

DETEKSI MALFORMASI UTERUS MELALUI CITRA HISTEROSALPINGOGRAFI MENGUNAKAN *DEEP LEARNING*

Muhammad Syarif Baital¹, Andani Achmad², Hazriani³

¹STMIK Catur Sakti Kendari, ²Universitas Hasanuddin Makassar, ³Universitas Handayani

Makassar

e-mail: ¹syarif.baital@gmail.com, ²andani@unhas.ac.id, ³hazriani@handayani.ac.id

Penelitian ini menyoroti tingginya angka kejadian malformasi uterus yang berkontribusi terhadap infertilitas, keguguran, serta komplikasi perinatal, sehingga diperlukan metode diagnostik yang lebih presisi. Tujuan utama dari penelitian ini adalah merancang model deep learning berbasis Convolutional Neural Network (CNN) dengan arsitektur ResNet untuk mendeteksi berbagai jenis malformasi uterus melalui citra Histerosalpingografi (HSG) serta menilai tingkat akurasi dalam mengklasifikasikan enam jenis malformasi, yaitu unicornuate, bicornuate, didelphys, septate, arcuate, dan uterus normal. Dataset yang digunakan mencakup 1.800 citra yang terbagi secara merata ke dalam enam kategori. Model ResNet Baseline menunjukkan performa terbaik dengan tingkat akurasi, presisi, recall, dan F1-Score sebesar 100% pada data latih sebesar 90%.

Kata Kunci : CNN, Deep Learning, Malformasi Uterus, HSG, ResNet.

I. PENDAHULUAN

Malformasi uterus adalah kelainan bentuk rahim yang dapat memengaruhi uterus, serviks, atau vagina. *American Fertility Society* (AFS) mengklasifikasikannya ke dalam beberapa jenis, seperti agenesis mullerian, uterus unicornuate, dan lainnya. Insiden malformasi sulit ditentukan karena sering teridentifikasi saat pemeriksaan medis. Sekitar 57-63% penderita masih dapat mengandung dan melahirkan. Gejala umum meliputi gangguan menstruasi, nyeri panggul, masalah kesuburan, dan keguguran berulang [1], [2].

Malformasi uterus dapat menyebabkan infertilitas, keguguran, dan komplikasi perinatal [1], [3]. Angka kejadiannya sulit dipastikan karena sering tidak menimbulkan gejala sebelum kehamilan. Diperkirakan terjadi pada 1-2 per 1.000 perempuan, dengan sekitar 60% masih dapat melahirkan bayi normal. Risiko persalinan prematur dan keguguran lebih tinggi, dengan insiden aborsi 28% dan kelahiran prematur 20% pada kasus parsial. Secara global, prevalensinya sekitar 4-7%, tetapi

bisa lebih tinggi pada individu dengan infertilitas atau riwayat keguguran berulang. Data spesifik di Indonesia masih terbatas [2].

Penerapan *deep learning* dalam kedokteran, terutama radiologi, semakin berkembang untuk diagnosis berbasis citra, peningkatan kualitas gambar, dan pembuatan laporan. Namun, penggunaannya dalam praktik klinis masih terbatas. Oleh karena itu, ahli radiologi perlu beradaptasi dengan perkembangan teknologi ini, termasuk meningkatkan pemahaman tentang kecerdasan buatan (AI) dan pembelajaran mesin [4], [5], [6].

Deep learning dengan *Convolutional Neural Network* (CNN) banyak digunakan dalam berbagai bidang, terutama pengenalan gambar [4], [5]. CNN dirancang untuk menganalisis citra secara efisien dengan mengekstrak fitur secara otomatis [3], [6], berbeda dari metode pembelajaran mesin konvensional yang memerlukan ekstraksi fitur manual [7].

Beberapa studi *deep learning* telah diterapkan dalam analisis citra medis, seperti deteksi lesi, klasifikasi, dan segmentasi. Segmentasi objek sangat penting dalam pendekatan radiomik dan radiogenomik. Meskipun segmentasi manual oleh dokter berpengalaman memiliki nilai tinggi, proses ini bisa melelahkan, memakan waktu, dan subjektif [5], [8].

II. METODE PENELITIAN

A. Rancangan Penelitian

Secara garis besar penelitian ini dibagi menjadi 5 tahap, yaitu pengumpulan/akuisisi citra, perbaikan kualitas citra, segmentasi citra, proses ekstraksi ciri, serta identifikasi dan klasifikasi

Data yang dipakai merupakan citra histerosalpingograf (HSG) uterus yang diambil dari sebuah Laboratorium Klinik. Data yang diperoleh ada 1.800 citra HSG uterus dan dibagi menjadi 300 citra di setiap kelas. Dalam pengujian performa model, pembagian data merupakan langkah penting untuk mengevaluasi kemampuan generalisasi model. Konfigurasi yang digunakan adalah 90:10, 80:20, dan 70:30, di mana angka pertama

menunjukkan persentase data yang digunakan untuk melatih model, sementara angka kedua menunjukkan persentase data untuk validasi dan pengujian.

Pre-processing dilakukan untuk mengoptimalkan kualitas gambar. Tahapan *pre-processing* pada penelitian ini terbagi kedalam 2 tahap yaitu data asli dan data augmentasi.

Tahapan *training* melakukan proses pembelajaran pada citra yang kemudian hasil keluarannya berupa model yang akan disimpan untuk digunakan dalam proses pengujian. Pada penelitian ini metode yang digunakan merupakan salah satu cabang algoritma *deep learning* yaitu CNN dengan arsitektur ResNet.

Input citra pada model CNN mempergunakan citra dengan ukuran 224x224. Kemudian citra masukan akan diproses melalui proses *convolution* dan proses *pooling*. Lapisan konvolusi berjumlah sebanyak 16 lapisan dan memiliki jumlah filter yang berbeda. Selanjutnya dilakukan proses *flatten*. Proses ini disebut tahapan *fully connected layer*.

Fungsi aktivasi yang digunakan pada *fully connected layer* yaitu *flatten*, *dropout* untuk membatasi jumlah neuron yang masuk agar mencegah *overfitting*, dan ditambah parameter *batch normalization* untuk menormalkan sumbu aliran. Selepas citra melewati *fully connected layer*, maka hasil keluarannya digunakan sebagai *input* fungsi aktivasi *softmax* untuk memprediksi kemungkinan dari data hasil training atas objek citra.

Tahap pengujian yaitu berupa proses klasifikasi malformasi uterus dengan menguji data citra uji dan membandingkannya dengan model hasil *training* data citra latih yang disimpan pada *database*. Citra yang telah diambil akan diproses oleh algoritma CNN hingga kemudian menghasilkan keluaran sistem berupa informasi jenis malformasi uterus.

