

# Implementasi Arsitektur MobileNetV2 Berbasis Citra untuk Deteksi Penyakit Dropsy dan Popeye pada Ikan Cupang

Fadhilah Rafi Musyaffa\*, Adzhal Arwani Mahfudh, Moh Hadi Subowo

Fakultas Sains dan Teknologi, Sistem Informasi Teknologi Informasi, Universitas Islam Negeri Walisongo, Semarang, Indonesia

Email: <sup>1,\*</sup>2208096110@student.walisongo.ac.id, <sup>2</sup>adzhal@walisongo.ac.id, <sup>3</sup>hadi.subowo@walisongo.ac.id

Email Penulis Korespondensi: 2208096110@student.walisongo.ac.id

Submitted: 10/02/2026; Accepted: 05/03/2026; Published: 06/03/2026

**Abstrak**—Identifikasi penyakit pada ikan cupang berdasarkan gejala visual masih sedikit menjadi tantangan, khususnya bagi para pemula yang belum memiliki pengalaman dalam mengenali karakteristik suatu penyakit. Penelitian ini bertujuan untuk mengimplementasikan arsitektur MobileNetV2 berbasis citra sebagai sistem pendukung diagnosis untuk mendeteksi penyakit *dropsy* dan *popeye* pada ikan cupang yang telah menunjukkan gejala visual. Dataset yang digunakan terdiri dari 600 citra ikan cupang yang terbagi ke dalam tiga kelas, yaitu sehat, *dropsy*, dan *popeye*, masing-masing sebanyak 200 citra, yang diperoleh dari internet. *Preprocessing* data dilakukan dengan penyesuaian rasio citra, normalisasi, serta augmentasi untuk meningkatkan variasi data. Metode *transfer learning* diterapkan dengan membekukan sebagian besar lapisan ekstraktor fitur MobileNetV2 dan melakukan penyesuaian pada beberapa lapisan akhir. Evaluasi model dilakukan menggunakan *5-Fold Cross Validation* untuk memastikan stabilitas dan reproduktibilitas eksperimen. Model terbaik dari setiap *fold* kemudian digabungkan menggunakan metode *ensemble* berdasarkan rata-rata probabilitas untuk meningkatkan performa prediksi pada data uji. Hasil eksperimen menunjukkan rata-rata akurasi *5-Fold Cross Validation* sebesar 74,71% dengan deviasi standar  $\pm 4,57\%$ , serta nilai *Macro-F1* sebesar  $\pm 74,43\%$ . Penerapan metode *ensemble* menghasilkan akurasi pengujian sebesar 85,56% dengan performa klasifikasi yang seimbang pada ketiga kelas. Visualisasi Grad-CAM menunjukkan bahwa model mampu memfokuskan perhatian pada area citra yang relevan dengan gejala penyakit. Hasil penelitian ini menunjukkan bahwa arsitektur MobileNetV2 efektif digunakan sebagai alat bantu diagnosis penyakit ikan cupang berbasis citra.

**Kata Kunci:** Transfer Learning; MobileNetV2; Ikan Cupang; Deteksi Penyakit; Citra Digital

**Abstract**—The identification of diseases in betta fish based on visual symptoms remains a challenge, particularly for beginners who lack experience in recognizing disease characteristics. This study aims to implement an image-based MobileNetV2 architecture as a diagnostic support system to detect dropsy and popeye diseases in betta fish that have already exhibited visual symptoms. The dataset used in this study consists of 600 betta fish images divided into three classes: healthy, dropsy, and popeye, with 200 images in each class, collected from the internet. Data preprocessing was conducted through image ratio adjustment, normalization, and data augmentation to increase data variability. A transfer learning approach was applied by freezing most layers of the MobileNetV2 feature extractor and fine-tuning several of the final layers. Model evaluation was performed using 5-Fold Cross Validation to ensure experimental stability and reproducibility. The best model from each fold was then combined using an ensemble method based on average probability to improve prediction performance on the test dataset. Experimental results show that the average 5-Fold Cross Validation accuracy reached 74.71% with a standard deviation of  $\pm 4.57\%$ , while the Macro-F1 score achieved  $\pm 74.43\%$ . The ensemble approach produced a test accuracy of 85.56% with balanced classification performance across all classes. Grad-CAM visualizations indicate that the model is able to focus on image regions relevant to disease symptoms. These findings demonstrate that the MobileNetV2 architecture is effective as an image-based diagnostic support tool for betta fish diseases.

**Keywords:** Transfer Learning; MobileNetV2; Betta Fish; Disease Detection; Digital Images

## 1. PENDAHULUAN

Penyakit pada ikan cupang (*Betta splendens*) seperti *dropsy* dan *popeye* (*exophthalmia*) merupakan kondisi yang dapat berdampak serius terhadap kesehatan dan kelangsungan hidup ikan jika tidak ditangani dengan tepat. Identifikasi penyakit ini umumnya dilakukan berdasarkan gejala visual, seperti pembengkakan tubuh, sisik yang mengembang, atau mata yang menonjol, yang sering diasosiasikan dengan berbagai gangguan infeksi maupun non-infeksi. Namun, proses identifikasi tersebut masih menjadi tantangan, terutama bagi pembudidaya dan penghobi yang belum memiliki pengalaman atau pengetahuan memadai dalam mengenali karakteristik setiap penyakit. Kesalahan dalam mendiagnosis penyakit dapat berujung pada penanganan yang tidak sesuai, sehingga berpotensi memperburuk kondisi ikan [1], [2], [3]. Oleh karena itu, diperlukan sebuah sistem pendukung diagnosis berbasis teknologi yang mampu membantu pengguna dalam mengidentifikasi jenis penyakit ikan cupang secara objektif dan konsisten berdasarkan citra visual yang menampilkan gejala penyakit, sehingga dapat menjadi alat bantu yang mudah diakses dan digunakan oleh masyarakat luas.

Pemanfaatan *deep learning*, khususnya arsitektur *Convolutional Neural Network* (CNN), telah banyak diterapkan dalam bidang pengolahan citra untuk membantu proses klasifikasi dan deteksi penyakit pada berbagai domain. Sejumlah penelitian menunjukkan bahwa CNN mampu mengekstraksi fitur visual secara otomatis dan memberikan performa yang baik. Penelitian oleh Suhendar, dkk. pada tahun 2023 dan Alamsyah, dkk. pada tahun 2025 menunjukkan keberhasilan CNN dalam mendeteksi penyakit tanaman berdasarkan citra daun dengan tingkat akurasi yang tinggi, yakni 98,67% dan 97,5% secara berurutan [4], [5]. Pendekatan serupa juga diterapkan pada bidang kesehatan hewan, seperti deteksi kondisi kesegaran ikan dan penyakit unggas menggunakan citra visual, yang menunjukkan bahwa pola visual tertentu dapat dimanfaatkan sebagai indikator kondisi biologis pada hewan [6], [7].

Selain itu, penggunaan arsitektur CNN yang ringan seperti MobileNet semakin populer karena mampu memberikan keseimbangan antara akurasi dan efisiensi komputasi, sehingga cocok untuk diimplementasikan pada sistem dengan keterbatasan sumber daya [8].

Penelitian terkait deteksi penyakit pada ikan berbasis citra juga mulai berkembang seiring meningkatnya kebutuhan otomasi dalam bidang akuakultur [9]. Beberapa studi memanfaatkan *computer vision* dan *machine learning* untuk mengklasifikasikan kondisi kesehatan ikan berdasarkan ciri visual. Penelitian oleh Tamut, dkk. pada tahun 2025 menggunakan CNN untuk mengidentifikasi penyakit pada ikan air tawar dan melaporkan akurasi yang sangat tinggi dalam klasifikasi penyakit ikan air tawar [10]. Studi lain oleh Hwang, dkk. pada tahun 2025 menerapkan *deep learning* untuk mendeteksi penyakit pada ikan budidaya berdasarkan citra tubuh ikan, namun masih menghadapi keterbatasan pada jumlah dan variasi dataset [11]. Selain itu, penelitian Peh, dkk. pada tahun 2025 menunjukkan bahwa sebagian besar studi masih berfokus pada ikan konsumsi, sementara penelitian mengenai ikan hias masih sangat terbatas [12].

Dalam konteks arsitektur *deep learning*, beberapa penelitian terbaru mulai mengadopsi model *lightweight* seperti MobileNet untuk klasifikasi citra dengan keterbatasan sumber daya komputasi. Penelitian oleh Daş pada tahun 2025 menunjukkan bahwa MobileNetV2 mampu memberikan keseimbangan yang baik antara akurasi dan efisiensi komputasi melalui penggunaan *depthwise separable convolution*, serta *inverted residual blocks* [13]. Studi lain oleh Somoal, dkk. pada tahun 2025 dan Rastogi, dkk. pada tahun 2025 menerapkan MobileNetV2 dengan metode *transfer learning* pada klasifikasi citra medis dan melaporkan performa yang kompetitif dengan arsitektur yang lebih kompleks [14], [15].

