JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Klasifikasi Kanker Payudara Berbasis Deep Learning Menggunakan Vision Transformer dengan Teknik Augmentasi Data Citra

Muhamad Salman Ardiyansyah*, Fajri Rakhmat Umbara, Melina

Fakultas Sains dan Informatika, Program Studi Teknik Informatika, Universitas Jenderal Achmad Yani, Cimahi, Indonesia
Email: msalman21@if.unjani.ac.id, fajri.rakhmat@lecture.unjani.ac.id, melina@lecture.unjani.ac.id
Email Penulis Korespondensi: msalman21@if.unjani.ac.id
Submitted **17-05-2025;** Accepted **13-06-2025;** Published **30-06-2025**

Abstrak

Kanker payudara menempati posisi utama dalam penyebab kematian perempuan di dunia. Deteksi dini melalui analisis citra mammografi memainkan peran penting dalam meningkatkan peluang kesembuhan. Namun, interpretasi citra mammografi secara manual membutuhkan keahlian tinggi dan rentan terhadap kesalahan. Penelitian ini bertujuan untuk mengembangkan model klasifikasi kanker payudara dengan gambar mammografi menggunakan arsitektur *Vision Transformer* (ViT) tanpa *transfer learning*. Dataset yang digunakan adalah *Digital Database for Screening Mammography* (DDSM) yang terdiri dari dua kategori, yaitu jinak dan ganas. Untuk mengatasi ketidakseimbangan kelas, digunakan teknik *undersampling* dan augmentasi data (*flipping*, *rotation*, *cropping*, dan *noise injection*). Gambar kemudian dinormalisasi dan diubah ukurannya menjadi 224×224 piksel untuk disesuaikan dengan arsitektur ViT. Model dilatih selama lima *epoch* dengan *batch size* 16. Evaluasi dilakukan pada data uji menggunakan tujuh metrik, yaitu akurasi, presisi, *recall*, F1-*score*, *Matthews Correlation Coefficient* (MCC), *Cohen's Kappa Score*, dan *Area Under Curve* (AUC). Hasil menunjukkan bahwa model mencapai akurasi 92,50%, presisi 90,48%, *recall* 95,00%, F1-*score* 92,68%, MCC 85,11%, *Kappa Score* 85,00%, dan AUC 95,75%. Hasil ini menunjukkan bahwa ViT efektif digunakan dalam klasifikasi gambar mammografi dan berpotensi sebagai sistem pendukung diagnosis kanker payudara secara otomatis.

Kata Kunci: Augmentasi data; evaluasi model; klasifikasi kanker payudara; mammografi; vision transformer

Abstract

Breast cancer ranks among the leading causes of death in women worldwide. Early detection through mammographic image analysis plays a crucial role in increasing survival rates. However, manual interpretation of mammograms requires expert knowledge and is prone to errors. This study aims to develop a breast cancer classification model using mammography images based on the Vision Transformer (ViT) architecture without employing transfer learning. The dataset used is the Digital Database for Screening Mammography (DDSM), consisting of two categories: benign and malignant. To address class imbalance, undersampling and data augmentation techniques (flipping, rotation, cropping, and noise injection) were applied. All images were normalized and resized to 224×224 pixels to match the ViT input requirements. The model was trained for five epochs with a batch size of 16. Evaluation on the test data was conducted using seven metrics: accuracy, precision, recall, F1-score, Matthews Correlation Coefficient (MCC), Cohen's Kappa Score, and Area Under the Curve (AUC). The results show that the model achieved an accuracy of 92.50%, precision of 90.48%, recall of 95.00%, F1-score of 92.68%, MCC of 85.11%, Kappa Score of 85.00%, and AUC of 95.75%. These findings indicate that the Vision Transformer is highly effective for mammographic image classification and holds potential as a reliable tool for automated breast cancer diagnosis support.

Keywords: Data augmentation; model evaluation; breast cancer classification; mammography; vision Transformer;

1. PENDAHULUAN

Kanker payudara merupakan salah satu masalah kesehatan prioritas di seluruh dunia dan di Indonesia. Perubahan yang terjadi pada penderita kanker payudara memerlukan penanganan khusus, meliputi aspek perawatan fisik, perawatan psikologis, dukungan sosial, dan kebutuhan spiritual [1]. Menurut laporan terbaru Organisasi Kesehatan Dunia (WHO) pada tahun 2018, jumlah kasus baru kanker payudara mencapai 58.256 kasus atau setara dengan 16,7% dari total kasus kanker yang terdiagnosis sebanyak 348.809 kasus. Secara global, kanker payudara merupakan kanker yang paling umum terjadi pada wanita, terhitung hampir seperempat dari seluruh kasus kanker pada kelompok ini. Sekitar 1,15 juta orang baru didiagnosis setiap tahunnya. Di Indonesia, sebagian besar kasus kanker payudara (lebih dari 80% dari seluruh kasus) terdeteksi pada stadium lanjut sehingga mempersulit proses pengobatan. Penyakit ini menyoroti pentingnya upaya pencegahan dan deteksi dini sebagai alat yang efektif dalam pengobatan kanker [2].

Diagnosis dini kanker payudara adalah salah satu pendekatan terbaik untuk mencegah penyebaran penyakit lebih lanjut. Penggunaan *machine learning* untuk melakukan proses identifikasi membuatnya lebih cepat dan akurat [3]. Penggunaan *machine learning* dalam layanan kesehatan memungkinkan proses identifikasi menjadi lebih cepat dan akurat, mendukung pengolahan data kesehatan yang optimal dan membuka banyak kemungkinan baru dalam perawatan kesehatan [4]. Dalam penerapan *deep learning* untuk diagnosis kanker payudara, berbagai arsitektur model telah menunjukkan hasil yang menjanjikan seperti *Vision Transformer* (ViT) yang memanfaatkan mekanisme perhatian (*self-attention*) untuk memahami hubungan spasial global pada gambar, termasuk tumor pada mammogram. Pada proses pembagian gambar menjadi '*patch*', model ini dapat menangkap detail penting tanpa kehilangan informasi global, sehingga menghasilkan performa yang lebih andal dibandingkan dengan *Convolutional Neural Network* (CNN) [5].

Beberapa penelitian terdahulu yang mengkaji metode pada diagnosis kanker payudara seperti dalam penelitian [6], menunjukkan bahwa belakangan ini metode *machine learning* semakin sering digunakan untuk meningkatkan akurasi dalam mendeteksi kanker payudara. Salah satu pendekatan yang dikembangkan adalah model *Vision Transformer*, yang menggunakan teknik *transfer learning*. Model ini dirancang untuk mengatasi kelemahan metode CNN, yang biasanya

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

hanya fokus pada sebagian kecil gambar dan membutuhkan proses komputasi yang rumit. Penelitian [7] menunjukkan bahwa penggunaan data augmentasi dalam klasifikasi citra semakin umum digunakan untuk meningkatkan akurasi model berbasis *machine learning*. Salah satu pendekatan yang diterapkan dalam penelitian ini adalah penggunaan CNN untuk klasifikasi citra apel. Model ini dirancang untuk mengatasi permasalahan *overfitting* yang sering terjadi akibat jumlah dataset yang terbatas. Untuk meningkatkan performa CNN, penelitian ini menerapkan teknik augmentasi data, seperti *flipping, cropping, rotation,* dan *noise injection*.

