

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Padi dengan nama latin *Oryza Sativa* merupakan salah satu komoditas pangan yang penting untuk dunia terutama negara-negara Asia. Hal tersebut dapat terlihat pada statistik konsumsi beras di dunia pada tahun 2021/2022 dengan urutan pertama China sekitar 154.890 ton metrik. Disusul oleh India, Vietnam, Bangladesh, dan Indonesia yang menempati posisi kelima [1]. Pada tahun 2021, Indonesia sendiri mengalami penurunan produksi beras dan luas panen sekitar 0.14 juta ton untuk produksi beras dan 0.25 hektar untuk luas panen dengan membandingkan data dari tahun 2020 [2].

Penurunan produksi padi pada tahun 2021 di Indonesia, disebabkan oleh berbagai faktor. Dilansir dari CNBC Indonesia [2], penurunan produksi padi disebabkan oleh kemarau tinggi pada bulan Agustus sampai September, bencana banjir pada awal tahun, erupsi gunung semeru, dan serangan hama. Dilansir pada artikel mengenai salah satu penyakit pada padi, dikatakan bahwa Rice Blast bertanggung jawab atas hilangnya hasil produksi padi sekitar 10-30% di dunia [4]. Penurunan produksi padi akibat faktor-faktor tersebut, tentu menimbulkan berbagai jenis penyakit pada tanaman padi. Pada website Rice Knowledge Bank [5], terdapat sekitar 19 penyakit yang dapat menyerang tanaman padi. Penyakit-penyakit tersebut menimbulkan berbagai gejala yang berujung terjadinya penurunan kualitas tanaman padi bahkan sampai kepada gagalnya panen.

Penyakit tanaman padi yang beragam jenisnya, dapat menyulitkan petani dalam melakukan identifikasi terhadap penyakit tersebut. Identifikasi penyakit pada tanaman padi oleh kebanyakan petani, dilakukan secara manual. Petani perlu mengandalkan pengalaman, pengetahuan, serta kemampuan visual untuk mengidentifikasi penyakit tanaman padi. Namun cara ini kurang efisien, memakan waktu, dan memungkinkan kesalahan identifikasi [6]. Dikarenakan jenis penyakit tanaman padi yang banyak, berdasarkan penelitian yang dilakukan oleh R.

Manavalan [7], petani memiliki pengetahuan yang kurang dalam identifikasi penyakit tanaman padi. Kurangnya pengetahuan serta banyaknya jenis penyakit tanaman padi, dapat menimbulkan telatnya penanggulangan dan bahkan kesalahan dalam mengatasi penyakit pada tanaman padi.

Peningkatan teknologi dan pengetahuan yang pesat, memunculkan berbagai cara yang lebih efektif, tepat, dan cepat dalam mengklasifikasi penyakit pada tanaman padi. Teknologi tersebut adalah *deep learning* dan juga *ensemble learning*. *Deep learning* dapat digunakan untuk mengekstrak fitur, analisa serta mempelajari gambar agar dapat mengklasifikasi penyakit pada tanaman padi. Sedangkan untuk *ensemble learning* memiliki fungsi yang sama dengan *deep learning*, tetapi dapat memberikan performa yang lebih baik. *Deep learning* dan *ensemble learning* dapat diimplementasikan pada ponsel para petani sehingga dapat digunakan untuk klasifikasi penyakit pada tanaman padi dengan lebih cepat dan tepat. Berdasarkan penelitian yang dilakukan oleh R. Manavalan [7], padi yang sudah matang sulit untuk disembuhkan. Sehingga dapat dikatakan klasifikasi dengan menggunakan teknologi tersebut kurang berguna. Namun, jika melihat dari sudut pandang yang lain, petani dapat mengklasifikasi penyakit padi yang sudah matang dan dapat digunakan sebagai pembelajaran bagi petani ke depannya untuk mempersiapkan diri dalam menghadapi penyakit serupa, misal perbaikan dalam praktik budidaya, penggunaan pupuk dan pestisida yang tepat. Tidak hanya bagi petani, penggunaan *deep learning* pada penelitian ini juga dapat berguna bagi peneliti di bidang ilmu pertanian. Penelitian ini dapat digunakan menjadi referensi untuk analisis maupun diterapkan pada teknologi lain seperti *Internet of Things* untuk peningkatan kualitas padi.

