



**PROPOSAL PENELITIAN
KUALIFIKASI**

**PENGEMBANGAN MODEL DIAGNOSIS PADA SMEAR-
NEGATIVE PULMONARY TUBERCULOSIS BERBASIS
DEEP LEARNING**

**NUR AZIZAH
NPM: 99217012**

**PROGRAM DOKTOR TEKNOLOGI INFORMASI
UNIVERSITAS GUNADARMA
AGUSTUS 2021**

DAFTAR ISI

Bab I. PENDAHULUAN

- 1.1. Latar Belakang Masalah
- 1.2. Batasan Masalah
- 1.3. Tujuan Penelitian
- 1.4. Kontribusi Penelitian

Bab II. TINJAUAN PUSTAKA

- 2.1. Tuberkulosis
- 2.2. Deep Learning
- 2.3. Convolutional Neural Network (CNN)
- 2.4. Region Convolutional Neural Network (R-CNN)
- 2.5. Faster Region Convolutional Neural Network (Faster R-CNN)
- 2.6. Penelitian Terdahulu

Bab III. METODOLOGI PENELITIAN

- 3.1. Tahapan Penelitian
 - 3.1.1. Pengumpulan Data Set
 - 3.1.2. Sistem Yang di Bangun
 - 3.1.3. Evaluasi
- 3.2. Rencana Kerja

DAFTAR PUSTAKA

BAB I. PENDAHULUAN

1.1 Latar Belakang

Salah satu organ manusia yang sangat penting dalam sistem pernapasan adalah paru-paru. Fungsi utama paru-paru adalah untuk proses respirasi yang bertugas memompa udara yang masuk ke dalam tubuh. Kesehatan organ paru-paru sangat penting, karena jika organ ini mengalami gangguan maka akan berpengaruh terhadap kesehatan tubuh lainnya. Jenis penyakit yang menyerang organ paru-paru salah satunya adalah Tuberkulosis (TB). TB merupakan penyakit yang paling banyak dialami masyarakat dan termasuk ke dalam salah satu penyakit penyebab kematian tertinggi di Dunia (World Health Organization, 2020). Penyakit TB disebabkan oleh bakteri *Mycrobacterium tuberculosis* yang tergolong pada bakteri yang menular. Dari data World Health Organization (2019), Indonesia adalah Negara yang menduduki peringkat ketiga dalam jumlah kasus penyakit TB dengan jumlah sebanyak 842.000 atau 46 persen dari total kasus yang ada. Dalam Pedoman Nasional Pengendalian Tuberkulosis Indonesia, dilaporkan bahwa salah satu penyebab meningkatnya beban TB adalah kegagalan program TB. Salah satu penyebab kegagalan ini adalah karena penemuan kasus atau diagnosis TB yang tidak standar dan tidak memadainya tatalaksana diagnosis (PNPT, 2014)

Penelitian terdahulu menyatakan bahwa diagnosis TB sulit dilakukan (Imianvan, 2012; Alvarez-Uria dkk, 2012). Pertama, karena gejala TB, seperti demam, batuk, batuk berdarah, batuk berdarah dan penurunan berat badan, mirip dengan gejala kanker paru (M.Batt dkk, 2012), pneumonia (Uzoka dkk, 2011) dan juga Infeksi Saluran Pernafasan Akut (ISPA), asma serta Penyakit Paru Obstruksi Kronis (PPOK) (ISTC, 2006). Kedua, pada pasien anak dikarenakan gambaran klinis TB tidak spesifik dan foto paru sulit diinterpretasi (Kusuma, 2007) dan dahak sulit didapat (WHO, 2015). Ketiga, pada kasus jumlah kuman yang sedikit (paucibacillary TB) (Kusuma, 2007; Nesredin, 2012). Keempat, pada kasus TB ekstra paru (extra pulmonary TB) karena pasien seringkali tidak menunjukkan gejala TB (Radzi dkk, 2011; Jain, 2011; Bahadori, 2012). Kelima, pada pasien yang positif HIV, hasil chest radiography bisa jadi atypical karena infeksi lainnya (Swai dkk, 2011). Keenam, pada kasus TB paru BTA negatif (Smear- Negative Pulmonary Tuberculosis, disingkat SNPT) (Mello dkk, 2006; Benfu dkk, 2009; Santiago dkk, 2013; Muvunyi, 2006).

Diagnosis TB paru dengan BTA negatif dan kultur positif dapat ditegakkan dengan kultur dahak, tetapi memerlukan 6 sampai 8 minggu. Selain itu, peralatan untuk kultur jarang ditemui di negara berkembang (WHO, 2015; Mello dkk, 2006; Muvunyi, 2006), sehingga penggunaan kultur terbatas dan jarang disarankan. Metode seperti nucleic acid amplification tests, dapat memberikan hasil yang lebih cepat, namun memerlukan biaya tinggi dan peralatannya tidak banyak ditemukan di negara berkembang. Sehingga pada kondisi terbatasnya peralatan, diagnosis SNPT ditegakkan dengan gejala, hasil pemeriksaan fisik dan hasil pemeriksaan penunjang (minimal dengan pemeriksaan chest radiography, disingkat CXR). Namun hasil pemeriksaan CXR antara tuberkulosis aktif dan tidak aktif sulit untuk dibedakan. Untuk itu sangat diperlukan diagnosis akhir yang tepat untuk mengurangi biaya

dan kemungkinan terjadinya kasus underdiagnose dan overtreat (Bhatt dkk, 2012; Muvunyi, 2006).

