

BAB III

METODE PENELITIAN

3.1. Studi Pustaka

Studi pustaka merupakan metode pengumpulan teori serta informasi yang dilakukan peneliti. Studi pustaka dilakukan dengan cara membaca, mengutip, mengolah dan mengekstrak teori dan informasi yang relevan mengenai pembuatan sistem. Sumber studi pustaka yang digunakan oleh peneliti berasal dari buku, jurnal dan situs online terkait dan membahas penelitian mengenai klasifikasi, deep learning dan algoritma CNN. Sumber-sumber yang telah dikumpulkan tersebut, kemudian dipelajari oleh penulis untuk diterapkan pada penelitian ini.

3.2. Populasi dan Sampel

Populasi dalam penelitian ini adalah gambar kue tradisional yang terdiri dari gambar kue dadar gulung, kue klepon, kue lapis, kue lumpur, kue risoles dan kue serabi yang diambil dari laman <https://atapdata.ai> dan kue cucur, kue onde-onde yang dikumpulkan secara mandiri melalui penelusuran *google search*. Kemudian total sampel yang digunakan dalam penelitian ini adalah 1845 data gambar dan 8 kelas yang sesuai dengan kategori gambar kue.

3.3. Definisi Variabel Penelitian

Variabel yang digunakan dalam penelitian ini adalah kue dadar gulung, kue cucur, kue klepon, kue lapis, kue lumpur, kue kue onde-onde, kue risoles dan kue serabi. Dalam pengujian nantinya, data gambar ini akan dibagi kedalam 3 kelompok yaitu data train, data test dan data validasi.

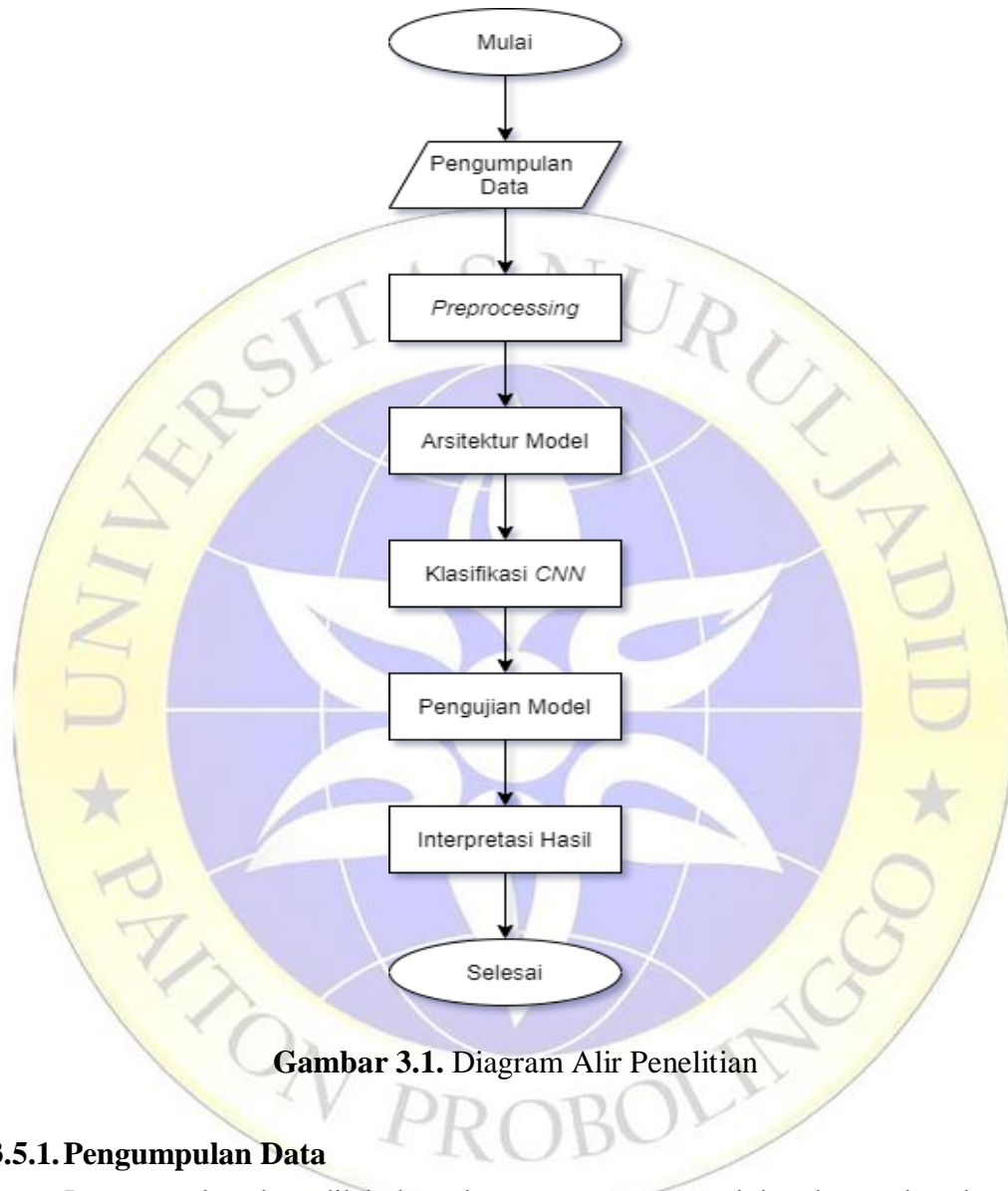
3.4. Metode Analisis

Metode analisis pada penelitian ini menggunakan *deep learning* dengan algoritma *Convolutional Neural Networks* (CNN), bahasa pemrograman, *tools* dan *software* yang digunakan dalam penelitian ini adalah sebagai berikut:

