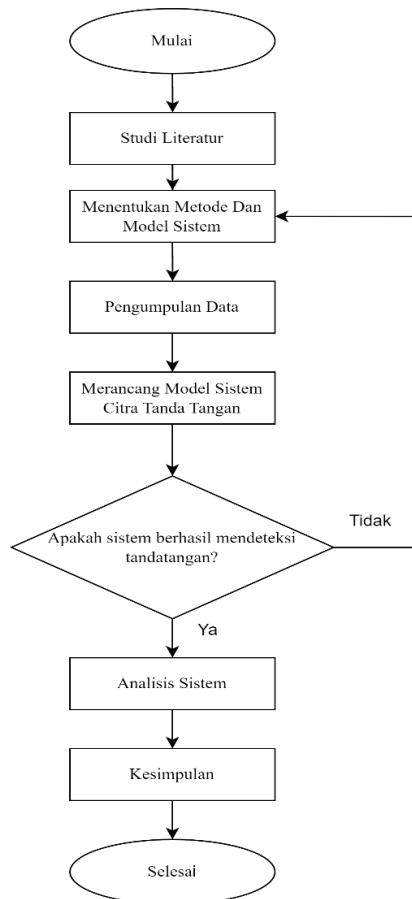


BAB III METODE PENELITIAN

Pada penelitian bab 3 ini akan dilakukan penjabaran teori-teori yang dilakukan dalam penelitian. Penelitian ini menggunakan data berupa citra tanda tangan dengan jumlah 500 gambar serta menggunakan metode PCA dan K-NN untuk pengenalan citra tanda tangan. Penelitian ini menggunakan metodologi dengan beberapa tahapan sebagai berikut.

3.1 ALUR PENELITIAN

Penelitian tugas akhir ini memiliki beberapa tahapan, agar mempermudah pada proses penelitian, maka dibuat alur penelitian seperti pada gambar 3.1 dibawah ini.



Gambar 3. 1 Diagram Penelitian

Gambar 3.1 menggambarkan proses penelitian yang diawali dengan melakukan kajian pustaka, ini merupakan tahap pertama dalam mencari berbagai informasi dan referensi untuk dijadikan acuan utama dalam melaksanakan penelitian ini. Dalam penelitian ini dilakukan beberapa kajian pustaka diantaranya yaitu Pada penelitian Riza Firdaus Ardiansyah dengan topik yang memiliki judul “Pengenalan Pola Tanda Tangan Dengan Menggunakan Metode *Principal Component Analysis (Pca)*”. Pada penelitian Dadan Nugraha dengan topik yang memiliki judul “Implementasi Algoritma Pca Pada Pengenalan Pola Tanda Tangan Dengan Menggunakan Bahasa Pemrograman *Phyton*”. Pada penelitian Mochamad Kurniawan dengan topik yang memiliki judul “Implementasi Shape Feature dan K-Nearest Neighbor untuk Klasifikasi Tanda Tangan”. Pada penelitian Lukito Hutomo dengan topik yang memiliki judul “Pengenalan Tanda Tangan Menggunakan Metode *Principal Component Analysis* dan *K-Nearest Neighbor*”.

Kemudian lanjut pada pengumpulan data dilakukan dengan mengumpulkan data-data tanda tangan dengan jumlah 500 tanda tangan dari mahasiswa D3TT08A yang berjumlah 5 orang. Dengan ketentuan tanda tangan di buat pada kertas putih dengan pulpen berwarna hitam dan di foto dengan jarak 10 cm dari gambar dengan ukuran gambar 3x4, selanjutnya gambar di kumpulkan menjadi satu pada drive yang sudah di sediakan. Kemudian selanjutnya dataset dimana data yang sudah dikumpulkan kemudian dibagi menjadi 2 yaitu data *training* dan data *testing*. Data *training* merujuk pada kumpulan data yang digunakan untuk melatih atau mengajar suatu model atau algoritma dalam pembelajaran mesin (*machine learning*), sedangkan data *testing* adalah kumpulan data yang digunakan untuk menguji performa atau mengukur kinerja model atau algoritma setelah melalui proses pelatihan menggunakan data *training*. Setelah pengumpulan dataset selesai, dilakukan perancangan sistem sesuai dengan prosedur yang telah direncanakan. Proses desain dilakukan di *Google Collab* dengan bahasa pemrograman yang digunakan adalah bahasa pemrograman *Python*.

