

ARTICLE

Deteksi Kanker Payudara Hasil Citra Mammografi menggunakan Metode *Convolutional Neural Network* (CNN) Arsitektur ResNet-50

Breast Cancer Detection from Mammography Images using Convolutional Neural Network (CNN) Method ResNet-50 Architecture

Ammanda Rosnatul Nelda.A¹, Rudi Setiawan², Muhammad Afif Hendrawan*³, Doni Bowo Nugroho⁴, dan Ahmad Novian Rahman Hakim⁵

¹Teknik Biomedis, Fakultas Teknologi Industri, Institut Teknologi Sumatera, Lampung Selatan, Lampung, Indonesia

² Teknik Biomedis, Fakultas Teknologi Industri, Institut Teknologi Sumatera, Lampung Selatan, Lampung, Indonesia

³ Teknik Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Malang, Malang, Jawa Timur, Indonesia

⁴ Teknik Biomedis, Fakultas Teknologi Industri, Institut Teknologi Sumatera, Lampung Selatan, Lampung, Indonesia

⁵ PT. C-Tech Labs Edwar Technology

*Penulis Korespondensi: afif.hendrawan@polinema.ac.id

(Disubmit 22-11-23; Diterima 23-3-24; Dipublikasikan online pada 30-3-24)

Abstrak

Kanker payudara merupakan salah satu jenis kanker yang paling umum di dunia dan menjadi masalah serius bagi kesehatan wanita, sehingga sangat penting bagi setiap perempuan untuk melakukan pengecekan rutin secara dini untuk meminimalkan angka kematian. Pemeriksaan ini dapat dilakukan dengan menggunakan mammografi dan analisis citra mammografi. Penelitian ini mengusulkan deteksi kanker payudara berbasis *Convolutional Neural Network* (CNN) dengan menggunakan parameter tekstur dan bentuk dari gambar mammografi untuk menentukan kelasnya. Terdapat dua kelas yang dilatih menggunakan CNN yaitu kanker dan non-kanker. Penelitian ini membandingkan kinerja model arsitektur ResNet-50 asli dengan arsitektur ResNet-50 yang dimodifikasi pada layer terakhir melalui penambahan *fully connected layer*, *batch normalization*, dan *dropout*. Hasil penelitian menunjukkan bahwa model ini mampu mendeteksi kanker payudara dengan akurasi tinggi. Model terbaik pada penelitian ini menggunakan arsitektur ResNet-50 yang dimodifikasi, dengan parameter *optimizer Adam*, *batch size* 8, fungsi aktivasi *sigmoid*, *learning rate* 0,0005, dan 30 *epoch*. Model ini mencapai nilai rata-rata keseluruhan dengan *accuracy* sebesar 97,00%, *precision* sebesar 97,00%, *recall* sebesar 97,00%, dan *f1-score* sebesar 97,00%.

Kata kunci: kanker payudara; *convolutional neural network*; gambar mammografi; arsitektur ResNet-50

Abstract

Breast cancer is one of the most common types of cancer in the world and a serious health issue for women, making it crucial for every woman to undergo regular early screenings to minimize mortality rates. This screening can be performed using mammography and mammogram image analysis. This study proposes a breast cancer detection method based on Convolutional Neural Network (CNN), using texture and shape parameters from mammogram images to determine the class. There are two classes trained using CNN, namely cancer and non-cancer. This study compares the performance of the original ResNet-50 architecture model with the modified ResNet-50 architecture at the last layer through the addition of fully connected layers, batch normalization, and dropout. The results show that this model is capable of detecting breast cancer with high accuracy. The best model in this study uses a modified ResNet-50 architecture, with the Adam optimizer, batch size of 8, sigmoid activation function, learning rate of 0.0005, and 30 epochs. This model achieved an overall average with an accuracy of 97.00%, precision of 97.00%, recall of 97.00%, and f1-score of 97.00%.

Keywords: breast cancer; *convolutional neural network*; mammography images; ResNet-50 architecture

1. Pendahuluan

Kanker adalah pertumbuhan sel abnormal yang bisa cepat menyebar ke bagian tubuh lain [1]. Kanker merupakan penyebab kematian kedua terbanyak di dunia setelah penyakit kardiovaskular. Berdasarkan data dari *World Health Organization* (WHO), jumlah kasus kanker meningkat dari 12,7 juta pada tahun 2018 menjadi 14,2 juta pada tahun 2020. Diperkirakan jumlah ini akan terus bertambah hingga mencapai 26 juta kasus pada tahun 2030, dengan total kematian sebanyak 17 juta orang [2]. Kanker memiliki beragam jenis, salah satu kanker dengan tingkat kematian tertinggi adalah kanker payudara.

