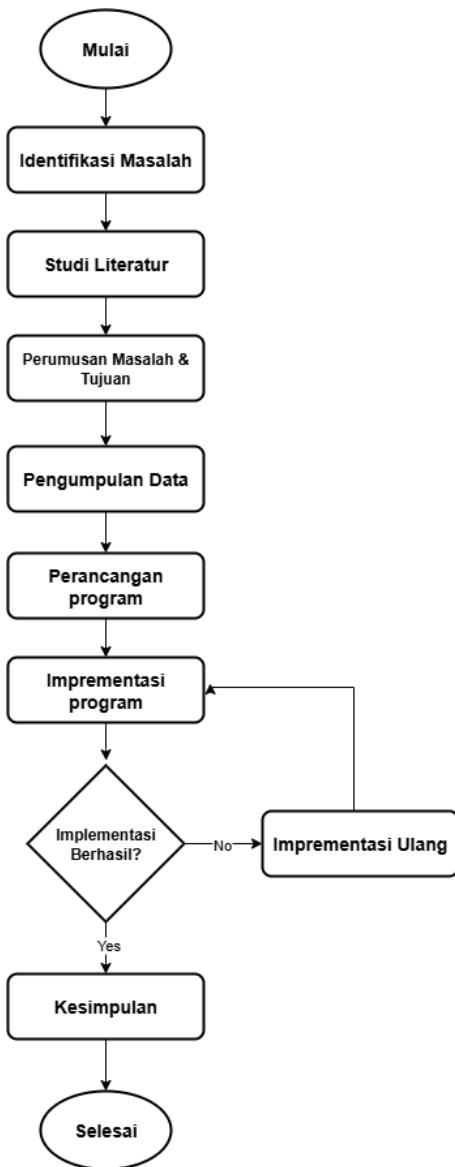
Pengumpulan data untuk penelitian ini dilakukan dengan teknik observasi menggunakan media internet dengan mengacu pada sumber jurnal yang relevan dengan topik penelitian. Seluruh proses pengumpulan data dilaksanakan oleh satu orang, yang bertanggung jawab untuk mencari, mengidentifikasi, dan menganalisis sumber-sumber informasi yang mendukung.

### **1.6.4 Analisis Data**

Penelitian ini mengimplementasikan bagaimana fitur kurasi otomatis berbasis Vision Transformer (ViT) di website galeri digital. Pengguna dapat dengan lebih mudah dalam melakukan kurasi karya seni berdasarkan aliran tertentu. Selain itu, fitur kurasi otomatis ini mempercepat proses pencarian dan penjelajahan, sehingga pengguna tidak perlu menghabiskan banyak waktu untuk mencari karya seni yang sesuai dengan preferensi mereka.

### **1.6.5 Prosedur Penelitian**

Diagram alir di atas menunjukkan tahapan-tahapan kunci dalam proses penelitian pengembangan model Vision Transformer (ViT) untuk kurasi karya seni 2D. Dimulai dari studi literatur untuk mengkaji teknologi dan metode yang relevan, dilanjutkan dengan perancangan model serta platform galeri digital. Setelahnya, data karya seni dikumpulkan untuk melatih model ViT yang akan diimplementasikan ke dalam sistem galeri.



**Gambar 1.1** Diagram alur penelitian

## **1.7 Sistematika Penulisan**

Di dalam penulisan Tugas Akhir ini sistematika penulisan diatur dan disusun sebagai berikut:

### **BAB I PENDAHULUAN**

Pada bab ini membahas tentang Latar Belakang, Rumusan Masalah, Batasan Masalah, Tujuan dan Manfaat Penelitian, Metodologi Penelitian dan Sistematika Penulisan.

### **BAB II TINJAUAN PUSTAKA**

Pada bab ini berisi pembahasan beberapa teori dan penelitian terdahulu yang mendukung materi pokok bahasan pada Tugas Akhir ini.

### **BAB III ANALISIS DAN PERANCANGAN**

Bab ini akan membahas permasalahan yang ada dan analisa masalah serta pembahasan tentang sistem yang akan dibangun.

### **BAB IV PEMBAHASAN**

Bab ini membahas tentang proses perancangan dan konfigurasi sistem serta implementasi dan juga pembahasan tentang petunjuk teknis penggunaan sistem.

### **BAB V PENUTUP**

Bab ini berisi kesimpulan dan saran dari materi yang telah dibahas dalam Tugas Akhir ini.