Penggunaan *deep learning* sudah banyak dilakukan untuk mengklasifikasi berbagai penyakit pada tanaman termasuk tanaman padi. Penelitian yang dilakukan oleh Paddy Doctor (2022) [6] menggunakan metode *deep learning* dengan memakai *pre-trained* model yaitu VGG16 dan MobileNet untuk mengklasifikasi penyakit tanaman padi menggunakan dataset yang mereka ambil sendiri dengan akurasi terbaik yaitu 93,83%. Namun, penelitian ini tidak menjelaskan secara detail analisa model yang digunakan dan hanya menggunakan *pre-trained* model dan

model CNN buatan sendiri. Terdapat juga penelitian klasifikasi penyakit padi lainnya yang dilakukan oleh Fakhri Habib Hawari, et al [8]. Penelitian ini menggunakan model CNN buatan sendiri untuk mengklasifikasi penyakit pada tanaman padi dengan akurasi *testing* 86%. Namun, akurasi yang didapat masih bisa ditingkatkan lagi dengan mencoba metode *ensemble*. Dari kedua penelitian tersebut membuktikan bahwa penggunaan *deep learning* untuk klasifikasi penyakit tanaman padi dapat dilakukan baik menggunakan CNN buatan sendiri maupun *pre-trained* model dengan akurasi yang baik. Sedangkan untuk penggunaan *ensemble learning* dapat terlihat pada penelitian yang dilakukan oleh Odd Virgantara Pura, et al [14], yang menggabungkan *pre-trained* model yaitu DenseNet dan MobileNet dengan teknik *ensemble*. Performa yang didapat lebih baik daripada penggunaan *pre-trained* model itu sendiri. Sehingga dari penelitian tersebut, penggunaan *ensemble learning* dapat meningkatkan performa dari penggunaan *deep learning*. Namun dikarenakan dataset yang digunakan hanya terdiri dari 4 kelas, bisa menjadi faktor utama yang menyebabkan performa meningkat.

Dari latar belakang, permasalahan, serta beberapa penelitian yang telah dipaparkan, penulis akan meneliti performa setiap model yang digunakan pada penelitian ini dan membandingkannya, baik antara model yang diteliti dan penelitian terdahulu. Model yang digunakan adalah dua *pre-trained* model yaitu EfficientNetV2B0 dan MobileNetV3-Large serta dua model yang dihasilkan dari penggunaan dua metode *ensemble* yaitu *average ensemble* dan *concatenation ensemble*. Alasan penggunaan kedua *pre-trained* model tersebut adalah karena memiliki parameter yang sama-sama kecil dengan performa yang bagus [10] [11] dan tidak memerlukan *resource* yang besar atau model ringan ketika diimplementasikan pada sistem *progressive web app*. Selain itu penggunaan kedua model ini dilatarbelakangi pada penelitian terdahulu yang memberikan saran untuk mencoba menggunakan model terbaru untuk meningkatkan performa. Selanjutnya, alasan penggunaan kedua metode *ensemble* tersebut adalah karena merupakan metode yang sederhana, tidak perlu melakukan konfigurasi parameter, kompatibilitas dengan berbagai model, serta adanya *library* keras yang menyediakan *function* untuk memudahkan penulis dalam menggunakan metode

*ensemble*. Selain itu, penggunaan dua metode *ensemble* ini juga didasari oleh penelitian terdahulu yang berhasil membuktikan peningkatan performa yang dibandingkan dengan *pre-trained* model.