Selanjutnya pada penelitian ini menggunakan 2 metode yaitu PCA untuk melakukan ekstraksi fitur dari data *training* dan data *testing*. Kemudian yang kedua metode K-NN pada tahap ini, setelah mendapatkan data latih dan data uji yang sudah di lakukan ekstraksi fitur menggunakan *Principal Component Analysis*, data

tersebut diklasifikasi menggunakan *K-Nearest Neighbor* untuk mengenali tanda tangan data uji berdasarkan data *training*.

3.2 ALAT DAN BAHAN

Untuk membangun sistem Pengenalan Citra Tanda Tangan Menggunakan Metode *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (K-NN), diperlukan beberapa alat dan bahan. Selanjutnya akan dijelaskan alat dan bahan yang akan digunakan dalam penelitian ini.

3.2.1 Perangkat Keras (*Hardware*)

Dalam penelitian ini menggunakan satu *hardware* untuk menjalankan *software* yang akan digunakan yaitu dengan menggunakan laptop untuk pengujian. Laptop adalah perangkat yang dapat mengolah data, menerima data, serta menghasilkan sebuah informasi sebagai keluaran. Laptop diperlukan dalam proses pengolahan data yang telah diperoleh oleh penulis dan juga digunakan dalam proses pembuatan sistem pengenalan citra tanda tangan yang akan dirancang, serta diperlukan dalam proses pengambilan data. Dalam penelitian ini juga dibutuhkan *hardware* pendukung yaitu kamera *handphone*. Spesifikasi *hardware* yang digunakan pada penelitian ini yaitu:

1. *AMD Ryzen 3 3250U @ 2.60 GHz*
2. *Operating System Windows 11*
3. RAM 4,00 GB
4. *Radeon Graphics*
5. Kamera *Handphone*

3.2.2 Perangkat Lunak (*Software*)

Dibutuhkan beberapa *software* dengan sistem operasi *Windows* dengan *tools* dan *software* dalam penelitian ini.

Tabel 3. 1 *Software*

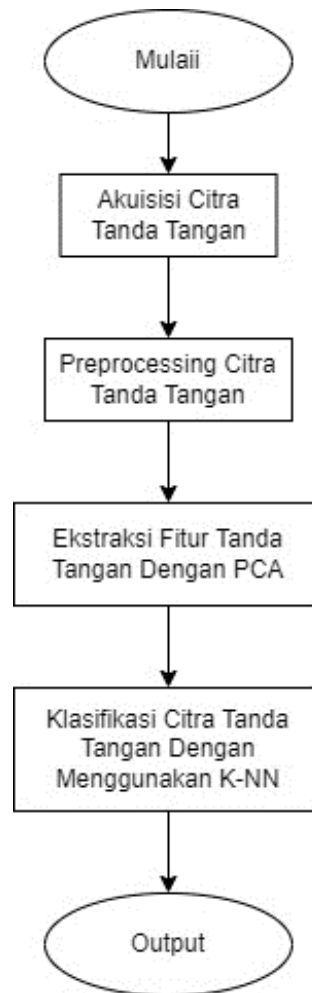
No.	Nama <i>Software</i>	Fungsi
1.	<i>Windows 11</i>	Sistem Operasi
2.	<i>Google Chrome</i>	Menjelajah Situs <i>Web</i>
3.	<i>Google Colaboratory</i>	<i>Jupyter Notebook Versi Cloud</i>

3.2.3 Dataset

Dataset adalah kumpulan data yang terorganisir secara sistematis dan biasanya digunakan untuk analisis atau pengolahan data. *Dataset* dapat berupa himpunan nilai numerik, teks, gambar, suara, atau kombinasi dari semuanya. Dalam penelitian ini digunakan *dataset* sebanyak 500.

3.3 Perancangan Sistem

Berikut ini adalah ringkasan tentang perancangan sistem untuk mengidentifikasi tanda tangan menggunakan metode *Principal Componen Analysis* (PCA) dan *K-Nearest Neighbor* (k-NN) yang dijelaskan dalam gambar *flowchart* 3.2.



