



Klasifikasi Kanker Paru-Paru Dengan *Deep Neural Networks* Dan Data CT-Scan

Dafa Rizqi Setiawan¹, Firdha Rofika Bryliana², Krisna Nur Aedi Aripin³, Sabita Adelia Wardani⁴

Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email : ^{1*}dafarizqis07@gmail.com, ²brylianafz@gmail.com, ³krisnanuraa@gmail.com,

⁴sabitaaw@gmail.com

(* : coresponding author)

Abstrak - Gambaran Umum: Dalam beberapa tahun terakhir, jumlah kematian akibat kanker meningkat pesat di seluruh dunia. Di antara beberapa jenis kanker, kanker paru-paru merupakan salah satu kasus yang paling banyak ditemui. Kanker paru-paru umumnya menyerang pria karena berbagai faktor, antara lain paparan asap rokok, polusi udara jangka panjang, dan senyawa karsinogenik seperti radon, asbestos, arsenik, tar batubara, dan knalpot solar. Deteksi dini kanker memudahkan pelacakan pengobatan dan proses penyembuhan, sehingga menyelamatkan jutaan nyawa di seluruh dunia setiap tahunnya. Diperlukan cara yang mudah untuk mengidentifikasi kanker paru-paru melalui pemindaian tomografi komputer (CT) paru-paru pasien. Salah satu metode pengenalan dan klasifikasi citra adalah metode convolutional neural network (CNN).

Kata Kunci : Kanker Paru, Jaringan Syaraf Dalam, CT Scan

Abstract - Overview: In recent years, the number of cancer deaths has increased rapidly worldwide. Among several types of cancer, lung cancer is one of the most common cases. Lung cancer generally attacks men due to various factors, including exposure to cigarette smoke, long-term air pollution, and carcinogenic compounds such as radon, asbestos, arsenic, coal tar, and diesel exhaust. Early detection of cancer makes it easier to track treatment and healing processes, saving millions of lives worldwide each year. An easy way is needed to identify lung cancer through computed tomography (CT) scans of the patient's lungs. One of the methods for image recognition and classification is the convolutional neural network (CNN) method.

Keywords: Lung Cancer, Deep Neural Network, CT Scan

1. PENDAHULUAN

Kanker paru-paru merupakan penyebab keganasan paling umum di seluruh dunia, menyumbang 13 persen dari seluruh kematian akibat kanker. Selain itu, kanker paru-paru menyumbang sepertiga dari seluruh kematian akibat kanker pada pria. Menurut Organisasi Kesehatan Dunia (WHO), kanker paru-paru merupakan kanker yang paling banyak menyerang pria Indonesia dan menempati urutan kelima pada wanita.

Paru-paru merupakan organ penting yang berperan dalam sistem pernapasan dan peredaran darah tubuh manusia. Merokok merupakan penyebab utama kanker paru-paru, namun perokok pasif diperkirakan memiliki risiko rendah terkena kanker paru-paru. Gambaran medis paru-paru penderita kanker mungkin menunjukkan bintik-bintik putih menggumpal, atau disebut nodul paru. Deteksi nodul manual memerlukan waktu sekitar 15 hingga 20 menit per gambar medis. Tentu saja, hal ini sangat merepotkan bila mengambil gambar medis dalam jumlah besar.

Berdasarkan hal tersebut, banyak penelitian yang dilakukan untuk mengenali gambar medis pada kasus kanker secara otomatis. Seiring berkembangnya teknologi, pembelajaran mesin, khususnya pembelajaran mendalam, juga sedang dikembangkan. Dikatakan bahwa penggunaan deep learning untuk klasifikasi gambar memiliki akurasi yang lebih tinggi dibandingkan machine learning. Hal ini dibuktikan dengan penelitian Mohammad Farid Naufal yang membandingkan kinerja algoritma K-Nearest Neighbors (KNN), Support Vector Machine (SVM), dan Convolutional Neural Network (CNN).

Berdasarkan hasil penelitian tersebut, CNN, salah satu algoritma deep learning, memperoleh hasil akurasi yang lebih baik dibandingkan KNN dan SVM, dengan akurasi sebesar 0,942. Mendeteksi kanker paru-paru secara akurat pada gambar CT memerlukan ahli radiologi yang



terlatih. Biaya yang diperlukan relatif tinggi sehingga masyarakat kelas bawah dan menengah tidak mampu menanggung biaya yang diperlukan.

