

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Paru-paru merupakan suatu organ yang penting bagi manusia. Oleh karena itu kita harus menjaganya dari berbagai penyakit. Penyakit paru-paru merupakan penyakit yang serius dan jika tidak ditangani dengan cepat maka akan berakibat fatal dapat menyebabkan kecacatan atau bahkan kematian. Beberapa penyakit paru-paru yang akan dibahas diantaranya adalah covid-19, *pneumonia* virus, *pneumonia* bakteri, dan tuberkulosis. Penyakit covid-19, merupakan penyakit menular yang disebabkan oleh virus Sars-Cov-2. Penyebaran virus ini dapat melalui droplet (percikan air) dan juga airborne (melalui udara), sehingga virus ini dapat menyebar dan menginfeksi manusia [1]. Penyakit pada paru-paru merupakan penyakit nomor 3 yang menyebabkan kematian setelah penyakit jantung dan stroke, salah satu penyakit paru-paru tersebut yaitu penyakit *pneumonia* [2]. Kejadian pneumonia cukup tinggi di dunia, yaitu sekitar 15% -20%. Pada usia lanjut angkakejadian pneumonia mencapai 25 -44 kasus per 1000 penduduk setiap tahun. Insiden pneumonia komunitas akan semakin meningkat seiring dengan penambahan usia, dengan 81,2% kasus terjadi pada usia lanjut. Penderita pneumonia komunitas usia lanjut memiliki kemungkinan lima kali lebih banyak untuk rawat inap dibandingkan dengan penderita pneumonia komunitas usia dewasa [3]. Penyakit Tuberkulosis (TB) merupakan penyakit infeksi kronis menular yang masih merupakan masalah kesehatan masyarakat di dunia. Indonesia salah satu negara yang berkontribusi besar dalam menyumbang kasus TB di dunia. Data WHO (2020) menunjukkan bahwa pada tahun 2018-2022 sebanyak 40 juta jiwa terancam TBC. Indonesia menempati urutan kedua (8,5%) setelah India dan china [4]. Kematian akibat tuberkulosis pada

populasi dengan status HIV negatif adalah 44/100.000 penduduk [5]. Diagnosa penyakit paru-paru menggunakan citra *x-ray* sering dilakukan karena terbukti akurat, namun terdapat faktor-faktor yang dapat menyebabkan citra *x-ray* menjadi buruk. Faktor-faktor tersebut diantaranya yaitu peralatan yang tidak mumpuni, kesalahan operator, dan kelainan pasien. Hal tersebut dapat menyebabkan citra *x-ray* memiliki detail yang tidak memadai serta kontras dan pencahayaan yang rendah sehingga terdapat kemungkinan hal tersebut juga dapat mempengaruhi akurasi [6].

Pada penelitian Tugas Akhir ini dilakukan klasifikasi citra *x-ray* dengan 5 kelas berbeda, yakni normal, covid-19, *pneumonia* bakteri, *pneumonia* virus dan tuberkulosis. Untuk meningkatkan akurasi pada proses klasifikasi, Peneliti mengusulkan penggunaan *image enhancement* dengan metode *Gamma Correction* dan *Contrast Limited Adaptive Histogram Equalization* (CLAHE) pada proses *pre-processing*. Arsitektur yang digunakan adalah *knowledge destination* dengan *Vision Transformer* (ViT) sebagai *teacher* dan *mobile-net* sebagai *student*. Arsitektur dipilih dikarenakan ViT memiliki kelebihan tingkat akurasi yang tinggi dengan waktu komputasi yang lebih sebentar, model arsitektur tanpa jaringan konvolusional, dan khasiat transformer dengan dengan *path* kecil. Seperti yang dijelaskan oleh Alexey pada [7]. Pada tugas akhir ini peneliti menganalisis perbandingan antara hasil proses klasifikasi dengan pembaharuan citra menggunakan metode *image enhancement* dengan citra asli. Proses *image enhancement* dilakukan bertujuan untuk memperbaiki citra *x-ray* agar kualitas gambar yang dihasilkan lebih baik dari sebelumnya.

## 1.2 Penelitian Sebelumnya

Penelitian mengenai penanganan penyakit paru-paru menggunakan *x-ray* telah banyak dilakukan menggunakan pendekatan *deep learning* [8], [9]. Pada survey [10] mengenai penelitian indentifikasi penyakit covid-19 pada citra *x-ray* mendapatkan akurasi paling tinggi dengan 2 kelas yaitu 99,05%, untuk nilai akurasi prediksi model dengan 4 kelas mendapatkan nilai akurasi 91,08%. Pada penelitian