Penelitian ini menggunakan klasifikasi 6 kelas; unicornuate, didelphys, bicornuate, septate, arcuate, dan normal. Kerangka kerja klasifikasi malformasi uterus menggunakan *transfer learning* CNN. Sistem dimulai dengan arsitektur CNN yang sudah dilatih menggunakan *dataset* Imagenet. Kemudian *pre-trained* CNN tersebut dipindahkan ke arsitektur yang dikembangkan untuk dilatih menggunakan *dataset* malformasi uterus.

Eksperimen dilakukan di lingkungan Google Colab dan model CNN (ResNet) dikembangkan dengan pustaka Torch.

B. Instrumen Penelitian

1. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) digunakan dalam visi komputer untuk mengekstrak fitur dari gambar. CNN terdiri dari lapisan konvolusi yang menggunakan kernel dan filter untuk menangkap pola dalam gambar, serta lapisan pooling yang mengurangi dimensi data. Terdapat dua jenis pooling: *max-pooling* (mengambil nilai maksimum) dan *average pooling* (mengambil rata-rata). Setelah melalui proses ini, fitur dari gambar diubah menjadi vektor menggunakan *flatten layer*, kemudian

diteruskan ke *fully connected layer* atau *dense layer* untuk menghasilkan *output* akhir [7], [9], [10].

2. Depth Residual Learning

ResNet (*Residual Network*) adalah arsitektur deep learning yang dikembangkan oleh Microsoft Research untuk mengatasi degradasi kinerja pada jaringan yang sangat dalam. Berbeda dengan VGG yang menggunakan deep convolutional network sederhana, ResNet memperkenalkan *residual connections* yang memungkinkan aliran gradien lebih lancar. Pendekatan ini membantu mengatasi masalah seperti vanishing gradients dan degradasi performa pada pelatihan jaringan yang dalam [4], [10].

3. Image Augmentation Optimization

Augmentasi gambar digunakan untuk meningkatkan pelatihan model dengan menciptakan variasi baru dari gambar dalam dataset tanpa mengubah kelas aslinya. Teknik ini mencakup manipulasi seperti pergeseran, pembalikan, rotasi, perubahan kecerahan, dan zoom. Augmentasi membantu algoritma *deep learning* mempelajari fitur gambar lebih efektif dan biasanya diterapkan pada *dataset* pelatihan, bukan pada *dataset* pengujian atau validasi [10], [11]. Dalam penelitian ini, *Torch deep learning library* digunakan, dengan fungsi augmentasi seperti *ImageDataGenerator* untuk melakukan transformasi gambar secara acak.

4. Dropout Regularization

Dropout adalah teknik untuk mencegah *overfitting* dalam *neural network* dengan menonaktifkan neuron secara acak selama pelatihan dengan probabilitas tertentu. Metode ini membantu jaringan menghindari pembelajaran berlebihan dan meningkatkan kinerjanya dalam *supervised learning* [10], [12]. Dalam penelitian ini, *Torch deep learning library* digunakan untuk mengimplementasikan *dropout* dalam model ResNet dengan nilai 0.5.

5. Transfer Learning

Transfer learning adalah metode dalam *machine learning* di mana model yang telah dilatih pada satu masalah digunakan kembali untuk masalah terkait. Teknik ini mengurangi waktu pelatihan dan meningkatkan generalisasi model. Bobot pada lapisan yang digunakan kembali dapat menjadi titik awal untuk pelatihan pada masalah baru, terutama jika *dataset* awal memiliki banyak data berlabel [7], [10]. Dalam penelitian ini, model ResNet dikembangkan dengan *transfer learning* dan diunduh menggunakan *Torch library* di Google Colab.

6. Konfigurasi Model ResNet

Arsitektur ResNet untuk gambar *grayscale* memodifikasi lapisan *convolutional* pertama agar menerima 1 kanal warna. Setelah itu, *output* diproses melalui *max-pooling* dan serangkaian blok residual yang terdiri dari *convolutional layers*, *batch normalization*, dan *residual connections*. Jumlah saluran meningkat bertahap

(64 → 128 → 256 → 512) dengan setiap *max-pooling layer*. Aktivasi *SoftMax* digunakan pada *dense layer*, sementara optimasi dilakukan dengan Adam. Untuk mencegah *overfitting*, *early stopping* diterapkan dengan *patience* 5 dan *verbose* 1. Model dilatih selama 50 *epoch* dengan *input* gambar berukuran 224 × 224 piksel. Konfigurasi rinci dari semua model ResNet disajikan dalam Tabel 1.

Tabel 1.
Konfigurasi model arsitektur ResNet

Hyper parameter	Baseline	Aug	Aug + Regu
Activation Function	Softmax	Softmax	Softmax
Epoch	50	50	50
Weight	ImageNet	ImageNet	ImageNet
BatchSize	8, 16, 32, 64, 128	8, 16, 32, 64, 128	8, 16, 32, 64, 128
Loss Function	Categorical Cross entropy	Categorical Cross entropy	Categorical Cross entropy
Optimizer	Adam	Adam	Adam
Augmentation	Yes	Yes	Yes
Regularization	Yes	No	No
Drop Out	Yes (0.5)	Yes	No
Early Stopping	Yes (Patience 5, verbose 1)	Yes (Patience 5, verbose 1)	Yes (Patience 5, verbose 1)
Data Split	90:10, 80:20, 70:30	90:10, 80:20, 70:30	90:10, 80:20, 70:30
Scheduller			
Leaming Rate	Yes	Yes	Yes
Transfer Learning	Yes	Yes	Yes

7. Model Evaluation

Setelah klasifikasi, model dievaluasi menggunakan metrik akurasi, presisi, *recall*, dan F1-Score. Akurasi mengukur persentase klasifikasi yang benar. Presisi menunjukkan proporsi prediksi yang benar dalam suatu kelas, sementara *recall* mengukur seberapa baik model mendeteksi semua contoh positif. F1-Score adalah rata-rata harmonis dari presisi dan *recall*, dengan nilai mendekati 1 menunjukkan akurasi yang lebih baik. Perhitungan metrik ini mengikuti rumus yang ditentukan dalam persamaan (1), (2), (3), dan (4).

$$accuracy = \frac{\text{data diklasifikasikan benar}}{\text{total data}} \quad (1)$$

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

$$precision = \frac{TP}{TP+FP} \quad (3)$$

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (4)$$

Di mana TP adalah *true positive*, FN adalah *false negative*, dan FP adalah *false positive*.