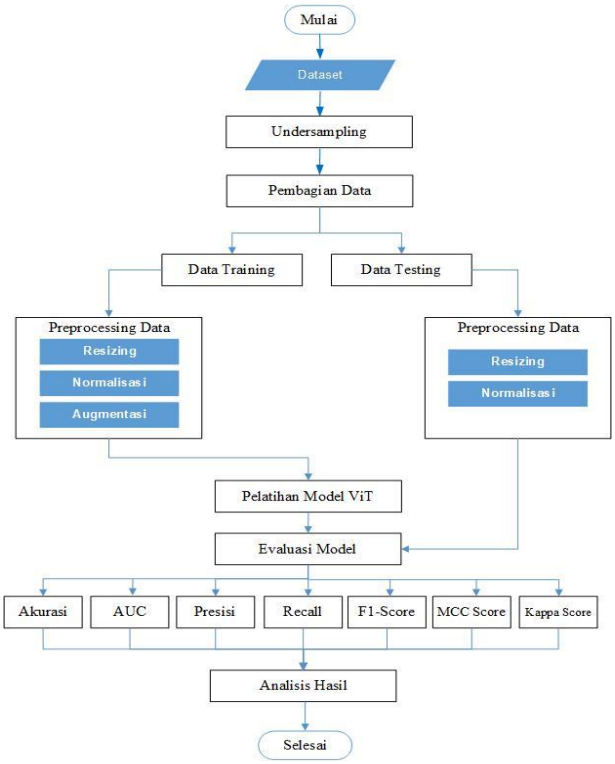
Meskipun ViT mampu menangkap hubungan spasial global dalam gambar, penggunaannya masih menghadapi tantangan berupa kompleksitas komputasi yang tinggi. Hal ini disebabkan oleh mekanisme *self-attention* dalam ViT, yang membutuhkan perhitungan hubungan antar setiap elemen dalam gambar, sehingga meningkatkan beban komputasi dibandingkan dengan model CNN tradisional [8]. Untuk mengatasi tantangan tersebut, digunakan teknik augmentasi data. Augmentasi data adalah metode yang efektif untuk memperbesar keragaman data tanpa menambah data secara fisik, sehingga model dapat belajar dari variasi gambar yang lebih banyak tanpa memerlukan sumber daya komputasi yang lebih besar [9]. Teknik augmentasi ini mencakup, namun tidak terbatas pada, rotasi gambar, *flipping, zooming, cropping*, dan perubahan kecerahan atau kontras. Penggunaan augmentasi data, model dapat diperkenalkan dengan lebih banyak variasi gambar selama pelatihan, yang memungkinkan ViT untuk belajar secara lebih efektif tanpa perlu meningkatkan jumlah data secara signifikan. Ini membantu mengurangi *overfitting* dan mempercepat proses pelatihan dengan tetap menjaga efisiensi komputasi yang optimal.

Berbagai penelitian terdahulu telah mengkaji penggunaan model *deep learning* seperti CNN dalam klasifikasi kanker payudara, dan menunjukkan hasil yang cukup baik. Namun, sebagian besar masih terbatas pada model konvensional yang memiliki keterbatasan dalam memahami hubungan spasial global antar bagian gambar [10]. Selain itu, teknik augmentasi data memang banyak digunakan, namun penerapannya masih dominan pada model CNN dan belum secara luas dievaluasi dalam konteks ViT, khususnya untuk citra mammografi. Beberapa penelitian juga belum secara komprehensif membandingkan pengaruh berbagai jenis augmentasi terhadap performa ViT dengan menggunakan metrik evaluasi lengkap. Berdasarkan hal tersebut, terdapat celah penelitian *(research gap)* dalam integrasi ViT dengan variasi teknik augmentasi pada domain medis, khususnya mammografi.

Oleh karena itu, penelitian ini bertujuan untuk mengkaji efektivitas ViT dalam klasifikasi kanker payudara dengan dukungan augmentasi data, serta mengevaluasi kinerjanya menggunakan tujuh metrik utama. Harapannya, hasil penelitian ini dapat memberikan kontribusi terhadap pengembangan sistem deteksi dini berbasis citra yang lebih akurat, serta menjadi acuan bagi penelitian selanjutnya yang ingin menggabungkan ViT dan teknik augmentasi dalam bidang medis.

2. METODOLOGI PENELITIAN

Berikut adalah alur metode penelitian yang dilakukan untuk klasifikasi kanker payudara menggunakan *Vision Transformer*, yang ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

2.1 Tahapan Penelitian

Penelitian ini dilakukan melalui serangkaian tahapan sistematis guna membangun model klasifikasi kanker payudara berbasis *Vision Transformer*. Tahapan pertama adalah pengumpulan dataset, di mana digunakan dataset *Digital Database for Screening Mammography* (DDSM) yang berisi citra mammogram beserta label klasifikasinya. Setelah itu dilakukan *undersampling* terhadap kelas mayoritas untuk mengatasi ketidakseimbangan kelas dan memastikan distribusi data yang seimbang antara kategori jinak dan ganas. Tahap berikutnya adalah pembagian data menjadi data latih dan data uji dengan proporsi 80:20 menggunakan teknik *stratified split*, untuk menjaga distribusi kelas yang seimbang dalam masing-masing subset. Selanjutnya, dilakukan *preprocessing* data, yang mencakup beberapa proses penting: *resizing* gambar ke ukuran 224x224 piksel, normalisasi nilai piksel ke rentang [0,1], serta augmentasi data seperti *flipping, rotation, cropping*, dan *noise injection* guna meningkatkan keragaman data dan mencegah *overfitting*. Tahap inti penelitian adalah pembangunan model *Vision Transformer*. Citra hasil *preprocessing* diubah menjadi *patch*, lalu diproses melalui mekanisme *self-attention* dalam *encoder Transformer* untuk menghasilkan representasi fitur. Model dilatih dengan data hasil augmentasi, dan diuji dengan data uji untuk mengevaluasi performanya. Terakhir, dilakukan evaluasi model menggunakan tujuh metrik klasifikasi, yaitu akurasi, presisi, *recall, F1-score, Matthews Correlation Coefficient* (MCC), *Kappa Score*, dan *Area Under Curve* (AUC). Evaluasi ini bertujuan untuk mengukur efektivitas model dalam mendeteksi kanker payudara pada citra mammogram secara komprehensif.

2.2 Pengumpulan Dataset

Digital Database for Screening Mammography (DDSM) merupakan kumpulan data citra mammografi yang telah digunakan secara luas dalam penelitian dan pengembangan sistem deteksi kanker payudara berbasis komputer. Dataset ini mencakup ribuan gambar mammogram beserta informasi klinis yang terkait dengan setiap gambar. DDSM pertama kali dikembangkan melalui kerja sama berbagai institusi penelitian, termasuk *University of Florida, Massachusetts General Hospital, dan Sandia National Laboratories* pada akhir 1990-an [11].

