

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Coronavirus disease 2019 (COVID-19) adalah penyakit menular yang disebabkan oleh virus bernama Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) [1]. Pada tanggal 31 Desember 2019, World Health Organization (WHO) mendapatkan informasi pertama tentang penyakit yang dimaksud sebagai kluster pneumonia dengan sebab yang belum diketahui di kota Wuhan, Republik Rakyat Tiongkok [2]. Virus penyebab penyakit baru tersebut awalnya dikenal dengan nama 2019 novel Coronavirus (nCoV), hingga ditetapkan sebagai klasifikasi dan penamaan virus yang dimaksud sebagai virus SARS-CoV-2 oleh International Committee of Taxonomy of Viruses (ICTV) [3]. Kemudian pada 30 Januari 2020, WHO mendeklarasikan wabah COVID-19 sebagai *Public Health Emergency of International Concern* (PHEIC) atau darurat kesehatan global [4]. Selanjutnya karena wabah sudah menjalar secara luas ke berbagai negara di dunia, WHO akhirnya menetapkan bahwa COVID-19 adalah sebuah pandemi global per tanggal 11 Maret 2020 [5]. Karena tingkat keterjangkitannya yang tinggi, maka dibutuhkan diagnosis dengan hasil yang akurat untuk menentukan apakah seseorang positif mengidap COVID-19. Deteksi awal bisa dilakukan dengan mengacu pada tes dengan teknik *swab sampling* seperti rRT-PCR dan Ag-RDT. Tetapi, untuk mendapatkan kepastian dan juga sebagai acuan proses perawatan lebih lanjut, harus dilakukan pencitraan rongga dada sebagai pendukung diagnosis, baik dari hasil X-ray maupun hasil CT scan. Dan masalah baru muncul dengan meningkatnya pasien yang membutuhkan diagnosis berdasarkan pencitraan paru-paru terutama pada penderita COVID-19.

Proses diagnosis paru-paru, termasuk dari CT toraks pada penelitian ini, normalnya dilakukan oleh dokter spesialis radiologi. Tetapi seiring meningkatnya jumlah pasien atau terduga pasien COVID-19, departemen radiologi rumah sakit juga mengalami peningkatan kebutuhan akan diagnosis menggunakan citra CT toraks tersebut, khususnya pada pasien dan terduga pengidap COVID-19. Namun,

dengan memanfaatkan pembelajaran mesin, komputer dapat membantu tugas diagnosis yang dimaksud. Penelitian terkait deteksi penyakit paru-paru secara umum sebelumnya pernah dilakukan terhadap penyakit kanker, stroke, dan tumor. Pada penelitian tersebut digunakan metode klasifikasi Support Vector Machine (SVM) dengan Gray-Level Co-occurrence Matrix (GLCM) sebagai metode ekstraksi fiturnya. *Classifier* yang dihasilkan mendapatkan akurasi sebesar 89% [6]. Selanjutnya penelitian dengan menggunakan SVM dan GLCM juga pernah dilakukan pada penyakit pneumonia, namun hasil terbaiknya hanya mencapai 62.66% [7]. Pada penelitian ini, SVM juga akan digunakan sebagai metode klasifikasi, hanya saja pada bagian ekstraksi fitur akan digunakan sebuah arsitektur CNN yang bernama VGG-16. Algoritma tersebut pernah digunakan untuk klasifikasi citra X-ray dan menghasilkan akurasi sebesar 75.25% [8].

VGG-16 pada penelitian lain menghasilkan akurasi yang sangat baik ketika dibandingkan dengan beberapa arsitektur *pretrained* CNN lain, dimana kesemua arsitektur yang digunakan menghasilkan tingkat akurasi serupa yaitu lebih dari 90% [9]. Pada penelitian tersebut, digunakan sebuah dataset citra CT toraks yang didapatkan dan juga tersedia secara daring [10]. Pada penelitian ini, dataset tersebut juga akan digunakan untuk keperluan klasifikasi. Data CT toraks yang digunakan terbagi menjadi dua kelas, yaitu positif COVID-19 dan negatif COVID-19. Tahapan utama untuk menghasilkan model klasifikasi adalah dengan ekstraksi fitur terhadap data lalu dilakukan pelatihan model dari fitur yang didapat. Pada proses ekstraksi fitur, VGG-16 digunakan untuk menghasilkan fitur yang dapat menjadi data masukan pada model. Lalu proses pelatihan SVM dilakukan terhadap data fitur yang telah diekstraksi untuk menghasilkan model *classifier* yang mampu melakukan klasifikasi pada data baru yang ingin diperiksa. Adapun model yang dihasilkan akan diuji dan diukur performanya menggunakan metode *confusion matrix*, sehingga didapat nilai akurasi sebagai ukuran keberhasilan model dalam melakukan prediksi. Harapannya, model SVM yang dihasilkan mampu melakukan prediksi COVID-19 dengan tingkat akurasi yang baik.