Kanker payudara adalah tumor ganas yang tumbuh dan berkembang abnormal di payudara [3]. Kondisi ini diawali dengan benjolan yang ada pada kelenjar susu atau jaringan penunjang payudara, yang bisa menyebar ke bagian tubuh lain. Berdasarkan data dari *Global Cancer Observatory* 2020, bahwa saat ini tercatat ada 396.914 kasus kanker baru di Indonesia, dengan jumlah kasus kanker payudara ada 68.858 kasus atau sekitar 16,6% [2].

Dengan tingginya jumlah kasus kanker payudara, setiap perempuan dianjurkan melakukan pemeriksaan rutin untuk mendeteksi kanker lebih awal guna mengurangi angka kematian. Pemeriksaan ini bisa dilakukan melalui mammografi, alat radiologi yang digunakan oleh ahli untuk mendeteksi kanker dan memerlukan keterampilan khusus seorang radiolog [4]. Mammografi saat ini dianggap sebagai metode terbaik untuk mendeteksi kanker payudara secara dini, sehingga meningkatkan peluang kesembuhan pasien [5].

Saat ini, analisis citra mammografi dapat dilakukan dengan menggunakan metode *Convolutional Neural Networks* (CNN). Algoritma ini dapat mempelajari dan mengekstrak fitur gambar dengan kompleksitas dan detail yang tinggi, mendeteksi pola yang mungkin sulit dilihat oleh manusia, sehingga mampu mengklasifikasikan dan mendiagnosis kanker payudara [6]. Model CNN banyak digunakan dalam berbagai aplikasi, termasuk segmentasi gambar medis, deteksi objek, dan pemrosesan bahasa alami seperti analisis teks [7].

Arsitektur CNN secara otomatis dan hierarkis dapat mempelajari fitur dari data gambar [8]. Pengolahan data citra menggunakan metode CNN dilakukan secara bertahap dan berulang-ulang melalui serangkaian *convolutional layer* (konvolusi), *pooling layer* (penggabungan), dan *fully connected layer* (terhubung secara penuh)[9]. Salah satu contoh arsitektur pada metode CNN adalah *Residual Neural Network* (ResNet) yang dikembangkan oleh Microsoft pada tahun 2015 [10].

Arsitektur ResNet (*Residual Neural Network*) ini terdiri dari berbagai jenis lapisan, termasuk lapisan 18, 34, 50, 101, dan 152 [11]. ResNet berhasil meraih kemenangan di ILSVRC 2015 [12]. Keunggulan utamanya adalah kemampuannya mengatasi masalah *vanishing gradient*, yang sering terjadi dalam jaringan saraf dalam, melalui penerapan *shortcut connections* atau *skip connections* [13]. Prinsip kerja ResNet adalah membangun jaringan yang lebih dalam untuk mengatasi masalah gradien yang menghilang, tanpa memerlukan pengaturan khusus pada lapisan-lapisannya [14].

Oleh karena itu, peneliti melakukan pengembangan dan evaluasi model arsitektur ResNet-50 untuk mendeteksi kanker payudara secara efektif. Penelitian ini memanfaatkan dataset INbreast yang diambil dari Centro Hospitalar de S. Joao (CHSJ), Breast Center, yang terletak di Porto, Portugal. Dataset ini terdiri dari berbagai citra mammografi yang telah dilabeli, memungkinkan analisis yang lebih mendalam mengenai keberadaan kanker payudara.

Tujuan utama penelitian ini adalah untuk mengidentifikasi dengan akurasi tinggi apakah citra yang dianalisis menunjukkan adanya struktur keberadaan kanker payudara atau tidak. Dengan menggunakan arsitektur ResNet-50, peneliti berupaya memanfaatkan keunggulan jaringan saraf dalam untuk mempelajari fitur-fitur kompleks dalam citra, yang sering kali sulit dideteksi oleh pengamat manusia.