2.3 Undersampling

Undersampling adalah metode yang digunakan untuk menangani masalah *imbalanced dataset* dengan mengurangi jumlah data pada kelas mayoritas. Salah satu teknik *undersampling* yang banyak digunakan adalah *Edited Nearest Neighbors* (ENN) yang menghapus sampel dari kelas mayoritas yang labelnya tidak serupa dengan sejumlah data yang berdekatan. Teknik lain yang sering digunakan adalah *TomekLinks*, yang mengurangi sampel dengan cara menghapus pasangan data yang berada dekat antara kelas minoritas dan mayoritas. Meskipun teknik ini efektif dalam mengurangi ketidakseimbangan, namun dapat menyebabkan hilangnya informasi penting dari kelas mayoritas yang berpotensi menurunkan kinerja model klasifikasi [12].

2.4 Pembagian Data

Pembagian dataset dalam *deep learning* merupakan langkah penting untuk memastikan model dapat belajar secara optimal dan mampu menggeneralisasi dengan baik pada data baru. Dalam beberapa penelitian, pembagian dataset bisa hanya terdiri dari *training* dan *testing* tanpa *validation set* terpisah. Hal ini dilakukan untuk memaksimalkan jumlah data yang digunakan dalam pelatihan serta mengupayakan peningkatan kapabilitas model guna memahami pola dari dataset yang terbatas [13].

2.5 *Preprocesing Data*

Preprocessing data adalah tahapan penting sebelum data mining dilakukan. Teknik-teknik *preprocessing* meliputi pembersihan data untuk menghilangkan noise dan ketidakkonsistenan, integrasi data dari berbagai sumber, reduksi data untuk menyederhanakan ukuran data, serta transformasi data seperti normalisasi untuk meningkatkan akurasi dan efisiensi algoritma data mining [14].

a. *Resizing* Gambar
   Salah satu tahap penting dalam *preprocessing* citra adalah *resizing*, yaitu proses penyesuaian ukuran gambar agar seragam untuk memudahkan tahap analisis dan klasifikasi berikutnya. *Resizing* dilakukan dengan menetapkan resolusi standar pada seluruh dataset citra sehingga seluruh gambar memiliki dimensi yang sama. Gambar yang berukuran lebih kecil dari standar akan diperbesar (*enlarge*), sedangkan gambar yang lebih besar akan diperkecil (*shrink*). Proses ini tidak hanya membantu menyederhanakan pemrosesan, tetapi juga meningkatkan konsistensi input model dalam sistem berbasis visi komputer [15].
b. Normalisasi
   Normalisasi data adalah tahap pembuatan sejumlah variabel dengan rentang nilai serupa, tidak ada yang terlalu kecil dan terlalu besar hingga mampu menyebabkan analisis statistik menjadi semakin mudah [16].
c. Augmentasi Data
   Augmentasi data yakni proses memperluas dataset pelatihan secara artifisial dengan membuat salinan data yang telah dimodifikasi. Tujuan utamanya adalah untuk meningkatkan keberagaman data agar model pembelajaran mesin tidak hanya terpaku pada pola sempit yang terdapat dalam data asli, mengurangi risiko *overfitting* yang mana model hanya bekerja optimal terhadap data pelatihan namun mengalami kegagalan pada data baru, serta menjembatani kesenjangan antara data pelatihan dan kondisi dunia nyata dengan menciptakan variasi sintetis. Teknik augmentasi dapat berupa

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

rotasi, pergeseran, penambahan noise, maupun transformasi lain yang disesuaikan dengan jenis data seperti gambar, teks, atau data tabular [17]. Teknik augmentasi yang digunakan:

1. *Flipping* (Membalik gambar horizontal/vertikal)
   Teknik augmentasi data yang umum digunakan dalam *preprocessing* citra adalah *flipping*, yaitu proses membalikkan gambar secara horizontal maupun vertikal untuk menciptakan variasi baru dari data asli tanpa mengubah makna visual utama [18].
2. *Rotation* (Memutar gambar dengan sudut acak)
   Rotasi gambar adalah salah satu teknik augmentasi data yang ditujukan guna mengupayakan peningkatan keberagaman dataset tanpa menambahkan data baru secara manual. Teknik ini dilakukan dengan memutar gambar pada sudut tertentu sehingga objek dalam citra dapat dikenali oleh model dalam berbagai orientasi [18].
3. *Cropping* (Memotong bagian gambar tertentu)
   Salah satu teknik *preprocessing* penting dalam pengolahan citra adalah *cropping*, yaitu proses memotong bagian gambar untuk memfokuskan area tertentu yang relevan terhadap objek yang dianalisis [19].
4. *Noise Injection* (Menambahkan noise acak ke gambar)
   Noise injection adalah proses menambahkan *noise* buatan ke dalam data pelatihan guna mengkaji sejauh mana model *machine learning* mampu bertahan terhadap data yang tidak sempurna atau terkontaminasi [20].

2.6 Pelatihan Model ViT

Vision Transformer (ViT) adalah model yang digunakan untuk mengklasifikasikan gambar dengan cara memecah gambar menjadi potongan kecil atau *patch*. Potongan ini diubah menjadi bentuk fitur yang kemudian diproses oleh bagian utama model, yaitu *Transformer encoder*, untuk menghasilkan hasil akhir. Keunggulan ViT terletak pada penggunaan *self-attention layer*, yang membuat model ini mampu memahami hubungan keseluruhan dalam gambar, termasuk bagian-bagian kecil yang sering tidak terdeteksi oleh metode lainnya [21].

2.7 Evaluasi Model

Evaluasi model merupakan proses fundamental pada proses pengembangan sistem klasifikasi, karena digunakan dalam menjalan pengukuran terkait seberapa jauh model bisa memperkirakan data uji secara akurat [6].

a. Akurasi
   Melaksanakan pengukuran proporsi prediksi yang tepat pada semua data. Akurasi yang semakin tinggi, artinya semakin baik model mengklasifikasikan semua kategori dengan tepat, baik positif maupun negatif [6]. Rumus akurasi diberikan pada Persamaan (1).

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

b. Presisi
   Menilai seberapa banyak prediksi positif yang benar. Semakin tinggi presisi, semakin sedikit kesalahan prediksi positif yang dilakukan model, menghindari terlalu banyak false positive [6]. Rumus presisi dapat dilihat pada Persamaan (2).

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

c. *Recall*
   Melaksanakan pengukuran kapabilitas model untuk menemukan seluruh kasus positif sesungguhnya. *Recall* tinggi artinya model efektif menangkap semua data positif, walaupun mungkin dengan mengorbankan beberapa kesalahan [6]. Rumus *recall* diberikan pada Persamaan (3).

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

d. F1-*Score*
   Gabungan harmonis dari presisi dan *recall*. F1-*score* berguna saat data tidak seimbang, untuk menilai keseimbangan antara kemampuan mendeteksi positif dan ketepatan prediksi positif [6]. Rumus *F1-Score* diberikan pada Persamaan (4).

$$F1 - Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (4)$$

e. ***Matthews Correlation Coefficient*** **(MCC)**
   Mengukur korelasi antara prediksi dan label aktual. MCC mempertimbangkan *true positive, true negative, false positive,* serta *false negative*, sangat cocok bagi data tidak seimbang [6]. Rumus MCC diberikan pada Persamaan (5).