[11], mendeteksi 2 kelas dan 3 kelas klasifikasi menggunakan CNN arsitektur *mobilenetV2* dengan jumlah masing masing dataset untuk 2 kelas 728 citra mendapatkan tingkat akurasi 96,78%, untuk 3 kelas 1.442 citra mendapatkan tingkat akurasi 94,72%. Penelitian [8], mampu mendapatkan tingkat akurasi mencapai 98% dengan menggunakan arsitektur CNN ResNet50, ResNet101, ResNet152, InceptionV3 dan Inception-ResNetV2 dengan mengklasifikasikan empat kelas (normal, pneumonia virus, pneumonia bakteri, dan covid-19). Pada [9], melakukan penelitian klasifikasi citra *x-ray* menggunakan arsitektur CNN CovNet dengan 3 kelas (normal, *pneumonia* bakteri, dan *pneumonia* virus) mendapatkan tingkat akurasi tertinggi di angka 90,02%. Pada tahun 2020, juga dilakukan penelitian pada [12] menggunakan 2 kelas yaitu paru-paru normal dan pneumonia dengan menggunakan metode SVM dan ekstraksi GLCM (Gray Level Co-Occurrence). Penelitian ini menggunakan dataset yang didapatkan dari website dengan total 5853 citra rontgen paru-paru. Penelitian ini bertujuan untuk mengklasifikasikan hasil X-Ray apakah terdapat pneumonia atau tidak pada hasil X-Ray tersebut dan mendapatkan akurasi sebesar 62,66% untuk dataset testing sebanyak 600 dan 59,2% untuk dataset testing sebanyak 750. Penelitian selanjutnya yaitu pada [13]. Penelitian ini menggunakan 2 kelas yaitu normal dan pneumonia dengan menggunakan metode KNN dan ekstraksi *Gabor Filter* serta *Wiener Filter* dan mendapatkan performa yang baik pada fitur ini. Penelitian ini mendapatkan akurasi sebesar 91,15% untuk nilai precision pada paru-paru normal, dan akurasi sebesar 93,63% untuk nilai recall pada paru-paru Pneumonia. Penelitian ini menggunakan dataset pada website *Kaggle* dan mendapatkan 5863 data yang terdiri dari gambar yang terkena *pneumonia* dan yang tidak terkena *pneumonia*. Namun yang digunakan hanya 3140 yang akan dipilih secara acak. Penelitian lain menggunakan dataset ChestX-ray8 untuk mengklasifikasi berbagai jenis penyakit toraks dengan metode CNN [14]. Berdasarkan penelitian yang dilakukan, kasus pneumonia dengan model usulan ResNet-50 memperoleh akurasi 63,33%. Pada penelitian [15] dua model convolutional neural network digunakan

yaitu, CNNet dan ResNet yang digunakan untuk mengklasifikasikan citra X-Ray paru-paru dengan penyakit pneumonia, COVID-19, dan normal. Dengan mengatur dataset ke dalam tiga folder (pelatihan 60%, tes 20%, validasi 20%) yang berisi subfolder untuk setiap kategori gambar (300 gambar COVID-19, 350 normal, dan 350 gambar pneumonia virus). Berdasarkan pengujian yang telah dilakukan, akurasi model arsitektur ResNet dan CNNet masing-masing adalah 97,61% dan 94,28%. Didapatkan bahwa performa model ResNet lebih baik daripada model CNNet dalam mengklasifikasikan COVID-19 dan pneumonia. Pada penelitian [16] dua model convolutional neural network digunakan yaitu, Xception dan VGG-16 untuk mendiagnosis pneumonia. Hasil pengujian yang dilakukan pada penelitian tersebut menunjukkan bahwa VGG-16 network melebihi Xception network dengan akurasi masing-masing 87% dan 82%. Namun, Xception network mencapai hasil yang lebih baik dalam mendeteksi kasus pneumonia. Hasilnya, dapat disimpulkan bahwa setiap model memiliki kemampuan khusus pada kumpulan data yang sama. Penelitian selanjutnya pada tahun 2020 [17] dengan menggunakan metode K-Means Clustering dan Support Vector Machine. Penelitian ini mengelompokkan bakteri TBC ke dalam tiga kelas, yaitu positif, negatif dan scanty. Penelitian ini menghasilkan tingkat akurasi sebesar 88% untuk K-Means Clustering dan sebesar 96% untuk Support Vector Machine. Penelitian selanjutnya pada [18] tentang klasifikasi citra berupa X-Ray paru-paru secara otomatis dengan berbagai model, seperti menggunakan Pre-trained model Convolutional Neural Network yang berfokus pada penyakit malar Tuberculosis dan Normal. Dengan menggunakan beberapa model (ResNet50, DenseNet121, MobileNet, Xception, InceptionV3, InceptionResNetV2) dan didapatkan hasil akurasi tertinggi sebesar 91,57DenseNet121. Akan tetapi, model ini hanya menggunakan dua kelas saja dan jumlah dataset yang sedikit.

### 1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka dapat ditarik kesimpulan masalah yang akan dibahas pada Tugas Akhir adalah sebagai berikut:

1. Sulitnya menganalisis *x-ray* dalam mendiagnosis jenis penyakit paru-paru jika dilakukan secara manual.
2. Kualitas citra *x-ray* kurang baik sehingga memperlambat proses diagnosis penyakit.
3. Teknik *image enhancement* dengan model *gamma correction* dan CLAHE yang digunakan apakah efektif dalam meningkatkan hasil akurasi proses klasifikasi.