Nilai ini harus mendekati 1, menunjukkan prediksi yang sempurna, nilai di bawah 0.5 menyiratkan tebakan acak [10], [11].

III. HASIL DAN PEMBAHASAN

A. Hasil Penelitian.

1. ResNet Baseline

1) ResNet Baseline BatchSize 8

Training model yang dilakukan untuk menguji model ResNet Baseline BatchSize 8.

Tabel 2.
Classification Report training ResNet Baseline BatchSize 8.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	9	1h 10m	Arcuate
	1	1	1	9		Bicomuate
	1	1	1	9		Didelphys
	1	1	1	9		Normal
	1	1	1	9		Septate
	1	1	1	9		Unicomuate
				54		Accuracy
	1	1	1	54		Macro Avg
	1	1	1	54		Weight Avg
80%	1	1	1	24	3h 26m	Arcuate
	0.92	0.96	0.94	24		Bicomuate
	1	1	1	24		Didelphys
	1	1	1	24		Normal
	0.96	0.92	0.94	24		Septate
	1	1	1	24		Unicomuate
			0.98	144		Accuracy
	0.98	0.98	0.98	144		Macro Avg
	0.98	0.98	0.98	144		Weight Avg
70%	1	1	1	39	2h 45m	Arcuate
	0.83	0.97	0.89	39		Bicomuate
	1	1	1	39		Didelphys
	1	1	1	39		Normal
	0.97	0.79	0.87	39		Septate
	1	1	1	39		Unicomuate
			0.96	234		Accuracy
	0.97	0.96	0.96	234		Macro Avg
	0.97	0.96	0.96	234		Weight Avg

Model menunjukkan akurasi tinggi di berbagai *TrainSize*, dengan akurasi sempurna (1.00) pada *TrainSize* 90%, namun sedikit menurun pada *TrainSize* 80% (0.98) dan 70% (0.96). *TrainSize* yang lebih besar meningkatkan stabilitas klasifikasi, tetapi juga memperpanjang waktu pelatihan. *TrainSize* 90% membutuhkan 1 jam 10 menit, sementara *TrainSize* 80% lebih dari 3 jam, menunjukkan trade-off antara efisiensi waktu dan kinerja. *TrainSize* 90% memiliki F1-Score sempurna (1.00) di semua kelas, tetapi kurang optimal untuk evaluasi generalisasi, sedangkan *TrainSize* 80% dan 70% lebih efisien tetapi memiliki F1-Score terendah 0.87.

Jika prioritas utama adalah akurasi dan stabilitas, maka model dengan *TrainSize* 90% merupakan pilihan terbaik.

2) ResNet Baseline BatchSize 16

Training model yang dilakukan untuk menguji model ResNet Baseline BatchSize 16.

Tabel 3.

Classification Report training ResNet Baseline BatchSize 16.

Train Size	Precision	Recall	F1	Support	Duration	
90%	1	1	1	11	1h 1m	Arcuate
	1	1	1	11		Bicornuate
	1	1	1	11		Didelphys
	1	1	1	11		Normal
	1	1	1	11		Septate
	1	1	1	11		Unicornuate
				66		Accuracy
	1	1	1	66		Macro Avg
	1	1	1	66		Weight Avg
80%	1	1	1	23	1h 5m	Arcuate
	0.85	1	0.92	23		Bicornuate
	1	1	1	23		Didelphys
	1	1	1	23		Normal
	1	0.83	0.9	23		Septate
	1	1	1	23		Unicornuate
				138		Accuracy
	0.98	0.97	0.97	138		Macro Avg
	0.98	0.97	0.97	138		Weight Avg
70%	1	1	1	40	1h 15m	Arcuate
	1	0.93	0.96	40		Bicornuate
	1	1	1	40		Didelphys
	1	1	1	40		Normal
	0.93	1	0.96	40		Septate
	1	1	1	40		Unicornuate
				240		Accuracy
	0.99	0.99	0.99	240		Macro Avg
	0.99	0.99	0.99	240		Weight Avg

Model dengan TrainSize 90% mencapai akurasi dan F1-Score sempurna (1.00), menunjukkan performa terbaik dalam identifikasi kategori. Namun, TrainSize 80% dan 70% mengalami sedikit penurunan akurasi (0.97 dan 0.99) serta F1-Score lebih rendah pada kategori tertentu, seperti Septate (0.90) dan Bicornuate serta Septate (0.96). Waktu pelatihan meningkat dengan penurunan TrainSize, dengan TrainSize 90% membutuhkan 1 jam 1 menit, sementara TrainSize 70% memakan waktu 1 jam 15 menit. Meskipun TrainSize 90% unggul dalam akurasi dan kestabilan, kurangnya variasi hasil menyulitkan evaluasi generalisasi, sedangkan TrainSize lebih kecil tetap akurat tetapi mengalami kelemahan dalam klasifikasi kategori spesifik akibat keterbatasan data pelatihan.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 90% merupakan pilihan terbaik.

3) ResNet Baseline BatchSize 32

Training model yang dilakukan untuk menguji model ResNet Baseline BatchSize 32.

Tabel 4.

Classification Report training ResNet Baseline BatchSize 32.

Train Size	Precision	Recall	F1	Support	Duration	
90%	1	1	1	11	1h 25m	Arcuate
	0.85	1	0.92	11		Bicornuate
	1	1	1	11		Didelphys
	1	1	1	11		Normal
	1	0.82	0.9	11		Septate
	1	1	1	11		Unicornuate
				66		Accuracy

80%	0.97	0.97	0.97	66	1h 25m	Macro Avg
	0.97	0.97	0.97	66		Weight Avg
	1	1	1	25		Arcuate
	0.83	1	0.91	25		Bicornuate
	1	1	1	25		Didelphys
	1	1	1	25		Normal
	1	0.8	0.89	25		Septate
	1	1	1	25		Unicornuate
				150		Accuracy
70%	0.98	0.97	0.97	150	1h 25m	Macro Avg
	0.98	0.97	0.97	150		Weight Avg
	1	0.97	0.99	39		Arcuate
	0.91	1	0.95	39		Bicornuate
	1	1	1	39		Didelphys
	1	1	1	39		Normal
	1	0.92	0.96	39		Septate
	1	1	1	39		Unicornuate
				234		Accuracy
	0.98	0.98	0.98	234		Macro Avg
	0.98	0.98	0.98	234		Weight Avg

Model ResNet Baseline menunjukkan akurasi tinggi (97–98%) di semua ukuran TrainSize dengan waktu pelatihan sekitar 1 jam 25 menit. Model ini mengklasifikasikan Didelphys, Normal, dan Unicornuate dengan sempurna (F1-Score = 1.0), tetapi memiliki kesulitan pada kelas Septate dan Bicornuate (F1-Score 0.89–0.96). Keunggulannya adalah akurasi stabil dan efisiensi waktu pelatihan, meskipun perlu perbaikan dalam membedakan fitur kelas tertentu.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 90% merupakan pilihan terbaik.