Dataset yang digunakan pada penelitian ini, menggunakan dataset sekunder yang diambil pada penelitian terdahulu yaitu dataset Paddy Doctor [9]. Latar belakang penggunaan dataset ini adalah tempat pengambilan gambar padi berada di India bagian selatan yang masuk ke iklim tropis mirip seperti iklim di Indonesia. Sehingga memiliki penyakit tanaman padi yang mirip dan dimungkinkan hasil penelitian dapat diimplementasikan di Indonesia. Dataset terdiri dari sembilan kelas penyakit tanaman padi, yaitu: Bacterial Leaf Blight, Bacterial Leaf Streak, Bacterial Panicle Blight, Blast, Brown Spot, Dead Heart, Downy Mildew, Hispa, dan Tungro. Penulis juga akan menambah satu kelas tanaman padi yang tidak berpenyakit atau normal. Dataset yang didapat bersifat *imbalanced* sehingga akan melewati proses *balancing* yang akan menghasilkan dataset dengan teknik *undersampling* dan dataset dengan teknik *oversampling*. Penelitian ini juga menggunakan dua ukuran gambar yang berbeda yaitu  $224 \times 224$  *pixel* dan  $256 \times 256$  *pixel*. Alasan penggunaan dua ukuran gambar tersebut karena pada penelitian Paddy Doctor, menggunakan ukuran gambar  $256 \times 256$  *pixel* dan untuk ukuran *default* pada model *pre-trained* yang digunakan adalah  $224 \times 224$  *pixel*. Sehingga pada penelitian ini akan melihat pengaruh ukuran mana yang lebih baik, yang akan diuji pada dua *pre-trained* model dan dua teknik *ensemble* serta juga akan melihat pengaruh dari penggunaan dua teknik *balancing* dataset yang berbeda.

Penelitian tidak hanya meneliti performa setiap percobaan model, tetapi juga meneliti kecepatan rata-rata waktu prediksi ketika model diimplementasi ke sistem Progressive Web App (PWA). Penggunaan PWA telah banyak diimplementasi pada website misalnya Youtube, Tokopedia, dan lainnya. Dikarenakan keuntungan utama penggunaan PWA adalah aksesibilitas lintas *platform* yang membuat PWA dapat diakses pada Android, iOS, Windows, dan website. Dengan adanya sistem PWA yang dibuat, penulis dapat meneliti kecepatan rata-rata waktu prediksi untuk melihat model yang memiliki kecepatan prediksi terbaik. Sehingga dapat dijadikan referensi untuk penelitian lainnya dalam menentukan model yang ingin dipakai.

Alasan lain penggunaan PWA adalah penulis tidak perlu berurusan dengan App Store yang memerlukan biaya agar dapat diinstal pada ponsel secara publik.

## 1.2 Identifikasi Masalah

Berdasarkan latar belakang yang sudah penulis paparkan, identifikasi masalah yang didapat terdiri dari beberapa poin yaitu:

- 1.2.1. Apakah penggunaan metode *pre-trained deep learning* dan *ensemble learning* dapat meningkatkan performa pada dataset klasifikasi penyakit tanaman padi?
- 1.2.2. Apakah metode *balancing* dataset yaitu *undersampling* dan *oversampling* mampu meningkatkan performa model dan pengaruh terhadap biaya komputasi serta waktu pelatihan?
- 1.2.3. Apakah ukuran gambar yaitu *224x224 pixel* dan *256x256 pixel* berpengaruh terhadap model yang akan diuji?
- 1.2.4. Apakah perbedaan model, dataset, dan ukuran gambar mempengaruhi waktu pemrosesan dalam klasifikasi penyakit tanaman padi pada sistem *progressive web app*?