### 1.4 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah untuk merancang sebuah arsitektur yang digunakan untuk proses klasifikasi citra *x-ray* paru-paru dengan 5 kelas yang diusulkan peneliti (*covid-19*, *pneumonia virus*, *pneumonia bakteri*, tuberkulosis, dan normal)

Adapun yang akan didapatkan :

1. Merancang dan mengimplementasi sistem yang dapat melakukan klasifikasi berbagai penyakit paru-paru (*tuberkulosis*, *pneumonia virus*, *pneumonia bakteri*, dan *covid-19*) dan normal menggunakan metode *knowledge distillation*.
2. Merancang dan mengimplementasi sistem *image pre-processing* menggunakan metode *gamma correction* dan *Contrast Limited Adaptive Histogram Equalization* (CLAHE).
3. Menganalisa pengaruh performa sistem klasifikasi terhadap hasil akurasi dengan penggunaan *image enhancement* pada citra *x-ray*

Manfaat dari penelitian ini membantu dokter/ahli radiologi dalam hal mendiagnosis penyakit paru-paru secara dini karena penggunaan sistem berbasis komputasi.

### 1.5 Batasan Masalah

Untuk memudahkan dan membatasi pembahasan masalah yang ada pada tugas akhir ini, maka diberikan batasan-batasan sebagai berikut:

1. Metode yang digunakan dalam proses klasifikasi adalah *knowledge distillation*.
2. *Image enhancement* yang digunakan *gamma correction* dan CLAHE.
3. Klasifikasi berdasarkan 5 kelas yaitu 4 penyakit paru-paru (tuberkulosis, *pneumonia* virus, *pneumonia* bakteri, dan covid-19) dan normal.
4. Format file dataset menggunakan \*.jpg dan \*.png.
5. Dataset bersifat publik.
6. Perancangan sistem menggunakan *platform Google Colab* dengan bahasa pemrograman *python* dan *framework tensorflow*.

### 1.6 Metode Penelitian

#### 1. Studi Literatur

Pada tahap studi literatur dilakukan dengan mempelajari penyakit paru-paru dan metode klasifikasi *knowledge distillation* serta *image enhancement gamma correction* dan CLAHE.

#### 2. Pengumpulan data

Pada tahap ini peneliti mengumpulkan dataset dari berbagai sumber yang bersifat publik yang bertujuan untuk menjadikan sampel dataset penelitian. Dataset diambil dari internet seperti *kaggle*. Dataset yang dipakai berjumlah 5

kelas yaitu normal, covid-19, *pneumonia* bakteri, *pneumonia* virus, dan tuberkulosis. Total citra *x-ray* pada dataset berjumlah 12.337.

### 3. Perancangan sistem

Pada tahap ini dilakukan kegiatan menganalisis dan merancang kebutuhan peneliti dalam merancang sistem untuk menyelesaikan permasalahan. Sistem yang digunakan menggunakan metode *knowledge distillation* dengan arsitektur vit sebagai *teacher* dan *mobilenet* sebagai student.

### 4. Implementasi sistem

Pada tahap ini dilakukan implementasi pada metode *knowledge distillation* yang diawali dengan proses *image enhancement* menggunakan metode *gamma correction* dan CLAHE menggunakan *platform google colabs*.

### 5. Pengujian

Dalam tahap ini dilakukan analisis performa yang dihasilkan oleh arsitektur yang diusulkan oleh peneliti dalam proses klasifikasi citra *x-ray* paru-paru terhadap 4 penyakit dan 1 normal.

### 6. Kesimpulan

Penelitian ini dilakukan dengan proses klasifikasi 5 kelas menggunakan metode *knowledge distillation* dengan membandingkan dataset asli dan dataset dengan penambahan *image enhancement*.

## 1.7 Sistematika Penulisan

Untuk sistematika penulisan tugas akhir adalah :

### 1. BAB II Konsep Dasar

Pada bab 2 dibahas tentang dasar teori dari citra digital, *x-ray*, covid-19, *pneumonia*, tuberkulosis, *image enhancement*, CLAHE, *gamma correction*, *knowledge distillation*, *vision transformer*, dan *mobile-net*.

## 2. BAB III Model Sistem dan Perancangan

Pada bab 3 ini membahas tentang deskripsi sistem, desain sistem, pelatihan model, dan parameter performansi sistem.

## 3. BAB IV Hasil dan Analisis

Pada bab 4 ini membahas tentang hasil pengujian dari beberapa parameter yang dilakukan pada penelitian.

## 4. BAB V Kesimpulan dan Saran

Pada bab 5 ini membahas tentang kesimpulan yang dapat diambil dari penelitian ini dan saran apa yang akan dilakukan untuk mengembangkan penelitian selanjutnya.