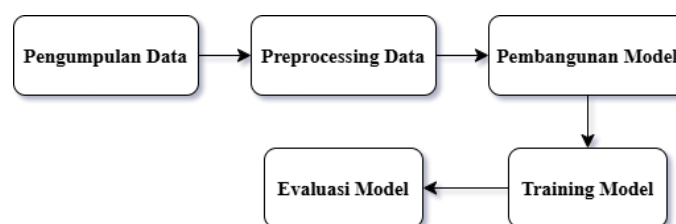
Berdasarkan penjabaran di atas, penelitian klasifikasi penyakit ikan hias berbasis citra menggunakan *convolutional neural network* (CNN) telah berkembang dalam beberapa tahun terakhir, namun sebagian besar studi masih mengevaluasi performa model menggunakan satu skema pembagian data, seperti *train–test* split, tanpa mengkaji stabilitas model secara statistik. Penerapan *k-fold cross-validation* pada klasifikasi penyakit ikan hias, khususnya pada ikan cupang, masih jarang dilaporkan. Selain itu, pemanfaatan *ensemble* model yang dibangun dari model-model hasil *cross-validation* untuk meningkatkan *robustnes* klasifikasi penyakit ikan hias berbasis citra juga belum banyak dikaji. Dengan demikian, masih terdapat celah penelitian pada evaluasi stabilitas model CNN berbasis *transfer learning* serta integrasi *ensemble* model dalam klasifikasi penyakit ikan cupang berbasis citra.

Berdasarkan permasalahan dan celah penelitian yang telah diuraikan, penelitian ini bertujuan untuk mengimplementasikan arsitektur MobileNetV2 berbasis citra sebagai sistem pendukung diagnosis dalam mendeteksi penyakit *dropsy* dan *popeye* pada ikan cupang yang telah menunjukkan gejala visual. Penelitian ini dirancang untuk membantu pemula yang belum memiliki pengalaman dalam mengenali karakteristik penyakit ikan cupang, sehingga dapat memperoleh informasi diagnosis secara lebih objektif dan konsisten.

Kontribusi utama penelitian ini terletak pada penerapan arsitektur MobileNetV2 berbasis *transfer learning* untuk klasifikasi penyakit *dropsy* dan *popeye* pada ikan cupang menggunakan citra visual. Selain itu, penelitian ini mengevaluasi kestabilan performa model secara lebih komprehensif menggunakan skema *5-Fold Cross Validation*, sehingga konsistensi kinerja model terhadap variasi data dapat dianalisis secara sistematis. Untuk meningkatkan performa klasifikasi, penelitian ini juga menerapkan pendekatan *ensemble soft voting* berbasis rata-rata probabilitas dari model-model hasil proses *cross-validation*. Di samping itu, aspek interpretabilitas model dianalisis menggunakan *Grad-CAM* untuk mengidentifikasi area citra yang berkontribusi terhadap keputusan klasifikasi model.

## 2. METODOLOGI PENELITIAN

Tahapan penelitian diawali dengan pengumpulan dataset citra ikan cupang yang telah menunjukkan gejala penyakit, kemudian dilanjutkan dengan *preprocessing* data untuk memastikan kualitas dan konsistensi citra sebelum digunakan dalam pelatihan model. Selanjutnya, dilakukan pembangunan dan pelatihan model menggunakan skema *5-Fold Cross Validation* guna memperoleh hasil yang stabil dan dapat direproduksi [16]. Model terbaik dari setiap *fold* kemudian digabungkan menggunakan metode *ensemble*. Tahap akhir penelitian meliputi evaluasi kinerja model menggunakan beberapa metrik klasifikasi serta analisis visualisasi *Gradient-weighted Class Activation Mapping* (Grad-CAM) untuk memahami fokus perhatian model terhadap area gejala penyakit pada citra [16]. Secara umum, alur penelitian ini disajikan dalam bentuk diagram yang ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

Gambar 1 menunjukkan alur penelitian yang dimulai dengan tahap pengumpulan data, *preprocessing* data, pembangunan model, *training* model dan diakhiri dengan evaluasi model.

## 2.1 Pengumpulan Data

Tahap pengumpulan data merupakan langkah awal dalam penelitian ini yang bertujuan untuk memperoleh kumpulan citra ikan cupang yang dapat mewakili kondisi sehat maupun kondisi penyakit yang diteliti. Dataset yang digunakan terdiri dari 600 citra ikan cupang yang dibagi ke dalam tiga kelas, yaitu sehat, *dropsy*, dan *popeye*, dengan masing-masing kelas berjumlah 200 citra. Seluruh citra diperoleh dari sumber terbuka di internet, seperti Reddit, Bettafish.org, dan Google Images. Sumber-sumber tersebut dipilih karena hingga saat ini dataset publik yang secara khusus memuat citra penyakit ikan cupang masih sangat terbatas, sehingga sulit ditemukan pada platform dataset umum seperti Kaggle. Selain itu, citra yang berasal dari komunitas pemelihara ikan cupang umumnya diambil dalam kondisi nyata sehari-hari, dengan perbedaan sudut pengambilan gambar, pencahayaan, latar belakang, serta variasi warna dan bentuk ikan, sehingga lebih mencerminkan kondisi di lapangan.

Citra yang digunakan merupakan citra yang secara visual telah menunjukkan gejala penyakit *dropsy* dan *popeye*. Oleh karena itu, penelitian ini tidak berfokus pada deteksi penyakit pada tahap dini, melainkan pada pengenalan jenis penyakit berdasarkan gejala yang sudah tampak pada citra. Pengumpulan data melalui pemotretan ikan secara langsung tidak dilakukan karena keterbatasan jumlah sampel dan pertimbangan etika, mengingat pengambilan data penyakit akan membutuhkan pemeliharaan ikan dalam jumlah besar serta menunggu munculnya penyakit tertentu, yang berpotensi berdampak pada kesejahteraan ikan. Penggunaan data dari sumber terbuka dipilih sebagai alternatif yang lebih efisien dan etis.

Proses pemberian label pada dataset dilakukan berdasarkan pengamatan visual terhadap gejala penyakit yang terlihat jelas pada citra. Pelabelan dilakukan oleh peneliti bersama seorang peternak ikan cupang yang telah berpengalaman lebih dari lima tahun dalam menangani berbagai kondisi kesehatan ikan cupang. Penentuan label didasarkan pada ciri-ciri yang umum dikenal oleh peternak, seperti sisik yang mengembang dan perut membesar pada penyakit *dropsy*, serta mata yang menonjol pada penyakit *popeye*. Citra dengan kualitas yang kurang baik atau gejala yang tidak jelas tidak digunakan dalam penelitian. Untuk menjaga kualitas data, setiap citra diperiksa secara manual guna memastikan tidak terdapat citra yang sama pada kelas atau kelompok data yang berbeda. Selanjutnya, dataset dibagi secara manual ke dalam data latih/validasi, dan data uji untuk mencegah terjadinya *data leakage*, terutama akibat kemiripan visual antar citra yang berasal dari ikan yang sama.

Meskipun dataset telah disusun dengan cermat, penelitian ini memiliki beberapa keterbatasan metodologis yang perlu diperhatikan. Label pada dataset ditetapkan berdasarkan pengenalan gejala penyakit yang tampak secara visual pada citra, sehingga hasil klasifikasi merepresentasikan identifikasi berbasis gejala, bukan diagnosis klinis yang dikonfirmasi melalui pemeriksaan medis atau laboratorium. Penelitian ini tidak melibatkan dokter ikan atau tenaga medis veteriner sebagai validator klinis, sehingga kemungkinan ketidakakuratan label tidak dapat sepenuhnya dihilangkan. Oleh karena itu, sistem klasifikasi yang dikembangkan diposisikan sebagai alat pendukung pengambilan keputusan (*decision support system*) bagi pemelihara atau peternak ikan, dan tidak dimaksudkan untuk menggantikan diagnosis medis atau konsultasi dengan ahli kesehatan ikan.

## 2.2 Preprocessing Data

Tahap *preprocessing* data dilakukan untuk memastikan seluruh citra memiliki format dan kualitas yang seragam sebelum digunakan dalam proses pelatihan model. Tahapan ini bertujuan untuk mengurangi variasi yang tidak relevan pada data citra, meningkatkan kemampuan generalisasi model, serta meminimalkan potensi bias akibat perbedaan ukuran, rasio, dan pencahayaan citra [17], [18]. Proses *preprocessing* pada penelitian ini terdiri dari *preprocessing* manual dan *preprocessing* berbasis kode program.

### 2.2.1 Preprocessing Manual Citra

*Preprocessing* manual dilakukan menggunakan perangkat lunak Adobe Photoshop untuk memastikan citra berfokus pada objek utama dan bukan latar belakang, dan menyesuaikan rasio citra menjadi 1:1. Penyesuaian ini dilakukan dengan menambahkan tepi transparan pada sisi citra tanpa melakukan *cropping* pada objek utama, sehingga bentuk ikan cupang tetap terjaga dan tidak terdistorsi. Langkah ini bertujuan untuk memudahkan proses *resize* pada tahap selanjutnya serta menjaga konsistensi komposisi visual pada seluruh dataset.

### 2.2.2 Augmentasi Data

Untuk meningkatkan variasi data dan mengurangi risiko *overfitting*, dilakukan proses augmentasi data menggunakan transformasi citra berbasis kode program. Augmentasi ini bertujuan untuk mensimulasikan variasi kondisi nyata yang mungkin terjadi pada saat pengambilan data, sehingga model tidak hanya bergantung pada pola visual yang terbatas pada dataset asli. Dengan adanya variasi buatan ini, diharapkan model mampu melakukan generalisasi yang lebih baik ketika dihadapkan pada data baru yang belum pernah dilihat sebelumnya [19]. Teknik augmentasi yang diterapkan meliputi.

