

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Paru – paru merupakan salah satu organ penting dalam tubuh, Fungsi utama paru-paru adalah membantu oksigen dari udara yang manusia hirup masuk ke dalam sel darah merah. Sel darah merah kemudian membawa oksigen ke seluruh tubuh untuk digunakan dalam sel yang ditemukan dalam tubuh manusia. Paru-paru juga membantu tubuh membuang gas CO₂ saat bernapas (Cotes dkk., 2009).

Penyakit paru – paru merupakan penyakit yang serius dan jika tidak ditangani dengan cepat maka akan berakibat fatal bahkan dapat menyebabkan kematian (Anggara dkk., 2016). Pneumonia merupakan salah satu penyakit yang terjadi di paru – paru. Pneumonia menginfeksi sekitar 450 juta orang setiap tahunnya.

Balita yang baru lahir sampai anak dengan umur 6 tahun lebih rentan terhadap penyakit (Werdiningsih dkk., 2019). Sistem diagnosis penyakit anak usia dini juga memakan waktu dan rentan terhadap kesalahan (Werdiningsih dkk., 2019). Deteksi penyakit anak usia dini sangat penting karena dapat membantu dalam menurunkan angka kematian Balita di Indonesia (Yanto dkk., 2017).

Pneumonia adalah penyakit infeksi yang merupakan penyebab utama kematian pada balita di dunia. Pneumonia adalah diagnosis yang paling umum yang dijumpai pada kasus penyakit paru-paru pada balita (Gie dkk., 2017). Berdasarkan data dari *World Health Organization* (WHO), Pneumonia menyumbang 15% dari semua kematian anak di bawah 5 tahun dan menewaskan 808.694 anak di tahun 2017. Penyebab pneumonia paling umum yang terjadi pada balita dikarenakan bakteri *Streptococcus pneumoniae*. Bakteri *Streptococcus pneumoniae* dapat ditularkan melalui udara, melalui kontak langsung, baik dengan sentuhan kulit, lendir, atau luka maupun merupakan hasil turunan dari ibu yang terinfeksi penyakit pneumonia (Weiser dkk., 2018).

Diagnosis terhadap pneumonia bisa dilakukan dengan beberapa cara, yaitu dengan anamnesis, pemeriksaan fisik, dan juga pemeriksaan penunjang. Pemeriksaan penunjang yang paling sering dilakukan adalah melalui pencitraan,

yaitu foto rontgen dada. Namun, mendeteksi penyakit Pneumonia melalui foto rontgen dada merupakan hal yang sulit oleh ahli radiologi dan harus dianalisa secara manual (Gordienko dkk., 2019). Dalam proses analisa, ada beberapa faktor yang dapat menyulitkan ahli radiologi seperti hasil foto yang kurang jelas, terjadinya tumpang tindih dengan diagnosis lainnya atau ada kesamaan ciri – ciri dengan penyakit lainnya yang mengakibatkan terjadinya kebingungan (Rajpurkar dkk., 2017). Faktor – faktor tersebut dapat mengakibatkan variabilitas yang cukup besar di antara ahli radiologi dalam diagnosis pneumonia (Neuman dkk., 2012).

Machine learning telah terbukti berhasil sebagai alat untuk memperkuat keakuratan diagnostik pasien pneumonia yang dirawat di rumah sakit, mengidentifikasi pasien yang cocok untuk perawatan di rumah dan mengurangi biaya perawatan kesehatan serta memprediksi kematian pada pasien yang dirawat di rumah sakit (Naydenova dkk., 2016). *Data Mining* adalah sebuah bidang yang banyak berkembang dari *Machine Learning*. *Data mining* merupakan bidang ilmu yang digunakan untuk menangani masalah pengambilan informasi dari *database* dan menggabungkan teknik dari statistik, *machine learning*, visualisasi data, dan pengenalan pola (Werdiningsih dkk., 2020). Salah satu metode yang digunakan pada data mining adalah klasifikasi. Klasifikasi merupakan salah satu teknik dari *supervised machine learning* yaitu teknik menetapkan kelas target ke objek atau kelompok yang berbeda (Kumar dkk., 2017). Klasifikasi terdiri dari dua langkah yaitu konstruksi model, yang digunakan untuk mengevaluasi dataset pelatihan dari sebuah *database* dan penggunaan model, dimana model yang dibangun digunakan untuk klasifikasi (Kumar dkk., 2017).