1.2. Perumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan, maka rumusan masalah penelitian ini yaitu: apakah penggunaan metode VGG-16 dan SVM dapat menghasilkan tingkat akurasi yang baik dalam mendeteksi penyakit COVID-19 dari data citra CT toraks.

1.3. Maksud dan Tujuan

Maksud dari penelitian ini yaitu untuk melakukan klasifikasi dengan metode pembelajaran mesin menggunakan fitur-fitur yang diekstraksi dari dataset citra CT toraks. Adapun tujuan dari penelitian ini adalah untuk menghasilkan *classifier* dengan tingkat akurasi yang baik dalam membedakan penderita dan non-penderita penyakit COVID-19.

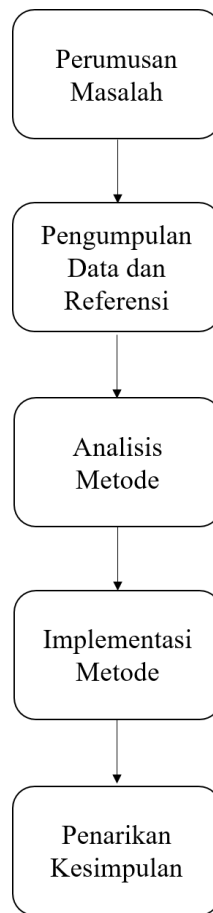
1.4. Batasan Masalah

Pembatasan masalah dalam penelitian ini bertujuan untuk menghindari penyimpangan dan pelebaran pokok masalah. Adapun batasan-batasan masalah dalam penelitian ini diantaranya:

1. Dataset citra yang digunakan berjumlah.
2. Data citra sebelumnya dilakukan ekstraksi fitur menggunakan VGG-16 dengan bantuan program Orange Data Mining, sehingga menghasilkan vektor fitur yang dapat digunakan untuk klasifikasi.
3. Model klasifikasi dibangun dengan bahasa pemrograman Python 3.

1.5. Metodologi Penelitian

Jenis penelitian yang digunakan pada penelitian ini adalah penelitian kuantitatif dengan pendekatan deskriptif. Penelitian dilakukan dengan melakukan analisis data lalu merekayasanya hingga menghasilkan suatu nilai ukur yang objektif. Adapun alur penelitian yang dilakukan pada penelitian ini dapat dilihat pada gambar 1.1.



Gambar 1.1 - Alur Penelitian

Berikut penjelasan tahapan-tahapan penelitian yang dilakukan pada penelitian ini:

1.5.1. Perumusan Masalah

Proses perumusan masalah dilakukan untuk mendapatkan inti masalah yang akan diselesaikan pada penelitian ini. Dari perumusan masalah maka dapat dirumuskan juga kerangka dasar penelitian ini. Mulai dari tahapan apa saja yang akan dilakukan beserta tujuannya, hingga akhirnya penarikan kesimpulan. Perumusan juga dilakukan berdasarkan data yang didapat dari tahapan selanjutnya.

1.5.2. Pengumpulan Data dan Referensi

Tahapan ini terbagi menjadi dua, yaitu pengumpulan data dan pengumpulan referensi. Pengumpulan data adalah tahap untuk mengoleksi data yang digunakan untuk kebutuhan analisis dan implementasi penelitian. Data didapatkan dengan mengacu pada data yang dikumpulkan dalam sebuah jurnal ilmiah tentang data citra CT toraks COVID-19. Sedangkan pengumpulan referensi dan literatur yang menjadi dasar dan penunjang penelitian sumber-sumbernya berasal dari buku, artikel, jurnal, proceeding, website, dan tugas akhir lain yang berkaitan dengan penelitian yang dilakukan.