2. Metode

2.1 Bahan dan Alat

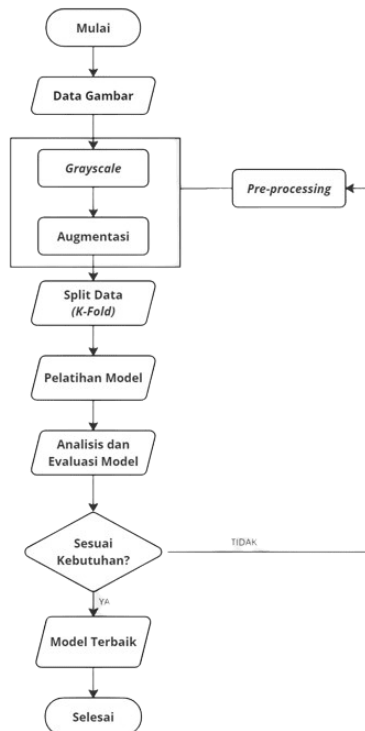
Penelitian ini menggunakan data sekunder yang diambil dari dataset INbreast, yang berisi citra mammografi dari Centro Hospitalar de S. Joao (CHSJ), Breast Center, di Porto, Portugal. Dataset INbreast terdiri dari 410 gambar, yang mencakup 343 gambar kanker dan 67 gambar non-kanker. Dari 410 gambar tersebut, data dikumpulkan dari 115 kasus kanker payudara, di mana 90 kasus berasal dari wanita yang mengalami masalah pada kedua payudara, sehingga terdapat 4 gambar untuk setiap kasus. Sementara itu, 25 kasus lainnya berasal

dari wanita yang telah menjalani mastektomi, di mana setiap kasus hanya memiliki 2 gambar.

Deteksi awal kanker payudara dilakukan dengan menggunakan *Google Collab Pro*, sebuah platform berbasis cloud yang mendukung bahasa pemrograman *Python* yang dirancang untuk mendukung berbagai kebutuhan analisis data. Platform ini sangat ideal untuk analisis citra mammografi, yang sering kali memerlukan pemrosesan data besar dan algoritma pembelajaran mendalam.

2.2 Tahapan Penelitian

Adapun tahapan-tahapan jurnal ini berdasarkan diagram alir berikut ini:



Gambar 2. 1 Diagram alir pengolahan dataset citra metode CNN

Berdasarkan Gambar 2.1 Penelitian ini dibagi menjadi empat tahap utama. Tahap pertama melibatkan pengumpulan data, diikuti oleh tahap *preprocessing* yang mencakup konversi ke grayscale dan augmentasi data. Selanjutnya, data dibagi menggunakan teknik *k-fold cross-validation* untuk memastikan validitas hasil citra, dengan perbandingan untuk data pelatihan, validasi, dan pengujian adalah 60:20:20. Teknik ini membagi dataset menjadi 5 *fold*, di mana setiap *fold* digunakan baik sebagai data pengujian maupun data pelatihan. Setelah itu, dilakukan pengujian model untuk mengidentifikasi data citra. Tahap terakhir adalah evaluasi model untuk menilai kinerja model CNN berbasis arsitektur ResNet-50 dengan menggunakan metrik seperti *accuracy*, *presisi*, *recall*, dan *f1-score*.

2.3 Analisis Data

Proses analisis data untuk merancang model CNN dengan arsitektur ResNet-50 dalam mendeteksi kanker payudara dilakukan di *Google Collab Pro*, platform cloud yang mendukung *Python* untuk analisis citra mammografi. Selama proses *training*, penulis menyesuaikan *hyperparameter* untuk meningkatkan kinerja model, dengan rentang nilai pengujian *hyperparameter* yang tercantum di Tabel 2.1.

Tabel 2. 1 Pengujian *hyperparameter*

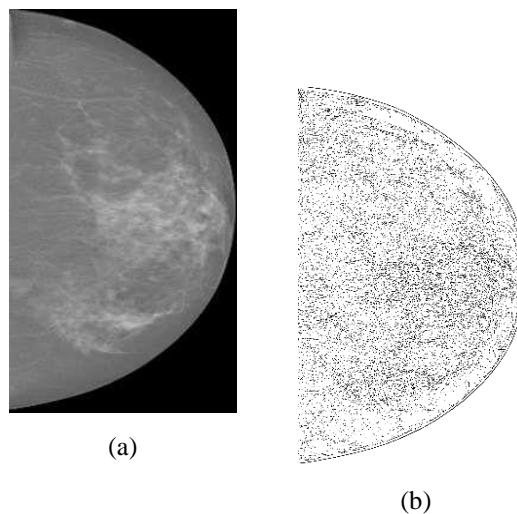
Parameter	Variasi
<i>Batch Size</i>	8,16
<i>Epoch</i>	15,30
<i>Learning Rate</i>	0.001, 0.0005
<i>Optimizer</i>	<i>Adam</i>
<i>Activation function</i>	<i>Sigmoid</i>
<i>Target size</i>	224