4) ResNet Baseline BatchSize 64

Training model yang dilakukan untuk menguji model ResNet Baseline BatchSize 64.

Tabel 5.

Classification Report training ResNet Baseline BatchSize 64.

Train Size	Precision	Recall	F1	Support	Duration	
90%	1	1	1	13	1h 25m	Arcuate
	0.87	1	0.93	13		Bicornuate
	1	1	1	13		Didelphys
	1	1	1	13		Normal
	1	0.85	0.92	13		Septate
	1	1	1	13		Unicornuate
				78		Accuracy
	0.98	0.97	0.97	78		Macro Avg
	0.98	0.97	0.97	78		Weight Avg
80%	1	1	1	20	1h 25m	Arcuate
	0.9	0.95	0.93	20		Bicornuate
	1	1	1	20		Didelphys
	0.95	1	0.98	20		Normal
	0.95	0.9	0.92	20		Septate
	1	0.95	0.97	20		Unicornuate
				120		Accuracy
	0.97	0.97	0.97	120		Macro Avg
	0.97	0.97	0.97	120		Weight Avg
70%	1	1	1	38	1h 25m	Arcuate
	1	0.87	0.93	38		Bicornuate
	1	1	1	38		Didelphys
	1	1	1	38		Normal
				38		Septate
	0.88	1	0.94	38		

1	1	1	38	Unicornuate
		0.98	228	Accuracy
0.98	0.98	0.98	228	Macro Avg
0.98	0.98	0.98	228	Weight Avg

Dari hasil pelatihan menggunakan model ResNet dengan ukuran BatchSize 64 pada berbagai ukuran *TrainSize*, terlihat bahwa ukuran *TrainSize* memengaruhi akurasi, serta performa model (F1-Score). Semua model menunjukkan akurasi mendekati 100%. F1-Score terbaik konsisten pada kelas arcuate dan didelphys, sementara F1-Score terendah cenderung muncul pada kelas dengan data yang sulit diklasifikasikan (Septate dan Bicornuate). Model dengan *TrainSize* 90% menghasilkan performa lebih stabil. Sebaliknya, model dengan *TrainSize* 70% lebih efisien, namun lebih rentan terhadap ketidakseimbangan data.

Jika prioritas utama adalah akurasi dan stabilitas, maka *TrainSize* 90% merupakan pilihan terbaik.

5) ResNet Baseline BatchSize 128

Training model yang dilakukan untuk menguji model ResNet Baseline BatchSize 128.

Tabel 6.
Classification Report training ResNet Baseline BatchSize 128.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	10	42m	Arcuate
	0.91	1	0.95	10		Bicornuate
	1	1	1	10		Didelphys
	1	1	1	10		Normal
	1	0.9	0.95	10		Septate
	1	1	1	10		Unicornuate
			0.98	60		Accuracy
	0.98	0.98	0.98	60		Macro Avg
	0.98	0.98	0.98	60		Weight Avg
	1	1	1	21	1h 17m	Arcuate
	0.88	1	0.93	21		Bicornuate
	1	1	1	21		Didelphys
80%	1	1	1	21		Normal
	1	0.86	0.92	21		Septate
	1	1	1	21		Unicornuate
			0.98	126		Accuracy
	0.98	0.98	0.98	126		Macro Avg
	0.98	0.98	0.98	126		Weight Avg
	1	1	1	34	1h 23m	Arcuate
	1	0.85	0.92	34		Bicornuate
	1	1	1	34		Didelphys
	1	1	1	34		Normal
	0.87	1	0.93	34		Septate
	1	1	1	34		Unicornuate
			0.98	204		Accuracy
	0.98	0.98	0.98	204		Macro Avg
	0.98	0.98	0.98	204		Weight Avg

Pelatihan model ResNet baseline dengan BatchSize 128 menghasilkan akurasi tinggi (98%) di semua ukuran *TrainSize*. Waktu pelatihan meningkat saat *TrainSize* berkurang, dari 42 menit untuk *TrainSize* 90% hingga 1 jam 23 menit untuk *TrainSize* 70%. F1-Score tertinggi (1.00) ditemukan pada beberapa kelas, sedangkan kelas Bicornuate dan Septate memiliki skor terendah (0.92).

Model dengan *TrainSize* 90% menunjukkan stabilitas dan performa terbaik dengan waktu pelatihan tercepat.

Jika prioritas utama adalah akurasi dan stabilitas, maka *TrainSize* 90% merupakan pilihan terbaik.

Dengan demikian, pada model ResNet Baseline pilihan terbaik jika mempertimbangkan akurasi, stabilitas, dan efisiensi waktu adalah BatchSize 16 *TrainSize* 90%.

2. ResNet dengan Augmentasi

1) ResNet dengan Augmentasi BatchSize 8

Training model yang dilakukan untuk menguji model ResNet dengan Augmentasi BatchSize 8.

Tabel 7.
Classification Report training ResNet dengan Augmentasi BatchSize 8.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	10	5h 23m	Arcuate
	0.9	0.9	0.9	10		Bicornuate
	1	1	1	10		Didelphys
	1	1	1	10		Normal
	1	0.9	0.95	10		Septate
	0.91	1	0.95	10		Unicornuate
			0.97	60		Accuracy
	0.97	0.97	0.97	60		Macro Avg
	0.97	0.97	0.97	60		Weight Avg
	0.96	1	0.98	24	9h 24m	Arcuate
	0.74	0.96	0.84	24		Bicornuate
	1	0.96	0.98	24		Didelphys
80%	0.96	1	0.98	24		Normal
	1	0.67	0.8	24		Septate
	1	1	1	24		Unicornuate
			0.93	144		Accuracy
	0.94	0.93	0.93	144		Macro Avg
	0.94	0.93	0.93	144		Weight Avg
	1	1	1	40	9h 10m	Arcuate
	0.88	0.9	0.89	40		Bicornuate
	1	1	1	40		Didelphys
	1	0.97	0.99	40		Normal
	0.88	0.88	0.88	40		Septate
	1	1	1	40		Unicornuate
			0.96	240		Accuracy
	0.96	0.96	0.96	240		Macro Avg
	0.96	0.96	0.96	240		Weight Avg

Model dengan *TrainSize* 90% menunjukkan performa terbaik dengan akurasi, precision, recall, dan F1-Score tertinggi (0.97) serta waktu pelatihan tercepat (5 jam 23 menit). *TrainSize* 80% masih memiliki akurasi baik (0.93) tetapi dengan waktu pelatihan lebih lama (9 jam 24 menit) dan sedikit penurunan F1-Score pada beberapa kelas. *TrainSize* 70% menunjukkan akurasi 0.96 dengan waktu pelatihan 9 jam 10 menit, tetapi mengalami keterbatasan dalam generalisasi akibat data yang lebih sedikit.