## BAB II

### TINJAUAN PUSTAKA

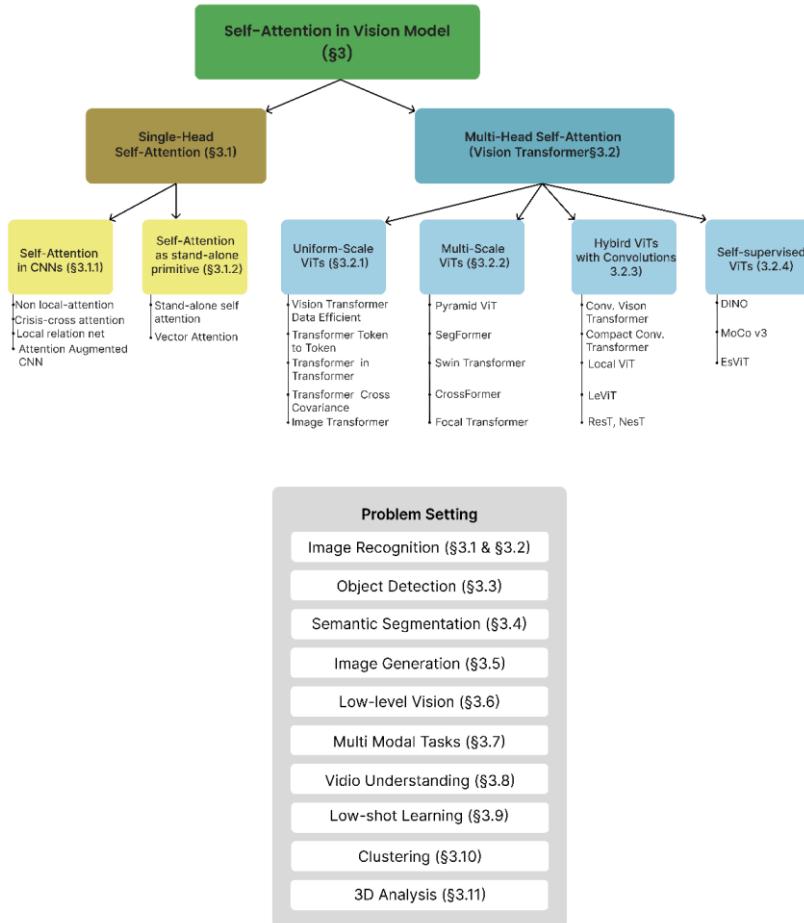
#### **2.1 Penelitian Terdahulu**

Penelitian terdahulu mengenai penggunaan teknologi untuk kurasi seni telah banyak dilakukan dengan berbagai pendekatan, termasuk penggunaan algoritma berbasis pembelajaran mesin dan visi komputer. Dalam subbab ini, akan dibahas beberapa studi yang relevan terkait dengan pengembangan kurasi seni berbasis digital, khususnya yang menggunakan model-model jaringan saraf dalam serta teknologi *Vision Transformer (ViT)*.

##### **2.1.1 *Transformers in Vision: A Survey***

Penelitian oleh Khan dkk tahun 2022 membahas *mengenai A taxonomy of self-attention design space*. *Taxonomy* ini membahas mengenai kehandalan *self-attention* dalam *vision model* berdasarkan masalah yang ditangani. Dalam penelitian ini Khan dkk membagi *self-attention* menjadi 2 jenis yakni *single-head self-attention* dan *multi-head attention*. Digambarkan pada penelitian ini, *multi-head attention (vision transformer)* merupakan metode yang cocok digunakan untuk *problem solving* yang berkaitan dengan *image recognition*. Model ini menerapkan arsitektur Transformer asli dengan sedikit perubahan pada urutan "patch" gambar yang telah diflatkan menjadi vektor (Khan et al., 2022). Dari Penelitian Khan dkk dapat disimpulkan bahwa ViT merupakan metode atau bisa disebut algoritma yang cocok dalam urusan mengolah gambar. Dalam bagan *taxonomy* yang disajikan bisa dilihat bahwa ViT masuk kedalam *problem setting image recognition*. Pemilihan literatur ini bertujuan untuk memberikan landasan teoritis bahwa penggunaan Vision Transformer (ViT) merupakan metode yang tepat untuk penelitian ini. Hal tersebut sejalan dengan temuan Khan dkk., yang menunjukkan bahwa ViT memiliki kemampuan unggul dalam mengenali pola visual dan sangat cocok digunakan untuk tugas klasifikasi gambar. Dengan mempertimbangkan hasil tersebut, penelitian ini mengadopsi ViT sebagai model utama untuk

mengklasifikasikan lima kelas aliran seni lukis, dengan harapan memperoleh akurasi dan kinerja yang optimal.



**Gambar 2.1** A taxonomy of self-attention design space

### 2.1.2 Artwork Style Recognition Using Vision Transformers and MLP Mixer

Penelitian Yang dilakukan oleh Iliadis dkk tahun 2021 ini membahas mengenai vision transformer dan *MLP mixer* dalam melakukan klasifikasi gambar terutama dalam kasus ini adalah gaya seni. Dalam penelitian ini dibantu *dataset* lukisan Wikiart yang dikumpulkan dari repositori Github ArtGan. Penelitian ini bertujuan untuk menunjukkan hasil akurasi pengklasifikasian antara *Vision Transformer (ViT)* dengan *MLP mixer* dalam menganalisis gaya seni. Manfaat penelitian ini adalah untuk menemukan metode yang paling baik untuk digunakan dalam pengklasifikasian gambar dan menentukan tolak ukur minimum dalam prediksi akurasi untuk studi di masa depan. Hasil dari penelitian ini *Vision Transformers (ViT)* dan *MLP Mixer* berhasil diterapkan pada dataset WikiArt dalam tugas pengenalan gaya artistik. Capaian akurasi *Vision Transformers (ViT)* dalam penelitian ini mencapai lebih dari 39% untuk 21 jenis gaya seni

**Tabel 2.1** Hasil klasifikasi aliran seni dengan a ViT

Class	Accuracy %
<i>Abstract Expressionism</i>	29.6
<i>Art Nouveau</i>	25.3
<i>Baroque</i>	48.3
<i>Color Field Painting</i>	65.5
<i>Cubism</i>	21.3
<i>Early Renaissance</i>	34.1
<i>Expressionism</i>	28.0
<i>Fauvism</i>	17.5
<i>High Renaissance</i>	8.5
<i>Impressionism</i>	65.0
<i>Mannerism Late Renaissance</i>	18.1
<i>Minimalism</i>	49.4
<i>Naive Art / Primitivism</i>	15.4