Gambar 3.2 Diagram Perancangan Sistem

Pada gambar 3.2 menjelaskan perancangan sistem dimana dimulai dengan mengakuisi citra tanda tangan. Proses akuisi gambar merupakan proses pengambilan atau pengumpulan data gambar dari tanda tangan baik secara elektronik

maupun manual. Pada penelitian ini, proses mengakuisi citra tanda tangan dilakukan dengan cara *scanning* gambar tanda tangan. Proses *scanning* dilakukan menggunakan kamera *smartphone* masing masing responden.

Proses pengambilan gambar dilakukan dengan cara meletakkan tanda tangan pada kertas hvs putih menggunakan bolpoint warna hitam, dan kemudian mengambil citra tanda tangan dari jarak sekitar 15 cm. Setiap tanda tangan diambil gambarnya hingga jumlah yang dibutuhkan terpenuhi. Proses pengambilan data ini memakan waktu lebih dari satu jam.

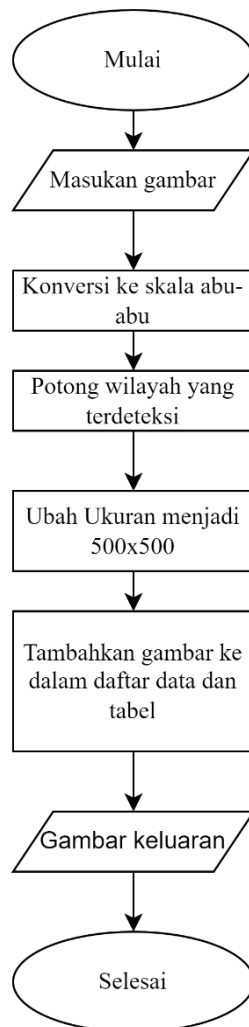
Kemudian melakukan tahap *preprocessing*, dimana pada tahap *preprocessing* ini mempersiapkan citra agar siap diproses oleh model. Pada tahap ini dilakukan pembagian folder menjadi 2 yaitu train dan test dimana dataset sebanyak 5 responden dengan jumlah masing-masing 100 tanda tangan. Untuk jumlah semua dataset sebanyak 500 tanda tangan dimana untuk pembagiannya 80 % data *training* dan 20 % data *testing*. Sehingga pada penelitian ini menggunakan data train sebanyak 400 citra tanda tangan dan data test sebanyak 100 citra tanda tangan.

Lalu setelah tahap *preprocessing* kemudian mengekstraksi data dalam bentuk gambar dari sebuah folder dan mengubahnya menjadi *array* agar bisa dibaca oleh *tensorflow*. Untuk melakukan ini, digunakan *image data generator* yang berfungsi untuk menghasilkan data dalam sebuah folder dan memungkinkan penentuan generator atau augmentasi.

Setelah proses *preprocessing* dan ekstraksi selesai, langkah berikutnya adalah menentukan model yang akan digunakan untuk sistem identifikasi citra tanda tangan, serta metode yang akan diterapkan dalam penelitian ini. Dalam penelitian ini, peneliti memilih metode PCA (*Principal Component Analysis*) dan K-NN (*K-Nearest Neighbor*) karena PCA dan K-NN dirancang khusus untuk mengolah dan memproses data dua dimensi seperti data citra. Selain itu, peneliti juga mempertimbangkan manfaat dan keunggulan sistem yang akan digunakan dibandingkan dengan sistem lain yang ada. Pada tahap ini sangat penting karena pada akhirnya bertujuan untuk identifikasi citra tanda tangan dengan akurat dalam penelitian ini.

Kemudian tahap tak kalah penting yaitu tahap melakukan training dan testing untuk menguji sistem yang telah peneliti buat. Pada tahap ini proses pengujian atau

testing data untuk melakukan pengujian model atau program dengan tujuan apakah model atau program hasilnya sesuai ekspektasi yang diinginkan oleh peneliti. Pada tahap ini ada beberapa hyperparameter yang diubah salah satunya yaitu jumlah nilai K. Proses tahap pengujian dimulai dengan melatih data pada data latih (*testing*) untuk memungkinkan model belajar dari data tersebut dan menghasilkan akurasi serta nilai recall dan precision pada data latih. Setelah itu masuk ke tahap akhir yaitu testing atau pengujian dimana pada tahap ini sistem atau model yang dibangun dievaluasi untuk menilai ketepatan dan keakuratan sistem identifikasi citra tanda tangan sesuai yang diharapkan. Hasil akhir pengujianya banyak melibatkan berbagai variasi parameter diantaranya *accuracy*, *reccal*, dan *preccision*.