1.5.3. Analisis Metode

Tahapan analisis dilakukan salah satunya untuk menentukan seperti apa jalannya metode-metode yang ditentukan pada penelitian ini. Dimulai dari analisis data masukan yang isinya menentukan bagaimana data selanjutnya dapat dilakukan proses ekstraksi fitur. Untuk menuju tahapan tersebut maka perlu dilakukan *preprocessing* berupa pengubahan ukuran dan warna. Analisis berikutnya adalah analisis ekstraksi fitur. Pada bagian ini dijabarkan bagaimana data yang formatnya telah diseragamkan akan melalui jaringan VGG-16 untuk menghasilkan vektor fitur yang dapat digunakan untuk pembelajaran mesin. Lalu analisis pelatihan adalah ketika vektor fitur masukan yang didapat digunakan untuk melatih model klasifikasi. Model inilah yang nanti akan diuji pada tahap berikutnya.

Selain analisis proses, dilakukan juga analisis kebutuhan, baik itu fungsional maupun non-fungsional. Pada kebutuhan non-fungsional didefinisikan lingkungan yang digunakan untuk mengembangkan dan menjalankan sistem. Termasuk didalamnya kebutuhan perangkat keras dan perangkat lunak. Sedangkan pada kebutuhan fungsional ditentukan fungsional-fungsional apa saja yang ada pada sistem yang dapat dilakukan interaksi oleh pengguna fungsional-fungsional tersebut didefinisikan didefinisikan dengan UML.

1.5.4. Implementasi Metode

Implementasi metode dilakukan dengan menuliskan kode program dalam Jupyter Notebook dengan bahasa Python. Implementasi dilakukan dengan

mempertimbangkan fungsionalitas yang telah ditentukan. Implementasi *machine learning* pada Python dilakukan dengan menggunakan *library* Scikit-learn dengan bantuan fitur olah data menggunakan Pandas. Selanjutnya model klasifikasi yang dihasilkan dari proses implementasi ini akan diuji dan diukur performanya.

1.5.5. Penarikan Kesimpulan

Pada tahap ini dilakukan pengambilan kesimpulan berdasarkan hasil pengujian yang didapatkan setelah melalui proses analisis dan implementasi yang telah dilakukan. Lalu juga akan dikemukakan saran untuk perbaikan dari kesimpulan yang didapat.

1.6. Sistematika Penulisan

Sistematika penulisan disusun untuk dijadikan kerangka acuan penulisan laporan penelitian. Sistematika penulisan laporan ini secara umum adalah sebagai berikut:

BAB 1 PENDAHULUAN

Bab 1 berisikan penjelasan tentang apa yang mendasari penelitian dan mengapa penelitian ini dilakukan. Lalu perumusan masalah dilakukan berdasarkan latar belakang yang telah diuraikan. Kemudian penjabaran singkat akan maksud dan tujuan dari penelitian ini dan juga pembatasan masalahnya. Terakhir berisikan Metodologi penelitian yang digunakan dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab 2 berisikan teori-teori yang melandasi penelitian ini, mulai dari teori yang berkenaan dengan masalah terkait, hingga teori tentang metode-metode dan *tools* yang digunakan. Semua berisikan hal yang berkaitan dengan penelitian yang dilakukan.

BAB 3 ANALISIS KEBUTUHAN ALGORITMA

Bab 3 berisikan proses-proses analisis pada penelitian, dimulai dari analisis masalah, analisis data masukan, analisis ekstraksi fitur, dan analisis pelatihan

metode SVM. Lalu juga ada analisis fungsional dan non-fungsional untuk tahap implementasi.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab 4 berisikan implementasi algoritma dengan menggunakan *library* yang menunjang dengan menggunakan bahasa Python, untuk dilakukan proses pelatihan data hingga menjadi model klasifikasi. Model yang dihasilkan selanjutnya diuji dan diukur menggunakan metrik akurasi. Sehingga dapat ditarik kesimpulan dari nilai akurasi terbaik yang didapatkan dari model yang telah dilatih.

BAB 5 KESIMPULAN DAN SARAN

Berisi kesimpulan yang didapat dari hasil pengujian serta saran untuk perbaikan agar akurasi klasifikasi dapat menjadi lebih baik.