Dalam penelitian sebelumnya, diagnosis tuberkulosis telah dilakukan dengan menggunakan beberapa jenis input. Input-input tersebut antara lain berupa citra digital dari sampel jaringan (Osman dkk, 2010), citra kuman TB (Santiago dkk, 2013), aroma sputum dengan electronic nose (Kodogiannis, 2013), suara batuk (Tracey dkk, 2013), gelombang suara paru (Becker dkk, 2013), data rekam medis (T. Ucar, 2011) dan darah (M. Lauria, 2016). Sayangnya, citra digital dari sampel jaringan sulit didapat karena mahal dan memerlukan peralatan yang tidak dimiliki oleh setiap layanan kesehatan di negara berkembang (R. Khans dkk, 2006). Demikian pula halnya dengan citra bakteri TB. Data berupa citra bakteri TB diperoleh dengan menggunakan mikroskop fluoresensi digital (A. Goyal dkk, 2015). Aroma sputum, suara batuk, gelombang suara paru dan darah bukan merupakan standar penemuan kasus dan diagnosis TB (PNPT, 2014).

Dalam mendiagnosis penyakit ini umumnya alat atau teknik diagnosis yang dilakukan adalah melakukan *rontgen* dengan sinar x pada bagian dada pasien, data citra yang dihasilkan dari teknik ini disebut sebagai citra *chest x-ray* (CXR), teknik ini dipilih karena teknik ini ekonomis dan mudah digunakan (C. Qin dkk, 2018). Pada CXR dapat dilihat gambaran permukaan dalam dada pasien yang digunakan sebagai bahan pertimbangan oleh seorang ahli untuk mengetahui apakah seorang pasien terjangkit penyakit pneumonia atau tuberkulosis.

Seiring pesatnya kemajuan teknologi terutama di bidang *deep learning* yang merupakan bagian dari metode *machine learning*, permasalahan pada citra seperti pengklasifikasian terus dilakukan pengembangan menjadi lebih dinamis untuk berbagai permasalahan, dimana dalam membedakan suatu objek pada citra dapat dilakukan dengan bantuan komputasi mesin tanpa campur tangan manusia. Sehingga prediksi yang dihasilkan oleh model *deep learning* dalam hal mendeteksi suatu penyakit dapat membantu seorang ahli dalam mendiagnosis kemungkinan penyakit yang diderita oleh pasien.

Menurut Pedoman dan Penatalaksanaan Tuberkulosis Indonesia (PNPT, 2014), proses diagnosis TB diawali dengan Penemuan Suspek TB (TB screening) menggunakan data gejala dan pemeriksaan fisik. Kemudian dilakukan pemeriksaan dahak (mikroskopis langsung) untuk menghasilkan diagnosis akhir. Bila hasil pemeriksaan mikroskopis positif, maka pasien terdiagnosis TB Paru BTA positif. Bila hasil pemeriksaan mikroskopis negatif, maka dilakukan pemeriksaan penunjang tambahan, minimal dengan pemeriksaan foto toraks untuk menentukan apakah pasien terdiagnosis TB Paru BTA negatif.

Penelitian ini akan menggunakan teknik klasifikasi pada *deep learning* dengan menggunakan algoritma *Faster Region Convolutional Neural Network* (Faster R-CNN) untuk menghasilkan model diagnosis awal dan akhir TB, khususnya pada kasus dengan BTA Negatif. *Faster Region Convolutional Neural Network* (R-CNN) adalah sebuah metode yang berbasis *deep learning object detection* yang biasa digunakan untuk mendeteksi objek. Penelitian terdahulu yang relevan dengan algoritma penulis gunakan adalah Yilin Xie et.al (2020) dalam penelitian yang berjudul “Computer-Aided System for the Detection of Multicategory Pulmonary Tuberculosis in

Radiographs”. Dalam penelitian ini dirancang suatu sistem untuk mendiagnosis tuberculosis menggunakan algoritma *Faster R-CNN* dengan mendeteksi multikategori pada radiografi dada.

Pada penelitian ini, mengusulkan pendekatan deep learning menggunakan metode *Faster Region Convolutional Neural Network* (*Faster R-CNN*) untuk mendeteksi SNPT melalui citra chest X-Ray. Evaluasi dilakukan untuk mengetahui performa metode yang diusulkan berupa precision, recall, F-1, dan akurasi. Dengan metode yang kami usulkan diharapkan dapat memberikan gambaran yang lebih komprehensif dalam tujuan membantu dan melengkapi model diagnosis Smear-Negative Pulmonary Tuberculosis (SNPT) awal dan akhir pasien.

Dari beberapa literatur yang dijadikan referensi, belum ditemukan penelitian yang melakukan diagnosis awal dan akhir dengan menggunakan metode *Faster Region Convolutional Neural Network* (*Faster R-CNN*) pada *deep learning*.

1.2 Batasan Masalah

Penelitian ini berusaha mengembangkan teknik klasifikasi pengembangan model diagnosis Pada Smear-Negative Pulmonary Tuberculosis (SNPT) Berbasis Deep Learning. Sehingga penelitian ini difokuskan pada:

1. Bagaimana Deep Learning pada pengembangan diagnosis SNPT menggunakan *Faster RCNN* dapat menghasilkan diagnosis awal dengan tingkat akurasi tinggi?
2. Bagaimana Deep Learning pada pengembangan diagnosis SNPT menggunakan *Faster RCNN* dapat menghasilkan diagnosis akhir dengan tingkat akurasi tinggi?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk membuat dua model klasifikasi, yaitu: Model yang dapat menghasilkan diagnosis awal bagi semua pasien dan diagnosis akhir bagi pasien dengan BTA negatif. Model ini diharapkan dapat menjadi referensi dalam pengembangan diagnosis tuberculosis menggunakan deep learning sesuai dengan kondisi di Indonesia. Manfaat untuk jangka panjangnya adalah bila diagnosis lebih tepat diketahui, maka diharapkan dapat membantu mengurangi kemungkinan penyebaran kuman TB. Hal ini akan mengurangi munculnya kasus baru dan akan meningkatkan jumlah *Case Detection Rate*. Selain itu juga dapat mengurangi jumlah pasien yang mendapatkan tindakan dan terapi obat yang berlebihan dan tidak tepat sasaran (*over-treat* atau *over-diagnose*).

