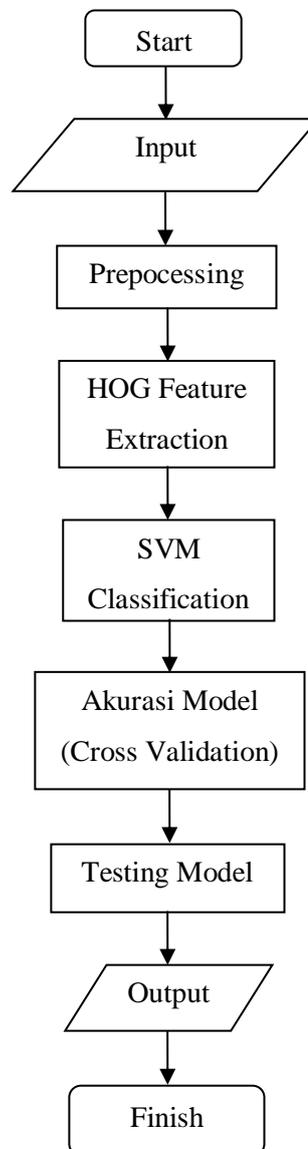


BAB III METODOLOGI PENELITIAN

3.1 Alur Penelitian

Alur penelitian merupakan konsep atau tahapan kerja pada suatu penelitian. Penelitian ini menggunakan metode *Machine Learning* dengan algoritma *Support Vector Machine* untuk melakukan klasifikasi dan *Histogram of Oriented Gradients* sebagai fitur ekstraksi citra. Berikut representasi diagram alir dari penelitian ini.



Gambar 3.1 Alur Penelitian

3.1.1 *Input Image*

Tahapan pertama pada penelitian ini adalah melakukan *input dataset*, yaitu berupa *image* atau citra. *Dataset* citra yang digunakan dalam penelitian ini, didapatkan dari *Kaggle*. *Dataset* berisi sebanyak 829 data yang terbagi dalam 30 kelas tanaman pertanian, yaitu *sugarcane* (tebu), almond, pisang, cardamom, ceri, cabe, *clove*, kelapa, kopi, kapas, ketimun, kacang, gram, jowar, jute, lemon, maize, *mustard-oil*, *olive-trees*, pepaya, *pearl-millet*, nanas, *rice*, kedelai, *sunflower*, teh, *tobacco-plant*, *tomato*, *vigna-radiati* dan *wheat*. *Dataset* dari *Kaggle* ini dilakukan pengolahan kembali sehingga *dataset* hanya memiliki 2 kelas saja, yaitu *sugarcane* dan *others*. Seluruh sampel citra ini akan dilakukan proses *resize*, konversi RGB menjadi citra abu, atau yang biasa dikenal dengan *image processing*, dan kemudian citra diekstraksi dengan fitur HOG agar dapat dijadikan sebagai parameter masukan pada SVM pelatihan. Karena SVM linier hanya dapat memprediksi objek penelitian, berdasarkan hasil fitur HOG dan model SVM pada proses pelatihan [5].

3.1.2 *HOG Feature Extraction*

Setelah dilakukan *image processing*, tahap berikutnya adalah melakukan ekstraksi fitur dengan *Histogram of Oriented Gradients* (HOG). Langkah pertama pada HOG adalah melakukan konversi citra dari RGB ke *gray*. Setelah melakukan konversi citra RGB ke *gray*, selanjutnya adalah penentuan jumlah *bin*, *pixel per cell*, dan *cells per block* (*orientations* = 9, *pixels_per_cell* = (8,8), *cells_per_block* = (2,2)) [10].

3.1.3 *SVM Classification*

Tahap selanjutnya pada penelitian ini, setelah melakukan ekstraksi fitur HOG yaitu melakukan klasifikasi dengan algoritma *Support Vector Machine* (SVM). Pada tahapan inilah algoritma *Support Vector Machine* bekerja sebagai *classifier* untuk menentukan objek yang dideteksi adalah

jenis tanaman pertanian apa menggunakan SVM terlatih. Namun sebelum melakukan klasifikasi, satu hal yang penting dilakukan yaitu, pembagian data (*Data Splitting*). *Data Splitting* adalah membagi *dataset* menjadi set pelatihan (*training set*) dan set pengujian (*testing set*). Tujuan dari proses ini adalah untuk memastikan bahwa model yang dibangun tidak hanya “menghafal” data yang telah dilihatnya, tetapi juga mampu melakukan generalisasi pada data yang belum pernah dilihat sebelumnya.

Setelah dilakukan pembagian data, ada satu proses lagi yang perlu dilakukan, yaitu normalisasi fitur. Normalisasi fitur ini dilakukan menggunakan *StandardScaler* dari *Scikit-learn*. Selanjutnya, setelah melalui 2 proses tersebut, data gambar akhirnya melakukan klasifikasi *Support Vector Machine* (SVM) dengan *kernel linear* dan penyetelan parameter terbaik menggunakan *GridSearchCV*.

3.1.4 Cross Validation

Cross Validation atau validasi silang merupakan suatu *model validation* yang digunakan untuk memprediksi atau memperkirakan keakuratan model saat dijalankan (*running*). *k-fold cross validation* merupakan salah satu dari *cross validation* yang bekerja dengan cara memecah data menjadi beberapa bagian dengan ukuran yang sama dan secara berulang melakukan pelatihan dan pengujian dengan bagian yang berbeda [11].