Jika prioritas utama adalah akurasi dan stabilitas, maka *TrainSize* 90% merupakan pilihan terbaik.

2) ResNet dengan Augmentasi BatchSize 16

Training model yang dilakukan untuk menguji model ResNet dengan Augmentasi BatchSize 16.

Tabel 8.

Classification Report training ResNet dengan Augmentasi BatchSize 16.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	0.92	0.92	0.92	12	1h 23m	Arcuate
	0.79	0.92	0.85	12		Bicomuate
	1	0.92	0.96	12		Didelphys
	1	1	1	12		Normal
	0.82	0.75	0.78	12		Septate
	1	1	1	12		Unicomuate
			0.92	72		Accuracy
	0.92	0.92	0.92	72		Macro Avg
	0.92	0.92	0.92	72		Weight Avg
80%	1	0.73	0.84	26	2h 21m	Arcuate
	0.77	0.88	0.82	26		Bicomuate
	0.93	1	0.96	26		Didelphys
	0.96	1	0.98	26		Normal
	0.92	0.92	0.92	26		Septate
	1	1	1	26		Unicomuate
			0.92	156		Accuracy
	0.93	0.92	0.92	156		Macro Avg
	0.93	0.92	0.92	156		Weight Avg
70%	0.9	0.92	0.91	39	2h 29m	Arcuate
	0.88	0.92	0.9	39		Bicomuate
	0.97	0.9	0.93	39		Didelphys
	0.91	1	0.95	39		Normal
	0.94	0.74	0.83	39		Septate
	0.86	0.95	0.9	39		Unicomuate
			0.91	234		Accuracy
	0.91	0.91	0.9	234		Macro Avg
	0.91	0.91	0.9	234		Weight Avg

Pelatihan model ResNet dengan augmentasi dan BatchSize 16 menunjukkan akurasi tertinggi (92%) pada TrainSize 90% dan 80%, serta 91% pada TrainSize 70%. F1-Score terbaik (1.00) diperoleh pada kelas Normal untuk TrainSize 90% dan 80%, sedangkan F1-Score terendah (0.78) terjadi pada kelas Septate dengan TrainSize 90%. Waktu pelatihan meningkat seiring dengan penurunan TrainSize, dengan TrainSize 70% memerlukan waktu terlama (2 jam 29 menit).

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 90% merupakan pilihan terbaik.

3) ResNet dengan Augmentasi BatchSize 32

Training model yang dilakukan untuk menguji model ResNet dengan Augmentasi BatchSize 32.

Tabel 9.

Classification Report training ResNet dengan Augmentasi BatchSize 32.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	11	1h 47m	Arcuate
	0.75	0.82	0.78	11		Bicomuate
	1	0.73	0.84	11		Didelphys
	1	1	1	11		Normal
	0.83	0.91	0.87	11		Septate
	0.92	1	0.96	11		Unicomuate
			0.91	66		Accuracy
	0.92	0.91	0.91	66		Macro Avg
	0.92	0.91	0.91	66		Weight Avg
80%	0.89	0.86	0.87	28	2h 23m	Arcuate
	0.9	1	0.95	28		Bicomuate
	0.93	1	0.97	28		Didelphys
	0.97	1	0.98	28		Normal

70%	1	0.86	0.92	28	1h 58m	Septate
	0.93	0.89	0.91	28		Unicomuate
			0.93	168		Accuracy
	0.94	0.93	0.93	168		Macro Avg
	0.94	0.93	0.93	168		Weight Avg
	0.95	0.95	0.95	42		Arcuate
	0.91	1	0.95	42		Bicomuate
	0.95	1	0.98	42		Didelphys
	1	0.88	0.94	42		Normal
70%	0.88	0.86	0.87	42	1h 58m	Septate
	1	1	1	42		Unicomuate
			0.95	252		Accuracy
	0.95	0.95	0.95	252		Macro Avg
	0.95	0.95	0.95	252		Weight Avg

Model dengan TrainSize 70% menunjukkan performa terbaik dalam presisi, recall, dan F1-Score (macro dan weighted average 0.95). Sementara itu, TrainSize 90% memiliki akurasi sedikit lebih rendah (0.91) tetapi menawarkan waktu pelatihan yang lebih cepat. Jika prioritasnya adalah performa klasifikasi yang tinggi, TrainSize 70% lebih unggul. Namun, jika efisiensi waktu juga menjadi faktor penting, TrainSize 90% tetap menjadi pilihan kompetitif dengan akurasi yang andal.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 90% merupakan pilihan terbaik.

4) ResNet dengan Augmentasi BatchSize 64

Training model yang dilakukan untuk menguji model ResNet dengan Augmentasi BatchSize 64.

Tabel 10.

Classification Report training ResNet dengan Augmentasi BatchSize 64.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	8	1h 18m	Arcuate
	0.7	0.88	0.78	8		Bicomuate
	1	0.88	0.93	8		Didelphys
	1	1	1	8		Normal
	0.86	0.75	0.8	8		Septate
	1	1	1	8		Unicomuate
			0.92	48		Accuracy
	0.93	0.92	0.92	48		Macro Avg
	0.93	0.92	0.92	48		Weight Avg
80%	0.96	1	0.98	26	1h 19m	Arcuate
	0.96	0.88	0.92	26		Bicomuate
	0.84	1	0.91	26		Didelphys
	1	0.96	0.98	26		Normal
	0.88	0.88	0.88	26		Septate
	1	0.88	0.94	26		Unicomuate
			0.94	156		Accuracy
	0.94	0.94	0.94	156		Macro Avg
	0.94	0.94	0.94	156		Weight Avg
70%	0.86	0.95	0.9	38	1h 34m	Arcuate
	0.81	1	0.89	38		Bicomuate
	0.94	0.87	0.9	38		Didelphys
	0.85	0.89	0.87	38		Normal
	0.89	0.87	0.88	38		Septate
	0.96	0.68	0.8	38		Unicomuate
			0.88	228		Accuracy
	0.89	0.88	0.88	228		Macro Avg
	0.89	0.88	0.88	228		Weight Avg

TrainSize 90% adalah pilihan terbaik karena memiliki

akurasi 92%, stabilitas tinggi, dan waktu pelatihan yang efisien. Model ini menunjukkan performa andal dengan risiko overfitting rendah dan generalisasi yang baik. Sebaliknya, TrainSize 70% memiliki akurasi cukup tinggi (88%), tetapi stabilitasnya lambat, risiko overfitting lebih tinggi, dan waktu pelatihan lebih lama, sehingga kurang optimal untuk aplikasi nyata.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 90% merupakan pilihan terbaik.