1.4 Kontribusi Penelitian

Kontribusi dalam penelitian ini adalah:

- a. Profile data diagnosis
- b. Model klasifikasi diagnosis awal
- c. Model klasifikasi diagnosis akhir untuk pasien dengan BTA negatif

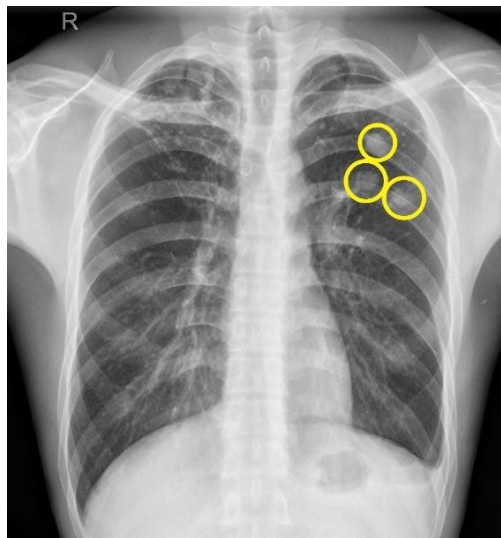
BAB II. TINJAUAN PUSTAKA

2.1. Tuberkulosis

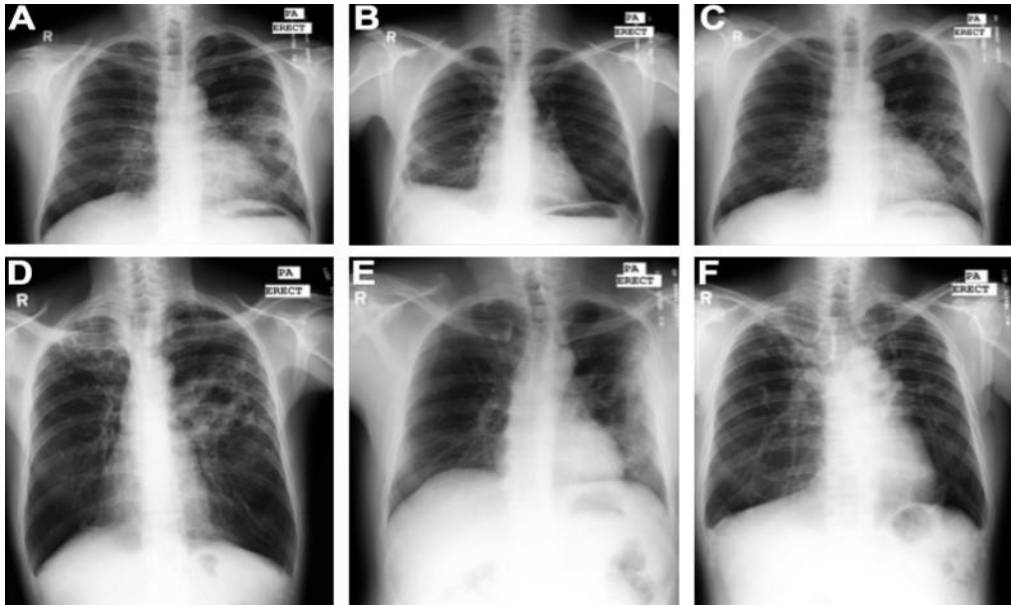
Tuberkulosis (Tuberculosis, disingkat Tbc), atau Tb (singkatan dari "Tubercle bacillus") merupakan penyakit menular yang umum, dan dalam banyak kasus bersifat mematikan. Penyakit ini disebabkan oleh berbagai strain mikobakteria, umumnya *Mycobacterium tuberculosis* (disingkat "MTb" atau "MTbc"). Tuberkulosis biasanya menyerang paru-paru, namun juga bisa berdampak pada bagian tubuh lainnya. Tuberkulosis menyebar melalui udara ketika seseorang dengan infeksi TB aktif batuk, bersin, atau menyebarkan butiran ludah mereka melalui udara. Infeksi TB umumnya bersifat asimtomatik dan laten. Namun hanya satu dari sepuluh kasus infeksi laten yang berkembang menjadi penyakit aktif. Bila Tuberkulosis tidak diobati maka lebih dari 50% orang yang terinfeksi bisa meninggal (Muchtar, Herman, & Yulistini, 2018).

Foto *rontgen* pasien dengan penyakit Tuberkulosis dapat diidentifikasi dengan cara mencari adanya *coin lesion* atau luka yang berbentuk bulat akibat adanya infeksi bakteri, ukuran diameter dari *coin lesion* berkisar antara 1 sampai 5cm (Smithuis, 2014). *Coin lesion* pada foto *rontgen* paru ditandai dengan area yang berwarna keabu-abuan dan berbentuk bulat seperti yang ditunjukkan pada gambar 2.1

3.



Gambar 2.1 Foto *rontgen* penyakit tuberculosi
(<http://openi.nlm.nih.gov>)



Gambar 2.2 abnormal CXR dari MC dataset

Gambar 2.2 merupakan contoh-contoh foto *rontgen* abnormal dalam dataset MC. Foto *rontgen* A memiliki *infiltrat kavitas* di sebelah kiri dan *infiltrat* halus di paru kanan bawah. Foto *rontgen* B adalah contoh TB *pleura*. Perhatikan bahwa sudut *costophrenic* kanan tumpul menunjukkan *efusi* sedang. Foto *rontgen* C memiliki *infiltrat* di kedua paru-paru. Foto *rontgen* D menunjukkan *infiltrat* yang tidak teratur di paru-paru kiri dengan *kavitasi* dan jaringan parut pada *apex* kanan. Foto *rontgen* E menunjukkan *infiltrat perifer* di paru-paru kiri. Foto *rontgen* F menunjukkan tanda-tanda TB, yang ditunjukkan oleh *retraksi hila bilateral superior*, yang lebih jelas di sebelah kanan (Jaeger, dkk., 2013).