Setelah penulis menguji berbagai kombinasi *hyperparameter* sesuai Tabel 2.1, model terbaik dipilih berdasarkan akurasi data pelatihan, sambil memastikan tidak terjadi *overfitting* dengan mengevaluasi akurasi data validasi.

Model kemudian diuji dengan data pengujian, dan hasilnya divisualisasikan dalam *confusion matrix*, yang menunjukkan jumlah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dengan informasi ini, nilai *accuracy*, *recall*, *presisi*, dan *f1-score* dihitung untuk menilai performa model secara menyeluruh dalam diagnosis medis.

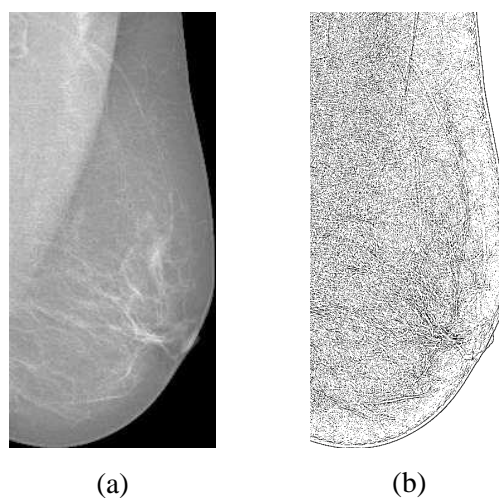
3. Hasil dan Pembahasan

Sebelum dilakukan pengujian, dilakukan preprocessing data yang mencakup *grayscale* dan augmentasi data. Proses *grayscale* mengubah gambar berwarna menjadi gambar dengan satu kanal warna saja, yaitu hitam-putih, yang berfungsi untuk menyederhanakan representasi visual tanpa mengurangi informasi penting yang dibutuhkan oleh model. Gambar 3.1 merupakan perbandingan kanker payudara setelah dan sebelum *grayscale*.



Gambar 3. 1 Perbandingan kanker payudara proses *grayscale*
(a) sebelum *grayscale*. (b) setelah *grayscale*

Selanjutnya untuk Gambar 3.2 merupakan perbandingan non kanker payudara setelah dan sebelum *grayscale*.



Gambar 3. 2 non kanker payudara pada proses *grayscale*
(a) sebelum *grayscale*. (b) setelah *grayscale*

Selain itu, augmentasi data diterapkan untuk meningkatkan variasi data yang ada, sehingga model dapat lebih tahan terhadap *overfitting*. Teknik augmentasi yang diterapkan adalah *rotation*, *flipping*, HSV (*Hue*, *Saturation*, *Value*). Setelah augmentasi data, langkah selanjutnya adalah membagi data menggunakan metode *K-Fold Cross*

Validation. Teknik ini membagi data ke dalam beberapa subset yang digunakan untuk melatih dan menguji model klasifikasi secara bergantian.

Selanjutnya, dilakukan pengujian terhadap beberapa parameter yang ditentukan pada Tabel 2.1, sehingga didapatkan model terbaik dengan menggunakan Resnet-50 hasil modifikasi melalui parameter *optimizer* Adam, *batch size* 8, *activation function sigmoid*, pada *learning rate* 0.0005, *epoch* 30, dan pada *fold* ke-3 yang dapat dilihat pada Tabel 3.1.

