

PENINGKATAN ROBUSTNESS MODEL RESNET50 UNTUK KLASIFIKASI UBUR-UBUR MENGGUNAKAN AUGMENTASI DATA CUTMIX DAN MIXUP

Satrio Wicaksono¹; Dian Ade Kurnia²; Yudhistira Arie Wijaya³; Bani Nurhakim⁴; Denni Pratama⁵.

Program Studi Teknik Informatika¹
Program Studi Manajemen Informatika²
Program Studi Sistem Informasi³
Program Studi Komputerisasi Akuntansi^{4,5}

STMIK IKMI Cirebon
<https://ikmi.ac.id/page/18/?lang=de>
satriojagatprakasa@gmail.com

(*) Corresponding Author : satriojagatprakasa@gmail.com
Published : 31 Desember 2025

Abstract—This study aims to enhance the robustness of the ResNet50 model for jellyfish classification by applying MixUp and CutMix augmentation techniques to underwater images affected by visual degradation such as color cast, noise, scattering, and low contrast. The research employs transfer learning with fine-tuning on the final 30 layers of the model and compares two augmentation schemes, namely basic augmentation and the MixUp–CutMix combination, supported by early stopping to ensure stable training. The findings show that the baseline model experienced overfitting and achieved only 41.11% validation accuracy, whereas the integration of MixUp and CutMix increased accuracy to 47.22% and reduced validation loss from 1.4373 to 1.3127. These improvements indicate that mixed-sample augmentation enriches synthetic data variation and strengthens the feature representations learned by the model, thereby improving resilience to noise and distribution shifts in underwater images. This study concludes that MixUp and CutMix are effective approaches for enhancing the performance of classification models on limited and diverse datasets, and suggests that advanced augmentation techniques can serve as practical solutions to improve model generalization without requiring large-scale data acquisition.

Keyword : ResNet50; classification; jellyfish; MixUp; CutMix

Abstrak—Penelitian ini bertujuan meningkatkan robustness model ResNet50 untuk klasifikasi ubur-ubur melalui penerapan teknik augmentasi MixUp dan CutMix pada kondisi citra bawah laut yang mengalami degradasi visual seperti color cast, noise, scattering, dan low contrast. Penelitian menggunakan pendekatan transfer learning dengan fine-tuning pada 30 lapisan akhir model serta membandingkan dua skema augmentasi, yaitu augmentasi dasar dan kombinasi MixUp–CutMix yang didukung mekanisme early stopping untuk menjaga stabilitas pelatihan. Hasil penelitian menunjukkan bahwa model baseline mengalami overfitting dan hanya mencapai akurasi validasi 41,11%, sedangkan integrasi MixUp dan CutMix meningkatkan akurasi menjadi 47,22% dan menurunkan validation loss dari 1,4373 menjadi 1,3127. Peningkatan ini mengindikasikan bahwa strategi mixed-sample augmentation memperkaya variasi data sintesis dan memperkuat representasi fitur yang dipelajari model, sehingga meningkatkan ketahanan terhadap noise dan pergeseran distribusi citra bawah laut. Penelitian ini menyimpulkan bahwa penggunaan MixUp dan CutMix merupakan pendekatan efektif untuk meningkatkan performa model klasifikasi pada dataset yang terbatas dan bervariasi, serta memberikan implikasi bahwa augmentasi tingkat lanjut dapat diterapkan sebagai solusi praktis untuk memperbaiki generalisasi model tanpa memerlukan penambahan data berskala besar.

Kata Kunci : ResNet50; klasifikasi; ubur-ubur; MixUp; CutMix

INTRODUCTION

Perkembangan teknologi visi komputer dan deep learning semakin mendorong pemanfaatan citra digital untuk tujuan klasifikasi, identifikasi,

serta pemantauan lingkungan. Salah satu pendekatan yang banyak digunakan adalah *Convolutional Neural Network* (CNN), sebuah arsitektur jaringan saraf dalam yang dirancang untuk mengekstraksi pola visual secara otomatis dan bertahap dari data citra Zhao et al. (2022). CNN bekerja melalui pemetaan bertingkat yang mampu mempelajari fitur sederhana hingga kompleks tanpa memerlukan rekayasa fitur manual, sehingga memberikan keunggulan signifikan dibandingkan metode konvensional berbasis ekstraksi fitur hand-crafted (Kattenborn et al, 2021).