Hal ini secara tidak langsung membuat tanda-tanda awal kanker paru sulit dideteksi sehingga pengobatan kanker paru pun semakin sulit. Oleh karena itu, perlu adanya otomatisasi deteksi kanker paru-paru dan stadium kanker dari gambar CT scan. Pengembangan metode berbantuan komputer untuk mendeteksi kanker ganas secara akurat berpotensi mengurangi biaya dan membuat proses pengobatan dan pemulihan menjadi lebih berhasil. Salah satu pendekatan yang berhasil adalah penggunaan jaringan saraf tiruan (JST), yang terinspirasi oleh jaringan saraf manusia.

Konsep ini kemudian dikembangkan lebih lanjut dengan jaringan saraf dalam (DNN). Metode DNN yang mencapai hasil terpenting dalam pengenalan gambar hingga saat ini adalah jaringan saraf konvolusional (CNN). CNN pertama kali diperkenalkan kepada kami oleh Yann LeCun dan kawan-kawan. CNN memungkinkan ekstraksi fitur hierarki menggunakan banyak lapisan konvolusional dan pengumpulan maksimal. Namun CNN, seperti teknik deep learning lainnya, memiliki kelemahan yaitu proses pelatihan model memakan waktu. Pada tugas akhir ini, Anda akan menerapkan model CNN untuk mengklasifikasikan derajat kanker pada gambar CT scan paru.

Arsitektur CNN yang digunakan adalah CanNet. Model yang dibuat diharapkan mampu melakukan klasifikasi citra dengan akurasi tinggi dan proses pembelajaran model yang cepat. Penelitian ini menggunakan data CT scan yang dianalisis menggunakan teknik pembelajaran mendalam seperti CNN dan variasinya untuk mengklasifikasikan gambar sebagai kanker atau non-kanker. Pendekatan ini memungkinkan identifikasi pola kompleks dalam data citra medis, sehingga memberikan keuntungan dalam diagnosis kanker paru-paru.

2. METODOLOGI PENELITIAN

2.1 Pengumpulan Data Data

CT scan diperoleh dari Cancer Imaging Archive (TCIA) yang berisi gambar paru pasien terdiagnosis kanker paru dan pasien sehat. Kumpulan data ini berisi ribuan gambar yang diklasifikasikan ke dalam dua kategori: gambar positif (kanker) dan negatif (bebas kanker).

2.2 Preprocessing Data

Sebelum gambar CT scan diterapkan pada model deep learning, terlebih dahulu dilakukan beberapa tahapan preprocessing, antara lain:

Normalisasi Intensitas gambar dinormalisasi untuk mengurangi noise dan meningkatkan kualitas gambar.

Ubah ukuran Gambar diubah ukurannya menjadi 224 x 224 piksel agar sesuai dengan input standar model pembelajaran mendalam.

Augmentasi Data Augmentasi dilakukan untuk meningkatkan variasi data menggunakan teknik seperti rotasi, mirroring, dan zooming, yang membantu dalam proses generalisasi model.

2.3 Arsitektur Model

Dalam penelitian ini, kami menguji beberapa model jaringan saraf dalam, antara lain:

- Convolutional Neural Networks (CNN) CNN digunakan untuk mengekstrak fitur dari gambar yang menjadi dasar model deteksi kanker paru-paru.
- VGG-16 dan VGG-19 Kedua model ini terkenal dalam klasifikasi citra medis dan memiliki arsitektur yang lebih dalam sehingga dapat menangkap detail visual dengan lebih baik.
- ResNet-50 Model ini menggunakan pembelajaran sisa. Ini membantu mengatasi masalah hilangnya gradien di jaringan dalam dan cocok untuk gambar resolusi tinggi seperti CT scan.
- Inception V3 Model ini menggunakan arsitektur Inception untuk melakukan pemrosesan paralel untuk berbagai ukuran filter dan membantu menghindari overfitting.



