

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Paru-paru adalah salah satu organ terpenting dalam sistem pernapasan yang berfungsi untuk memompa udara ke dalam tubuh [1]. Salah satu penyakit yang menyerang paru-paru adalah *tuberculosis* atau yang sering dikenal dengan sebutan TBC. Penyakit ini disebabkan oleh *mycobacterium tuberculosis* atau bakteri tahan asam yang menginfeksi paru-paru lebih sering daripada bagian tubuh manusia lainnya [2]. Peningkatan kasus tuberkulosis setiap tahunnya disebabkan karena masyarakat yang kurang mengetahui bahaya serta cara pengobatan dan pencegahan penyakit ini. Tindakan pencegahan untuk pengobatan penyakit paru-paru dapat dilakukan melalui beberapa pemeriksaan kesehatan apabila seseorang memiliki gejala yang mengarah pada masalah paru-paru. Salah satu pemeriksaan yang dilakukan adalah pemeriksaan *x-ray* atau *rontgen*. *X-Ray* adalah suatu metode pencitraan medis yang menggunakan radiasi elektromagnetik untuk mengambil gambar bagian dalam tubuh manusia. [3]. Hasil dari tes ini membantu dokter dalam membuat diagnosis penyakit pasien. Ada situasi ketika dokter meragukan diagnosis bahkan setelah melihat hasil *x-ray*. Oleh karena itu diperlukan suatu metode yang dapat membantu meyakinkan dokter untuk mengidentifikasi penyakit pasien berdasarkan hasil *x-ray*. Tujuannya adalah untuk menghindari hal-hal yang tidak diinginkan seperti diagnosis yang rentan terhadap kesalahan dan membutuhkan bantuan ahli [4].

Metode pemrosesan gambar dapat diterapkan untuk mengembangkan sistem yang mampu mengklasifikasikan penyakit berdasarkan citra *chest x-ray* seperti *Convolutional Neural Network* (CNN), *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN) dan berbagai metode lainnya. Metode-metode tersebut telah menunjukkan akurasi yang cukup baik dalam mengenali citra *chest x-ray* yang menunjukkan adanya infeksi tuberkulosis pada paru-paru.

Penelitian Ade Clinton Sitepu, dkk tahun 2022 tentang klasifikasikan kasus TBC dan *non-TBC* menggunakan model CNN dengan arsitektur LeNet tanpa atensi ganda (*Dual Attention*) *Convolutional Block Attention Module*

(CBAM). Nilai F1- *score* dalam deteksi tuberkulosis menggunakan citra *X-ray* tanpa atensi ganda adalah 80% sedangkan dengan atensi ganda adalah 84%. [5]. Penelitian Favorisen R, dkk tahun 2020 tentang prediksi jumlah penderita penyakit tuberkulosis menggunakan metode SVM dengan 3 kernel yaitu, *Linear*, *Gaussian*, dan *Polynomial*. Dengan hasil nilai rata-rata sebesar 51.43%, 58.53% dan 36.03%[6].

Berdasarkan permasalahan di atas dan beberapa penelitian sebelumnya, Penelitian ini akan membahas klasifikasi penyakit tuberkulosis berdasarkan citra *chest x-ray* dengan menggunakan metode CNN untuk mengatasi masalah dalam proses diagnosa oleh dokter yang rentan terhadap kesalahan dan harus dilakukan oleh dokter spesialis. Model CNN yang akan digunakan memiliki tujuh *layer* konvolusi dan *output* yang dihasilkan adalah akurasi, presisi, *recall* dan F1-Score untuk keakuratan model dalam melakukan klasifikasi citra *chest x-ray* dengan bantuan *Computer Aided Diagnosis* (CAD).

Penelitian sebelumnya telah menggunakan algoritma dengan empat *layer* konvolusi. Jumlah filter yang digunakan adalah 32 untuk lapisan konvolusi 1 dan 2 serta 64 untuk lapisan konvolusi 3 dan 4. Hasil terbaik yang diperoleh menunjukkan hasil optimal dengan parameter *epoch* 50, pengoptimasi ADAM, ukuran citra 100x100 piksel, ukuran kernel 3x3, dan skenario data 80%:20%. Model tersebut mencapai akurasi sebesar 85,4%.

Salah satu kekurangan dari penelitian tersebut adalah jumlah *layer* konvolusi yang terbatas hanya empat *layer*. Model CNN dengan lebih banyak *layer* konvolusi biasanya dapat mengekstrak fitur-fitur yang lebih kompleks dan abstrak dari gambar, yang berpotensi meningkatkan kinerja klasifikasi. Meskipun model mencapai akurasi yang cukup baik dengan mencapai 85,40% tetapi masih ada potensi untuk meningkatkan tingkat akurasi lebih lanjut. Dengan memperluas model dan melakukan eksperimen lebih lanjut pada parameter dan skenario data dapat meningkatkan performa model. Secara keseluruhan meskipun model telah mencapai akurasi yang baik ada potensi untuk meningkatkan kinerjanya dengan eksplorasi lebih lanjut terhadap arsitektur model dan parameter yang digunakan.

