

# PENDETEKSIAN PLAGIASI DESAIN VISUAL MENGGUNAKAN ANALISIS CITRA DENGAN METODE SIFT

Muhammad Utbah Husnuth Thoriq<sup>1\*</sup>, Basuki Rahmat<sup>2</sup>

<sup>12</sup>*Program Studi Informatika, Universitas Pembangunan Nasional Veteran Jawa Timur, Indonesia*

\*Corresponding Author: 21081010131@upnjatim.ac.id

## Abstract

Plagiarism of visual designs, including logos, posters, and illustrations, has become a significant issue in the digital era. This study aims to develop a plagiarism detection system based on image analysis that is able to identify direct plagiarism and minor modifications to visual designs. This system integrates the Scale-Invariant Feature Transform (SIFT) method for local feature extraction and FLANN (Fast Library for Approximate Nearest Neighbors) for feature matching. The research dataset consists of a collection of original images (training data) and test images suspected of being plagiarized. The evaluation process is carried out by adjusting the threshold using the Precision, Recall, and F1-Score metrics to determine the level of system accuracy in detecting plagiarism. The results of the study show that the system is able to detect plagiarism with high accuracy, especially in images that have significant local feature similarities. The optimal threshold obtained from the evaluation allows the system to differentiate between plagiarism and non-plagiarism adaptively. This research contributes to protecting intellectual property rights in the creative industry, as well as opening up opportunities for the development of a broader image-based plagiarism detection system.

**Keywords:** Plagiarism Detection, Image Analysis, SIFT, FLANN, Threshold

## Abstrak

Plagiasi desain visual, termasuk logo, poster, dan ilustrasi, telah menjadi isu yang signifikan di era digital. Penelitian ini bertujuan untuk mengembangkan sistem pendeteksian plagiasi berbasis analisis citra yang mampu mengidentifikasi plagiasi langsung maupun modifikasi minor pada desain visual. Sistem ini mengintegrasikan metode Scale-Invariant Feature Transform (SIFT) untuk ekstraksi fitur lokal dan FLANN (Fast Library for Approximate Nearest Neighbors) untuk pencocokan fitur. Dataset penelitian terdiri dari koleksi gambar asli (data latih) dan gambar uji yang diduga plagiat. Proses evaluasi dilakukan dengan penyesuaian ambang batas menggunakan metrik Precision, Recall, dan F1-Score untuk menentukan tingkat akurasi sistem dalam mendeteksi plagiasi. Hasil penelitian menunjukkan bahwa sistem mampu mendeteksi plagiasi dengan akurasi tinggi, khususnya pada gambar yang memiliki kemiripan fitur lokal signifikan. Threshold optimal yang diperoleh dari evaluasi memungkinkan sistem untuk membedakan antara plagiasi dan non-plagiasi secara adaptif. Penelitian ini memberikan kontribusi dalam melindungi hak kekayaan intelektual dalam industri kreatif, serta membuka peluang untuk pengembangan sistem pendeteksian plagiasi berbasis citra yang lebih luas.

**Kata Kunci:** Pendeteksian Plagiasi, Analisis Citra, SIFT, FLANN, Threshold

## I. INTRODUCTION

### A. Latar belakang

Plagiasi desain visual, seperti logo, poster, dan ilustrasi, telah menjadi isu signifikan di era digital. Dalam dunia yang semakin terhubung secara global, penyebaran karya desain menjadi lebih mudah, tetapi di sisi lain, risiko plagiasi juga meningkat. Plagiasi ini tidak hanya merugikan secara finansial tetapi juga merusak kredibilitas dan hak kekayaan intelektual para desainer.

Sistem pendeteksian plagiasi yang ada saat ini sebagian besar berfokus pada teks, sedangkan plagiasi berbasis gambar masih kurang mendapatkan perhatian yang memadai. Hal ini terutama menjadi tantangan ketika desain visual telah dimodifikasi secara minor, seperti perubahan skala, rotasi, warna, atau penggabungan elemen desain. Oleh karena itu, diperlukan sistem pendeteksian plagiasi berbasis analisis citra yang mampu mengidentifikasi plagiasi langsung maupun modifikasi minor.