5) ResNet dengan Augmentasi BatchSize 128

Training model yang dilakukan untuk menguji model ResNet dengan Augmentasi BatchSize 128.

Tabel 11.
Classification Report training ResNet dengan Augmentasi BatchSize 128.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	11	2h	Arcuate
	0.92	1	0.96	11		Bicornuate
	1	1	1	11		Didelphys
	1	0.91	0.95	11		Normal
	1	0.91	0.95	11		Septate
	0.92	1	0.96	11		Unicornuate
			0.97	66		Accuracy
	0.97	0.97	0.97	66		Macro Avg
	0.97	0.97	0.97	66		Weight Avg
80%	1	1	1	26	1h 46m	Arcuate
	1	0.87	0.93	26		Bicornuate
	1	1	1	26		Didelphys
	1	1	1	26		Normal
	0.88	1	0.94	26		Septate
	1	1	1	26		Unicornuate
			0.98	156		Accuracy
	0.98	0.98	0.98	156		Macro Avg
	0.98	0.98	0.98	156		Weight Avg
70%	0.97	0.97	0.97	38	2h 24m	Arcuate
	0.85	0.76	0.81	38		Bicornuate
	0.94	0.87	0.9	38		Didelphys
	1	0.97	0.99	38		Normal
	0.71	0.92	0.8	38		Septate
	0.89	0.82	0.85	38		Unicornuate
			0.89	228		Accuracy
	0.89	0.89	0.89	228		Macro Avg
	0.89	0.89	0.89	228		Weight Avg

Model menunjukkan performa terbaik pada TrainSize 80% dengan akurasi 0.98 dan keseimbangan optimal antara akurasi, presisi, serta F1-Score. TrainSize 70% mengalami sedikit penurunan performa, terutama pada kelas Septate (F1-Score 0.80), serta waktu pelatihan lebih lama (2 jam 24 menit). Secara keseluruhan, TrainSize 80% memberikan generalisasi yang lebih kuat dibandingkan TrainSize 70%, yang kurang konsisten antar kelas.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 80% merupakan pilihan terbaik.

Dengan demikian, pada model ResNet dengan augmentasi pilihan paling optimal berdasarkan prioritas akurasi dan stabilitas adalah BatchSize 128 TrainSize 80%.

3. ResNet dengan Augmentasi dan Regularisasi

1) ResNet dengan Augmentasi dan Regularisasi

BatchSize 8

Training model yang dilakukan untuk menguji model ResNet dengan augmentasi dan regularisasi BatchSize 8.

Tabel 12.
Classification Report training ResNet dengan Augmentasi dan regularisasi BatchSize 8.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	10	2h 1m	Arcuate
	0.62	0.8	0.7	10		Bicornuate
	0.91	1	0.95	10		Didelphys
	1	1	1	10		Normal
	0.83	0.5	0.62	10		Septate
	1	1	1	10		Unicornuate
			0.88	60		Accuracy
	0.89	0.88	0.88	60		Macro Avg
	0.89	0.88	0.88	60		Weight Avg
80%	1	1	1	20	4h 4m	Arcuate
	0.88	0.75	0.81	20		Bicornuate
	1	1	1	20		Didelphys
	1	1	1	20		Normal
	0.78	0.9	0.84	20		Septate
	1	1	1	20		Unicornuate
			0.94	120		Accuracy
	0.94	0.94	0.94	120		Macro Avg
	0.94	0.94	0.94	120		Weight Avg
70%	0.97	0.97	0.97	36	5h 20m	Arcuate
	0.87	0.72	0.79	36		Bicornuate
	0.97	1	0.99	36		Didelphys
	1	1	1	36		Normal
	0.78	0.89	0.83	36		Septate
	0.97	0.97	0.97	36		Unicornuate
			0.93	216		Accuracy
	0.93	0.93	0.92	216		Macro Avg
	0.93	0.93	0.92	216		Weight Avg

TrainSize 70% memberikan F1-score lebih tinggi (0.92) dibandingkan TrainSize 90% (0.88), namun dengan waktu pelatihan lebih lama (5 jam 20 menit dan 2 jam 18 menit). Kelas Normal diklasifikasikan dengan sempurna, sementara Bicornuate dan Septate menunjukkan performa lebih rendah. Model menunjukkan peningkatan dalam menangkap pola pada data yang lebih beragam, meskipun ada tantangan dalam klasifikasi kategori tertentu.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 80% merupakan pilihan terbaik.

2) ResNet dengan Augmentasi dan Regularisasi BatchSize 16.

Training model yang dilakukan untuk menguji model ResNet dengan augmentasi dan regularisasi BatchSize 16.

Tabel 13.
Classification Report training ResNet dengan Augmentasi dan regularisasi BatchSize 16.

Train Size	Precision	Recall	F1	Support	Durasi	
90%	1	1	1	13	2h 55m	Arcuate
	0.86	0.92	0.89	13		Bicornuate
	1	1	1	13		Didelphys
	1	1	1	13		Normal
	1	0.85	0.92	13		Septate
	0.93	1	0.96	13		Unicornuate
			0.96	78		Accuracy

80%	0.96	0.96	0.96	78	Macro Avg
	0.96	0.96	0.96	78	Weight Avg
	0.96	0.96	0.96	23	Arcuate
	1	0.91	0.95	23	Bicomuate
	0.96	1	0.98	23	Didelphys
	1	0.96	0.98	23	Normal
	0.92	1	0.96	23	Septate
	1	1	1	23	Unicomuate
			0.97	138	Accuracy
	0.97	0.97	0.97	138	Macro Avg
70%	0.97	0.97	0.97	138	Weight Avg
	1	0.95	0.97	40	Arcuate
	0.83	0.85	0.84	40	Bicomuate
	0.95	1	0.98	40	Didelphys
	1	0.95	0.97	40	Normal
	0.89	0.82	0.86	40	Septate
	0.91	1	0.95	40	Unicomuate
			0.93	240	Accuracy
	0.93	0.93	0.93	240	Macro Avg
	0.93	0.93	0.93	240	Weight Avg

Model ResNet dengan augmentasi dan regularisasi (BatchSize 16) menunjukkan performa tinggi, dengan akurasi dan weighted average F1-score mendekati 1.00 di berbagai TrainSize. Semua kelas diklasifikasikan dengan presisi, recall, dan F1-score yang konsisten tinggi. Waktu pelatihan berkisar antara 2–3 jam, menunjukkan efisiensi model meskipun TrainSize bervariasi.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 80% adalah pilihan terbaik.