2.2 Deep Learning

Deep Learning yang juga dikenal sebagai *deep structured learning* atau *hierarchical learning* merupakan bagian dari metode *Machine Learning*. Berdasarkan cara pembelajarannya dibagi menjadi 2 yaitu : *Supervised Learning* (misal : klasifikasi) dan *Unsupervised Learning* (misal : analisis pola). Arsitektur *Deep Learning* banyak digunakan dalam bidang *Computer Vision*, *speech recognition*, *natural language processing*, *audio recognition*, *social network filtering*, *machine translation*, *bioinformatics*, dan lain-lain (LeCun, Bengio, & Hinton, 2015).

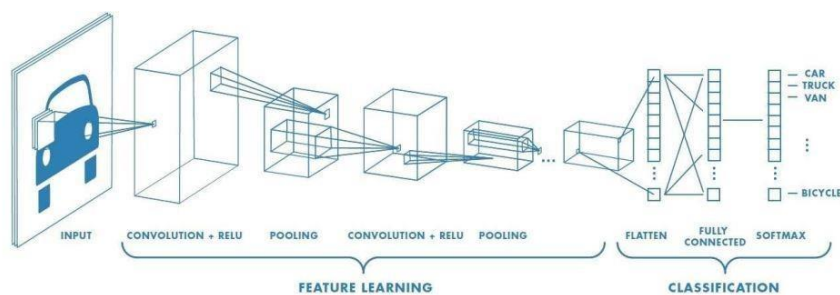
Perkembangan teknologi dan ilmu pengetahuan saat ini mendorong para peneliti untuk dapat menyelesaikan permasalahan yang ada, permasalahan yang ada diantaranya adalah sulitnya untuk mendapatkan informasi dan mengekstraksi fitur abstrak dari data mentah. *Deep learning* dapat melakukan ekstraksi fitur abstrak yang didapat dari data mentah dengan baik (Goodfellow, Bengio, & Courville, 2016)

Deep Learning menggunakan aliran yang berlapis-lapis dari unit pemroses yang non linear, setiap lapisan berturut-turut menggunakan output dari lapisan sebelumnya sebagai input (Deng & Yu, 2014). Namun *Deep Learning* membutuhkan performa hardware yang sangat baik, dikarenakan banyaknya perhitungan yang harus di proses dan data yang besar.

2.3. Convolutional Neural Network (CNN)

Convolutional Neural Network adalah sebuah *neural network* yang digunakan untuk menganalisis gambar, mendeteksi objek dan mengenali objek pada *image*, yang melibatkan berbagai parameter untuk mencirikan jaringan mulai dari pencirian objek sampai dengan pengklasifikasian objek pada saraf tiruan. *Convolutional neural network* dibuat untuk mengurangi jumlah parameter dan menyederhanakan arsitektur jaringan dalam melakukan pengenalan objek (Karapathy, 2018).

CNN merupakan *neural network* yang tidak terlalu berbeda dengan *neural network* lainnya. CNN terdiri dari *neuron* yang memiliki *weight*, *bias* dan *activation function*, namun arsitektur dari *Convolutional Neural Network* dibagi menjadi dua bagian besar, yaitu *Feature Extraction Layer* dan *Fully-Connected Layer* (*Multilayer Perceptron*).



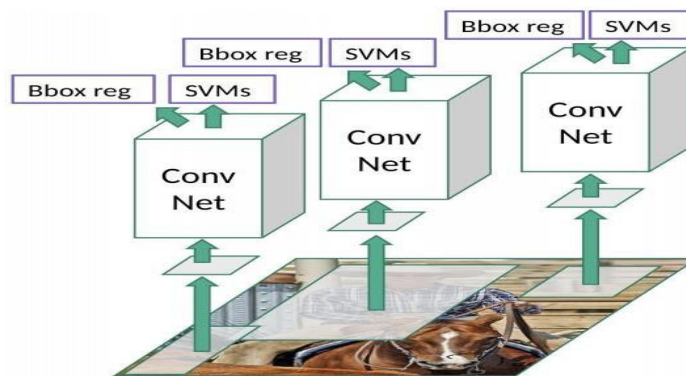
Gambar 2.3. Arsitektur Umum dari CNN (Karapathy, 2018)

2.4 Region Convolutional Neural Network (R-CNN)

Region Convolutional Neural Network (R-CNN) adalah sebuah metode yang berbasis *deep learning object detection* yang biasa digunakan untuk pendeteksian objek. R-CNN menggunakan algoritma *selective search* untuk membuat proposal dari *image*, dimana gambar yang di *input* akan dikelompokkan menjadi 2000 *region* yang dipilih berdasarkan tekstur, intentsitas, dan warna. Hal tersebut dilakukan untuk menutupi kelemahan CNN yang membagi *region* gambar dengan skala *region* besar yang membuat proses pengidentifikasian menjadi lebih lambat.

Kelemahan pada R-CNN :

1. Proses training data yang relatif lambat karena menggunakan 2000 *region proposal* untuk setiap *image*.
2. Tidak dapat diimplementasikan untuk klasifikasi *real time* karena membutuhkan waktu sekitar 47 – 50 detik untuk proses per-*image*.
3. Hanya bisa menggunakan algoritma *selective search* pada proses pengenalanya, tidak dapat menggunakan algoritma lain selain *selectivesearch*.



Gambar 2.4. Arsitektur Umum dari R-CNN (Gandhi., 2018)

2.5 Faster Region Convolutional Neural Network (Faster R-CNN)

Faster R-CNN adalah sebuah metode yang berbasis *deep learning object detection* yang biasa digunakan untuk pendeteksian objek yang dikembangkan dari algoritma *R-CNN* untuk menutupi kelemahan yang ada pada *R-CNN*. Kelebihan pada *Faster R-CNN* adalah *Faster R-CNN* menggunakan *RPN*, dimana *RPN* adalah sebuah *neural network* yang menggantikan peran *selective search* untuk mengajukan *region*, peran *selective search*

digantikan karena prosesnya yang lambat dalam mengolah image yaitu sekitar 2 sekon per gambar (Ren, et al., 2017). *RPN* berfungsi untuk menghasilkan beberapa *bounding box* dimana setiap *box* memiliki 2 skor probabilitas apakah pada lokasi tersebut terdapat objek atau tidak, dengan adanya *RPN* pemrosesan tidak dilakukan berulang ulang seperti yang dilakukan pada *R-CNN* dan membuat keseluruhan model dapat di-*train* secara *end-to-end*.