Penelitian ini bertujuan untuk mengembangkan sistem pendeteksian plagiasi desain visual menggunakan metode Scale-Invariant Feature Transform (SIFT) untuk ekstraksi fitur lokal dan FLANN (Fast Library for Approximate Nearest Neighbors) untuk pencocokan fitur. Evaluasi sistem dilakukan dengan mengukur kinerjanya menggunakan metrik seperti Precision, Recall, dan F1-Score. Penelitian ini diharapkan dapat memberikan kontribusi nyata dalam melindungi hak kekayaan intelektual dalam industri kreatif dan membuka peluang pengembangan lebih lanjut di bidang pendeteksian plagiasi berbasis citra.

#### *B. Rumusan Masalah*

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara mengembangkan sistem pendeteksian plagiasi desain visual yang efektif menggunakan metode analisis citra?
2. Bagaimana kinerja sistem dalam mendeteksi plagiasi langsung maupun modifikasi minor pada desain visual?
3. Berapa ambang batas optimal yang diperlukan untuk membedakan plagiasi dan non-plagiasi secara akurat?

#### *C. Tujuan*

Tujuan dari penelitian ini adalah:

1. Mengembangkan sistem pendeteksian plagiasi desain visual berbasis analisis citra menggunakan metode SIFT dan FLANN.
2. Mengevaluasi kinerja sistem dalam mendeteksi plagiasi desain visual menggunakan metrik evaluasi seperti Precision, Recall, dan F1-Score.
3. Menentukan ambang batas optimal untuk membedakan plagiasi dan non-plagiasi secara adaptif.

#### *D. Manfaat*

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

- 1) *Manfaat Akademik*
  - Menyediakan referensi baru dalam bidang pengolahan citra dan pendeteksian plagiasi desain visual.
  - Memberikan dasar untuk penelitian lebih lanjut di bidang analisis citra dan perlindungan hak kekayaan intelektual.
- 2) *Manfaat Praktis*
  - Membantu industri kreatif dalam melindungi hak kekayaan intelektual desainer dari plagiasi.
  - Memberikan solusi teknologi yang dapat digunakan oleh platform desain atau agensi untuk memverifikasi orisinalitas karya visual.
- 3) *Manfaat Teknologi*
  - Menyediakan framework sistem pendeteksian plagiasi berbasis citra yang adaptif dan extensible.
  - Membuka peluang untuk pengembangan sistem pendeteksian plagiasi yang lebih luas di berbagai platform digital.

## II. LITERATURE REVIEW

### *A. Pendeteksian Plagiasi*

Plagiasi adalah tindakan menyalin atau menggunakan karya orang lain tanpa memberikan atribusi yang tepat, yang dapat terjadi pada berbagai jenis karya, termasuk teks, gambar, dan desain visual. Dalam konteks desain visual, plagiasi sering kali melibatkan penggunaan elemen desain yang identik atau modifikasi kecil untuk menyamarkan tindakan tersebut. Sistem pendeteksian plagiasi berbasis citra telah

berkembang untuk menangani tantangan ini, terutama dengan memanfaatkan metode analisis citra yang dapat mendeteksi kemiripan pada tingkat fitur lokal maupun global.

#### B. Analisis Citra untuk Pendeteksian Plagiasi

Analisis citra merupakan cabang dari pengolahan citra digital yang fokus pada ekstraksi informasi dari gambar. Dalam pendeteksian plagiasi desain visual, analisis citra bertujuan untuk mengidentifikasi kemiripan antara gambar uji dan gambar referensi. Proses ini mencakup beberapa langkah utama:

- Ekstraksi Fitur: Mengidentifikasi karakteristik penting dari gambar, seperti tepi, sudut, atau pola.
- Pencocokan Fitur: Membandingkan fitur yang diekstraksi dari gambar uji dengan gambar referensi untuk menemukan kesamaan.
- Evaluasi Kemiripan: Mengukur tingkat kemiripan menggunakan metrik tertentu, seperti jumlah fitur yang cocok.