2.4 Pelatihan dan evaluasi model

Data dibagi menjadi data pelatihan, validasi, dan pengujian dengan perbandingan 70: 15: 15. Model ini dilatih menggunakan pengoptimal penurunan gradien, yang menyesuaikan parameter seperti kecepatan pembelajaran dan ukuran batch untuk memaksimalkan performa. Proses pelatihannya menggunakan teknik penghentian dini untuk mencegah overfitting.

2.5 Metrik Evaluasi

Evaluasi kinerja model dilakukan dengan menggunakan metrik berikut.

- Akurasi: Mengukur persentase prediksi yang benar.
- Akurasi dan Sensitivitas. Indikator-indikator ini penting untuk mengidentifikasi kasus kanker secara akurat dan menghindari prediksi negatif yang salah.
- Skor F1 adalah rata-rata harmonik antara presisi dan sensitivitas dan mengevaluasi keseimbangan antara presisi dan sensitivitas.
- Matriks kebingungan.
- Jelaskan performa model dalam format tabel, yang menunjukkan prediksi yang benar dan salah untuk setiap kategori.

2.6 Metode Optimasi

Pengujian dilakukan dengan memvariasikan parameter optimasi untuk setiap arsitektur. Pengujian dilakukan menggunakan arsitektur CanNet, LeNet, VGG16, dan VGG19 dengan parameter epoch 1000. Pengujian dijalankan menggunakan data pelatihan. Dari hasil pengujian pada arsitektur CanNet, metode optimasi SGD mencapai akurasi tertinggi sebesar 91%. Tiga optimasi lainnya yaitu RMSProp, Adagrad, dan Adam tidak menunjukkan perubahan akurasi yang signifikan sejak epoch pertama.

2.7 Runtime Program

Pengujian ini menghitung runtime untuk setiap arsitektur jaringan. Pengujian dijalankan dengan parameter epoch 1000. Waktu pelaksanaan dihitung dari awal pelatihan hingga perhitungan klasifikasi.

Hasil runtime arsitektur CNN ditunjukkan pada Tabel 1

Tabel 1. Rata-Rata Hasil Eksekusi Arsitektur CNN

Arsitektur	Waktu Pelatihan Model
CanNet	121981 detik (33,88 jam)
Lenet	38903 detik (10,8 jam)
VGG16	33474 detik (9.89 jam)
VGG19	36028 detik (10 jam)

Dari hasil pengujian, arsitektur yang paling lama digunakan adalah arsitektur CanNet yang mana membutuhkan waktu 121981 detik; Jadi sekitar 34 jam. Arsitektur VGG16 membutuhkan waktu paling sedikit, yaitu 33.474 detik, atau sekitar 10 jam.

Hal ini karena ukuran kernel matriks konvolusi pertama dalam arsitektur CanNet besar (yaitu 50x50). Dapat disimpulkan bahwa ukuran kernel yang besar mempengaruhi waktu pelatihan arsitektur CNN.

2.8 Metode analisis data

Perangkat lunak yang digunakan dalam penelitian ini adalah perangkat lunak Python dan TensorFlow sebagai perpustakaannya. Metode analisis data yang digunakan pada penelitian ini

adalah metode convolutional neural network yang bertujuan untuk mengklasifikasikan citra kanker paru.

2.9 Pembagian data

Data dibagi menjadi dua kategori tergantung jenisnya: data asli dan data CLAHE. Selanjutnya, kami membagi kedua jenis data tersebut masing-masing menjadi data pelatihan dan data pengujian. Data gambar terlebih dahulu dibagi menjadi data latih dan data uji dengan perbandingan 90:. Kemudian sebagian data pelatihan yang sebelumnya dibagi dengan rasio 90: 10 digunakan untuk data validasi. Deep learning Data gambar yang diolah pada tahap sebelumnya diklasifikasikan menggunakan deep learning.

3. ANALISA DAN PEMBAHASAN

Pada bagian ini berisi temuan, hasil serta pembahasan dari topik penelitian.