Dalam beberapa tahun terakhir, kecerdasan buatan telah berkembang pesat baik industri dan akademisi. Banyak aplikasi seperti klasifikasi objek, pemrosesan bahasa, dan pengenalan ucapan, yang dimana sebelumnya tidak dapat dilakukan oleh manusia, namun saat ini dapat terwujud (Sahiner dkk., 2019). *Machine learning* adalah bagian dari kecerdasan buatan dan *deep learning* merupakan salah satu bidang dari *machine learning*. *Deep Learning* mencoba untuk memproses data dalam jumlah besar seperti gambar, suara, dan teks dengan menggunakan *Deep Neural Networks* (DNNs) (Yu dkk., 2011). Di dalam *Deep Learning* ada tiga arsitektur dasar jaringan saraf yang berkinerja baik pada berbagai jenis data yaitu *Feed Forward Network* (FFNN), *Recurrent Neural Network* (RNN), dan

Convolutional Neural Network (CNN) (Singh dkk., 2017). CNN memiliki kemampuan yang kuat dalam pengenalan gambar dan telah terbukti sebagai alat yang baik untuk menilai karakteristik batas dan warna. Metode CNN paling umum digunakan dalam analisis citra biomedis karena kemampuannya yang luar biasa dalam menganalisis informasi spasial khususnya di bidang pengolahan gambar (Cao dkk., 2018).

Penelitian ini menggunakan data sekunder yang diambil dari website Kaggle.com. Foto rontgen dada (anterior-posterior) diperoleh dari kohort retrospektif pasien anak berusia satu hingga lima tahun dari *Guangzhou Women and Children's Medical Center*, Guangzhou, China. Semua pencitraan rontgen dada dilakukan sebagai bagian dari perawatan klinis rutin pasien. Untuk analisis foto rontgen dada, semua foto awalnya disaring untuk menghapus semua foto yang berkualitas rendah atau tidak dapat dibaca. Setelah itu foto akan dianalisa oleh dua dokter ahli untuk mengelompokkan foto paru – paru yang normal dan paru – paru yang terdeteksi terdapat pneumonia. Beberapa metode pada analisis dan diagnosa pneumonia dapat diterapkan dengan bantuan *deep learning* seperti *Recurrent Neural Network* (RNN), *Deep Neural Network* (DNN), *Deep Belief Network* (DBN) dan *Stacked Auto-Encoder* (SAE) (Cao dkk., 2018). Salah satu contoh metode yang dapat membantu adalah metode *Convolutional Neural Network* (Cao dkk., 2018).

Metode CNN dipercaya lebih efektif dan lebih baik dibandingkan dengan metode klasifikasi lainnya. Salah satu contoh penelitian yang dapat membuktikannya adalah penelitian yang dilakukan oleh (Ahmed dkk., 2019). Pada penelitian ini membahas tentang prediksi terhadap penyakit Leukemia dengan membandingkan penggunaan metode CNN, Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), dan Decision Tree (DT). Dari penelitian tersebut didapatkan kesimpulan bahwa metode CNN memiliki akurasi yang lebih tinggi yaitu 88,25% sehingga model CNN memiliki kinerja yang lebih baik daripada algoritma pembelajaran mesin terkenal lainnya.

Beberapa penelitian terkait penggunaan metode CNN terhadap prediksi atau diagnosa pada penyakit paru-paru telah banyak dilakukan. Penelitian pertama yang dilakukan oleh (Beig dkk., 2019). Pada penelitian ini membahas tentang perbandingan metode CNN dengan *Support Vector Machine* untuk membedakan adenokarsinoma pada kanker paru sel kecil dari granuloma. Metode evaluasi yang