#### a. *ResizeWithPadding* (224)

Transformasi ini digunakan untuk menyesuaikan ukuran citra menjadi  $224 \times 224$  piksel agar sesuai dengan ukuran masukan standar yang dibutuhkan oleh arsitektur MobileNetV2 [20]. Pada tahap ini, *padding* digunakan untuk mempertahankan rasio aspek objek utama dalam citra sehingga bentuk objek tidak mengalami distorsi. Selain itu,

*padding* juga berfungsi untuk mengakomodasi ketidaksempurnaan kecil yang mungkin terjadi pada tahap *preprocessing* manual di awal, yakni proporsi yang tidak sesuai.

b. *Random Horizontal Flip*

Transformasi ini melakukan pembalikan citra secara horizontal dengan probabilitas tertentu. Teknik ini ditujukan untuk menambah variasi orientasi objek dalam dataset, sehingga model tidak bergantung pada satu arah tampilan tertentu [21], [22]. Dengan adanya pembalikan horizontal, model diharapkan mampu mengenali objek dengan orientasi yang beragam tanpa kehilangan akurasi klasifikasi.

c. *Random Rotation* ( $\pm 20^\circ$ )

Rotasi acak hingga  $\pm 20$  derajat diterapkan untuk meningkatkan ketahanan model terhadap variasi sudut pengambilan gambar yang umum terjadi pada kondisi pengambilan data di lapangan [21], [22]. Dengan demikian, model menjadi lebih *robust* terhadap pergeseran sudut pandang kamera dan tidak sensitif terhadap kemiringan kecil pada citra.

d. *Color Jitter* (0.2)

Transformasi ini memodifikasi nilai *brightness*, *contrast*, dan *saturation* citra dengan tingkat variasi sebesar 0,2. Transformasi ini bertujuan untuk mensimulasikan perbedaan kondisi pencahayaan dan karakteristik warna yang dapat muncul akibat variasi sumber cahaya atau lingkungan pengambilan gambar [21], [22].

### 2.2.3 Transformasi Tensor dan Normalisasi

Sebelum proses pelatihan dimulai, semua gambar yang telah melalui tahap augmentasi diubah dari format gambar biasa (*PIL Image*) menjadi bentuk data numerik yang disebut *tensor*. Proses ini dilakukan menggunakan fungsi *ToTensor()*. Perubahan ini diperlukan karena model *deep learning*, seperti yang digunakan dalam *PyTorch*, hanya dapat mengolah data dalam bentuk angka. Selain itu, fungsi *ToTensor()* juga secara otomatis mengubah nilai piksel gambar dari skala 0–255 menjadi 0–1, sehingga perhitungan selama pelatihan menjadi lebih stabil [23], [24].

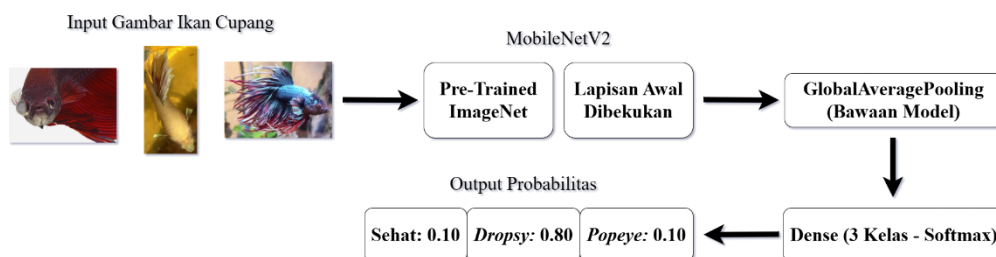
Setelah gambar berbentuk *tensor*, langkah selanjutnya adalah normalisasi. Pada tahap ini, nilai piksel disesuaikan menggunakan nilai rata-rata (mean) dan standar deviasi yang sama dengan yang digunakan saat melatih model MobileNetV2 sebelumnya. Tujuan normalisasi ini adalah agar data yang dimasukkan sesuai dengan pola data yang sudah dikenal oleh model. Dengan data yang lebih seragam, proses pelatihan menjadi lebih stabil, risiko kesalahan perhitungan dapat dikurangi, dan model dapat belajar dengan lebih cepat dan efektif [23], [24].

### 2.3 Pembangunan Model

Dalam penelitian ini, arsitektur MobileNetV2 dipilih karena dinilai cukup sesuai dengan karakteristik data dan keterbatasan sumber daya yang digunakan. MobileNetV2 merupakan pengembangan dari MobileNetV1 yang dirancang untuk meningkatkan kualitas ekstraksi fitur tanpa menambah beban komputasi secara signifikan. Peningkatan tersebut dicapai melalui penggunaan *inverted residual blocks* dan *linear bottleneck*, yaitu mekanisme yang membantu jaringan mempertahankan informasi penting saat memproses fitur dengan jumlah parameter yang relatif kecil [25] [26]. Dengan karakteristik ini, MobileNetV2 banyak digunakan pada penelitian dengan dataset berukuran terbatas karena mampu memberikan performa yang cukup stabil tanpa memerlukan model yang terlalu kompleks.

Dibandingkan dengan MobileNetV3, yang merupakan versi lebih baru dan dilengkapi dengan optimasi tambahan seperti *neural architecture search* dan modul *squeeze-and-excitation* [27], MobileNetV2 memiliki struktur yang lebih sederhana dan lebih mudah dilatih. Kompleksitas tambahan pada MobileNetV3 umumnya dirancang untuk mencapai efisiensi optimal pada perangkat tertentu, namun pada skenario pelatihan berulang seperti *k-Fold Cross Validation* dan penggunaan perangkat dengan sumber daya terbatas, model yang lebih sederhana sering kali lebih praktis dan konsisten. Oleh karena itu, MobileNetV2 dipilih sebagai kompromi antara efisiensi komputasi dan kemampuan ekstraksi fitur, sehingga dianggap memadai untuk mendukung proses klasifikasi citra pada penelitian ini tanpa meningkatkan beban pelatihan secara berlebihan [25] [26] [27].

Alur pembangunan model, mulai dari input citra ikan cupang hingga keluaran berupa probabilitas prediksi untuk setiap kelas, ditunjukkan pada Gambar 2 berikut.



**Gambar 2.** Proses *transfer learning* menggunakan arsitektur MobileNetV2 untuk klasifikasi citra ikan cupang

Seperti pada Gambar 2, metode yang digunakan dalam penelitian ini adalah klasifikasi citra berbasis *transfer learning* dengan memanfaatkan MobileNetV2 yang telah dilatih sebelumnya menggunakan dataset ImageNet. Model

MobileNetV2 digunakan sebagai ekstraktor fitur, di mana sebagian besar lapisan awalnya dibekukan untuk mempertahankan kemampuan dalam mengenali pola visual umum. Selanjutnya, fitur yang dihasilkan diproses melalui mekanisme *global average pooling* yang sudah menjadi bagian dari arsitektur MobileNetV2, kemudian diteruskan ke lapisan klasifikasi baru yang disesuaikan dengan tiga kelas keluaran, yaitu ikan sehat, *dropsy*, dan *popeye* [28], [29].

Untuk menyesuaikan model dengan karakteristik citra ikan cupang, diterapkan strategi *fine-tuning* dengan membuka kembali dua blok terakhir MobileNetV2 untuk dilatih ulang. Pendekatan ini memungkinkan model mempelajari ciri khusus dari data penelitian tanpa harus melatih seluruh jaringan dari awal, sehingga proses pelatihan menjadi lebih efisien dan risiko *overfitting* akibat keterbatasan jumlah data dapat dikurangi [28], [29].

## 2.4 Training Model

Tahap pelatihan model dilakukan untuk menyesuaikan parameter arsitektur MobileNetV2 agar mampu mengklasifikasikan citra ikan cupang ke dalam tiga kelas, yaitu sehat, *dropsy*, dan *popeye*. Proses pelatihan dirancang untuk menghasilkan model yang stabil serta memiliki kemampuan generalisasi yang baik terhadap data baru. Oleh karena itu, penelitian ini menerapkan pendekatan *5-Fold Cross Validation* guna mengurangi ketergantungan performa model terhadap satu skema pembagian data tertentu [30].

Konfigurasi dan parameter pelatihan model MobileNetV2 yang digunakan dalam penelitian ini dirangkum secara singkat pada Tabel 1.