Tabel 3.1 Model terbaik

<i>Optimizer</i>	<i>Batch size</i>	<i>K-fold</i>	<i>Train acc (%)</i>	<i>Val acc (%)</i>	<i>Test acc (%)</i>
Adam	8	1	84,20%	54,49%	79,19%
		2	85,30%	58,39%	80,49%
		3	86,37%	90,12%	90,11%
		4	86,70%	67,75%	84,78%
		5	85,21%	63,93%	90,49%

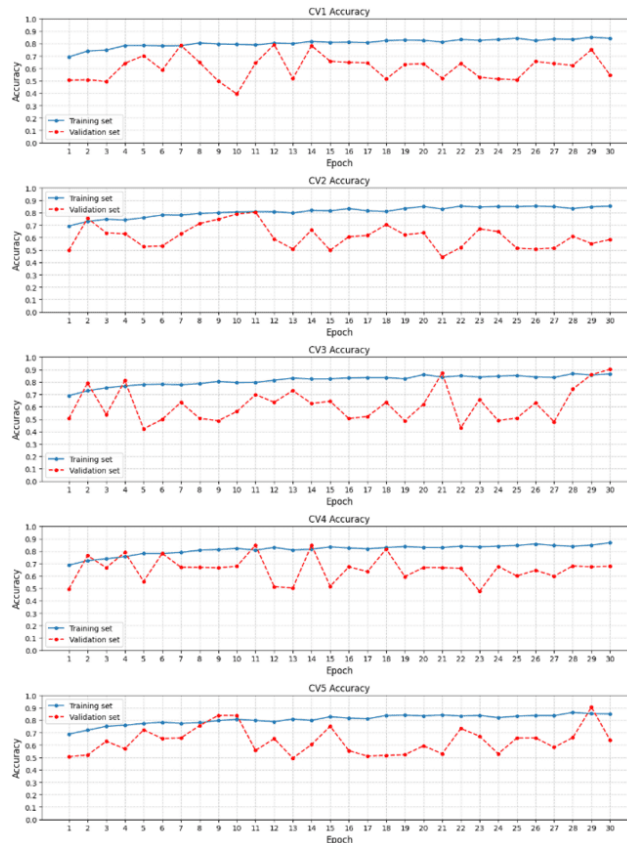
Berdasarkan Tabel 3.1, model terbaik diperoleh pada *fold* ke-3, dengan akurasi data *train* sebesar 86,37%, akurasi data validasi sebesar 90,12%, dan akurasi data *test* sebesar 90,11%. Hal ini menunjukkan bahwa model mampu belajar secara efektif dari data pelatihan yang digunakan. Meskipun akurasi data train tidak terlalu tinggi, hal ini mengindikasikan bahwa model tidak mengalami overfitting. Untuk melihat performa model ResNet-50 terbaik tersebut, hasilnya disajikan pada Tabel 3.2.

Tabel 3.2 Evaluasi model terbaik

<i>Optimizer</i>	<i>Batch size</i>	<i>K-fold</i>	<i>Prec (%)</i>	<i>Recall (%)</i>	<i>F1-score (%)</i>
Adam	8	1	91,79%	73,66%	81,73%
		2	62,63%	96,74%	76,03%
		3	82,87%	97,62%	89,64%
		4	82,93%	85,86%	84,37%
		5	85,14%	94,97%	89,79%

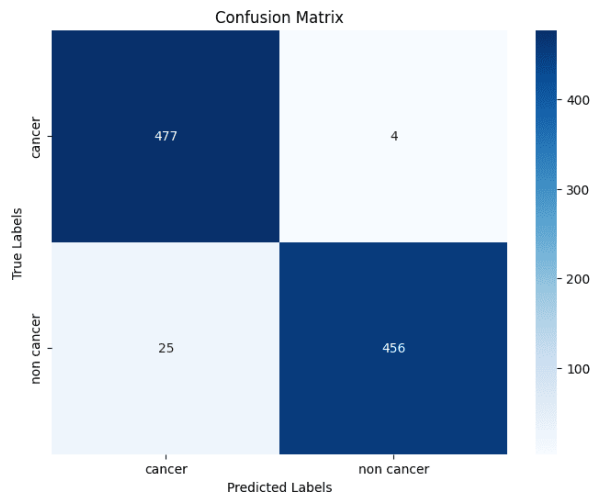
Berdasarkan Tabel 4.2, evaluasi model terbaik di *fold* ke-3 menunjukkan *precision* sebesar 82,87%, yang menandakan masih adanya *false positives*. *Recall* mencapai 97,62%, artinya model hampir mendeteksi semua *true positives*. Dengan *f1-score* 89,64%, ini mencerminkan keseimbangan yang baik antara *precision* dan *recall*, sehingga performa keseluruhan model tetap kuat.