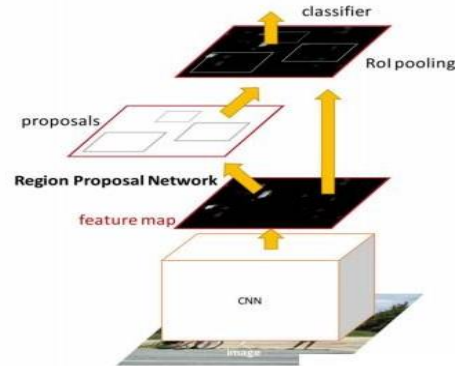
Cara Kerja *Faster R-CNN* :

1. *Convolution layer* membuat dan mengirimkan *feature map* kepada *RPN*.
2. *RPN* memproses *feature map* yang ada dan membuat *region* proposal serta membuat *bounding box* sebagian yang dianggap memiliki kemungkinan terdapat objek.
3. *R-CNN* melakukan klasifikasi terhadap proposal yang telah dibuat *RPN* dan menentukan apakah objek pada proposal merupakan objek yang ada pada model yang telah dilatih dan memberikan pelabelan pada objek.

Kelebihan *Faster R-CNN* :

1. Proses training data yang lebih cepat karena *proposal region* yang dibuat tidak sebesar pada *R-CNN* yang membuat proses pengerjaan lebih ringan.
2. Dapat diimplementasikan untuk klasifikasi *real time* karena membutuhkan waktu pemrosesan yang relatif singkat yaitu sekitar 1 – 2 detik.

Faster R-CNN:



Gambar 2.5. Arsitektur Umum dari *Faster R-CNN* (Gavrilescu., 2018)

Penjelasan arsitektur umum dari *Faster R-CNN*:

- a) *Convolutional Layer* : Pada tahap ini convolutional layer mempelajari bagian bagian penting yang dapat menjadi ciri khas dari objek tersebut serta membuat *feature map* dari objek yang telah ditangkap kamera *smartphone*. *Convolutional* adalah *layer* yang terdiri dari *neuron* yang tersusun sedemikian rupa sehingga membentuk sebuah *filter* dua dimensi dengan panjang dan tinggi (*pixels*).
- b) *Feature Map* : *Feature Map* adalah sebuah map yang dibuat oleh *convolutional layer* yang berisi informasi tentang representasi *vector* dari *image* yang ditangkap.
- c) *RPN (Region Proposal Network)* : *RPN* merupakan sebuah modul yang bekerja

untuk mengolah *feature map* yang telah dibuat pada *convolution layer* untuk memprediksi bagian yang dianggap sebagai objek dan melakukan prediksi *bounding box* dari objek tersebut, *RPN* dibagi menjadi 2 *convolution layer* dimana 1 *layer* bertanggung jawab untuk mendeteksi letak objek dan 1 *layer* berfungsi memprediksi *bounding box*.

- d) *ROI Pooling* : *ROI* merupakan *layer* yang bertanggung jawab untuk mengekstrak informasi *feature map* yang telah diproses oleh *RPN* untuk dikasifikasi pada *classification layer*.
- e) *Classification Layer* : *Classification layer* merupakan *layer* yang berfungsi untuk mengelompokkan objek yang telah dideteksi pada *RPN* dan melakukan pelabelan terhadap objek tersebut serta memberikan *bounding box* pada objek tersebut.

2.6. Penelitian Terdahulu

Penelitian-penelitian terkait Deep Learning pada pengembangan model diagnosis disajikan pada tabel 2.1

No.	Peneliti	Tahun	Judul	Hasil
1.	Rao, Pereira	2016	Convolutional neural networks for lung cancer screening in computed tomography (CT) scans	penelitian ini menggunakan <i>Convolutional Neural Network (CNN)</i> dengan <i>CanNet</i> sebagai arsitekturnya untuk mengklasifikasikan penyakit kanker paru- paru melalui <i>CT Scan</i> .
2	S. N. Kumar, et all	2021	Deep Learning Algorithms in Medical Image Processing for Cancer Diagnosis: Overview, Challenges and Future	penelitian ini menggunakan <i>Convolutional Neural Network (CNN)</i> dengan <i>SVM</i> sebagai arsitekturnya untuk mengklasifikasikan penyakit kanker payudara menghasilkan tingkat Accuracy— 99.38%
3	Muhammad Noor	2019	Implementasi <i>Deep</i>	penelitian ini

	Dwi Eldianto		<i>Learning</i> Pada Sistem Klasifikasi Penyakit Paru Berdasarkan Foto Rontgen Menggunakan Metode <i>Convolutional Neural Network</i> (CNN)	menggunakan salah satu metode deep learning yaitu Convolutional Neural Network yang akan diimplementasikan ke dalam sistem yang dapat mengklasifikasikan foto rontgen paru ke dalam tiga kelas : Normal, Pneumonia dan Tuberculosis. Hasil akurasi model terbaik sebesar 85.96% didapatkan setelah dilakukan <i>testing</i> pada model menggunakan <i>confusion matrix</i>
4.	ROTTY, BRENDA MONICA and Adrian, Angelia Melani and Sanger, Junaidy Budi	2020	Aplikasi Prediksi Penyakit Tuberculosis Menggunakan Algoritma Convolutional Neural Network (CNN)	Penelitian ini mendiagnosis penyakit TB dengan menggunakan gambar rontgen dengan menggunakan Algoritma Convolutional Neural Network (CNN) menghasilkan prediksi dengan hasil ketelitian untuk normal atau non-TB 78.0% dan abnormal 80.0%.
5.	Fidelius Hartanto Tanian	2020	Implementasi Algoritma Deep Convolutional Neural Network Untuk Diagnosa Penyakit Tuberculosis	Hasil terbaik dari penelitian ini merupakan sistem yang mampu mendiagnosa 84% gambar hasil rontgen data testing pada model yang dibuat dengan presisi rata-rata sebesar 85% dan recall rata-rata sebesar 85%.