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (5)$$

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

f. ***Kappa Score***
Mengevaluasi seberapa jauh akurasi prediksi model melebihi tebakan acak. *Kappa* memperhitungkan peluang kesepakatan acak sehingga lebih adil dalam menilai kinerja klasifikasi [6]. Rumus *Kappa* diberikan pada Persamaan (6).

$$Kappa = \frac{2 \times (TP \times TN) - (FP \times FN)}{(TP + FP) \times (FP + TN) \times (TP + FN) \times (FN + TN)} \quad (6)$$

Dimana,
True Positive (TP) : jumlah data positif yang diprediksikan benar.
True Negative (TN) : jumlah data negatif yang diprediksikan benar.
False Positive (FP) : jumlah data negatif yang salah diprediksikan sebagai positif.
False Negative (FN) : jumlah data positif yang salah diprediksikan sebagai negatif.

3. HASIL DAN PEMBAHASAN

3.1 Hasil

3.1.1 Pengumpulan Data

Data yang dipergunakan pada penelitian ini berasal dari *Digital Database for Screening Mammography* (DDSM), yang digunakan dalam penelitian [6], yaitu sebuah dataset publik yang secara luas digunakan dalam penelitian deteksi kanker payudara berbasis citra. Dataset ini menyediakan kumpulan gambar mammografi yang diperoleh melalui proses pemeriksaan medis dan telah disertai dengan label diagnosis yang memisahkan antara kondisi jinak (*benign*) dan ganas (*malignant*). Keunggulan utama dari DDSM adalah ketersediaan citra resolusi tinggi serta klasifikasi label yang valid berdasarkan penilaian ahli radiologi, sehingga menjadikannya relevan untuk penelitian klasifikasi kanker berbasis *deep learning*. Secara keseluruhan, DDSM terdiri dari lebih dari 13.000 gambar mammografi, yang dikelompokkan menjadi 2kategori besar, yakni kategori jinak sebanyak 5.970 gambar dan kategori ganas sebanyak 7.158 gambar. Setiap gambar mewakili hasil pemindaian payudara dan memiliki berbagai variasi dalam ukuran, kontras, dan tingkat kompleksitas visual, sehingga cocok untuk digunakan sebagai bahan pelatihan model berbasis *computer vision*. Pada tahap pengumpulan data, peneliti terlebih dahulu mengunduh dataset dalam format *zip* dari platform penyimpanan terbuka yang tersedia melalui Mendeley Data. Setelah proses ekstraksi, gambar-gambar diklasifikasikan kembali secara manual ke dalam dua folder utama, yaitu *Benign Masses* untuk gambar kategori jinak dan *Malignant Masses* untuk gambar kategori ganas. Proses ini dilakukan untuk menyederhanakan manajemen data serta mempermudah implementasi pemrosesan dalam kode program. Dataset ini menjadi fondasi utama dalam pelatihan model klasifikasi kanker payudara menggunakan pendekatan *Vision Transformer*. Data tersebut diharapkan dapat mendukung pengembangan sistem klasifikasi yang akurat dalam membedakan antara jinak dan ganas berdasarkan gambar mammografi. Gambar 2 menunjukkan hasil dari proses pengumpulan dataset.

```
[ ] # Hitung jumlah kasus (folder)
    benign_files = os.listdir(os.path.join(dataset_path, "Benign Masses"))
    malignant_files = os.listdir(os.path.join(dataset_path, "Malignant Masses"))

    print(f"Jumlah Benign Masses: {len(benign_files)}")
    print(f"Jumlah Malignant Masses: {len(malignant_files)}")

    Jumlah Benign Masses: 5970
    Jumlah Malignant Masses: 7158
```

Gambar 2. Hasil Pengumpulan Dataset

3.1.2 *Undersampling*

Setelah proses pengumpulan data, ditemukan bahwa dataset DDSM memiliki distribusi kelas yang tidak seimbang, di mana jumlah gambar dari kategori ganas (*malignant*) jauh lebih banyak dibandingkan dengan kategori jinak (*benign*). Ketidakseimbangan ini dapat menimbulkan masalah pada proses pelatihan model, karena model cenderung lebih "menghafal" pola dari kelas mayoritas serta mengesampingkan kelas minoritas. Hal ini dapat menimbulkan nilai akurasi yang tampak tinggi, namun performa aktual terhadap deteksi kasus minoritas menjadi rendah, terutama pada metrik seperti *recall* dan F1-*score*. Dalam menangani persoalan tersebut, penelitian ini menerapkan teknik *undersampling* pada kelas mayoritas, yaitu kategori ganas. Teknik *undersampling* dilakukan dengan cara memilih sejumlah data secara acak dari kelas mayoritas agar jumlahnya setara dengan jumlah data pada kelas minoritas. Namun, karena keterbatasan kapasitas komputasi (terutama memori/RAM), proses *undersampling* tidak hanya berfungsi untuk menyeimbangkan kelas, tetapi juga sebagai upaya untuk mengurangi beban pemrosesan selama pelatihan model. Dalam implementasinya, dipilih sebanyak 100 sampel gambar dari masing-masing kelas, sehingga total dataset yang digunakan untuk pelatihan dan pengujian model berjumlah 200 gambar mammografi. Pemilihan data dilakukan secara acak untuk kelas ganas agar tetap mewakili distribusi pola yang beragam dalam kategori tersebut. Diharapkan dengan data yang seimbang, model

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

dapat belajar secara adil terhadap kedua kelas dan menghasilkan performa yang lebih akurat serta representatif. *Undersampling* ini juga memudahkan proses evaluasi karena distribusi label yang setara memungkinkan pengukuran metrik seperti presisi dan recall dilakukan secara lebih proporsional. Langkah ini menjadi salah satu tahapan penting dalam memastikan model tidak bias terhadap kelas tertentu selama pelatihan. Hasil distribusi dataset setelah proses *undersampling* bisa disaksikan melalui Gambar 3.

```
[ ]  print(f"Setelah undersampling:")
     print(f" - Benign: {len(benign_files)} sampel")
     print(f" - Malignant: {len(malignant_files)} sampel")

     Setelah undersampling:
      - Benign: 5970 sampel
      - Malignant: 5970 sampel
```