Gambar 3. 3 Flowchart *Preprocessing*

Diagram flowchart gambar 3.3 menjelaskan bagaimana menggambarkan proses *preprocessing* dimana dimulai dengan menginputkan *image* kedalam sistem yang

telah dibuat. Kemudian melakukan proses *convert to grayscale* dimana proses mengonversi gambar berwarna ke citra skala abu-abu. Ini bisa dilakukan untuk menyederhanakan pemrosesan dan mengurangi dimensi warna. Lalu melakukan tahap *crop detected region* dengan mengidentifikasi dan memotong bagian yang terdeteksi dari gambar. Ini mungkin melibatkan penggunaan algoritma deteksi objek atau teknik lainnya untuk menemukan area tertentu yang relevan. Selanjutnya melakukan *resize image to 500 x 500* dengan mengubah ukuran gambar yang telah dipotong menjadi 500 x 500 piksel. Hal ini dapat dilakukan untuk membuat semua gambar memiliki dimensi yang konsisten. Tahap berikutnya melakukan *append image to the data list and label list* dengan cara menambahkan gambar yang telah diubah ukur ke dalam dua struktur data terpisah yaitu data list (daftar gambar) dan label list (daftar label atau kategori yang sesuai). Ini biasanya dilakukan sebagai persiapan untuk melatih model pembelajaran mesin. Langkah berikutnya proses output yang menghasilkan atau menampilkan atau menyimpan gambar yang telah diproses untuk memeriksa hasilnya. Ini dapat membantu dalam pemecahan masalah atau penyesuaian langkah-langkah pemrosesan.

3.4 Metode pengujian

Dalam penelitian ini hanya menggunakan satu metode pengujian yaitu dengan metode pengujian sistem.

3.4.1 Metode pengujian sistem

Pengujian sistem merupakan pemeriksaan terhadap keseluruhan sistem yang telah dirancang agar dapat mengetahui apakah sistem beroperasi secara akurat atau tidak. dengan keluaran *training* model yang akan diuji yaitu rata-rata *accuracy*, *precision* dan *recall* yang didapatkan [8]. Pada pengujian ini menggunakan metode KNN dimana untuk menentukan nilai K biasanya ditentukan dengan bilangan ganjil (3, 5, 7) sehingga proses klasifikasi tidak menunjukkan jumlah jarak yang sama. Dalam kasus di mana dua atau lebih kelas dianggap sama, nilai k akan menjadi k-1 (jarak antara satu kelas tidak terlalu dekat), nilai k-2 (jarak antara kelas yang sama lagi), dan seterusnya. Kelas memiliki nomor yang sama. Secara umum, dampak kebisingan pada klasifikasi akan berkurang seiring dengan nilai k yang lebih tinggi, tetapi batas kelas akan menjadi lebih samar begitu juga sebaliknya. Untuk hal ini ada pendekatan sederhana untuk menentukan nilai K yaitu : $k\sqrt{n}$, dimana n adalah jumlah dari sampel data yang ada. Contohnya terdapat 30 sampel data, gunakan

rumus untuk menentukan nilai k, jadi $K\sqrt{n} = \sqrt{30} = 5,47$ berarti nilai k = 5. Setelah penentuan nilai K, kemudian melakukan perhitungan yang dikalah penting yaitu nilai jarak *euclidean*, dimana *euclidean* merupakan ukuran jarak yang dapat digunakan untuk menentukan kemiripan objek dan mengklasifikasikan jarak berdasarkan kedekatan objek tetangga. Persamaan untuk menghitung euclidean sebagai berikut : $d(x,y) = \sqrt{\sum (\omega_j - \omega_i)^2}$, dimana $d(x,y)$ = jarak euclidean, ω_j = bobot citra uji, dan ω_i = bobot citra latih [10].