**Tabel 1.** Parameter Pelatihan Model MobileNetV2

Parameter	Nilai/Deskripsi
Arsitektur Model	MobileNetV2 ( <i>pre-trained</i> )
Jumlah Kelas	3 (Sehat, <i>Dropsy</i> , dan <i>Popeye</i> )
Metode Validasi	<i>5-Fold Cross Validation</i>
Jumlah Fold	5
<i>Epoch per Fold</i>	16
<i>Batch Size</i>	16
<i>Learning Rate</i>	$1 \times 10^{-4}$
<i>Optimizer</i>	<i>Adam</i>
<i>Scheduler</i>	<i>StepLR</i>
Lapisan yang Dilatih	Lapisan klasifikasi dan dua blok terakhir MobileNetV2
Kriteria Pemilihan Model Terbaik	Akurasi validasi tertinggi
<i>Seed</i>	Ditetapkan sama untuk <i>random</i> , <i>NumPy</i> , dan <i>PyTorch</i>
Pengaturan <i>CUDA</i>	Deterministik ( <i>benchmark</i> dinonaktifkan)
Strategi Penyimpanan Model	Model terbaik disimpan pada setiap <i>fold</i>
Penggunaan Model	Evaluasi individual <i>per fold</i> dan <i>ensemble</i> berbasis rata-rata probabilitas

Seperti ditunjukkan pada Tabel 1, proses pelatihan dilakukan dengan membagi dataset pelatihan ke dalam lima *fold*, di mana pada setiap iterasi satu *fold* digunakan sebagai data validasi dan empat *fold* lainnya sebagai data pelatihan. Setiap *fold* dilatih selama 16 *epoch* dengan ukuran *batch* sebesar 16 citra. *Learning rate* awal sebesar  $1 \times 10^{-4}$  digunakan untuk menjaga kestabilan proses *fine-tuning*, khususnya pada lapisan klasifikasi dan dua blok terakhir MobileNetV2 yang dibuka untuk dilatih kembali. Selama proses pelatihan, digunakan skema penyesuaian *learning rate* secara bertahap menggunakan *StepLR*, di mana nilai *learning rate* dikurangi secara periodik untuk membantu model melakukan pembaruan bobot yang lebih halus seiring bertambahnya *epoch*.

Untuk memastikan reproduktibilitas eksperimen, nilai *seed* yang sama diterapkan pada pustaka *random*, *NumPy*, dan *PyTorch*. Selain itu, pengaturan *CUDA backend* dibuat deterministik dengan menonaktifkan opsi *benchmark*, sehingga hasil pelatihan dapat direplikasi secara konsisten meskipun dengan konsekuensi penurunan kecepatan komputasi [31]. Pada setiap *fold*, model terbaik dipilih dan disimpan berdasarkan nilai akurasi validasi tertinggi. Seluruh model terbaik dari lima *fold* tersebut selanjutnya digunakan pada tahap pengujian dan *ensemble* untuk meningkatkan performa prediksi pada data uji [32].

## 2.5 Evaluasi Model

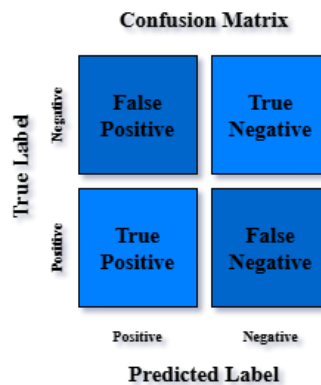
Tahap evaluasi dilakukan untuk mengukur kinerja dan kemampuan generalisasi model dalam mengklasifikasikan citra ikan cupang berdasarkan kondisi kesehatannya. Evaluasi model pada penelitian ini dilakukan melalui dua pendekatan, yaitu evaluasi individual pada setiap model hasil *5-Fold Cross Validation* serta evaluasi gabungan menggunakan metode *ensemble* pada data uji.

### 2.5.1 Evaluasi 5-Fold Cross Validation

Pada tahap ini, performa model dievaluasi pada setiap *fold* menggunakan metrik akurasi dan *Macro-F1 Score*. Penggunaan *Macro-F1 Score* bertujuan untuk memberikan penilaian yang seimbang terhadap seluruh kelas, terutama ketika distribusi kesalahan antar kelas tidak merata [33]. Nilai akurasi dan *Macro-F1* dari masing-masing *fold* kemudian dirata-ratakan untuk memperoleh gambaran performa model secara keseluruhan, disertai dengan perhitungan deviasi standar sebagai indikator stabilitas model [34].

### 2.5.2 Evaluasi Menggunakan Normalized Confusion Matrix

*Normalized Confusion Matrix* digunakan untuk menganalisis performa model klasifikasi dengan membandingkan hasil prediksi terhadap label aktual dalam bentuk persentase. Berbeda dengan *confusion matrix* konvensional yang menampilkan jumlah absolut, normalisasi memungkinkan distribusi prediksi pada setiap kelas diamati secara lebih proporsional, sehingga perbandingan performa antar kelas menjadi lebih jelas [35]. Bentuk umum *Confusion Matrix* ditunjukkan pada Gambar 3.



Gambar 3. *Confusion Matrix*

Pada Gambar 3, *Confusion Matrix* menyajikan representasi visual dari hubungan antara kelas aktual dan kelas hasil prediksi. Pada kasus klasifikasi multi-kelas, setiap baris pada *Confusion Matrix* merepresentasikan distribusi prediksi untuk satu kelas aktual. Elemen diagonal menunjukkan proporsi prediksi yang benar, sedangkan elemen di luar diagonal menggambarkan kesalahan klasifikasi antar kelas. Dalam penelitian ini, *Normalized Confusion Matrix* digunakan untuk mengidentifikasi pola kesalahan model dalam membedakan kondisi ikan cupang sehat, *dropsy*, dan *popeye*, sehingga evaluasi tidak hanya berfokus pada nilai akurasi keseluruhan, tetapi juga pada distribusi kesalahan antar kelas yang diamati secara lebih seimbang.

### 2.5.3 Evaluasi Menggunakan Metrik Klasifikasi

Selain menggunakan *Normalized Confusion Matrix*, kinerja model juga dievaluasi menggunakan metrik klasifikasi yang umum digunakan dalam analisis hasil prediksi. Metrik ini diperoleh dengan membandingkan hasil prediksi model dengan label sebenarnya pada data uji untuk setiap kelas, yaitu ikan sehat, *dropsy*, dan *popeye*. Melalui metrik ini, kemampuan model dalam mengenali masing-masing kondisi ikan dapat dinilai secara lebih terperinci, tidak hanya berdasarkan jumlah prediksi yang benar secara keseluruhan.

Evaluasi kinerja model dilakukan menggunakan metrik *precision*, *recall*, *F1-score*, dan *accuracy*, yang dirumuskan sebagai berikut.

$$Precision_k = \frac{TP_k}{TP_k + FP_k} \quad (1)$$

$$Recall_k = \frac{TP_k}{TP_k + FN_k} \quad (2)$$

$$F1 - score_k = 2 \times \frac{Precision_k \times Recall_k}{Precision_k + Recall_k} \quad (3)$$

$$Accuracy = \frac{\sum_k TP_k}{\sum_k (TP_k + FP_k + FN_k + TN_k)} \quad (4)$$

Selain itu, nilai rata-rata antar kelas dihitung menggunakan *macro average* dan *weighted average* sebagai berikut.

$$Macro Average = \frac{1}{K} \sum_{k=1}^K M_k \quad (5)$$

$$Weighted Average = \frac{\sum_{k=1}^K n_k \times M_k}{\sum_{k=1}^K n_k} \quad (6)$$

Pada rumus di atas,  $TP_k$  menyatakan jumlah data kelas  $k$  yang diprediksi dengan benar,  $FP_k$  adalah jumlah data yang diprediksi sebagai kelas  $k$  tetapi sebenarnya berasal dari kelas lain,  $FN_k$  adalah data kelas  $k$  yang tidak berhasil dikenali oleh model, dan  $TN_k$  merupakan data dari kelas lain yang juga diprediksi dengan benar.

*Precision* menunjukkan seberapa tepat prediksi model ketika menyatakan suatu citra termasuk ke dalam kelas tertentu. *Recall* menggambarkan kemampuan model dalam menemukan seluruh citra yang benar-benar termasuk dalam kelas tersebut. *F1-score* menggabungkan *precision* dan *recall* dalam satu nilai, sehingga memberikan gambaran

keseimbangan antara ketepatan dan kelengkapan prediksi. *Accuracy* menunjukkan tingkat ketepatan prediksi model secara keseluruhan terhadap seluruh data uji.

*Macro average* diperoleh dengan menghitung rata-rata nilai metrik dari setiap kelas tanpa memperhatikan jumlah data pada masing-masing kelas, sehingga setiap kelas memiliki bobot yang sama. Sementara itu, *weighted average* mempertimbangkan jumlah sampel pada tiap kelas, sehingga hasilnya lebih mencerminkan distribusi data yang sebenarnya. Penggunaan berbagai metrik ini memungkinkan evaluasi kinerja model secara menyeluruh, baik dari sisi akurasi umum maupun konsistensi model dalam mengenali setiap jenis penyakit ikan cupang.

### 2.5.3 Ensemble Model dengan Soft Voting

Untuk meningkatkan akurasi prediksi pada data uji, penelitian ini menggunakan metode *ensemble* dengan pendekatan *soft voting* [32]. Metode ini menggabungkan beberapa model terbaik yang diperoleh dari setiap *fold* pada proses *cross validation*. Setiap model menghasilkan nilai probabilitas untuk setiap kelas pada data uji, kemudian seluruh probabilitas tersebut digabungkan untuk menentukan hasil prediksi akhir. Dengan cara ini, keputusan klasifikasi tidak bergantung pada satu model saja, melainkan merupakan hasil gabungan dari beberapa model.