Gambar 3. Hasil Setelah *Undersampling*

3.1.3 Pembagian Data

Setelah data dikumpulkan dan diseimbangkan melalui proses *undersampling*, langkah berikutnya adalah melaksanakan pembagian dataset ke dalam 2 bagian utama, yaitu data latih *(training set)* serta data uji *(testing set)*. Pembagian ini dilakukan dengan tujuan untuk melatih model menggunakan sebagian data, kemudian menguji performanya terhadap data yang belum pernah disaksikan sebelumnya. Strategi ini umum digunakan dalam pengembangan model pembelajaran mesin guna melaksanakan peninjauan kapabilitas generalisasi model pada data baru. Dalam penelitian ini, pembagian data dilaksanakan memakai fungsi *train_test_split* dari pustaka *scikit-learn*, melalui rasio 80% untuk data latih dan 20% untuk data uji. Jumlah total data sebanyak 200 gambar (100 jinak dan 100 ganas), maka sebanyak 160 gambar dipergunakan dalam pelatihan serta 40 gambar digunakan untuk pengujian. Pembagian ini dilakukan secara acak, namun tetap mempertahankan proporsi label antara kedua kelas menggunakan parameter *stratify*. Baik pada data latih maupun data uji, distribusi jumlah gambar dari kategori jinak dan ganas tetap seimbang. Pentingnya pembagian data secara proporsional ini adalah untuk menghindari ketimpangan distribusi kelas pada tahap pelatihan dan pengujian. Jika data dibagi secara acak tanpa mempertimbangkan distribusi kelas, ada kemungkinan model hanya belajar dari satu kelas dominan dan gagal mendeteksi kelas lainnya, terutama dalam dataset kecil. Oleh karena itu, penggunaan *stratified sampling* menjadi solusi ideal dalam pembagian data yang seimbang dan adil. Langkah pembagian data ini juga berfungsi sebagai dasar evaluasi performa model. Semua hasil evaluasi dihitung berdasarkan prediksi model terhadap data uji yang tidak pernah digunakan dalam pelatihan, sehingga mencerminkan kemampuan model dalam menghadapi data dunia nyata. Gambar 4 menunjukkan hasil pembagian dataset ke dalam data latih serta data uji.

```
[ ]  # Acak dan bagi data
     train_files, test_files = train_test_split(
         all_files,
         test_size=0.2,
         random_state=42,
         stratify=[label for _, label in all_files]
     )
     print(f"Jumlah data train: {len(train_files)}")
     print(f"Jumlah data test: {len(test_files)}")

     Jumlah data train: 160
     Jumlah data test: 40
```

Gambar 4. Hasil Pembagian Data

3.1.4 *Preprocesing* Data

Tahap *preprocessing* data ialah proses fundamental yang di dalamnya memastikan bahwasanya data gambar mammografi siap dipergunakan pada proses pelatihan model *Vision Transformer*. *Preprocessing* dilakukan untuk menyamakan format data, meningkatkan kualitas input, serta memperkaya variasi data agar model dapat belajar secara optimal. Pada penelitian ini, *preprocessing* dilakukan setelah proses pembagian data, sehingga diterapkan baik pada data latih maupun data uji, dengan beberapa teknik augmentasi hanya diterapkan pada data latih. Langkah pertama adalah *resizing* gambar, yaitu mengubah ukuran setiap gambar menjadi 224 x 224 piksel. Ukuran ini dipilih dikarenakan kompatibel dengan dimensi input standar yang digunakan oleh arsitektur *Vision Transformer*. Selain itu, gambar mammografi dalam dataset DDSM memiliki resolusi yang cukup tinggi dan bervariasi, sehingga tanpa proses *resizing*, pelatihan model akan memerlukan sumber daya komputasi yang jauh lebih besar. Penyamaan resolusi input memungkinkan model dapat belajar lebih efisien dan cepat tanpa kehilangan terlalu banyak informasi penting dalam gambar. Langkah kedua adalah normalisasi nilai piksel, di mana seluruh nilai piksel gambar yang semula berada pada rentang 0–255 diubah menjadi rentang [0, 1]. Proses ini dilakukan dengan cara membagi nilai piksel dengan angka 255 dan dikonversi ke tipe data *float*32. Normalisasi penting untuk memastikan bahwa input memiliki skala yang konsisten, sehingga proses pembelajaran model menjadi lebih stabil. Tanpa normalisasi, nilai piksel yang besar dapat menyebabkan gradien dalam proses pelatihan menjadi tidak seimbang dan memengaruhi konvergensi model. Langkah ketiga adalah augmentasi data, yang diterapkan hanya pada data latih.

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

Tujuannya adalah untuk menambah variasi citra sehingga model tidak terlalu bergantung pada pola spesifik dari gambar tertentu. Teknik augmentasi yang digunakan meliputi:

a. *Flipping* : Membalik gambar secara *horizontal* dan *vertikal*. Teknik ini membantu model mengenali objek dari berbagai orientasi.
b. *Rotation* : Memutar gambar dengan sudut acak antara 10–30 derajat, agar model tidak hanya mengenali tumor dalam posisi tetap.
c. *Cropping* : Memotong bagian tengah gambar dan mengubahnya kembali ke ukuran 224x224 piksel, agar model belajar dari bagian-bagian kecil gambar.
d. *Noise Injection* : Menambahkan noise acak untuk menyimulasikan kondisi gambar dunia nyata yang mungkin memiliki gangguan visual.

Teknik augmentasi ini diimplementasikan secara manual menggunakan pustaka *OpenCV* dan *NumPy*, serta dikombinasikan dalam sebuah fungsi *augment_image()* yang menghasilkan beberapa versi gambar dari satu input asli. Dengan pendekatan ini, jumlah data pelatihan menjadi lebih bervariasi tanpa perlu mengunduh tambahan gambar dari luar. Setelah proses augmentasi dan normalisasi selesai dilakukan, tahap selanjutnya adalah memastikan bahwa format data sesuai dengan kebutuhan arsitektur model *Vision Transformer*. Secara default, gambar mammografi dalam dataset DDSM hanya memiliki satu kanal warna *(grayscale)*, sehingga bentuk awal dari data pelatihan dan pengujian adalah (jumlah_data, 224, 224). Namun, arsitektur ViT umumnya memerlukan input dengan tiga kanal warna (RGB) atau bentuk (jumlah_data, 224, 224, 3). Untuk mengatasi hal ini tanpa mengubah informasi visual dari gambar, dilakukan replikasi kanal *grayscale* sebanyak tiga kali, sehingga setiap gambar yang awalnya memiliki satu kanal disalin menjadi tiga kanal identik. Proses ini dilakukan dengan fungsi *np.stack*, yang menumpuk data pada axis terakhir agar membentuk array berdimensi empat. Sebagaimana ditampilkan dalam kode, bentuk *X_train* dan *X_test* yang semula (960, 224, 224) dan (40, 224, 224) diubah menjadi (960, 224, 224, 3) dan (40, 224, 224, 3). Langkah ini sangat penting agar data dapat diproses oleh model ViT tanpa kesalahan dimensi input, dan sekaligus mempertahankan karakteristik asli gambar *grayscale* tanpa konversi warna buatan. Penyesuaian ini memastikan data kompatibel dan tetap representatif secara visual. Gambar 5 memperlihatkan contoh hasil dari proses augmentasi data.

```
[ ]  # Cek bentuk data
     print(f"X_train shape: {X_train.shape}")
     print(f"y_train shape: {y_train.shape}")
     print(f"X_test shape : {X_test.shape}")
     print(f"y_test shape : {y_test.shape}")

     X_train shape: (960, 224, 224)
     y_train shape: (960,)
     X_test shape : (40, 224, 224)
     y_test shape : (40,)

[ ]  # Pastikan shape jadi (N, 224, 224, 3)
     X_train = np.stack([X_train]*3, axis=-1)
     X_test = np.stack([X_test]*3, axis=-1)

     print(f"New X_train shape: {X_train.shape}")
     print(f"New X_test shape : {X_test.shape}")

     New X_train shape: (960, 224, 224, 3)
     New X_test shape : (40, 224, 224, 3)
```