3) ResNet dengan Augmentasi dan Regularisasi BatchSize 32.

Training model yang dilakukan untuk menguji model ResNet dengan augmentasi dan regularisasi BatchSize 32.

Tabel 14.
Classification Report training ResNet dengan Augmentasi dan regularisasi BatchSize 32.

Train Size	Precision	Recall	F1	Support	Duration
90%	1	1	1	9	Arcuate
	1	0.89	0.94	9	Bicomuate
	1	1	1	9	Didelphys
	1	1	1	9	Normal
	0.9	1	0.95	9	Septate
	1	1	1	9	Unicomuate
			0.98	54	Accuracy
	0.98	0.98	0.98	54	Macro Avg
	0.98	0.98	0.98	54	Weight Avg
80%	1	1	1	25	Arcuate
	0.89	0.96	0.92	25	Bicomuate
	1	0.92	0.96	25	Didelphys
	0.81	1	0.89	25	Normal
	1	0.72	0.84	25	Septate
	0.96	1	0.98	25	Unicomuate
			0.93	150	Accuracy
	0.94	0.93	0.93	150	Macro Avg
	0.94	0.93	0.93	150	Weight Avg
70%	1	1	1	35	Arcuate
	0.94	0.97	0.96	35	Bicomuate
	1	1	1	35	Didelphys
	0.97	1	0.99	35	Normal
	0.97	0.91	0.94	35	Septate

1	1	1	35	Unicomuate
		0.98	210	Accuracy
0.98	0.98	0.98	210	Macro Avg
0.98	0.98	0.98	210	Weight Avg

Model menunjukkan akurasi tinggi (0.98) di semua skenario TrainSize. Confusion matrix mengindikasikan klasifikasi yang sangat baik tanpa kesalahan signifikan. Presisi, recall, dan F1-score mendekati sempurna, terutama pada TrainSize 90%. Waktu pelatihan meningkat seiring berkurangnya TrainSize (2 jam 2 menit pada 90% hingga 2 jam 46 menit pada 70%), tetapi akurasi tetap stabil, mencerminkan efektivitas metode yang diterapkan.

Jika akurasi dan stabilitas sama-sama penting, maka TrainSize 90% sebagai pilihan terbaik.

4) ResNet dengan Augmentasi dan Regularisasi BatchSize 64.

Training model yang dilakukan untuk menguji model ResNet dengan augmentasi dan regularisasi BatchSize 64.

Tabel 15.
Classification Report training ResNet dengan Augmentasi dan regularisasi BatchSize 64.

Train Size	Precision	Recall	F1	Support	Duration
90%	1	1	1	11	Arcuate
	0.91	0.91	0.91	11	Bicomuate
	1	1	1	11	Didelphys
	1	1	1	11	Normal
	1	0.82	0.9	11	Septate
	0.85	1	0.92	11	Unicomuate
			0.95	66	Accuracy
	0.96	0.95	0.95	66	Macro Avg
	0.96	0.95	0.95	66	Weight Avg
80%	1	0.96	0.98	25	Arcuate
	0.96	0.88	0.92	25	Bicomuate
	0.93	1	0.96	25	Didelphys
	1	1	1	25	Normal
	0.86	0.96	0.91	25	Septate
	1	0.92	0.96	25	Unicomuate
			0.95	150	Accuracy
	0.96	0.95	0.95	150	Macro Avg
	0.96	0.95	0.95	150	Weight Avg
70%	1	0.93	0.96	41	Arcuate
	0.9	0.9	0.9	41	Bicomuate
	1	0.85	0.92	41	Didelphys
	0.93	1	0.96	41	Normal
	0.87	0.83	0.85	41	Septate
	0.84	1	0.91	41	Unicomuate
			0.92	246	Accuracy
	0.92	0.92	0.92	246	Macro Avg
	0.92	0.92	0.92	246	Weight Avg

Model memiliki *precision*, *recall*, dan F1-score yang sangat baik, terutama pada kelas Arcuate, Didelphys, dan Normal, dengan nilai mendekati atau sama dengan 1. Namun, terdapat sedikit penurunan performa pada kelas Septate, khususnya pada skenario TrainSize 70%, yang mencerminkan tantangan dalam membedakan kelas ini. Waktu pelatihan semakin singkat seiring pengurangan ukuran TrainSize, dengan performa model tetap terjaga, menunjukkan efisiensi yang baik.

Jika prioritas utama adalah akurasi dan stabilitas, maka TrainSize 90% adalah pilihan terbaik.

5) ResNet dengan Augmentasi dan Regularisasi BatchSize 128

Training model yang dilakukan untuk menguji model ResNet dengan augmentasi dan regularisasi BatchSize 128.

Tabel 16.
Classification Report training ResNet dengan augmentasi dan regularisasi BatchSize 128.

Train Size	Precision	Recall	F1	Support	Durasi
90%	1	1	1	9	Arcuate
	1	0.78	0.88	9	Bicornuate
	0.82	1	0.9	9	Didelphys
	1	1	1	9	Normal
	0.89	0.89	0.89	9	Septate
	1	1	1	9	Unicornuate
			0.94	54	Accuracy
	0.95	0.94	0.94	54	Macro Avg
	0.95	0.94	0.94	54	Weight Avg
80%	1	0.92	0.96	25	Arcuate
	0.86	1	0.93	25	Bicornuate
	1	1	1	25	Didelphys
	0.96	1	0.98	25	Normal
	0.95	0.84	0.89	25	Septate
	1	1	1	25	Unicornuate
			0.96	150	Accuracy
	0.96	0.96	0.96	150	Macro Avg
	0.96	0.96	0.96	150	Weight Avg
70%	0.92	0.97	0.94	35	Arcuate
	0.65	0.91	0.76	35	Bicornuate
	1	0.89	0.94	35	Didelphys
	0.9	1	0.95	35	Normal
	1	0.37	0.54	35	Septate
	0.85	1	0.92	35	Unicornuate
			0.86	210	Accuracy
	0.89	0.86	0.84	210	Macro Avg
	0.89	0.86	0.84	210	Weight Avg

Model menunjukkan akurasi tinggi dengan peningkatan TrainSize. Pada TrainSize 70%, akurasi mencapai 86%, meningkat menjadi 94% dan 96% pada TrainSize 80% dan 90%. Model mampu mengenali pola dengan baik, terutama untuk kategori Normal dan Unicornuate dengan F1-score mendekati sempurna. Meskipun waktu pelatihan sedikit lebih lama pada TrainSize yang lebih besar, model tetap andal dalam klasifikasi anomali uterus.