## 1.3 Batasan Penelitian

Batasan dari penelitian ini adalah sebagai berikut:

- 1.3.1. Penelitian ini menggunakan dataset publik Paddy Doctor [9]
- 1.3.2. Penelitian ini memakai model yang telah dilatih atau *pre-trained* model yang disediakan oleh TensorFlow dan Keras
- 1.3.3. Sistem tidak dapat mengklasifikasi foto padi yang diambil pada malam hari.
- 1.3.4. Penelitian ini tidak menggunakan kelas non-padi untuk membedakan antara padi atau bukan. Sehingga penelitian hanya mampu membedakan 9 penyakit tanaman padi dan tanaman padi normal.
- 1.3.5. Sistem *progressive web app* yang dibuat hanya sebatas *prototype* dalam artian sistem tidak seluruhnya *deploy* di *Internet*.
- 1.3.6. Penelitian ini tidak melakukan validasi ke petani atau *experts* pada bidang pertanian.

#### 1.4 Tujuan Penelitian

Penelitian ini dilakukan dengan tujuan membandingkan performa beberapa model yang memiliki akurasi terbaik dalam mengklasifikasi penyakit pada tanaman padi baik dengan dataset *undersampling* maupun *oversampling* atau dengan ukuran gambar  $224 \times 224$  *pixel* dan  $256 \times 256$  *pixel*. Selain itu menentukan model yang memiliki performa kecepatan pemrosesan terbaik jika diimplementasikan pada sistem progressive web app.

#### 1.5 Manfaat Penelitian

Penelitian ini memiliki beberapa manfaat sebagai berikut:

1.5.1. Dapat digunakan sebagai sumber untuk penelitian lainnya sebagai tolak ukur model yang memiliki akurasi terbaik serta pemrosesan yang tercepat jika diimplementasikan pada *progressive web app*.

1.5.2. Dapat digunakan sebagai sumber acuan untuk jurnal lainnya dalam pengaruh akurasi baik dengan dataset *undersampling* maupun *oversampling*.

1.5.3. Dapat digunakan sebagai sumber acuan dalam penggunaan metode *ensemble learning* yang dapat meningkatkan akurasi model.

1.5.4. Meningkatkan penggunaan *progressive web app* sebagai salah satu alternatif dalam membangun aplikasi mobile dan website pada satu kode program.

1.5.5. Membantu petani dalam mengklasifikasi penyakit pada tanaman padi mereka untuk meningkatkan produksi padi serta dapat digunakan pada penelitian sejenis dan dapat diimplementasikan pada sistem lainnya seperti IoT.

#### 1.6 Sistematika Penulisan

Sistematika penulisan laporan penelitian disusun dan dibagi atas lima bab sebagai berikut.

1.6.1. BAB 1 PENDAHULUAN

Pada Bab ini berisi latar belakang masalah yang ingin diteliti, identifikasi permasalahan, batasan masalah yang akan diteliti, tujuan penelitian, manfaat penelitian, dan sistematika dalam penulisan laporan.

#### 1.6.2. BAB 2 LANDASAN TEORI

Pada Bab ini menjelaskan landasan penelitian terdahulu yang mendukung penelitian yang dilakukan penulis serta teori yang akan digunakan peneliti dalam meneliti permasalahan seperti Convolutional Neural Network (CNN), *Deep Learning*, *Ensemble Learning*, MobileNetV3-Large, EfficientNetV2B0, Karakteristik Penyakit Tanaman Padi, dan Implementasi pada sistem PWA.

#### 1.6.3. BAB 3 ANALISIS DAN PERANCANGAN SISTEM

Pada Bab ini berisi metode penelitian, perancangan modul, dan perancangan aplikasi.

#### 1.6.4. BAB 4 HASIL DAN ANALISA PENELITIAN

Pada Bab ini berisi hasil dari penelitian yang telah dilakukan serta menganalisa hasil penelitian.

#### 1.6.5. BAB 5 SIMPULAN DAN SARAN

Pada Bab ini berisi simpulan serta saran dari penelitian yang telah dilakukan yang nantinya dapat menjadi referensi untuk civitas akademika lainnya untuk mengembangkan penelitian mereka.