

BAB 3

METODOLOGI PENELITIAN

3.1. Desain Penelitian

Penelitian ini menggunakan metode *machine learning* untuk menganalisis dan membandingkan metadata EXIF dari foto asli dan foto yang dimanipulasi oleh AI. Model klasifikasi yang digunakan adalah *Random Forest*, dengan fitur utama yang dianalisis meliputi Make, Model, dan GPS. Data diperoleh dari kumpulan gambar asli dan gambar manipulasi.

3.1.1. Tahapan Penelitian

Berikut adalah tahapan utama dalam penelitian ini:

1. Pengumpulan Data

Dataset terdiri dari 80 foto, dengan kategori foto asli dan foto AI. Metadata EXIF diekstrak dari masing-masing gambar untuk digunakan sebagai fitur dalam model *machine learning*. Selain itu, 10 foto tambahan disertakan untuk pengujian, guna menguji kemampuan model dalam membedakan antara foto asli dan foto AI.

2. *Preprocessing Data*

Hanya metadata penting yang digunakan, yaitu *Make* (merk kamera), *model* (model kamera), dan *GPS* (lokasi pengambilan gambar) dan data yang kosong atau tidak relevan diisi atau dihapus.

3. *Feature Engineering*

Mengubah data kategori menjadi numerik menggunakan Label Encoding dan normalisasi data menggunakan *StandardScaler* agar distribusi fitur lebih seragam.

4. Pelatihan Model *Machine Learning*

Dataset dibagi menjadi training set (80%) dan testing set (20%) dan model *Random Forest* digunakan sebagai algoritma klasifikasi.

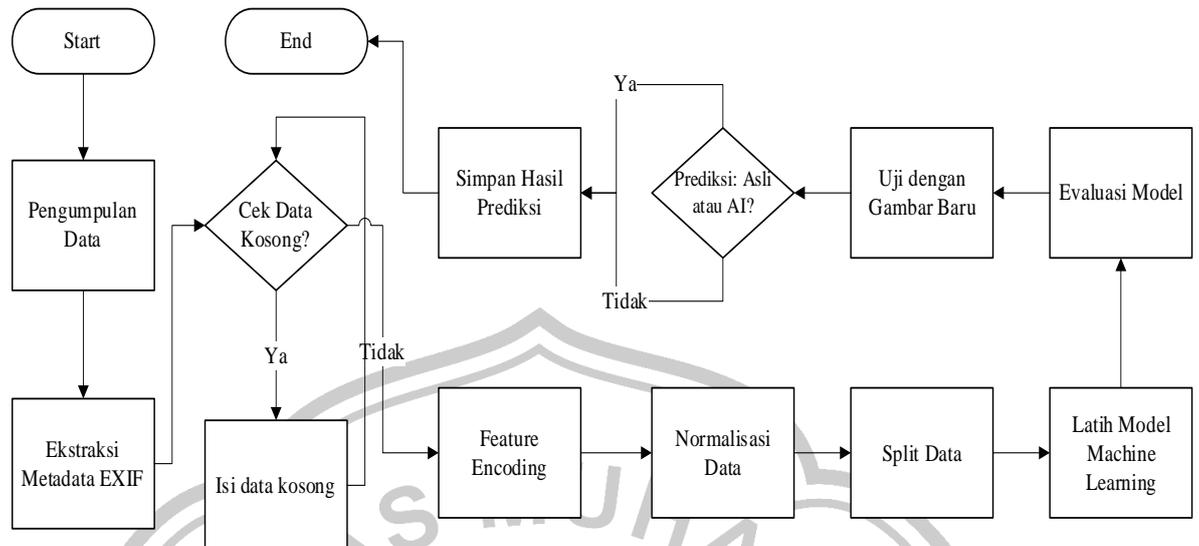
5. Evaluasi Model

Menggunakan akurasi, confusion matrix, dan classification report untuk menilai performa model.

6. Eksperimen dengan Data Baru

Menguji model dengan foto baru untuk memprediksi apakah foto tersebut

asli atau buatan AI.