#### C. Scale-Invariant Feature Transform (SIFT)

SIFT adalah algoritma yang dikembangkan oleh David Lowe untuk mendeteksi dan mendeskripsikan fitur lokal dalam gambar. Algoritma ini memiliki keunggulan dalam menangani transformasi skala, rotasi, dan perubahan pencahayaan, sehingga sangat cocok untuk mendeteksi plagiasi pada desain visual yang telah dimodifikasi.

##### 1) Proses Kerja SIFT

- Deteksi Keypoint: Mengidentifikasi titik-titik penting dalam gambar menggunakan Difference of Gaussian (DoG).
- Deskripsi Fitur: Membuat representasi deskriptor berdasarkan orientasi dan intensitas piksel di sekitar keypoint.
- Pencocokan Fitur: Membandingkan deskriptor dari dua gambar untuk menemukan keypoint yang serupa.

##### 2) Keunggulan SIFT

- Invarian terhadap transformasi geometris (skala, rotasi).
- Efektif dalam menangani noise dan perubahan pencahayaan.
- Cocok untuk berbagai jenis gambar, termasuk desain visual yang kompleks.

#### D. Fast Library for Approximate Nearest Neighbors (FLANN)

FLANN adalah pustaka yang dirancang untuk pencarian tetangga terdekat secara cepat dalam ruang berdimensi tinggi. Dalam pendeteksian plagiasi, FLANN digunakan untuk mencocokkan deskriptor fitur dari dua gambar dengan efisiensi tinggi.

##### 1) Keunggulan FLANN

- Kecepatan: Lebih cepat dibandingkan metode brute-force, terutama pada dataset besar.
- Efisiensi: Menggunakan algoritma approximate nearest neighbor untuk mempercepat proses pencocokan.
- Fleksibilitas: Mendukung berbagai jenis data dan dimensi fitur.

#### E. Evaluasi Sistem Pendeteksian Plagiasi

Evaluasi sistem pendeteksian plagiasi bertujuan untuk mengukur kinerja sistem dalam mendeteksi plagiasi secara akurat. Beberapa metrik evaluasi yang umum digunakan adalah:

- 1) *Precision*: Mengukur proporsi deteksi plagiasi yang benar dari seluruh prediksi plagiasi.
- 2) *Recall*: Mengukur kemampuan sistem dalam mendeteksi semua plagiasi yang benar.
- 3) *F1-Score*: Kombinasi dari precision dan recall untuk memberikan gambaran kinerja keseluruhan.

### III. RESEARCH METHOD

#### A. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan eksperimen berbasis analisis citra. Sistem pendeteksian plagiasi desain visual dikembangkan dan diuji menggunakan dataset gambar yang terdiri dari data latih (gambar asli) dan data uji (gambar yang diduga plagiat). Proses penelitian mencakup

beberapa tahap, yaitu pengumpulan data, preprocessing, ekstraksi fitur, pencocokan fitur, evaluasi sistem, dan analisis hasil.

#### B. Alur Penelitian

Tahapan penelitian ini digambarkan dalam alur berikut:

- 1) *Pengumpulan Dataset*: Mengumpulkan gambar asli sebagai data latih dan gambar yang diduga plagiat sebagai data uji.
- 2) *Preprocessing*: Melakukan normalisasi gambar dengan resizing dan konversi ke grayscale.
- 3) *Ekstraksi Fitur*: Menggunakan metode SIFT untuk mendeteksi dan mendeskripsikan fitur lokal pada gambar.
- 4) *Pencocokan Fitur*: Menggunakan FLANN untuk mencocokkan fitur antara gambar uji dan gambar latih.
- 5) *Evaluasi Sistem*: Mengukur kinerja sistem menggunakan metrik Precision, Recall, dan F1-Score dengan penyesuaian ambang batas.
- 6) *Analisis Hasil*: Menentukan efektivitas sistem dalam mendeteksi plagiasi dan menentukan ambang batas optimal.