Gambar 5. Hasil Augmentasi Data

3.1.5 Pelatihan Model ViT

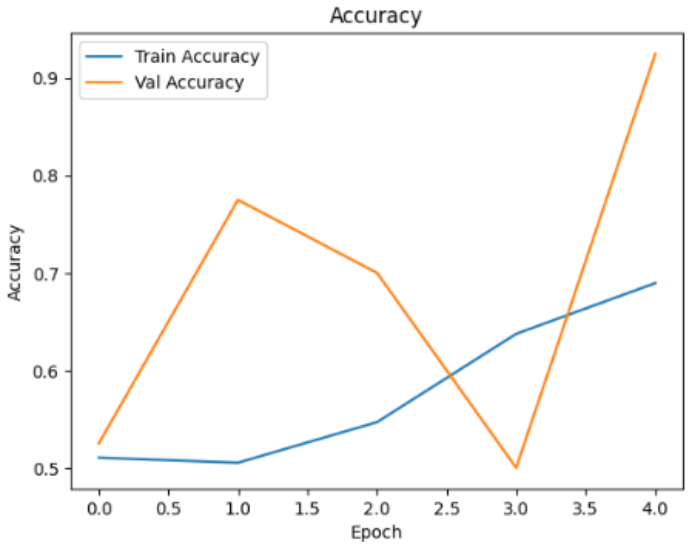
Setelah seluruh proses *preprocessing* selesai dilakukan, tahap berikutnya adalah melakukan pelatihan model menggunakan arsitektur *Vision Transformer*. ViT merupakan pendekatan berbasis transformer yang awalnya dikembangkan untuk pemrosesan bahasa alami, namun telah berhasil diadaptasi dalam bidang visi komputer. Dalam penelitian ini, ViT digunakan sebagai model utama untuk mengklasifikasikan gambar mammografi ke dalam dua kategori, yaitu jinak dan ganas. Arsitektur ViT bekerja dengan cara membagi gambar input menjadi beberapa bagian kecil *(patch)*, kemudian setiap *patch* diubah menjadi representasi *vektor* dan diproses melalui mekanisme *self-attention*. Dalam implementasinya, setiap gambar ukuran 224×224 piksel dibagi menjadi *patch* berukuran 16×16 piksel, yang kemudian dienkode menggunakan lapisan *PatchEncoder*. *Patch-patch* ini diproses oleh sejumlah blok *transformer encoder* yang terdiri dari *layer* normalisasi, *multi-head attention*, dan *multilayer perceptron* (MLP). *Output* dari proses ini diratakan dan dilewatkan ke lapisan *Dense* sebagai *classification head* untuk menghasilkan prediksi. Model ViT dilatih menggunakan *optimizer Adam*, yang dikenal efisien untuk tugas-tugas pembelajaran mendalam, serta fungsi *loss Categorical Crossentropy* yang sesuai untuk klasifikasi dua kelas. Parameter pelatihan yang digunakan disesuaikan dengan keterbatasan sumber daya komputasi. Proses pelatihan model ViT dilakukan selama 5 *epoch* dengan *batch size* sebesar 16 yang ditunjukkan pada Gambar 6.

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

```
Epoch 1/5
60/60 ━━━━━━━━━━━━━━━━━━━━ 129s 2s/step - accuracy: 0.5035 - loss: 4.1999 - val_accuracy: 0.5250 - val_loss: 0.6739
Epoch 2/5
60/60 ━━━━━━━━━━━━━━━━━━━━ 88s 1s/step - accuracy: 0.5065 - loss: 0.7629 - val_accuracy: 0.7750 - val_loss: 0.6448
Epoch 3/5
60/60 ━━━━━━━━━━━━━━━━━━━━ 142s 1s/step - accuracy: 0.5494 - loss: 0.9313 - val_accuracy: 0.7000 - val_loss: 0.6096
Epoch 4/5
60/60 ━━━━━━━━━━━━━━━━━━━━ 143s 2s/step - accuracy: 0.6209 - loss: 0.6444 - val_accuracy: 0.5000 - val_loss: 1.0486
Epoch 5/5
60/60 ━━━━━━━━━━━━━━━━━━━━ 141s 1s/step - accuracy: 0.6523 - loss: 0.6539 - val_accuracy: 0.9250 - val_loss: 0.2504
```

Gambar 6. Proses Pelatihan Model ViT dengan Lima *Epoch* dengan *batch size* sebesar 16

Gambar 7 menunjukkan grafik perbandingan nilai akurasi antara data pelatihan dan data validasi.



Gambar 7. Grafik Perbandingan Nilai Akurasi antara Data Pelatihan dan Data Validasi

Dalam Gambar 7, terlihat adanya tren peningkatan performa model dari waktu ke waktu, baik dari sisi akurasi pelatihan (*training accuracy*) maupun akurasi validasi (*validation accuracy*). Dalam *epoch* pertama, akurasi pelatihan masih rendah, yaitu sebesar 50.35%, sedangkan akurasi validasi mencapai 52.50%. Nilai ini menunjukkan bahwa pada tahap awal, model belum mampu secara optimal mengenali pola-pola pada data latih, namun memiliki potensi generalisasi awal terhadap data uji. Seiring berjalannya pelatihan, model menunjukkan perbaikan. Akurasi pelatihan mengalami peningkatan dengan cara bertahap sampai mencapai 65.23% pada *epoch* kelima, sementara akurasi validasi melonjak tajam menjadi 92.50%. Kenaikan ini menandakan bahwasanya model sukses memahammii fitur-fitur penting dari data yang diberikan, serta mampu melakukan prediksi yang cukup baik terhadap data yang tidak terlihat sebelumnya. Nilai *validation loss* juga menurun drastis dari 0.6739 menjadi 0.2504, yang mengindikasikan peningkatan stabilitas dan keandalan model dalam meminimalkan kesalahan prediksi pada data uji. Visualisasi grafik akurasi menunjukkan tren yang konsisten dengan hasil log pelatihan. Grafik memperlihatkan bahwa meskipun terjadi sedikit fluktuasi pada akurasi validasi, tren keseluruhan mengarah pada peningkatan yang signifikan. Ini menunjukkan bahwa model ViT mampu memanfaatkan proses augmentasi dan struktur *patch embedding* untuk menangkap fitur spasial dalam gambar mammografi dengan efektif.

3.1.6 Evaluasi Model

Setelah proses pelatihan selesai, tahap berikutnya ialah proses evaluasi performa model *Vision Transformer* memakai data uji yang terdiri dari gambar-gambar yang tidak diikutsertakan pada proses pelatihan. Evaluasi ini bertujuan untuk mengukur kemampuan model dalam mengklasifikasikan gambar mammografi ke dalam kategori jinak dan ganas secara objektif. Untuk memberikan gambaran yang menyeluruh, digunakan tujuh metrik evaluasi, yaitu akurasi, *Area Under Curve* (AUC), presisi, *recall*, F1-*score*, *Matthews Correlation Coefficient* (MCC), dan *Cohen's Kappa Score*. Berdasarkan hasil pengujian, model berhasil mencapai akurasi sebesar 92,50%, yang menandakan bahwasanya mayoritas prediksi yang dilakukan oleh model sesuai dengan label sebenarnya. Metrik AUC *Score* sebesar 95,75% memperlihatkan bahwasanya model mempunya kapabilitas yang sangat baik untuk membedakan antara kelas jinak dengan ganas, berdasarkan kurva *Receiver Operating Characteristic* (ROC). Nilai ini mencerminkan keseimbangan antara sensitivitas dan spesifisitas model. Presisi model tercatat sebesar 90,48%, yang menunjukkan bahwa sebagian besar prediksi positif (ganas) yang dihasilkan oleh model adalah benar. *Recall* mencapai 95,00%, yang berarti model mampu mengidentifikasi hampir seluruh gambar yang memang tergolong sebagai kasus ganas. Hasil ini sangat penting dalam konteks medis, karena *recall* yang tinggi menunjukkan sensitivitas tinggi terhadap kasus positif yang krusial dalam deteksi dini kanker. F1-*score* yang diperoleh sebesar 92,68% mencerminkan keseimbangan antara presisi dan *recall*. MCC *Score* sebesar 85,11% mengindikasikan adanya korelasi yang kuat antara prediksi dan label aktual. Terakhir, *Kappa Score* sebesar 85,00% menunjukkan taraf kesepakatan tinggi antara model dengan label aktual, setelah dikoreksi terhadap peluang