Tujuan dari penelitian ini adalah mencapai hasil klasifikasi yang akurat. Untuk mencapai tujuan tersebut digunakan sebuah dataset berisi 4200 citra *x-ray*

paru-paru yang diperoleh dari *National Library of Medicine* (NLM) yaitu TB *National Institute of Allergy and Infectious Diseases* (NIAID) Dataset [7]. Dataset tersebut terdiri dari dua kelas yaitu "Normal" dan "Positif Tuberkulosis" sehingga digunakan klasifikasi *categorical*. Sebelum melakukan proses *training*, dataset akan dibagi menjadi dua bagian yaitu *training set* dan *validation set*. *Training set* akan digunakan untuk melatih model sementara *validation set* akan digunakan untuk mengoptimasi model.

Dalam penelitian ini dilakukan perbandingan performa sistem model CNN dengan tujuh *layer* konvolusi dalam melakukan klasifikasi dengan menggunakan skenario yang berbeda. Skenario tersebut mencakup penggunaan *batch size* dan jenis *pooling* yang berbeda serta penambahan augmentasi data yang diharapkan dapat meningkatkan kinerja model. Dengan demikian penelitian ini bertujuan untuk mencari kombinasi terbaik dari berbagai skenario tersebut guna mencapai akurasi klasifikasi yang lebih tinggi pada citra *chest x-ray* dalam mengenali kasus Normal dan Positif Tuberkulosis.

1.2 RUMUSAN MASALAH

Rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana mengklasifikasi tuberkulosis melalui citra *x-ray* menggunakan metode CNN dengan tujuh *layer* konvolusi?
2. Bagaimana tingkat akurasi, Presisi, *Recall* dan *F1-Score* yang diperoleh pada klasifikasi tuberkulosis melalui citra *x-ray* menggunakan metode CNN dengan tujuh *layer* konvolusi?

1.3 BATASAN MASALAH

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penelitian menggunakan *dataset* yang diperoleh dari basis data terbuka untuk proses *training* dan *testing*.
2. Penelitian menggunakan *dataset* sebanyak 4200 gambar dengan pembagian dua kelas yaitu positif TB sebanyak 700 gambar dan 3500 gambar normal dengan format png.

3. Penelitian ini menggunakan metode CNN dengan tujuh *layer* konvolusi, sedangkan untuk penelitian sebelumnya menggunakan empat *layer* konvolusi.
4. Sistem klasifikasi *tuberculosis* menggunakan bahasa pemrograman *python* pada *Google Colaboratory* dengan penggunaan *Tensorflow* sebagai *Library Open Source*.
5. Pengujian data menggunakan *cross validation* yaitu 70% data sebagai *training* 10% data sebagai *validation* dan 20% data sebagai *testing*.
6. Skenario parameter yang digunakan yaitu *batch size* 32, 64 dan 128 dengan jenis *pooling* yaitu *max pooling* dan *average pooling*.

1.4 TUJUAN

Tujuan pada penelitian ini adalah sebagai berikut:

1. Melakukan perancangan sistem untuk klasifikasi penyakit tuberkulosis berdasarkan citra dengan metode CNN dengan tujuh *layer* konvolusi.
2. Mengetahui tingkat akurasi, Presisi, *Recall* dan *F1-Score* yang diperoleh pada klasifikasi penyakit tuberkulosis melalui citra *x-ray* menggunakan metode CNN dengan tujuh *layer* konvolusi.

1.5 MANFAAT

Manfaat dari penelitian ini yaitu diharapkan dapat membantu mempercepat klasifikasi penyakit dalam proses diagnosa menggunakan teknologi *Artificial Intellegence* pada bidang kesehatan. Serta membantu para dokter untuk mengurangi dan meminimalisir kesalahan diagnosis pada klasifikasi penyakit terutama pada hasil citra *chest x-ray*. Penelitian ini dilakukan untuk menindak lanjuti penelitian sebelumnya agar menjadi referensi model *deep learning* bagi penelitian selanjutnya. Dengan adanya klasifikasi tuberkulosis berdasarkan citra *chest x-ray* menggunakan CNN dengan tujuh *layer* konvolusi yang dibuat dapat meningkatkan keakuratan hasil diagnosis.

1.6 SISTEMATIKA PENULISAN

Untuk mempermudah pembahasan dalam penelitian ini, maka sistematika penulisan penelitian disusun sebagai berikut:

1. BAB I : PENDAHULUAN

Bab ini berisi mengenai latar belakang, perumusan masalah yang diangkat, batasan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan.

2. BAB II : DASAR TEORI

Bab ini membahas terkait kajian pustaka dan dasar teori yang bersangkutan dengan penelitian ini tentang *Artifitial Intellegence, Depp Learning, CNN, Convolutional layer, Pooling layer, Flatten, Fully connected layer, Relu, Sigmoid*, Definisi Citra, Definisi Citra *X-Ray*, Tuberkulosis, Augmentasi Data dan *Confusion matrix*

3. BAB III : METODE PENELITIAN

Bab ini meliputi pembahasan mengenai alat dan bahan, alur penelitian, perancangan sistem, pengujian sistem dan evaluasi.

4. BAB IV : HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil pembuatan model, hasil akurasi dari setiap iterasi dan hasil akurasi dari pengujian citra *chest x-ray*.

5. BAB V : KESIMPULAN DAN SARAN

Bab ini berisikan kesimpulan hasil dari pemodelan sistem klasifikasi dan saran untuk mengembangkan model untuk penelitian selanjutnya agar lebih baik lagi.

6. DAFTAR PUSTAKA

Halaman ini berisikan referensi yang sudah peneliti rujuk atau citasi pada kalimat dan gambar.

7. LAMPIRAN

Halaman ini berisikan sample dataset, hasil *training* model, serta keseluruhan *script* dalam pembuatan model.