Secara matematis, proses *soft voting* dapat dijelaskan sebagai berikut. Misalkan terdapat  $M$  model dan  $C$  kelas. Untuk setiap data uji  $x_i$ , setiap model menghasilkan probabilitas bahwa data tersebut termasuk ke dalam kelas ke- $c$ , yang dinyatakan sebagai  $p_{m,c}(x_i)$ . Nilai probabilitas gabungan untuk setiap kelas dihitung dengan merata-ratakan probabilitas dari seluruh model.

$$\bar{p}_c(x_i) = \frac{1}{M} \sum_{m=1}^M p_{m,c}(x_i) \quad (7)$$

Kelas dengan nilai probabilitas rata-rata tertinggi kemudian dipilih sebagai hasil prediksi akhir, yang dirumuskan sebagai.

$$\hat{y}_i = \arg \max_c \bar{p}_c(x_i) \quad (8)$$

Pada rumus di atas,  $p_{m,c}(x_i)$  menunjukkan tingkat keyakinan model ke- $m$  bahwa data  $x_i$  termasuk kedalam kelas ke- $c$ . Nilai  $\bar{p}_c(x_i)$  merupakan rata-rata tingkat keyakinan dari seluruh model untuk kelas tersebut. Prediksi akhir  $\hat{y}_i$  diperoleh dengan memilih kelas yang memiliki nilai probabilitas rata-rata paling tinggi.

Pendekatan ini membantu mengurangi kesalahan yang mungkin terjadi apabila hanya menggunakan satu model saja. Dengan menggabungkan hasil dari beberapa model yang dilatih pada data yang berbeda, metode *ensemble* dengan *soft voting* mampu menghasilkan prediksi yang lebih stabil dan lebih andal pada data uji [32].

### 2.5.4 Visualisasi Grad-CAM

Sebagai upaya untuk memahami cara kerja model dalam mengambil keputusan, penelitian ini menggunakan teknik *Gradient-weighted Class Activation Mapping* (Grad-CAM) sebagai metode interpretabilitas. Grad-CAM digunakan untuk memvisualisasikan bagian citra yang paling berpengaruh terhadap hasil prediksi model. Melalui visualisasi ini, dapat diamati area-area pada tubuh ikan yang menjadi perhatian utama model saat mengklasifikasikan kondisi ikan, sehingga proses prediksi tidak bersifat sepenuhnya sebagai *black box* [16].

Secara umum, Grad-CAM dihitung berdasarkan hubungan antara gradien keluaran model dan peta fitur (*feature map*) pada lapisan konvolusional terakhir. Misalkan  $A^k$  menyatakan peta fitur ke- $k$  pada lapisan konvolusional terakhir, dan  $y^c$  adalah skor keluaran model untuk kelas ke- $c$ . Bobot kepentingan untuk setiap peta fitur dihitung sebagai rata-rata gradien spasial berikut.

$$\alpha_k^c = \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (9)$$

Peta aktivasi Grad-CAM untuk kelas ke- $c$  kemudian diperoleh dengan mengombinasikan bobot tersebut dengan peta fitur, yang dirumuskan sebagai.

$$L_{Grad-CAM}^c = ReLU(\sum_k \alpha_k^c A^k) \quad (10)$$

Pada rumus di atas,  $\frac{\partial y^c}{\partial A_{ij}^k}$  menunjukkan seberapa besar pengaruh setiap piksel pada peta fitur  $A^k$  terhadap skor prediksi kelas ke- $c$ . Nilai  $\alpha_k^c$  merepresentasikan tingkat kepentingan peta fitur ke- $k$  dalam menentukan prediksi kelas tersebut. Peta Grad-CAM  $L_{Grad-CAM}^c$  kemudian diperoleh dengan menjumlahkan seluruh peta fitur yang telah diberi bobot dan menerapkan fungsi *ReLU* agar hanya kontribusi positif yang ditampilkan.

Hasil Grad-CAM divisualisasikan dalam bentuk *heatmap* yang di-*overlay* pada citra asli. Area dengan warna yang lebih intens menunjukkan bagian citra yang lebih berpengaruh terhadap keputusan model. Melalui visualisasi ini, peneliti dapat mengamati apakah model cenderung memberikan perhatian pada area tubuh ikan yang secara visual berkaitan dengan kondisi penyakit tertentu atau tidak. Dengan demikian, Grad-CAM membantu meningkatkan kepercayaan terhadap hasil prediksi model sekaligus memberikan gambaran intuitif mengenai dasar pengambilan keputusan model.

### 3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi metode yang telah dijelaskan pada bagian metodologi penelitian, sekaligus membahas temuan-temuan yang diperoleh. Pembahasan difokuskan pada evaluasi kinerja model dalam mendeteksi penyakit *dropsy* dan *popeye* pada ikan cupang berbasis citra menggunakan arsitektur MobileNetV2. Hasil yang ditampilkan tidak hanya berupa nilai numerik, tetapi juga didukung oleh visualisasi dan tabel, untuk memberikan pemahaman yang lebih komprehensif terhadap performa model.

#### 3.1 Hasil Pengumpulan Data

Hasil pengumpulan data pada penelitian ini menghasilkan sebuah dataset citra ikan cupang yang terdiri dari 600 citra yang terbagi ke dalam tiga kelas, yaitu ikan cupang sehat, *dropsy*, dan *popeye*. Masing-masing kelas memiliki jumlah citra yang sama, yaitu 200 citra per kelas, sehingga distribusi data bersifat seimbang. Kondisi ini penting untuk menghindari bias model terhadap kelas tertentu selama proses pelatihan dan evaluasi.

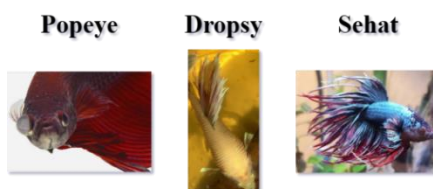
Citra yang dikumpulkan dipastikan memiliki variasi yang cukup beragam dari segi warna tubuh ikan, bentuk sirip, pencahayaan, serta sudut pengambilan gambar. Pada kelas *dropsy*, citra umumnya menampilkan gejala seperti perut membesar dan sisik yang terlihat mengembang, sedangkan pada kelas *popeye* tampak pembengkakan pada area mata ikan. Sementara itu, citra ikan sehat menunjukkan bentuk tubuh proporsional tanpa indikasi visual penyakit. Keberagaman karakteristik visual ini menjadi tantangan tersendiri bagi model dalam membedakan kelas, sekaligus mencerminkan kondisi nyata yang mungkin ditemui oleh pengguna di lapangan.

Dataset yang telah terkumpul kemudian dipisahkan secara manual menjadi data latih/validasi dan data uji untuk memastikan bahwa citra ikan yang sama tidak muncul pada kedua subset tersebut. Pendekatan ini dilakukan untuk mencegah terjadinya *data leakage* yang dapat menyebabkan nilai evaluasi menjadi tidak realistis. Pembagian dataset dilakukan dengan proporsi 85% untuk data latih dan validasi, serta 15% untuk data uji. Rincian distribusi dataset pada setiap kelas dan subset data disajikan pada Tabel 2.

**Tabel 2.** Jumlah dan Distribusi Dataset

Kelas	Jumlah Citra	Latih/Validasi (85%)	Uji (15%)
Sehat	200	170	30
<i>Dropsy</i>	200	170	30
<i>Popeye</i>	200	170	30
Total	600	510	90

Sebagai ilustrasi, beberapa contoh citra dari masing-masing kelas ditampilkan pada Gambar 4 untuk menunjukkan perbedaan visual antar kelas yang menjadi dasar proses klasifikasi.



**Gambar 4.** Contoh Dataset

Berdasarkan Tabel 2, setiap kelas memiliki jumlah citra yang sama dan dibagi dengan proporsi yang konsisten antara data latih/validasi dan data uji. Dengan total 510 citra digunakan untuk pelatihan dan validasi, model memiliki cukup data untuk mempelajari pola visual dari masing-masing kelas. Sementara itu, 90 citra data uji digunakan untuk mengukur performa model terhadap data yang benar-benar baru dan belum pernah dilihat sebelumnya.

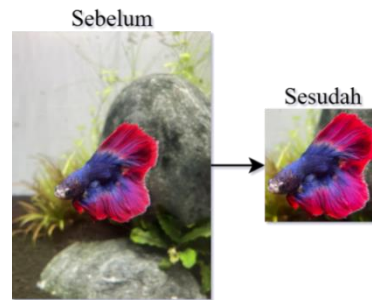
Gambar 4 memperlihatkan contoh citra dari kelas sehat, *dropsy*, dan *popeye*. Perbedaan visual yang ditampilkan pada gambar tersebut menjadi dasar bagi model dalam mempelajari ciri khas setiap kelas. Melalui kombinasi distribusi data yang seimbang dan variasi citra yang beragam, dataset ini diharapkan mampu mendukung proses pelatihan model yang adil, stabil, dan representatif terhadap kondisi nyata ikan cupang di lapangan.

#### 3.2 Hasil Preprocessing Data

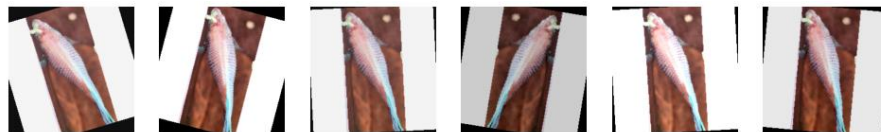
Tahap *preprocessing* data bertujuan untuk menyiapkan citra ikan cupang agar memiliki ukuran dan tampilan yang seragam sebelum digunakan dalam proses pelatihan model yang diperlihatkan pada gambar 5. Pada tahap awal, setiap citra disesuaikan secara manual agar memiliki rasio 1:1. Penyesuaian ini dilakukan tanpa memotong bagian penting dari tubuh ikan, sehingga bentuk dan ciri utama ikan tetap terjaga. Untuk menjaga proporsi objek, area kosong di sekitar citra diisi dengan tepi transparan. Pendekatan ini membantu mencegah perubahan bentuk ikan yang tidak diinginkan akibat proses penyesuaian ukuran.