Selanjutnya, ditampilkan grafik pengujian data *train* dan data validasi yang ada pada Gambar 3.1. Berdasarkan Gambar 3.1 dapat dilihat bahwa, kurva pelatihan yang menunjukkan peningkatan konsisten artinya, model terus melatih kemampuannya dalam mengurangi loss dan meningkatkan akurasi pada data latih. Sebaliknya, kurva validasi yang mengalami fluktuasi mencerminkan adanya variasi kinerja model saat diuji pada data yang tidak digunakan selama pelatihan. Perubahan ini biasanya disebabkan oleh ketidakstabilan dalam subset data validasi atau karena ukuran subset yang relatif kecil.



Gambar 3.3 Akurasi *training* model terbaik

Pada tahap terakhir, model akan di uji menggunakan data baru yang belum pernah digunakan sebelumnya pada proses *training*. Gambar 3.2 merupakan tampilan hasil data *testing* model menggunakan arsitektur ResNet-50.



Gambar 3.4 *Confusion matrix* model terbaik

Berdasarkan gambar 3.4 hasil evaluasi menggunakan *confusion matrix* menunjukkan bahwa prediksi yang benar untuk kasus kanker berjumlah 477 data, sementara hanya 4 data kanker yang diprediksi salah. Untuk non-kanker, terdapat 456 data yang diprediksi dengan benar dan 25 data yang salah prediksi.

Tabel 3.3 Evaluasi *confusion matrix*

Kelas	<i>Prec</i> (%)	<i>Recall</i> (%)	<i>F1-score</i> (%)	<i>Support</i>
Kanker	95.00	99.00	97.00	481
Non Kanker	99.00	95.00	97.00	481
Accuracy			97.00	962
Average	97.00	97.00	97.00	962

Berdasarkan Tabel 3.3 dapat dilihat hasil evaluasi model untuk data kanker memperoleh nilai akurasi sebesar 97.00%, nilai *precision* sebesar 95.00%, nilai *recall* 99.00%, dan nilai *f1-score* 97.00%.

Sebagai perbandingan dengan jurnal sebelumnya pada jurnal Setiawan, Wahyudi (2021) menyebutkan bahwa dengan menggunakan data yang berbeda tetapi klasifikasi 2 kelas yaitu jinak dan ganas melalui metode CNN didapatkan akurasi 83.62% untuk arsitektur ResNet50 dan 85.06% untuk arsitektur ResNet18 [6]. Sedangkan pada jurnal Himanish Shekhar (2023) melakukan pendeteksian kanker menggunakan data yang sama yaitu INBreast dengan 2 kelas klasifikasi jinak dan ganas melalui metode CNN arsitektur ResNet-50 menghasilkan *accuracy* sebesar 91.5%, *precision* sebesar 93.2%, *recall* sebesar 91.1% [15].

Dengan demikian, dalam penelitian ini model yang diterapkan pada dataset INbreast dengan 410 sampel menunjukkan kinerja yang sangat baik, dengan nilai *accuracy*, *precision*, *recall*, dan *f1-score* keseluruhan masing-masing sebesar 97.00%. Hasil ini sangat bagus dan terbaik jika dibandingkan dengan penelitian lain yang menggunakan dataset serupa.

4. Simpulan

Berdasarkan penelitian deteksi kanker payudara menggunakan metode CNN dengan arsitektur ResNet-50, dapat disimpulkan bahwa modifikasi arsitektur ResNet-50 dilakukan dengan menambahkan layer flatten, beberapa lapisan dense, batch normalization, serta dropout dengan aktivasi sigmoid. Modifikasi ini bertujuan untuk meningkatkan akurasi deteksi kanker payudara. Performa terbaik diperoleh pada model modifikasi yang menggunakan optimizer Adam, batch size 8, fungsi aktivasi sigmoid, learning rate sebesar 0.0005, dan 30 epoch. Dengan konfigurasi ini, model mencapai *accuracy*, *precision*, *recall*, dan F1-score sebesar 97.00%.