JURIKOM (Jurnal Riset Komputer), Vol. 12 No. 3, Juni 2025
e-ISSN 2715-7393 (Media Online), p-ISSN 2407-389X (Media Cetak)
DOI 10.30865/jurikom.v12i3.8619
Hal 241-250
https://ejurnal.stmik-budidarma.ac.id/index.php/jurikom

kesepakatan acak. Metrik evaluasi ini secara keseluruhan memperlihatkan bahwasanya model ViT yang dibangun pada penelitian ini mampu melakukan klasifikasi gambar mammografi secara efektif dan dapat dijadikan pendekatan yang potensial dalam mendukung sistem diagnosis kanker payudara. Gambar 8 memperlihatkan proses evaluasi model yang sudah dilatih.

```
[ ]  acc = accuracy_score(y_test, y_pred)
     auc = roc_auc_score(y_test, y_pred_prob[:, 1])
     prec = precision_score(y_test, y_pred)
     rec = recall_score(y_test, y_pred)
     f1 = f1_score(y_test, y_pred)
     mcc = matthews_corrcoef(y_test, y_pred)
     kappa = cohen_kappa_score(y_test, y_pred)

     print(f"Akurasi        : {acc * 100:.2f}%")
     print(f"AUC Score      : {auc * 100:.2f}%")
     print(f"Presisi        : {prec * 100:.2f}%")
     print(f"Recall         : {rec * 100:.2f}%")
     print(f"F1-Score       : {f1 * 100:.2f}%")
     print(f"MCC Score      : {mcc * 100:.2f}%")
     print(f"Kappa Score    : {kappa * 100:.2f}%")

     Akurasi        : 92.50%
     AUC Score      : 95.75%
     Presisi        : 90.48%
     Recall         : 95.00%
     F1-Score       : 92.68%
     MCC Score      : 85.11%
     Kappa Score    : 85.00%
```

Gambar 8. Evaluasi Model

3.2 Pembahasan

Hasil penelitian ini dibandingkan dengan studi yang dilaksanakan dalam penelitian [6], Kedua penelitian menggunakan pendekatan ViT dalam klasifikasi gambar mammografi dari dataset *Digital Database for Screening Mammography* (DDSM). Namun, terdapat perbedaan signifikan dalam metode, konfigurasi pelatihan, serta hasil evaluasi model. Pada penelitian ini, ViT dikembangkan dari awal tanpa menggunakan model pralatih (*pre-trained model*), dan dilatih pada dataset yang telah melalui proses *undersampling* dan augmentasi sederhana (*flipping, rotation, cropping*, dan *noise injection*). Model dilatih selama lima *epoch* menggunakan ukuran *batch* sebesar 16, mengingat keterbatasan kapasitas komputasi. Hasil evaluasi memperlihatkan bahwasanya model mendapatkan akurasi senilai 92,50%, AUC *Score* senilai 95,75%, presisi senilai 90,48%, *recall* senilai 95.00%, *F1-score* senilai 92,68%, serta nilai MCC dan *Kappa* masing-masing sebesar 85,11% dan 85,00%. Sebaliknya, penelitian [6] menggunakan pendekatan *transfer learning* dengan memanfaatkan ViT pralatih dari *ImageNet-21k*. Model dilatih ulang dengan konfigurasi lebih kompleks, termasuk augmentasi lanjutan seperti *sharpening* dan *gamma correction*, serta penggunaan GPU kelas tinggi (NVIDIA RTX 3090). Dengan 50 *epoch* dan *batch size* sebesar 64, model yang dihasilkan menunjukkan performa sempurna pada seluruh metrik evaluasi (akurasi, AUC, presisi, *recall*, *F1-score*, MCC, dan *kappa* sebesar 1,00). Perbedaan hasil yang signifikan antara kedua penelitian dapat dijelaskan oleh perbedaan strategi pelatihan, kapasitas komputasi, dan teknik augmentasi yang digunakan. Meskipun demikian, hasil yang dicapai dalam penelitian ini tetap menunjukkan bahwa ViT dapat memberikan performa klasifikasi yang sangat baik meskipun dilatih dari awal dan dijalankan dalam lingkungan dengan keterbatasan sumber daya. Hal ini mengindikasikan bahwa pendekatan ViT tetap efektif bahkan tanpa *transfer learning*, asalkan didukung oleh strategi *preprocessing* dan augmentasi yang tepat.

4. KESIMPULAN

Penelitian ini berhasil membangun model klasifikasi kanker payudara berbasis citra mammografi menggunakan arsitektur ViT. Berdasarkan eksperimen yang dilakukan, model ViT mampu mengklasifikasikan gambar ke dalam dua kategori, yaitu jinak dan ganas, dengan performa yang sangat baik meskipun dijalankan dalam lingkungan dengan keterbatasan komputasi. Model mencapai akurasi sebesar 92,50%, AUC sebesar 95,75%, presisi 90,48%, *recall* 95,00%, F1-*score* 92,68%, serta nilai MCC dan *Kappa Score* masing-masing sebesar 85,11% dan 85,00%. Berdasarkan rangkaian eksperimen yang dilakukan, model ViT mampu mengklasifikasikan gambar ke dalam dua kategori, yaitu jinak dan ganas, dengan performa yang sangat baik meskipun dijalankan dalam lingkungan dengan keterbatasan komputasi. Model mencapai akurasi sebesar 92,50%, AUC *Score* sebesar 95,75%, presisi sebesar 90,48%, *recall* sebesar 95.00%, *F1-score* sebesar 92,68%, serta nilai MCC dan *Kappa Score* masing-masing sebesar 85,11% dan 85,00%. Hasil ini menunjukkan bahwa ViT dapat mengenali fitur-fitur penting dari gambar mammografi meskipun dilatih dari awal (tanpa *transfer learning*), dan tetap efektif dengan jumlah data yang terbatas serta konfigurasi pelatihan yang sederhana. Seluruh tahapan penelitian, mulai dari pengumpulan data, *undersampling*, pembagian data, *preprocessing* (*resizing*, normalisasi, augmentasi), hingga pelatihan dan evaluasi model telah dilakukan secara sistematis. Pendekatan augmentasi yang digunakan terbukti mampu meningkatkan variasi data tanpa menambah beban dataset secara signifikan. Namun demikian,

penelitian ini memiliki beberapa keterbatasan. Jumlah data yang digunakan relatif kecil karena keterbatasan kapasitas memori, dan proses pelatihan dibatasi pada 5 *epoch*. Selain itu, tidak diterapkannya *transfer learning* dari model pralatih menyebabkan model membutuhkan waktu lebih lama untuk mencapai konvergensi optimal. Untuk penelitian selanjutnya, pengembangan dapat dilakukan dengan menggunakan dataset yang lebih besar, pelatihan dengan *epoch* lebih banyak, agar performa model dapat ditingkatkan secara signifikan dan mendukung pengembangan sistem deteksi dini kanker payudara yang lebih andal.