Jika prioritas adalah akurasi dan stabilitas, maka TrainSize 80% adalah pilihan terbaik.

Dengan demikian, pada model ResNet dengan Augmentasi dan Regularisasi pilihan paling optimal berdasarkan prioritas akurasi dan stabilitas adalah BatchSize 64 TrainSize 90%.

B. Hasil Pengujian.

Untuk menghitung akurasi sistem, digunakan perhitungan nilai *confussion matrix*. Dalam konteks *confusion matrix* multi-kelas, hanya fokus pada total dari *True Positive* yang benar untuk semua kelas.

Hasil pengujian menggunakan 3 model data dapat dilihat pada tabel 17.

Tabel 17.
Hasil Confussion Matrix pengujian model.

Model ResNet	Accuracy	Precision	Recall	F1-Score
Baseline BatchSize 16 TrainSize 90%	1 (100%)	1 (100%)	1 (100%)	1 (100%)
Augmentasi BatchSize 128 TrainSize 80%	0.89 (89%)	0.96 (96%)	0.95 (95%)	0.94 (94%)
Augmentasi dan Regularisasi BatchSize 64 TrainSize 90%	0.64 (64%)	0.75 (75%)	0.78 (78%)	Tidak didefinisikan

Berdasarkan tabel 17, jika prioritasnya adalah akurasi dan stabilitas, pilihan terbaik adalah model ResNet Baseline dengan BatchSize 16 dan TrainSize 90%. Alasannya adalah model ini memiliki akurasi 100%, lebih tinggi dibanding model lainnya. Stabilitas model ini baik dengan Precision 100%, recall 100%, dan F1-score 100% menunjukkan performa yang konsisten di berbagai metrik evaluasi.

IV. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil evaluasi sistem menggunakan *confusion matrix* untuk klasifikasi enam kelas (Arcuate, Bicornuate, Didelphys, Normal, Septate, dan Unicornuate), model ResNet baseline dengan BatchSize 16 dan TrainSize 90% menunjukkan kinerja terbaik dengan akurasi, presisi, *recall*, dan F1-score mencapai 100%. Model ResNet dengan augmentasi BatchSize 128 dan TrainSize 80% mengalami sedikit penurunan performa, dengan akurasi 89% dan F1-score 95%. Sementara itu, model ResNet dengan augmentasi serta regularisasi BatchSize 64 dan TrainSize 90% menunjukkan kinerja terendah dengan akurasi 64% dan F1-score yang tidak terdefinisi, mengindikasikan bahwa kombinasi augmentasi dan regularisasi dalam parameter ini tidak optimal untuk deteksi malformasi uterus.

B. Saran

Untuk meningkatkan performa model, perlu dilakukan eksplorasi lebih lanjut terhadap kombinasi *hyperparameter*, termasuk *BatchSize*, proporsi *TrainSize*, serta teknik augmentasi dan regularisasi yang lebih tepat. Selain itu, pengujian dengan arsitektur *deep learning* lain seperti *EfficientNet* atau *Vision Transformer* dapat dipertimbangkan untuk membandingkan efektivitas deteksi. Penggunaan *dataset* yang lebih besar dan beragam juga diharapkan dapat meningkatkan generalisasi model dalam mendeteksi malformasi uterus secara lebih akurat. Pertimbangan untuk mengevaluasi menggunakan metrik

lain seperti AUC-ROC atau Precision-Recall untuk analisis lebih mendalam terhadap performa model.

DAFTAR PUSTAKA

- [1] I. B. W. Dharma, "Legalitas Abortus Provocatus Sebagai Akibat Tindakan Pemerkosaan," *Kertha Wicaksana*, vol. 16, no. 1, pp. 45–50, 2022.
- [2] S. I. Setiawan, "Diagnosis dan Tatalaksana Uterus Bikornu," *Cermin Dunia Kedokt.*, vol. 48, no. 3, pp. 123–127, 2021.
- [3] Y. Yoshimasa and T. Maruyama, "Bioengineering of the Uterus," *Reprod. Sci.*, vol. 28, no. 6, pp. 1596–1611, Jun. 2021, doi: 10.1007/s43032-021-00503-8.
- [4] E. D and N. P. G. Bhavani, "An Effective DNN Based ResNet Approach for Satellite Image Classification," in *2023 4th International Conference on Smart Electronics and Communication (ICOSEC)*, Sep. 2023, pp. 1055–1062. doi: 10.1109/ICOSEC58147.2023.10276330.
- [5] T. Nakaura, T. Higaki, K. Awai, O. Ikeda, and Y. Yamashita, "A primer for understanding radiology articles about machine learning and deep learning," *Diagn. Interv. Imaging*, vol. 101, no. 12, pp. 765–770, 2020, doi: <https://doi.org/10.1016/j.diii.2020.10.001>.
- [6] S. A. Robertson, "Control of the immunological environment of the uterus," *Rev. Reprod.*, vol. 5, no. 3, pp. 164–174, 2000.
- [7] Y. Kurata *et al.*, "Automatic segmentation of the uterus on MRI using a convolutional neural network," *Comput. Biol. Med.*, vol. 114, p. 103438, 2019, doi: <https://doi.org/10.1016/j.compbiomed.2019.103438>.
- [8] F. Zafarani, F. Ahmadi, and G. Shahrzad, "Hysterosalpingographic features of cervical abnormalities: acquired structural anomalies," *Br. J. Radiol.*, vol. 88, no. 1052, p. 20150045, 2015, doi: 10.1259/bjr.20150045.
- [9] V. Jadeja, A. L. N. Rao, A. Srivastava, S. Singh, P. Chaturvedi, and G. Bhardwaj, "Convolutional Neural Networks: A Comprehensive Review of Architectures and Application," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, Sep. 2023, pp. 460–467. doi: 10.1109/IC3I59117.2023.10397695.
- [10] F. Zhipeng and H. Gani, "Cultural Events Classification using Hyper-parameter Optimization of Deep Learning Technique," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 5, 2021.
- [11] Z. Ahmed, M. S. Kareem, H. A. Khan, F. H. Jaskani, Z. Saman, and B. Mughal, "Detection of Uterine Fibroids in Medical Images Using Deep Neural Networks".
- [12] C. Irawan, E. D. Udayanti, and F. A. Nugroho, "Visualisasi dan Rekonstruksi 3D Citra Medis," *Semantik 2013*, vol. 3, no. 1, pp. 61–64, 2013.