Setelah itu, dilakukan proses augmentasi data untuk memperkaya variasi citra yang diperlihatkan pada gambar 6. Beberapa transformasi sederhana, seperti membalik gambar secara horizontal, memutar gambar secara acak, dan menyesuaikan tingkat warna serta pencahayaan, diterapkan pada citra. Proses ini bertujuan untuk membantu model

mengenali ikan cupang dalam berbagai kondisi tampilan yang mungkin ditemui pada data nyata, tanpa mengubah ciri utama dari masing-masing kelas penyakit.



**Gambar 5.** Contoh Citra Sebelum dan Sesudah Proses *Preprocessing* Manual



**Gambar 6.** Contoh Citra Setelah Augmentasi

Gambar 5 memperlihatkan hasil *preprocessing* manual, di mana citra ikan cupang telah disesuaikan ke dalam rasio 1:1 dengan penambahan tepi transparan di sekeliling objek. Langkah ini menghasilkan komposisi citra yang lebih konsisten tanpa menghilangkan informasi visual penting pada tubuh ikan.

Gambar 6 menunjukkan contoh citra sesudah proses augmentasi. Terlihat bahwa meskipun terjadi perubahan orientasi, sudut pandang, dan pencahayaan, ciri utama penyakit pada ikan tetap dapat dikenali. Hal ini menunjukkan bahwa augmentasi berhasil menambah variasi visual citra tanpa mengubah karakteristik penting yang menjadi dasar proses klasifikasi.

Setelah tahap augmentasi, seluruh citra diubah ukurannya menjadi  $224 \times 224$  piksel sesuai dengan kebutuhan arsitektur MobileNetV2 dan kemudian dinormalisasi agar nilai pikselnya berada pada rentang yang sesuai dengan model *pretrained*. Secara keseluruhan, tahapan *preprocessing* ini menghasilkan dataset yang seragam dari segi ukuran dan format, namun tetap memiliki variasi visual yang cukup untuk mendukung proses pembelajaran model dalam mengenali kondisi kesehatan ikan cupang secara lebih akurat.

### 3.3 Persiapan Model untuk Training

Setelah dataset melalui tahap *preprocessing*, langkah selanjutnya adalah mempersiapkan model agar siap digunakan dalam proses pelatihan. Pada penelitian ini, arsitektur MobileNetV2 yang telah dilatih sebelumnya (*pretrained*) digunakan sebagai dasar model klasifikasi. Penggunaan model *pretrained* memungkinkan pemanfaatan fitur visual umum yang telah dipelajari dari dataset skala besar, sehingga proses pelatihan dapat dilakukan secara lebih efisien meskipun jumlah dataset penelitian terbatas.

Pada tahap persiapan, lapisan *feature extractor* MobileNetV2 sebagian besar dibekukan (*freeze*) untuk mempertahankan bobot awal yang telah mampu mengekstraksi fitur dasar citra, seperti tepi, tekstur, dan pola bentuk. Namun demikian, dua blok terakhir dari arsitektur MobileNetV2 dibuka (*unfreeze*) untuk dilakukan *fine-tuning*. Strategi ini bertujuan agar model dapat menyesuaikan fitur tingkat tinggi dengan karakteristik spesifik citra ikan cupang dan gejala penyakit yang diteliti, tanpa menyebabkan *overfitting* yang berlebihan.

Lapisan klasifikasi akhir pada MobileNetV2 kemudian disesuaikan dengan jumlah kelas pada penelitian ini, yaitu tiga kelas. Lapisan ini berfungsi untuk menghasilkan probabilitas prediksi masing-masing kelas menggunakan fungsi aktivasi *softmax*. Konfigurasi ini memastikan bahwa model mampu memberikan keluaran berupa probabilitas yang dapat dimanfaatkan lebih lanjut pada tahap *ensemble*.

Selain konfigurasi arsitektur, pengaturan lingkungan pelatihan juga dilakukan untuk memastikan konsistensi hasil eksperimen. Penetapan nilai *seed* yang sama serta penggunaan mode deterministik pada pustaka *CUDA* bertujuan untuk menjaga agar proses pelatihan dapat direproduksi dengan hasil yang stabil pada setiap eksekusi.

### 3.4 Hasil Training dan Evaluasi Model

Proses pelatihan model dilakukan menggunakan pendekatan *5-Fold Cross Validation* untuk mengevaluasi performa model secara menyeluruh dan stabil. Pada setiap *fold*, model dilatih menggunakan data latih dan divalidasi menggunakan data validasi yang berbeda, sehingga setiap subset data memiliki kesempatan yang sama untuk digunakan sebagai data validasi. Pendekatan ini bertujuan untuk mengurangi ketergantungan hasil terhadap satu pembagian data tertentu dan memberikan estimasi performa yang lebih reliabel.

Hasil pelatihan menunjukkan bahwa nilai akurasi dan *Macro-F1 Score* bervariasi antar *fold*. Variasi ini mencerminkan perbedaan kompleksitas data pada masing-masing subset, namun secara umum model mampu

menunjukkan performa yang konsisten. Rincian nilai akurasi dan *Macro-F1 Score* pada setiap *fold* disajikan pada Tabel 3.

**Tabel 3.** Hasil Akurasi dan *Macro-F1 Score* pada *5-Fold Cross Validation*

<i>Fold</i>	Akurasi (%)	<i>Macro-F1</i> (%)
1	76,47	76,24
2	73,53	72,58
3	76,47	76,49
4	66,47	66,79
5	80,39	80,03
Rata-Rata	74,71	74,43
Deviasi Standar	4,57	4,49

Berdasarkan hasil tersebut, diperoleh rata-rata akurasi sebesar 74,71% dengan deviasi standar  $\pm 4,57\%$  serta rata-rata *Macro-F1 Score* sebesar 74,43% dengan deviasi standar  $\pm 4,49\%$ . Nilai deviasi standar ini menunjukkan bahwa performa model relatif stabil pada berbagai pembagian data selama proses *5-Fold Cross Validation*. Variasi performa yang masih berada dalam rentang tersebut mengindikasikan bahwa model tidak terlalu sensitif terhadap perbedaan subset data latih dan validasi, sehingga hasil evaluasi tidak bergantung pada satu pembagian data tertentu. Kondisi ini menunjukkan bahwa proses pelatihan mampu menangkap pola umum dari data tanpa mengalami fluktuasi performa yang ekstrem antar *fold*, serta tidak menunjukkan indikasi *overfitting* yang berlebihan meskipun ukuran dataset relatif terbatas.

Evaluasi lebih lanjut dilakukan pada data uji menggunakan model terbaik dari masing-masing *fold*. Untuk meningkatkan performa klasifikasi, diterapkan metode *ensemble* dengan pendekatan *soft voting*, di mana probabilitas prediksi dari seluruh model dirata-ratakan untuk menghasilkan keputusan akhir. Pendekatan *ensemble* ini menunjukkan peningkatan kinerja klasifikasi pada data uji dibandingkan penggunaan model tunggal, yang ditunjukkan oleh peningkatan nilai metrik evaluasi pada tahap pengujian.

### 3.5 Pengujian

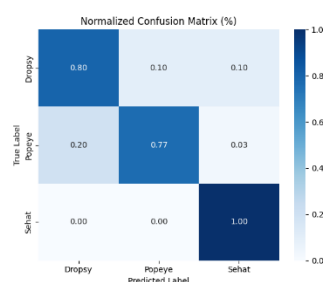
Tahap pengujian dilakukan menggunakan data uji yang sepenuhnya terpisah dari data pelatihan dan validasi. Data uji ini terdiri dari 90 citra ikan cupang yang dibagi secara seimbang ke dalam tiga kelas, yaitu ikan sehat, *dropsy*, dan *popeye*, dengan masing-masing kelas berjumlah 30 citra. Pengujian ini bertujuan untuk menilai sejauh mana model mampu mengenali dan mengklasifikasikan citra baru yang belum pernah digunakan selama proses pelatihan.

Pada tahap ini digunakan pendekatan *ensemble* dengan metode *soft voting*. Dalam pendekatan ini, model terbaik yang dihasilkan dari setiap *fold* memberikan nilai probabilitas untuk masing-masing kelas pada setiap citra uji. Nilai probabilitas tersebut kemudian dirata-ratakan untuk menentukan hasil prediksi akhir. Pendekatan ini membantu mengurangi kesalahan prediksi yang mungkin muncul jika hanya mengandalkan satu model, sehingga hasil klasifikasi menjadi lebih stabil dan andal.

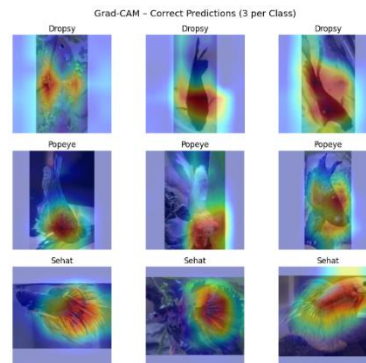
Hasil pengujian model *ensemble* menunjukkan tingkat akurasi sebesar 85,56%, yang mencerminkan peningkatan performa dibandingkan dengan rata-rata hasil pada tahap *cross validation*. Rincian kinerja model untuk setiap kelas ditampilkan pada Tabel 4 dalam bentuk nilai *precision*, *recall*, dan *F1-score*. Selain itu, hasil evaluasi juga disajikan dalam bentuk *Normalized Confusion Matrix* pada gambar 7 dan visualisasi Grad-CAM untuk membantu memahami pola prediksi model secara lebih mendalam pada gambar 8.