REFERENCES

[1] Della Zulfa Rifda, Zahroh Shaluhiyah, and Antono Surjoputro, "Studi Fenomenologi Pasien Kanker Payudara dalam Upaya Meningkatkan Kualitas Hidup : Literature Review," *Media Publ. Promosi Kesehat. Indones.*, vol. 6, no. 8, pp. 1495–1500, 2023, doi: 10.56338/mppki.v6i8.3513.

[2] K. Masyarakat, "Literatur review : pengetahuan remaja terhadap deteksi kanker payudara dengan cara pemeriksaan kanker payudara sendiri (sadari)," 2024.

[3] Ratna Septia Devi, Triando Hamonangan Saragih, and Mohammad Reza Faisal, "Seleksi Fitur Hybrid Grey Wolf Optimization dan Particle Swarm Optimization pada Distance Biased Naive Bayes untuk Klasifikasi Kanker Payudara," *J. Inform. Polinema*, vol. 10, no. 2, pp. 307–314, 2024, doi: 10.33795/jip.v10i2.4737.

[4] N. Afiatuddin, M. T. Wicaksono, V. R. Akbar, R. Rahmaddeni, and D. Wulandari, "Komparasi Algoritma Machine Learning dalam Klasifikasi Kanker Payudara," *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 889, 2024, doi: 10.30865/mib.v8i2.7457.

[5] B. Gheflati and H. Rivaz, "Vision Transformers for Classification of Breast Ultrasound Images," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2022-July, pp. 480–483, 2022, doi: 10.1109/EMBC48229.2022.9871809.

[6] G. Ayana *et al.*, "Vision-Transformer-Based Transfer Learning for Mammogram Classification," *Diagnostics*, vol. 13, no. 2, 2023, doi: 10.3390/diagnostics13020178.

[7] D. Tsalsabila Rhamadiyanti and Kusrini, "Analisa Performa Convolutional Neural Network dalam Klasifikasi Citra Apel dengan Data Augmentasi," *J. Kaji. Ilm. Inform. dan Komput.*, vol. 5, no. 1, pp. 154–162, 2024, doi: 10.30865/klik.v5i1.2023.

[8] X. T. Vo, D. L. Nguyen, A. Priadana, and K. H. Jo, "Efficient Vision Transformers with Partial Attention," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* , vol. 15141 LNCS, no. 4460, pp. 298–317, 2025, doi: 10.1007/978-3-031-73010-8_18.

[9] R. Z. Fadillah, A. Irawan, M. Susanty, and I. Artikel, "Data Augmentasi Untuk Mengatasi Keterbatasan Data Pada Model Penerjemah Bahasa Isyarat Indonesia (BISINDO)," *J. Inform.*, vol. 8, no. 2, pp. 208–214, 2021, [Online]. Available: https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/10768

[10] A. R. N. A, R. Setiawan, M. A. Hendrawan, D. B. Nugroho, and N. Rahman, "Deteksi Kanker Payudara Hasil Citra Mammografi menggunakan Metode Convolutional Neural Network (CNN) Arsitektur ResNet-50," no. 01, pp. 25–32, 2024, doi: 10.25047/jiitu.v1i01.5498.

[11] J. Jaradat, R. Amro, R. Hamamreh, A. Musleh, and M. Abdelgalil, "From Data to Diagnosis: Narrative Review of Open-Access Mammography Databases for Breast Cancer Detection," *High Yield Med. Rev.*, vol. 2, no. 1, pp. 1–10, 2024, doi: 10.59707/hymrpfnz8344.

[12] A. Indrawati, "Penerapan Teknik Kombinasi Oversampling Dan Undersampling Untuk Mengatasi Permasalahan Imbalanced Dataset," *JIKO (Jurnal Inform. dan Komputer)*, vol. 4, no. 1, pp. 38–43, 2021, doi: 10.33387/jiko.v4i1.2561.

[13] S. Yuliany, Aradea, and Andi Nur Rachman, "Implementasi Deep Learning pada Sistem Klasifikasi Hama Tanaman Padi Menggunakan Metode Convolutional Neural Network (CNN)," *J. Buana Inform.*, vol. 13, no. 1, pp. 54–65, 2022, doi: 10.24002/jbi.v13i1.5022.

[14] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*. Cham, Switzerland: Springer, 2015.

[15] Syefrida Yulina and Hoky Nawa, "Dataset Gambar Wajah untuk Analisis Personal Identification," *J. Appl. Comput. Sci. Technol.*, vol. 3, no. 2, pp. 193–198, 2022, doi: 10.52158/jacost.v3i2.427.

[16] M. R. Kusnaidi, T. Gulo, and S. Aripin, "Penerapan Normalisasi Data Dalam Mengelompokkan Data Mahasiswa Dengan Menggunakan Metode K-Means Untuk Menentukan Prioritas Bantuan Uang Kuliah Tunggal," *J. Comput. Syst. Informatics*, vol. 3, no. 4, pp. 330–338, 2022, doi: 10.47065/josyc.v3i4.2112.

[17] R. S. Wahono, *Data Mining Data mining*, vol. 2, no. January 2013. 2023. [Online]. Available: https://www.cambridge.org/core/product/identifier/CBO9781139058452A007/type/book_part

[18] J. Homepage, D. Saputra, and T. Y. Hadiwandra, "Indonesian Journal of Informatic Research and Software Engineering Classification Of Letters And Numbers In Bisindo Using The Convolutional Neural Network Method Klasifikasi Huruf Dan Angka Dalam Bisindo Menggunakan Metode Convolutional Neural Network," *Ijirse*, vol. 4, no. 2, pp. 88–95, 2024.

[19] E. Astiadewi, M. Martanto, A. R. Dikananda, dan D. Rohman, "Algoritma YOLOv8 untuk Meningkatkan Analisa Gambar dalam Mendeteksi Jerawat," *Jurnal Informatika Teknologi dan Sains (JINTEKS)*, vol. 7, no. 1, pp. 346–353, Feb. 2025.

[20] K. M. Al-Gethami, M. T. Al-Akhras, and M. Alawairdhi, "Empirical Evaluation of Noise Influence on Supervised Machine Learning Algorithms Using Intrusion Detection Datasets," *Secur. Commun. Networks*, vol. 2021, 2021, doi: 10.1155/2021/8836057.

[21] J. D. Halim, "Klasifikasi Sel Darah Putih Menggunakan Vision Transformer (ViT)," vol. 6, no. November, pp. 291–307, 2024.