BAB III. METODOLOGI PENELITIAN

3.1.Tahapan Penelitian

Penelitian ini berusaha mengembangkan model diagnosis Smear Negative Pulmonary Tuberculosis dengan metode Deep Learning menggunakan algoritma *Faster R-CNN* sebagai solusi dari masalah dan kekurangan dari teknik yang pernah dilakukan peneliti terdahulu yang dapat menghasilkan sebuah model diagnosis awal dan diagnosis akhir bagi pasien TB Negatif dengan menghasilkan tingkat akurasi tinggi. Rencana penelitian mencoba menggabungkan untuk mengembangkan deep learning dan mengembangkan metode yang dapat mengoptimalkan tingkat akurasi diagnosis.

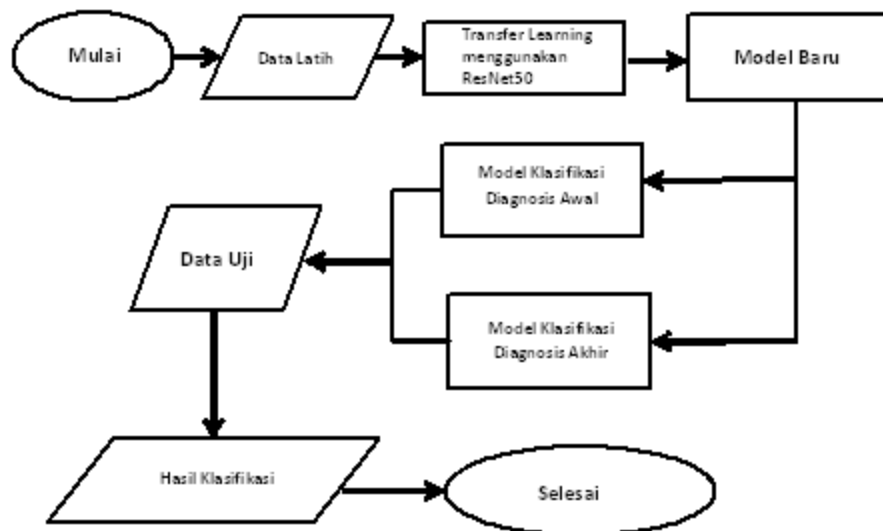
3.1.1. Pengumpulan Dataset

Data yang dipakai dalam studi ini diambil dari dua sumber terbuka terdiri dari data Tuberculosis melalui laman <https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset> dan data normal yang dapat diakses melalui <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>. Dari dua sumber ini, kami kemudian memformulasikan dataset gabungan dengan dua kelas yaitu Tuberculosis dan Normal yang masing-masing memiliki jumlah data latih 48 citra dan data validasi masing-masing 15 citra. Sedangkan untuk data uji, kelas Tuberculosis sebanyak 42 citra dan kelas normal sebanyak 234 citra.

3.1.2. Sistem yang dibangun

Setelah mengumpulkan dataset, langkah selanjutnya adalah membangun sistem untuk pelatihan dan mengevaluasi jaringan yang dibuat. Pada penelitian ini, kami menggunakan metode transfer learning, yang melakukan proses latih dengan menggunakan model yang sudah terlebih dahulu dilatih menggunakan dataset lain. Untuk kasus ini, kami menggunakan model ResNet yang sudah terlebih dahulu dilatih menggunakan ImageNet dataset (Deng, dkk, 2009). Diagram alir dari sistem yang kami bangun dapat dilihat di Gambar 3

Berikut adalah langkah-langkah yang dilakukan dalam penelitian ini:



Gambar 3.1. Diagram Alir dari Proses Latih dan Evaluasi

3.1.3. Evaluasi

Metriks yang digunakan untuk mengevaluasi model adalah akurasi, precision, recall dan F1-score. Untuk memahami metriks yang digunakan, sebelumnya akan didefinisikan terlebih dahulu true positive (TP), false positive (FP), false negative (FN) dan true negative (TN) seperti yang ditunjukkan pada confusion matrix Tabel 1. TP didefinisikan sebagai data positif yang diprediksi sebagai positif dan TN didefinisikan sebagai data negatif yang di prediksi sebagai negatif. Sedangkan FN merupakan kebalikan dari TP yaitu data positif yang diprediksi sebagai negative dan FP, kebalikan dari TN, yaitu data negatif yang diprediksi positif.

Table 1. Confusion Matrix

		Kelas sebenarnya	
		Positif	Negatif
Prediksi	Positif	True Positive (TP)	False Positive (FP)
	Negatif	False Negative (FN)	True Negative (TN)

Akurasi didefinisikan sebagai perbandingan jumlah data yang diprediksi secara benar terhadap total jumlah data. Precision menggambarkan perbandingan tp terhadap total data yang diprediksi positif. Recall didefinisikan sebagai perbandingan tp terhadap total data positif. Sedangkan F1 score adalah rata-rata harmonik antara precision dan recall. Persamaan 2, 3, 4, dan 5 merupakan rumus untuk precision, recall, F1-score, dan akurasi.

$$Precision = \frac{TP}{TP+} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - score = \frac{2 * Recall * Precision}{Recall + Preciso} \quad (4)$$

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

3.2. Rencana Kerja

Tabel 3.1. Rencana Pelaksanaan Penelitian

[illegible]