**Tabel 4.** Hasil Uji

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
<i>Dropsy</i>	0,80	0,80	0,80	30
<i>Popeye</i>	0,88	0,77	0,82	30
Sehat	0,88	1,00	0,94	30
<i>Accuracy</i>			0,86	90
<i>Macro Average</i>	0,86	0,86	0,85	90
<i>Weighted Average</i>	0,86	0,86	0,85	90



**Gambar 7.** *Normalized Confusion Matrix*



**Gambar 8.** Visualisasi Grad-CAM

Berdasarkan Tabel 4, kelas ikan sehat menunjukkan performa terbaik dengan nilai *recall* mencapai 100%, yang berarti seluruh citra ikan sehat berhasil dikenali dengan benar oleh model. Pada kelas *dropsy* dan *popeye*, masih terdapat sejumlah kesalahan klasifikasi. Hal ini terlihat pada *Normalized Confusion Matrix* pada gambar 7, di mana sebagian citra dari kedua kelas tersebut masih saling tertukar.

Meskipun demikian, hasil pengujian menunjukkan bahwa performa model relatif cukup seimbang pada ketiga kelas. Model tidak hanya mampu dalam mengenali satu kelas tertentu, tetapi juga mampu memberikan hasil klasifikasi yang cukup konsisten pada setiap kondisi. Visualisasi Grad-CAM pada Gambar 8 menunjukkan area citra yang menjadi fokus perhatian model saat melakukan prediksi, sehingga memberikan gambaran bagian tubuh ikan yang berkontribusi terhadap keputusan klasifikasi. Pada kelas *dropsy*, area dengan aktivasi tertinggi umumnya terpusat pada bagian perut dan sisik ikan. Pola ini sesuai dengan karakteristik biologis penyakit *dropsy* yang ditandai oleh pembengkakan perut dan sisik yang mengembang [3], sehingga perubahan visual pada area tersebut menjadi indikator utama bagi model. Sementara itu, pada kelas *popeye*, aktivasi Grad-CAM cenderung berfokus pada area mata. Hal ini sejalan dengan gejala khas *popeye* berupa mata yang menonjol atau membesar [3], sehingga perhatian model terhadap area tersebut menunjukkan bahwa proses prediksi didasarkan pada ciri visual yang relevan secara biologis. Temuan ini mengindikasikan bahwa model tidak hanya memanfaatkan informasi global citra, tetapi juga menangkap pola visual yang berkaitan langsung dengan gejala penyakit ikan cupang. Secara keseluruhan, hasil ini menunjukkan bahwa kombinasi *transfer learning* menggunakan MobileNetV2 dan pendekatan *ensemble* efektif untuk mendukung diagnosis penyakit ikan cupang berbasis citra.

### 3.6 Pembahasan

Berdasarkan hasil pengujian dan evaluasi yang telah dipaparkan pada sub-bab sebelumnya, dapat disimpulkan bahwa penerapan arsitektur MobileNetV2 berbasis citra mampu memberikan performa yang cukup baik dalam mendeteksi penyakit *dropsy* dan *popeye* pada ikan cupang yang telah menunjukkan gejala visual. Peningkatan akurasi dari rata-rata *5-Fold Cross Validation* sebesar 74,71% menjadi 85,56% pada pengujian data uji melalui metode *ensemble* menunjukkan bahwa penggabungan beberapa model hasil pelatihan mampu meningkatkan kemampuan generalisasi dan mengurangi variansi prediksi.

Performa klasifikasi yang relatif seimbang pada ketiga kelas, khususnya nilai *Macro-F1 Score* yang mendekati nilai akurasi, mengindikasikan bahwa model tidak mengalami bias terhadap kelas tertentu. Kelas ikan sehat menunjukkan performa terbaik, yang dapat disebabkan oleh karakteristik visual yang lebih konsisten dibandingkan kelas penyakit. Sementara itu, kesalahan klasifikasi yang masih terjadi antara kelas *dropsy* dan *popeye* umumnya disebabkan oleh kemiripan visual pada beberapa citra, seperti pembengkakan tubuh yang tidak terlalu ekstrem atau pembesaran mata yang kurang jelas akibat sudut pengambilan gambar dan pencahayaan.

Jika dibandingkan dengan penelitian-penelitian sebelumnya yang menerapkan metode *Convolutional Neural Network* untuk klasifikasi penyakit ikan atau organisme akuatik, penelitian ini menunjukkan beberapa perbedaan dan keunggulan. Pertama, penelitian ini secara spesifik berfokus pada ikan cupang dengan dua jenis penyakit yang umum ditemui oleh penghobi, namun masih jarang dibahas dalam literatur. Kedua, pendekatan *transfer learning* dengan MobileNetV2 dipadukan dengan *5-Fold Cross Validation* dan metode *ensemble*, yang belum banyak diterapkan secara bersamaan pada penelitian serupa. Pendekatan ini terbukti mampu meningkatkan stabilitas dan performa prediksi model.

Selain itu, penggunaan visualisasi Grad-CAM memberikan nilai tambah dari sisi interpretabilitas model. Hasil visualisasi menunjukkan bahwa model cenderung memfokuskan perhatian pada area tubuh ikan yang relevan dengan gejala penyakit, seperti bagian perut pada kasus *dropsy* dan area mata pada kasus *popeye*. Hal ini memperkuat keyakinan bahwa model tidak hanya melakukan klasifikasi berdasarkan pola acak, tetapi benar-benar mempelajari fitur visual yang berkaitan dengan kondisi penyakit.

Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan. Dataset yang digunakan bersumber dari internet sehingga variasi kualitas citra dan kondisi pengambilan gambar tidak sepenuhnya terkontrol. Selain itu, model hanya dilatih untuk mengenali penyakit yang telah menunjukkan gejala visual yang jelas, sehingga belum dapat

digunakan untuk deteksi dini. Oleh karena itu, penelitian selanjutnya dapat diarahkan pada pengumpulan dataset primer, penambahan jenis penyakit lain, serta eksplorasi arsitektur model atau teknik optimasi yang lebih lanjut.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa arsitektur MobileNetV2 berbasis citra berpotensi digunakan sebagai alat bantu diagnosis bagi pemula dalam mengenali penyakit ikan cupang.

#### 4. KESIMPULAN

Kesimpulan penelitian ini menunjukkan bahwa arsitektur MobileNetV2 berbasis *transfer learning* dapat dimanfaatkan sebagai sistem pendukung diagnosis untuk mengklasifikasikan penyakit *dropsy* dan *popeye* pada ikan cupang berdasarkan citra yang menampilkan gejala visual. Pendekatan *fine-tuning* terbatas pada lapisan akhir jaringan terbukti efektif dalam menyesuaikan model dengan dataset berukuran relatif kecil. Evaluasi menggunakan skema *5-Fold Cross Validation* menunjukkan performa model yang cukup stabil terhadap variasi pembagian data. Selain itu, penerapan metode *ensemble soft voting* mampu meningkatkan performa klasifikasi dibandingkan model tunggal, yang mengindikasikan peningkatan kemampuan generalisasi model. Visualisasi Grad-CAM juga menunjukkan bahwa model memfokuskan perhatian pada area tubuh ikan yang relevan dengan gejala penyakit, sehingga mendukung interpretabilitas hasil klasifikasi. Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Dataset citra yang diperoleh dari internet memiliki variasi kualitas, pencahayaan, sudut pengambilan, dan latar belakang yang tidak sepenuhnya terkontrol. Variasi ini meningkatkan representativitas kondisi nyata, namun menyebabkan dataset kurang terstandarisasi. Jumlah data latih yang terbatas serta fokus sistem pada penyakit dengan gejala visual yang jelas juga membatasi ruang lingkup penerapan model. Penelitian selanjutnya dapat difokuskan pada pengumpulan dataset primer yang lebih terstandarisasi, penambahan variasi penyakit, serta eksplorasi arsitektur model, teknik augmentasi data, dan strategi *fine-tuning* yang lebih optimal untuk meningkatkan akurasi dan reliabilitas sistem diagnosis berbasis citra.