#### C. Pengumpulan Data

Dataset penelitian terdiri dari:

- 1) *Data Latih*: Koleksi gambar asli yang mencakup berbagai desain visual, seperti logo, poster, dan ilustrasi.
- 2) *Data Uji*: Gambar yang diduga plagiat, termasuk gambar yang identik maupun yang telah dimodifikasi (rotasi, skala, atau perubahan elemen).

Dataset diperoleh dari sumber terbuka, yaitu *Movie Poster Dataset* yang tersedia pada laman Kaggle, serta dengan tambahan yang dibuat secara manual dengan mencari desain yang diduga melakukan plagiasi terhadap karya lain.

#### D. Preprocessing Data

Tahap preprocessing bertujuan untuk memastikan konsistensi data sebelum dilakukan analisis. Langkah-langkah preprocessing meliputi:

- 1) *Resizing*: Mengubah ukuran gambar menjadi 500x500 piksel untuk konsistensi dimensi.
- 2) *Grayscale Conversion*: Mengkonversi gambar ke format grayscale untuk mengurangi kompleksitas komputasi.

#### E. Ekstraksi Fitur dengan SIFT

Pada tahap ini, fitur lokal pada gambar diekstraksi menggunakan algoritma SIFT. Proses ini mencakup:

- 1) Deteksi keypoint menggunakan Difference of Gaussian (DoG).
- 2) Pembuatan deskriptor berdasarkan orientasi dan intensitas piksel di sekitar keypoint.
- 3) Penyimpanan deskriptor untuk proses pencocokan fitur.

#### F. Pencocokan Fitur dengan FLANN

Pencocokan fitur dilakukan dengan menggunakan FLANN. Langkah-langkahnya adalah:

- 1) Membandingkan deskriptor gambar uji dengan deskriptor setiap gambar latih.
- 2) Menggunakan ratio test untuk menyaring pencocokan yang buruk.
- 3) Menyimpan jumlah fitur yang cocok untuk setiap gambar latih.

#### G. Evaluasi Sistem

Evaluasi sistem dilakukan dengan mengukur kinerja pendeteksian plagiasi menggunakan metrik berikut:

- 1) *Precision*: Proporsi deteksi plagiasi yang benar dari seluruh prediksi plagiasi.
- 2) *Recall*: Kemampuan sistem dalam mendeteksi semua plagiasi yang benar.
- 3) *F1-Score*: Kombinasi Precision dan Recall untuk memberikan gambaran kinerja keseluruhan.
- 4) *Threshold Analysis*: Menentukan ambang batas optimal berdasarkan F1-Score tertinggi.

#### H. Alat dan Perangkat yang Digunakan

Penelitian ini menggunakan alat dan perangkat berikut:

- 1) *Google Colab*: Untuk implementasi dan eksekusi program.
- 2) *Python*: Sebagai bahasa pemrograman utama.

3) *Pustaka OpenCV*: Untuk analisis citra dan implementasi algoritma SIFT serta FLANN.

4) *Matplotlib*: Untuk visualisasi hasil pencocokan dan analisis evaluasi.

#### I. *Analisis Data*

Hasil dari sistem dianalisis untuk menentukan tingkat akurasi, efektivitas, dan ambang batas optimal. Proses analisis melibatkan:

1) Menghitung metrik evaluasi untuk setiap threshold.

2) Menentukan gambar latih dengan pencocokan fitur terbaik untuk setiap gambar uji.

3) Membandingkan hasil sistem dengan ground truth untuk mengukur kinerja secara keseluruhan.

#### IV. RESULTS AND DISCUSSION

The presentation of results should be simple and straightforward. This section reports the most important findings, including results of statistical analysis as appropriate and comparisons to other research results. Results given in figures should not be repeated in tables. This is where Authors should explain in words what he/she/they discovered in the research. It should be clearly laid out and in a logical sequence. This section should be supported suitable references.

#### V. Conclusion

Summarize sentences the primary outcomes of the study in a paragraph. Explain if the claims in this section supported by the results, and if they seem reasonable. Also, describe whether the result support or contradict previous theories, and explain how the research has moved the body of scientific knowledge forward.

#### ACKNOWLEDGMENT

The authors would like to thank...

## REFERENCES

[1]