DAFTAR PUSTAKA

- World Health Organization. (2020). *World Health Organization*. Retrieved April 13, 2020, from <https://www.who.int/newsroom/fact-sheets/detail/tuberculosis>
- C. Qin, D. Yao, Y. Shi, dan Z. Song, "Computer-Aided Detection in Chest Radiography Based on Artificial Intelligence: A Survey," *BioMedical Engineering Online*, vol. 17, no. 1, pp. 1–23, 2018. <https://doi.org/10.1186/s12938>
- Rao, P., Pereira, N. A., & Srinivasan, R. (2016). Convolutional neural networks for lung cancer screening in computed tomography (CT) scans. *2nd International Conference on Contemporary Computing and Informatics (IC3I)*, 489-493.
- Muchtar, N. H., Herman, D., & Yulistini. (2018). Gambaran Faktor Risiko Timbulnya Tuberkulosis Paru pada Pasien yang Berkunjung ke Unit DOTSRSUP Dr. M. Djamil Padang Tahun 2015 . *Jurnal Kesehatan Andalas*, 7(1),80-87.
- Kementrian Kesehatan Republik Indonesia. (2018). *InfoDatin*. Jakarta Selatan: Pusat Data dan Informasi
- Yilin Xie, et all (2020). Computer-Aided System for the Detection of Multicategory Pulmonary Tuberculosis in Radiographs. *Journal of Healthcare Engineering*, Volume 2020, Article ID 9205082, 12 pages, <https://doi.org/10.1155/2020/9205082>.
- Rusdah & Nindya Kusumwardhany (2018), Diagnosis Smear-Negative Pulmonary Tuberculosis: Studi Pendahuluan, *Jurnal TICOM* Vol. 6 No.3 Mei 2018
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 436–444
- Deng, L., & Yu, D. (2014). Deep Learning: Methods and Applications. In L. Deng, & D. Yu, *Deep Learning: Methods and Applications* (pp. 199-201)
<https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>
<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>
- Aditama, T.Y. dan Subuh, M., 2011. *Strategi Nasional Pengendalian TB di Indonesia 2010-2014* D. E. Mustikawati & A. Surya, eds., Dirjen P2PL Kemenkes RI.
- Aguiar, F.S., Almeida, L.L., Ruffino-Netto, A., Kritski, A.L., Mello, F.C.D.Q. dan Werneck, G.L., 2012. Classification and regression tree (CART) model to predict pulmonary tuberculosis in hospitalized patients. *BMC pulmonary medicine*, 12(1), pp.12 – 40.
- Ali, A., Elfaki, M. dan Jawawi, D.N.A., 2012. Using Naïve Bayes and Bayesian Network for Prediction of Potential Problematic Cases in Tuberculosis. *International Journal of Informatics and Communication Technology (IJ-ICT)*, 1(2), pp.63–71.
- Alvarez-Uria, G., Azcona, J.M., Midde, M., Naik, P.K., Reddy, S. dan Reddy, R., 2012. Rapid Diagnosis of Pulmonary and Extrapulmonary Tuberculosis in HIV-Infected Patients. Comparison of LED Fluorescent Microscopy and the GeneXpert MTB/RIF Assay in a District Hospital in India. *Tuberculosis research and treatment*, 2012, p.932862.
- Ansari, A.Q., Gupta, N.K. dan Ekata, 2012. Adaptive neurofuzzy system for tuberculosis. In *2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing*. IEEE, pp. 568–573.
- Asha, T., Natarajan, S. dan Murthy, K.N.B., 2011. A Data Mining Approach to the Diagnosis of Tuberculosis by Cascading Clustering and Classification. *Journal of computing*, 3(4).
- Asha, T., Natarajan, S. dan Murthy, K.N.B., 2012. Data Mining Techniques in the Diagnosis of Tuberculosis. In P.-J. Cardona, ed. *Understanding Tuberculosis - Global Experiences and Innovative Approaches to the Diagnosis*. InTech, pp. 333 – 352.