Gambar 3 1 Alur Penelitian

3.2. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini terdiri dari 80 gambar, yang terbagi ke dalam dua kategori:

1. Foto Asli:

Foto-foto ini diambil langsung dari kamera smartphone tanpa melalui proses pengeditan..



Gambar 3 2 Contoh Foto Asli

2. Foto Manipulasi AI:

Foto-foto ini adalah hasil dari manipulasi menggunakan perangkat lunak berbasis AI



Gambar 3.3 Contoh Foto Manipulasi

3.3. Analisis Data

3.3.1. Preprocessing Data

Data yang diperoleh dalam bentuk metadata EXIF perlu diproses sebelum digunakan dalam model machine learning. Langkah-langkah preprocessing data meliputi:

1. **Pemilihan Fitur:** Hanya metadata yang relevan, seperti *Make*, *Model*, dan *GPS*, yang digunakan dalam analisis.
2. **Penanganan Data Kosong:** Jika metadata tidak tersedia dalam suatu gambar (terutama pada gambar AI yang cenderung tidak memiliki metadata lengkap), nilai *Unknown* akan digunakan sebagai pengganti.
3. **Konversi Format Data:** Data dikonversi menjadi format numerik agar dapat digunakan dalam model machine learning.

3.3.2. Feature Engineering

Untuk mempersiapkan data dalam format yang dapat dipahami oleh model, dilakukan langkah-langkah berikut:

1. **Encoding Data Kategorikal:** Data yang bersifat kategorikal, seperti *Make* dan *Model*, dikonversi menjadi bentuk numerik menggunakan *Label Encoding*.
2. **Normalisasi Data:** Skala data dinormalisasi menggunakan metode *StandardScaler* untuk memastikan distribusi yang lebih seragam dan meningkatkan kinerja model

3.3.3. Pelatihan Model Machine Learning

Setelah preprocessing selesai, model *machine learning* dilatih menggunakan algoritma Random Forest. Tahapan pelatihan meliputi:

1. **Split Data:** Dataset dibagi menjadi **training set (80%)** dan **testing set (20%)** untuk menghindari overfitting.
2. **Pelatihan Model:** Model dilatih dengan data training menggunakan parameter default dengan jumlah estimators sebanyak 100 pohon keputusan.

3.3.4. Evaluasi Model

Untuk mengukur performa model, dilakukan evaluasi menggunakan beberapa metrik:

1. **Akurasi:** Mengukur persentase prediksi yang benar dari model.
2. **Confusion Matrix:** Menggambarkan jumlah prediksi yang benar dan salah dalam setiap kategori (foto asli vs foto AI).
3. **Classification Report:** Menyajikan metrik evaluasi seperti precision, recall, dan f1-score untuk setiap kelas.

3.3.5. Eksperimen dengan Data Baru

Setelah model dilatih dan dievaluasi, eksperimen dilakukan dengan menggunakan gambar baru yang belum pernah dipelajari oleh model. Metadata dari gambar baru diekstrak dan diproses dengan langkah yang sama seperti data sebelumnya, kemudian digunakan untuk prediksi oleh model. Hasil prediksi akan menunjukkan apakah gambar diklasifikasikan sebagai foto asli atau foto AI.

3.4. Implementasi dan Tools yang Digunakan

Untuk menjalankan penelitian ini, digunakan beberapa alat dan pustaka sebagai berikut:

1. Google Colab: Platform berbasis cloud untuk menjalankan kode Python dengan dukungan GPU.
2. Python: Bahasa pemrograman utama yang digunakan.
3. Pustaka yang digunakan:
 1. exifread: Untuk membaca metadata EXIF dari gambar.
 2. pandas: Untuk manipulasi dan analisis data.
 3. numpy: Untuk perhitungan numerik.
 4. scikit-learn: Untuk preprocessing data dan pelatihan model machine learning.
 5. matplotlib & seaborn: Untuk visualisasi data dan evaluasi model.