digunakan dalam penelitian ini adalah AUC (*Area Under the Curve*) yang dihitung dengan menjumlahkan area di bawah kurva ROC (*Receiver Operating Characteristics*). Kurva ROC adalah cara yang umum digunakan untuk memvisualisasikan kinerja *biner classifier*, yang berarti *classifier* dengan dua kelas output. Dari penelitian tersebut didapatkan kesimpulan bahwa metode CNN menghasilkan nilai AUC yaitu 0,76 yang dimana lebih tinggi daripada menggunakan metode Support Vector Machine yaitu 0.75. Penelitian kedua dilakukan oleh (Kido dkk., 2018). Pada penelitian ini membahas tentang evaluasi kinerja dari *Computer-Aided Diagnosis (CADx)* berbasis gambar dengan menggunakan CNN dan kinerja dari *Computer-Aided Detection (CADe)* berbasis gambar dengan menggunakan R-CNN untuk berbagai jenis kelainan paru-paru seperti nodul paru dan penyakit paru difus. Dari penelitian tersebut didapatkan kesimpulan bahwa CADx berbasis gambar dengan menggunakan CNN dapat membedakan berbagai macam kelainan paru seperti nodul paru dan penyakit paru difus. Penelitian ketiga dilakukan oleh (Nóbrega dkk., 2018). Pada penelitian ini membahas tentang perbandingan arsitektur CNN dan kombinasi dengan klasifikasi lain untuk klasifikasi nodul paru – paru melalui gambar CT-Scan. Arsitektur CNN yang digunakan dalam penelitian ini adalah VGG16, VGG19, MobileNet, Xception, InceptionV3, ResNet50, Inception-ResNet-V2, DenseNet169, DenseNet201, NASNetMobile dan NASNetLarge. Klasifikasi yang digunakan adalah Naive Bayes, KNN, SVM, MLP dan RF. Dari penelitian tersebut didapatkan kesimpulan bahwa kombinasi antara ResNet50 dan SVM RBF terpilih sebagai yang terbaik, dengan nilai AUC yang mencapai 93,19%, ACC 88,41%, F-score 78,83%, TPR 85,38% dan PPV 73,48%. Dari beberapa penelitian yang dilakukan tersebut, didapatkan hasil bahwa penggunaan metode CNN merupakan metode klasifikasi data yang paling efektif dalam memprediksikan atau mendiagnosa penyakit paru-paru pada pasien. Metode ini dapat memberikan kemudahan pada interpretasi model juga memberikan hasil diagnosa dengan lebih akurat.

Maka dari itu, penelitian ini menggunakan metode CNN untuk menganalisis diagnosa pada penyakit pneumonia pada paru – paru balita. Hasil dari penelitian ini diharapkan dapat memberikan manfaat yaitu memberikan alternatif diagnosa secara lebih cepat dan akurat, agar pasien penderita penyakit pneumonia mendapatkan penanganan secara cepat dan tepat.

1.2. Rumusan Masalah

Dari latar belakang yang telah dipaparkan, untuk mengidentifikasi penyakit pneumonia pada manusia berbasis *Convolutional Neural Network* dapat dirumuskan sebagai berikut:

1. Bagaimana melakukan preprocessing pada foto rontgen paru-paru normal dan paru – paru terinfeksi pneumonia dengan menggunakan metode augmentasi transformasi tradisional?
2. Bagaimana hasil ekstraksi fitur dari data foto rontgen paru-paru yang digunakan untuk inputan proses klasifikasi dalam mengidentifikasi pneumonia menggunakan metode *convolutional feature* jenis VGG16, VGG19, ResNet50, dan InceptionV3?
3. Bagaimana hasil evaluasi dari klasifikasi data foto rontgen paru-paru normal dan paru – paru terinfeksi pneumonia dengan metode *confusion matrix* dan ROC?

1.3. Tujuan Penelitian

Dari latar belakang dan rumusan masalah di atas, tujuan dari mengidentifikasi penyakit pneumonia pada manusia berbasis Convolutional Neural Network adalah sebagai berikut :

1. Melakukan preprocessing pada foto rontgen paru-paru normal dan paru – paru terinfeksi pneumonia dengan menggunakan metode augmentasi transformasi tradisional
2. Hasil ekstraksi fitur dari data foto rontgen paru-paru yang digunakan untuk inputan proses klasifikasi dalam mengidentifikasi pneumonia menggunakan metode *convolutional feature* jenis VGG16, VGG19, ResNet50, dan InceptionV3.
3. Hasil evaluasi dari klasifikasi data foto rontgen paru-paru normal dan paru – paru terinfeksi pneumonia dengan metode *confusion matrix* dan ROC.

1.4. Manfaat Penelitian

Manfaat dari penelitian ini adalah untuk membantu peneliti yang

akan melakukan penelitian menggunakan metode Convolutional Neural Network (CNN) di bidang medis dan membantu tenaga medis dalam meningkatkan diagnosa penyakit pneumonia pada balita.