- 1.1 *Python* versi 3
- 2.1 *Tensorflow* versi 2
- 3.1 Peramban *Google Chrome*
- 4.1 *Google Collab*

3.5. Tahapan Penelitian

Tahapan yang dilakukan pada penelitian ini dapat digambarkan dalam **Gambar 3.1** berikut:



Gambar 3.1. Diagram Alir Penelitian

3.5.1. Pengumpulan Data

Pengumpulan data dilakukan dengan cara mengunduh sekumpulan dataset kue tradisional dari laman <https://atapdata.ai> yang diunggah oleh *Remko Weingarten* pada tanggal 01-12-2020, dengan lisensi *Creative Commons*. Dalam dataset tersebut, terdapat 1845 data gambar kue tradisional dengan total ukuran 317,95 MB dan ukuran *pixel* dataset yang masih bawaan (*original*) atau belum dinormalisasikan. Ada dua kue tradisional yang tidak dipakai dalam dataset yaitu kue cucur dan kue putri salju.

Ada dua jenis dataset yang nantinya akan digunakan pada penelitian ini.

1. Dataset Original

Dataset original merupakan set data bawaan atau *default* berdasarkan unduhan dari situs <https://atapdata.ai>, dimana pada dataset ini terdiri dari 8 kue tradisional. Dari 8 kue ini, hanya akan digunakan 6 kue tradisional saja kue kastengel dan kue putri salju tidak akan dipakai.

2. Dataset Modifikasi

Dataset modifikasi merupakan set data yang dikumpulkan secara mandiri melalui penelusuran *google* untuk mengganti dua dataset yang tidak dipakai pada dataset original, kue tradisional yang diambil dari google ini yaitu kue cucur dan kue onde-onde.

Setelah pengumpulan dataset selesai dilakukan maka pada kelas atau label keseluruhan ada 8 jenis objek kue yaitu Kue Klepon, Kue Lumpur, Kue Cucur, Kue Onde-onde, Kue Serabi, Kue Dadar Gulung, Kue Lapis dan Kue Risoles.

3.5.2. Preprocessing

Tahapan *preprocessing* (pra-pemrosesan) berguna untuk menyiapkan data citra kue agar dapat digunakan dengan optimal. Yaitu dengan cara *me-resize* semua ukuran citra agar sama. Karena citra kue dari dataset yang telah dikumpulkan memiliki ukuran piksel yang berbeda-beda. Kemudian dilanjutkan dengan proses *rescaling* merubah nilai pixel antara 0-255 menjadi 0-1 agar mempermudah ketika proses *input training*.

3.5.3. Arsitektur Model CNN

Arsitektur model yang digunakan adalah model *MLP* dengan banyak lapisan. Pada *API Keras* terdapat fungsi untuk memanggil pembuatan model yaitu model *Sequential*. Dengan model *Sequential* cukup mudah untuk membuat jaringan syaraf dengan banyak lapisan (Keras, 2018).

3.5.4. Klasifikasi CNN

Data citra kue yang telah dinormalisasi berikutnya sudah siap untuk dilakukan proses klasifikasi dengan algoritma *CNN*. Proses klasifikasi dilakukan pada lapisan terakhir model. Citra *input* selanjutnya akan diproses melalui konvolusi dan proses *pooling* pada tahapan *feature learning*. Jumlah proses konvolusi berdasarkan lapisan konvolusi pada model *CNN*. Jumlah filter dan ukuran kernel memiliki ukuran berbeda-beda pada setiap proses konvolusi. Kemudian dilakukan proses *flatten* atau proses mengubah *feature map* hasil *pooling layer* ke dalam bentuk vektor. Proses ini biasa disebut dengan tahap *fully Connected layer*.

3.5.5. Pengujian Model

Pengujian model akan dilakukan setelah mendapatkan model dari proses *training* oleh *CNN* untuk mengevaluasi model yang dibuat agar dapat mengenali datasetnya. Proses pengujian model ini merupakan proses memasukkan file citra kue dari data testing yang sudah tersedia pada dataset ke dalam *fit()* model yang nantinya menghasilkan sebuah *output* dan dihitung bobot yang telah ditingkatkan. Kemudian akan dibandingkan dengan bobot model dari *data train*, sehingga model klasifikasi *data train* dapat mengklasifikasikan data uji tersebut ke label atau kelas yang sesuai.

3.5.6. Interpretasi Hasil

Pada bab interpretasi hasil merupakan pembahasan hasil akurasi dari beberapa pengujian dengan hasil akurasi. (Google Machine Learning Crash Course, 2020) Hasil akurasi merupakan salah satu metrik untuk mengevaluasi model klasifikasi. Sistem klasifikasi akan mendapat *output* hasil berupa akurasi dari proses pengujian model. Secara informal, akurasi adalah fraksi prediksi model kita yang benar. Secara formal, akurasi memiliki fungsi berikut:

$$Akurasi = \frac{PB}{Jumlah\ Data} \dots\dots\dots (3.1)$$

Keterangan:

$\sum PB$ = Prediksi Benar (jumlah data gambar yang berhasil di prediksi)

$\sum Jumlah\ Data$ = Jumlah total data gambar yang di prediksi.