Adapun saran dari penelitian adalah :

1. Pastikan dataset seimbang antara kelas 'kanker' dan 'non-kanker' serta eksplorasi teknik augmentasi data.
2. Lakukan penyesuaian *hyperparameter* dan fine-tuning pada lapisan ResNet-50
3. Gunakan *early stopping* dan regularisasi untuk mencegah *overfitting*

DAFTAR PUSTAKA

- [1] S. Ketut, "Kanker payudara: Diagnostik, Faktor Risiko dan Stadium," *Ganesha Med. J.*, vol. 2, no. 1, pp. 2–7, 2022.
- [2] W. Asmalinda *et al.*, "Deteksi Dini Kanker Payudara Menggunakan Pemeriksaan Payudara Sendiri (Sadari) (Early Detection of Breast Cancer Using Breast Self-Examination)," *J. Abdikemas*, vol. 4, no. 1, pp. 10–17, 2022, [Online]. Available: <https://doi.org/10.36086/j.abdikemas.v4i1>
- [3] A. Rizka, M. K. Akbar, and N. A. Putri, "CARCINOMA MAMMAE SINISTRA T4bN2M1 METASTASIS PLEURA," *AVERROUS J. Kedokt. dan Kesehat. Malikussaleh*, vol. 8, no. 1, p. 23, 2022, doi: 10.29103/averrous.v8i1.7006.
- [4] Z. Wang *et al.*, "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion with CNN Deep Features," *IEEE Access*, vol. 7, no. September 2021, pp. 105146–105158, 2019, doi: 10.1109/ACCESS.2019.2892795.
- [5] C. A. A. Baruna and I. B. T. W. Manuaba, "Ketepatan ultrasonografi dan mammografi dalam mendiagnosis wanita dengan kanker payudara di RSUP Sanglah Denpasar," *Intisari Sains Medis*, vol. 10, no. 3, pp. 684–687, 2019, doi: 10.15562/ism.v10i3.489.
- [6] W. Setiawan, "Klasifikasi Citra Histopatologi Kanker Payudara menggunakan Data Resampling Random dan Residual Network," vol. 01, pp. 70–77, 2021, doi: 10.21456/vol11iss1pp70-77.
- [7] M. Masud, A. E. Eldin Rashed, and M. S. Hossain, "Convolutional neural network-based models for diagnosis of breast cancer," *Neural Comput. Appl.*, 2022, doi: 10.1007/s00521-020-05394-5.
- [8] Q. A. Al-hajja, "Breast Cancer Diagnosis in Histopathological Images Using ResNet-50 Convolutional Neural Network," vol. 50, 2020.
- [9] S. A. Agrawal, V. D. Rewaskar, R. A. Agrawal, S. S. Chaudhari, and Y. Patil, "International Journal of

INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Advancements in NSFW Content Detection : A Comprehensive Review of ResNet-50 Based Approaches INTELLIGENT SYSTEMS AND APPLICATIONS IN Advancements in NSFW Content Detection : A C,” no. October, 2023.

- [10] B. M. Sujatmiko, E. Yudaningtyas, and P. M. Raharjo, “CONVOLUTION NEURAL NETWORK DENGAN DESAIN JARINGAN RESNET SEBAGAI METODE KLASIFIKASI TUMOR KULIT CONVOLUTION NEURAL NETWORK USING RESNET NETWORK DESIGN AS SKIN TUMOR,” vol. 11, no. 1, pp. 53–64, 2022.
- [11] F. Ben Nasr Barber and A. Elloumi Oueslati, “Human exons and introns classification using pre-trained Resnet-50 and GoogleNet models and 13-layers CNN model,” *J. Genet. Eng. Biotechnol.*, 2024, doi: 10.1016/j.jgeb.2024.100359.
- [12] Q. A. Fitroh and S. Uyun, “Deep Transfer Learning untuk Meningkatkan Akurasi Klasifikasi pada Citra Dermoskopi Kanker Kulit,” pp. 78–84, 2023.
- [13] R. Erwandi, “Klasifikasi Kanker Payudara Menggunakan Residual Neural Network,” *Indones. J. Comput.*, vol. 5, no. 1, pp. 45–52, 2020, doi: 10.21108/indojc.2020.5.1.373.
- [14] A. Ridhovan and A. Suharso, “PENERAPAN METODE RESIDUAL NETWORK (RESNET) DALAM KLASIFIKASI PENYAKIT PADA DAUN GANDUM,” *JIPi (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, 2022, doi: 10.29100/jipi.v7i1.2410.
- [15] H. S. Das, A. Das, A. Neog, S. Mallik, K. Bora, and Z. Zhao, “Breast cancer detection: Shallow convolutional neural network against deep convolutional neural networks based approach,” *Front. Genet.*, vol. 13, no. January, pp. 1–14, 2023, doi: 10.3389/fgene.2022.1097207.