- Asha, T., Natarajan, S. dan Murthy, K.N.B., 2010. Diagnosis of Tuberculosis using Ensemble methods. In *Computer Science and Information Tehcnology (ICCSIT), 2010 3rd IEEE International Conference on.* IEEE, pp. 409–412.
- Asha, T., Natarajan, S. dan Murthy, K.N.B., 2011. Effective Classification Algorithms to Predict the Accuracy of Tuberculosis-A Machine Learning Approach. *International Journal of Computer Science and Information Security (IJCSIS)*, 9(7), pp.89–94.
- Bakar, A.A. dan Febriyani, F., 2007. Rough Neural Network Model for Tuberculosis Patient Categorization. In *Proceedings of the International Conference on Electrical Engineering and Informatics*. Bandung, Indonesia, pp. 765–768.
- Benfu, Y., Hongmei, S., Ye, S., Xiuhui, L. dan Bin, Z., 2009. Study on the Artificial Neural Network in the Diagnosis of Smear Negative Pulmonary Tuberculosis. In *2009 World Congress on Computer Science and Information Engineering*. Los Angeles: IEEE Comput. Soc, pp. 584–588.
- Coaquira, F. dan Acuña, E., 2007. Applications of Rough Sets Theory in Data Preprocessing for Knowledge Discovery. In *Proceedings of the World Congress on engineering and Computer Science 2007*. San Francisco, USA, pp. 1–6.
- Colebunders, R. dan Bastian, I., 2000. A review of the diagnosis and treatment of smear-negative pulmonary tuberculosis. *International Journal of Tuberculosis and Lung Disease*, 4(2), pp.97–107.
- Dağ, H., Sayın, K.E., Yenidoğan, I., Albayrak, S. dan Acar, C., 2012. Comparison of Feature Selection Algorithms for Medical Data. In *2012 International Symposium on Innovations in Intelligent Systems and Applications (INISTA)*. Trabzon, pp. 1–5.
- Djam, X.Y. dan Kimbi, Y.H., 2011. A Decision Support System for Tuberculosis Diagnosis. *The Pacific Journal of Science and Technology*, 12(2), pp.410–425.
- El-solh, A.A., Hsiao, C., Goodnough, S., Serghani, J. dan Grant, B.J.B., 1999. Predicting Active Pulmonary Tuberculosis Using an Artificial Neural Network. *CHEST*, 116(4), pp.968–973.
- Elveren, E. dan Yumuşak, N., 2011. Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm. *Journal of medical systems*, 35(3), pp.329–32.
- Er, O., Temurtas, F. dan Tanrikulu, a. Ç., 2008. Tuberculosis Disease Diagnosis Using Artificial Neural Networks. *Journal of Medical Systems*, 34(3), pp.299–302.
- Han, J. dan Kamber, M., 2006. *Data Mining: Concepts and Techniques* Second., San Francisco, Canada: Morgan Kaufmann Publishers.
- Imianvan, A.A. dan Obi, J.C., 2012. Decision Support System for the Identification of Tuberculosis using Neuro Fuzzy logic. *Nigerian Annals of Natural Sciences*, 12(1), pp.12–20.
- Jain, A., 2011. Extra pulmonary tuberculosis: a diagnostic dilemma. *Indian journal of clinical biochemistry : IJCB*, 26(3), pp.269–73.
- Kemenkes, 2013. *Fakta Seputar Tuberkulosis : Pengendalian Tuberkulosis Indonesia*, Jakarta.
- Kusuma, H.M.S.C., 2007. Diagnostik Tuberkulosis Baru. *Sari Pediatri*, 8(4), pp.143–151.
- Mach, M.A. dan Abdel-Badeeh, M.S., 2010. Intelligent techniques for business intelligence in healthcare. In *2010 10th International Conference on Intelligent Systems Design and Applications*. IEEE, pp. 545–550.
- Mello, F.C.D.Q., Bastos, L.G. do V., Soares, S.L.M., Rezende, V.M., Conde, M.B., Chaisson, R.E., Kritski, A.L., Ruffino-netto, A. dan Werneck, G.L., 2006. Predicting smear negative pulmonary tuberculosis with classification trees and logistic regression : a cross-sectional study. *BMC Public Health*, 6(43), pp.1–8.

- Muvunyi, C.M. dan Masaisa, F., 2006. Diagnosis of Smear-Negative Pulmonary Tuberculosis in Low-Income Countries : Current Evidence in Sub-Saharan Africa with Special Focus on HIV Infection or AIDS. In P.-J. Cardona, ed. *Understanding Tuberculosis - Global Experiences and Innovative Approaches to the Diagnosis*. InTech, pp. 127–146.
- Nafar, Z. dan Golshani, A., 2006. Data Mining Methods for Protein-Protein Interactions. *2006 Canadian Conference on Electrical and Computer Engineering*, (May), pp.991–994.
- Nahar, J., Imam, T., Tickle, K.S. dan Chen, Y.-P.P., 2013. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications*, 40(4), pp.1086–1093.
- Nesredin, A., 2012. *Mining Patients' Data for Effective Tuberculosis Diagnosis: The Case of Menelik II Hospital*. Addis Ababa University.
- Nugroho, A.S., 2007. Pengantar Support Vector Machine. In pp. 1–6.
- PDPI, 2002. *Tuberkulosis Pedoman Diagnosis dan Penatalaksanaan di Indonesia*,
- Prasetyo, E., 2012. *Data Mining Konsep dan Aplikasi Menggunakan MATLAB* 1st ed. N. WK, ed., Penerbit ANDI Yogyakarta.
- Radzi, R.U.K.R.M., Mansor, W. dan Johari, J., 2011. Review of mycobacterium tuberculosis detection. In *2011 IEEE Control and System Graduate Research Colloquium*. IEEE, pp. 189–192.
- Santiago-Mozos, R., Perez-Cruz, F., Madden, M. dan Artes-Rodriguez, A., 2013. An automated screening system for Tuberculosis. *IEEE journal of biomedical and health informatics*, pp(99), pp.1–8.
- Santos, A.M. Dos, Pereira, B.B., Seixas, J.M., Mello, F.C.D.Q. dan Kritski, A.L., 2004. Neural Networks: An Application for Predicting Smear Negative Pulmonary Tuberculosis. In *Statistics in the Health Sciences*. pp. 275–287.
- Santos, A.M. Dos, Silva, A.C. dan Martins, L.D.O., 2007. A Neural-Bayesian Approach for Predicting Smear Negative Pulmonary Tuberculosis. In *Congresso Brasileiro de Redes Neurais*. Florianópolis, Santa Catarina, pp. 1–5.
- Srikant, R. dan Agrawal, R., 1995. Mining Quantitative Association Rules in Large Relational Tables.
- Swai, H.F., Mugusi, F.M. dan Mbwambo, J.K., 2011. Sputum smear negative pulmonary tuberculosis: sensitivity and specificity of diagnostic algorithm. *BMC research notes*, 4, p.475.
- Tuberculosis Coalition for Technical Assistance, 2006. *International Standards for Tuberculosis Care (ISTC)*, San Francisco, Canada.
- Uçar, T. dan Karahoca, A., 2011. Predicting existence of Mycobacterium tuberculosis on patients using data mining approaches. *Procedia Computer Science*, 3, pp.1404–1411.
- Uçar, T., Karahoca, A. dan Karahoca, D., 2013. Tuberculosis disease diagnosis by using adaptive neuro fuzzy inference system and rough sets. *Neural Computing and Applications*, 23(2), pp.471–483.
- Yuli Sun Hariyani dkk, 2020. Deteksi Penyakit Covid-19 Berdasarkan Citra X-Ray Menggunakan Deep Residual Network. *LKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*. Vol.8 No.2, Halaman 443 – 453, Mei 2020, ISSN(p): 2338-8323 | ISSN(e): 2459-9638