Tabel 2. Hasil Temuan Artikel Relevan

No.	Nama Peneliti dan Tahun	Metode yang dibahas	Tujuan Penelitian	Hasil yang didapat
1.	(Sri Widodo, 2022)	Studi literatur	Klasifikasi citra kanker dan arteri menggunakan metode convolutional neural network pada CT scan.	Ukuran input 60 x 60 piksel, nilai kecepatan pembelajaran 0,001, ukuran filter 3 x 3, jumlah epoch 30, data pelatihan 180, data pengujian 20, yang menghasilkan akurasi tertinggi sebesar 95%. Jadi adalah deteksi kanker menggunakan pembelajaran mendalam pada gambar CT scan dengan gambar yang dilokalisasi dengan sangat tepat.
2.	(Dzaky Abdillah Salafy dkk., 2023)	Deskripsi kualitatif	a. Klasifikasi organ paru normal dan organ paru yang terinfeksi sel kanker. b. Menggunakan augmentasi gambar CLAHE dengan kombinasi batas kliping dan kisi ubin yang tepat dapat	Berbagi data tidak hanya dapat memengaruhi performa model dalam mengklasifikasikan gambar, tetapi juga memiliki implikasi terhadap penggunaan CLAHE saat memperkaya gambar skala abu-abu, terutama gambar CT scan, untuk memberi informasi kepada

			mempengaruhi performa model dalam mengklasifikasikan gambar.	<p>model selama pelatihan.</p> <p>Kami merekomendasikan pelatihan dengan lebih banyak data gambar atau menggunakan kombinasi batas klip dan petak petak yang lebih beragam.</p> <p>Anda juga dapat menggunakan beberapa jenis koreksi gambar lainnya, seperti koreksi gamma dan transformasi wavelet.</p>
3.	(Aldo Vierisyah et al.)	Kualitatif Deskriptif	Membandingkan kinerja lima arsitektur (VCG16, VCG19, Resnet50, Resnet101, dan Xception) untuk meningkatkan akurasi klasifikasi kanker paru deskriptif secara kualitatif.	<p>Kami menemukan bahwa semua model memiliki kinerja akurasi yang tinggi, terutama model Resnet101 yang mencapai nilai akurasi 93,4%, sensitivitas 90,2, dan spesifisitas 95,1%, menjadikan model ini ditampilkan optimal untuk melakukan uji klasifikasi.</p> <p>Di sisi lain, nilai model VGG16, VGG19, dan Resnet50, khususnya model Xception, sangat rendah.</p> <p>b.Jika dilihat klasifikasi penyakit paru jinak, ganas dan normal, strukturnya berbeda.</p> <p>VGG16, VGG19, Resnet101, Resnet50, dan Xception dapat</p>

				mendeteksi jenis penyakit jinak, ganas, dan normal dengan kinerja yang baik dalam beberapa pengujian berdasarkan nilai hyperparameter yang diperoleh.
4.	(Andreas Galang Anugerah, 2018)	Convolutional Neural Network (CNN)	Pengembangan model CNN untuk mengklasifikasikan keganasan kanker paru pada gambar CT scan dengan tujuan untuk menciptakan klasifikasi yang akurat antara kelas ganas dan jinak.	kami menemukan bahwa arsitektur LeNet memiliki akurasi rata-rata tertinggi sebesar 81,4% dan teknik optimasi terbaik adalah gradien stokastik. Akurasi Keturunan (SGD) 91,4 D44
5.	(Yuli Martha Krist Damanik, 2023)	Metode Residual Network (ResNet) dengan tiga varian arsitektur: ResNet-18, ResNet-50, dan ResNet-101.	Dalam penelitian ini, kami menggunakan teknik pembelajaran mendalam, khususnya arsitektur Residual Network (ResNet), untuk melatih model yang menggunakan perubahan jumlah zaman untuk mengklasifikasikan stadium kanker paru-paru dari gambar CT scan.	Tujuan dari penelitian ini adalah menemukan model yang dapat mendeteksi kanker dengan akurasi tinggi. Berdasarkan hasil penelitian, ResNet-101 pada epoch 6 memberikan performa terbaik dengan akurasi validasi sebesar 99,50%, tingkat error terendah sebesar 0,3%, sensitivitas sebesar 98%, dan tingkat akurasi sebesar 99,2%.

4. KESIMPULAN

Berdasarkan hasil pengujian kinerja arsitektur convolutional neural network arsitektur CNN menggunakan arsitektur VGG16, VGG19, Resnet101, Resnet50, dan Xception dalam mengklasifikasikan kanker paru, jinak, ganas, dan normal. Hal ini dapat disimpulkan sebagai berikut.