#### REFERENCES

- [1] M. Duman, I. B. Saticioğlu, and J. M. Janda, "A Review of the Industrial Importance, Common Bacterial Diseases, and Zoonotic Risks of Freshwater Aquarium Fish," *VECTOR-BORNE ZOONOTIC Dis.*, vol. 24, no. 2, pp. 69–85, 2024, doi: 10.1089/vbz.2023.0094.
- [2] N. M. Mahmud, A. A. Ansary, F. Y. Ritu, N. A. Hasan, and M. M. Haque, "An Overview of Fish Disease Diagnosis and Treatment in Aquaculture in Bangladesh," *Aquac. J.*, vol. 5, no. 18, pp. 1–29, 2025, doi: 10.3390/aquacj5040018.
- [3] S. A. Smith, *Fish Diseases and Medicine*, 1st ed. Boca Raton: CRC Press, 2019. [Online]. Available: <https://www.routledge.com/Fish-Diseases-and-Medicine/Smith/p/book/9781498727860>
- [4] S. Suhendar, A. Purnama, and E. Fauzi, "Deteksi Penyakit Pada Daun Tanaman Ubi Jalar Menggunakan Metode Convolutional Neural Network," *J. Ilm. Inform. Glob.*, vol. 14, no. 3, pp. 62–67, 2023, doi: 10.36982/jiig.v14i3.3478.
- [5] A. A. S. Nur, L. Iwan, and P. Rio, "Implementasi CNN untuk Identifikasi Penyakit Daun Cabai," *J. Comput. Sci. Inf. Technol.*, vol. 6, no. 3, pp. 619–624, 2025, doi: 10.37859/coscitech.v6i3.9381.
- [6] Z. L. R. Nafi' and M. S. Hidayatullah, "Deteksi Kesehatan Ikan Nila Menggunakan Convolutional Neural Networks Berbasis Citra Digital," *J. Data Anal. Information, Comput. Sci.*, vol. 2, no. 2, pp. 192–198, 2025, doi: 10.70248/jdaics.v2i2.914.
- [7] A. Mustopa, A. Sasongko, H. M. Nawawi, S. K. Wildah, and S. Agustiani, "Deteksi Penyakit Ayam berdasarkan Citra Feses dengan Model EfficientNetV2L," *Sist. J. Sist. Inf.*, vol. 12, no. 3, pp. 715–725, 2023, doi: 10.32520/stmsi.v12i3.2807.
- [8] M. Lv *et al.*, "Si-CA MobileNet: A lightweight and efficient convolutional neural network for distracted driver detection," *Neurocomputing*, vol. 654, 2025, doi: 10.1016/j.neucom.2025.131281.
- [9] S. M. Roy *et al.*, "Application of artificial intelligence in aquaculture – Recent developments and prospects," *Aquac. Eng.*, vol. 111, 2025, doi: 10.1016/j.aquaeng.2025.102570.
- [10] H. Tamut, R. Ghosh, K. Gosh, and M. A. S. Siddique, "Enhancing Disease Detection in the Aquaculture Sector Using Convolutional Neural Networks Analysis," *Aquac. J.*, vol. 5, no. 6, pp. 1–19, 2025, doi: 10.3390/aquacj5010006.
- [11] S.-B. Hwang, H.-Y. Kim, C.-Y. Heo, H.-Y. Jeong, S.-J. Jung, and Y.-J. Cho, "Flatfish Lesion Detection Based On Part Segmentation Approach And Lesion Image Generation," *J. World Aquac. Soc.*, pp. 1–25, 2025, doi: 10.1111/jwas.70031Digital Object Identifier (DOI).
- [12] J. H. Peh and M. N. Azra, "A global review of ornamental fish and shellfish research," *Aquac. J.*, vol. 596, no. 1, 2025, doi: 10.1016/j.aquaculture.2024.741719.
- [13] D. B. Daş, "An Intelligent and Lightweight Approach Based on MobilenetV2 Architecture for Identifying Brain Tumors," *Sak. Univ. J. Comput. Inf. Sci.*, vol. 8, no. 2, pp. 392–399, 2025, doi: 10.35377/saucis...
- [14] M. G. Somoal and A. R. Dzikrillah, "Komparasi Mobilenetv2 Dengan Kustomisasi Transfer Learning Dan Hyperparameter Untuk Identifikasi Tumor Otak," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 1, pp. 229–240, 2025, doi: 10.25126/jtiik.2025129582.
- [15] D. Rastogi *et al.*, "Brain Tumor Detection and Prediction in MRI Images Utilizing a Fine-Tuned Transfer Learning Model Integrated Within Deep Learning Frameworks," *Life*, vol. 15, no. 327, pp. 1–37, 2025, doi: 10.3390/life15030327.
- [16] S. Murugesan, J. Chinnadurai, S. Srinivasan, S. K. Mathivanan, R. R. Chandan, and U. Moorthy, "Robust multiclass classification of crop leaf diseases using hybrid deep learning and Grad-CAM interpretability," *Sci. Rep.*, vol. 15, pp. 1–22, 2025, doi: 10.1038/s41598-025-14847-7.
- [17] D. Tribuana and H. Zainuddin, "Image Preprocessing Approaches Toward Better Learning Performance with CNN," *RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 8, no. 1, pp. 1–9, 2024, doi: 10.29207/resti.v8i1.5417.
- [18] K. R. Ummah, T. Karlita, R. Sigit, E. M. Yuniarno, I. K. E. Purnama, and M. H. Purnomo, "Effect Of Image Pre-Processing Method On Convolutional Neural Network Classification Of Covid-19 Ct Scan Images," *Int. J. Innov. Comput. Inf. Control*, vol. 18, no. 6, pp. 1895–1912, 2022, doi: 10.24507/ijicic.18.06.1895.



- [19] W. Zeng, “Image data augmentation techniques based on deep learning: A survey,” *Math. Biosci. Eng.*, vol. 21, no. 6, pp. 6190–6224, 2024, doi: 10.3934/mbe.2024272.
- [20] R. B. Pradana, C. Paramita, N. A. S. Winarsih, and R. A. Premunendar, “Lightweight Deep Learning Approach for Sugarcane Leaf Disease Classification Using MobileNetV2,” *J. Teknol. Inf. Dan Terap.*, vol. 12, no. 2, pp. 115–126, 2025, doi: 10/25047/jtit.v12i2.456.
- [21] M. M. Islam, M. B. Hossain, M. N. Akhtar, M. A. Moni, and K. F. Hasan, “CNN Based on Transfer Learning Models Using Data Augmentation and Transformation for Detection of Concrete Crack,” *Algorithms*, vol. 15, no. 287, pp. 1–17, 2022, doi: 10.3390/a15080287.
- [22] J. Chaki, *The Art of Deep Learning Image Augmentation: The Seeds of Success*. Singapore: Springer, 2025. doi: 10.1007/978-981-96-5081-1.
- [23] D. Murcia-gómez, I. Rojas-valenzuela, and O. Valenzuela, “Impact of Image Preprocessing Methods and Deep Learning Models for Classifying Histopathological Breast Cancer Images,” *Appl. Sci.*, vol. 12, no. 11375, pp. 1–18, 2022, doi: 10.3390/app122211375.
- [24] PyTorch, “Transforming images, videos, boxes and more.” Accessed: Feb. 02, 2026. [Online]. Available: <https://docs.pytorch.org/vision/main/transforms.html>
- [25] G. Habib, I. A. Malik, S. Sharma, S. Singh, and J. Kim, “Optimizing MobileNetV3 for multimodal eye gaze and emotion recognition via advanced pruning and quantisation techniques,” *Sci. Rep.*, vol. 15, no. 37766, pp. 1–18, 2025, doi: 10.1038/s41598-025-19617-z.
- [26] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” *Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2019, doi: 10.48550/arXiv.1801.04381.
- [27] A. Howard *et al.*, “Searching for MobileNetV3,” *Comput. Vis. Pattern Recognit.*, pp. 1314–1324, 2019, doi: 10.48550/arXiv.1905.02244.
- [28] M. S. A. M. Al-Gaashani, W. Xu, and E. Y. Obsie, “MobileNetV2-based deep learning architecture with progressive transfer learning for accurate monkeypox detection,” *Appl. Soft Comput.*, vol. 169, 2025, doi: 10.1016/j.asoc.2024.112553.
- [29] M. Kamal, M. M. Haque, R. N. Farabi, and M. Ibrahim, “From Scratch to Fine Tuning: Comparing Transfer Learning and CNN Training Strategies on Five Bangladesh-Centric Datasets,” *Res. Sq.*, pp. 1–16, 2026, doi: 10.21203/rs.3.rs-8546096/v1.
- [30] S. Rattanaphan and A. Briassouli, “Evaluating Generalization, Bias, and Fairness in Deep Learning for Metal Surface Defect Detection: A Comparative Study,” *Processes*, vol. 12, no. 3, pp. 1–32, 2024, doi: 10.3390/pr12030456.
- [31] A. E. Maxwell, M. S. Bester, and C. A. Ramezan, “Enhancing Reproducibility and Replicability in Remote Sensing Deep Learning Research and Practice,” *Remote Sens.*, vol. 14, no. 22, pp. 1–12, 2022, doi: 10.3390/rs14225760.
- [32] A. Mohammed and R. Kora, “A comprehensive review on ensemble deep learning: Opportunities and challenges,” *J. King Saud Univ. – Comput. Inf. Sci.*, vol. 35, no. 2, pp. 757–774, 2023, doi: 10.1016/j.jksuci.2023.01.014.
- [33] M. C. H. Lee and J. B. and J. Springael, “Performance Metrics for Multilabel Emotion Classification: Comparing Micro, Macro, and Weighted F1-Scores,” *Appl. Sci.*, vol. 14, no. 21, pp. 1–21, 2024, doi: 10.3390/app14219863.
- [34] I. Said Almuniri *et al.*, “Beyond peak accuracy: a stabilitycentric framework for reliable multimodal student engagement assessment,” *Sci. Rep.*, vol. 16, no. 5, pp. 1–21, 2026, doi: 10.1038/s41598-025-31215-7.
- [35] J. Erbani, P.-E. Portier, E. Egyed-Zsigmond, S. Ben Mokhtar, and D. Nurbakova, “On the Normalization of Confusion Matrices: Methods and Geometric Interpretations,” 2025, doi: 10.48550/arXiv.2509.04959.