Ukuran kernel matriks konvolusi mempengaruhi waktu pelatihan model CNN. Arsitektur CanNet dengan ukuran kernel 50x50 untuk matriks konvolusi pertama membutuhkan waktu paling lama yaitu kurang lebih 34 jam. Di sisi lain, arsitektur VGG16 memiliki waktu penyelesaian tercepat, yaitu sekitar 10 jam. VGG16 memiliki matriks konvolusi lebih banyak, tetapi ukuran kernel hanya 3x3.

Terlihat bahwa semua model mempunyai kinerja akurasi yang tinggi. Secara khusus, model Resnet101 mencapai nilai akurasi 93,4%, sensitivitas 90,2, dan spesifisitas 95,1%, menunjukkan bahwa model ini paling baik dalam melakukan klasifikasi. Sedang diuji. Di sisi lain, nilai model VGG16, VGG19, dan Resnet50, khususnya model Xception, sangat rendah.

Kami menyarankan untuk meningkatkan ukuran data gambar pelatihan atau menggunakan kombinasi batas klip dan petak petak yang lebih beragam. Anda juga dapat menggunakan jenis koreksi gambar lainnya, seperti: Koreksi gamma atau transformasi wavelet.

REFERENSI

- R.Moch Diar, R.Y.Fu'adah, K.Usman, "Klasifikasi penyakit paru-paru berdasarkan pengolahan citra sinar-X menggunakan jaringan saraf konvolusional," e-Proceedings of Engineering, vol.9, nein.2, S.476–484, 2022.
- Arifin, D., & Putranto, T. S. (2017). Pengembangan Model Sistem Pengelolaan Data Pemerintah Daerah Menggunakan Pendekatan Big Data. Institut Teknologi Sepuluh Nopember.
- A.Guidan MV, LMPD. Teknik pemrosesan gambar yang berbeda digunakan untuk mendeteksi stadium kanker paru-paru dari gambar CT scan. IOSR J Komputasi Eng 2014;16(5): 28–35.
- Sharma D, Gagandeep J. Identifikasi kanker paru-paru menggunakan teknik pengolahan gambar. Int Conf Comput Tech Artif Intell.
- Sri Widodo Vol 10, No 2 (2022) Klasifikasi kanker dan arteri dalam tomografi komputer menggunakan jaringan saraf konvolusional pembelajaran mendalam.
- Sahni P, Mittal N. Deteksi kanker payudara menggunakan teknik pengolahan citra. Catatan Kuliah Teknik 2019;(Maret): 813–23.
- Wirayudha, A., Yudistira, H.A., Wijaya, J.S., Yusuf, R. & Rosyani, P. (2023). Studi Kasus: Penerapan Jaringan Syaraf Konvolusional untuk Pengenalan Manusia. JRIIN: Jurnal Penelitian Inovasi Informatika, 2(5), 795-801.
- Ningsih, R., & Subekti, E. (2020). Analisis Pemanfaatan Teknologi Informasi dalam Menunjang E-Commerce di Kalangan UMKM. KLIK: Kumpulan Jurnal Ilmu Komputer, 8(3), 1514.
- F.M.Hana dan I.D.Maulida, "Analisis parameter penyetaraan histogram adaptif terbatas kontras (CLAHE) untuk identifikasi cetakan buku jari", Jurnal Fisika: Conference Series, IOP Publishing Ltd, Februari 2021 doi: 10.1088/1742-6596/1764/1/012049
- Pandu Wikaksono, D., Abdulroman, F., Rizky Maulana, M., Nuruddin, N., dan Roshani, P. (2024). Pengenalan gender secara real-time dengan pengenalan wajah menggunakan Convolution Neural Network dan OpenCV. JRIIN: Jurnal Riset Inovasi Informatika, 2(6), 869–873.
- Indah K., A., & Afifah, A. (2021). Pengaruh Kemajuan Teknologi Informasi Terhadap Produktivitas Kerja di Masa Pandemi COVID-19. Inohim: Jurnal Inovasi dan Humaniora, 5(2),