Dalam perkembangan arsitektur deep learning modern, ResNet50 muncul sebagai model yang stabil dan berkinerja tinggi berkat inovasi *residual learning* melalui mekanisme *skip connections*. Model dasar yang digunakan adalah ResNet50, yang dipilih karena kemampuannya dalam mengekstraksi fitur visual secara hierarkis melalui mekanisme residual learning, sehingga lebih stabil ketika dilatih pada data yang bervariasi Shafiq et al, (2022). Arsitektur residual ini memungkinkan model mencapai kedalaman jaringan yang lebih besar tanpa memunculkan masalah *vanishing gradients*, serta membuat proses pelatihan lebih efisien pada dataset yang kompleks (Furusho et al., 2020).

Dalam berbagai aplikasi, ResNet50 telah digunakan secara luas untuk klasifikasi objek visual pada beragam domain seperti pengenalan organisme laut, identifikasi spesies, analisis medis, hingga deteksi penyakit tanaman, yang menunjukkan versatilitas dan daya adaptasinya terhadap domain baru Zhou et al, (2023). Keberhasilan model ini mendorong penggunaannya dalam penelitian yang melibatkan citra dengan kompleksitas tinggi, terutama ketika kualitas citra bervariasi atau jumlah sampel tidak merata antar kelas.

Tantangan umum yang sering dijumpai dalam penelitian berbasis citra adalah ketersediaan data yang terbatas. Jumlah data yang sedikit dapat menyebabkan *overfitting*, yaitu kondisi ketika model terlalu menyesuaikan diri dengan data pelatihan sehingga gagal melakukan generalisasi pada data baru Xu et al, (2023). Dalam konteks ini, *transfer learning* menjadi strategi penting yang memungkinkan model memanfaatkan pengetahuan dari dataset besar seperti ImageNet untuk kemudian diadaptasi pada domain baru yang memiliki keterbatasan data Kim et al, (2022). Transfer learning terbukti meningkatkan akurasi dan mempercepat proses pelatihan, terutama ketika jumlah sampel dalam dataset target terbatas (Hossain et al, 2022).

Selain menggunakan transfer learning, penggunaan teknik *fine-tuning* diperlukan untuk mengadaptasi lapisan tertentu dari model pralatih

agar dapat mengekstraksi fitur spesifik yang sesuai dengan karakteristik dataset Li et al, (2023). Pendekatan ini memungkinkan model mempelajari pola domain target tanpa harus membangun arsitektur dari awal, yang tidak efisien dan berisiko gagal ketika jumlah data kecil.

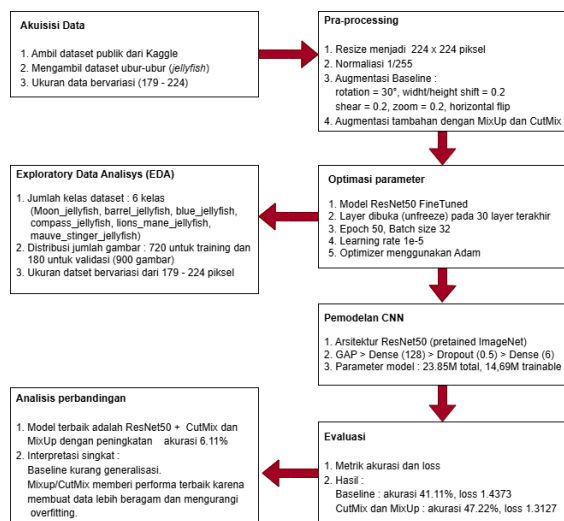
Masalah *overfitting* juga diatasi melalui strategi *data augmentation*, yaitu proses mengubah citra asli menjadi variasi baru tanpa mengubah labelnya. Data augmentation meningkatkan keragaman struktur visual dan memperkuat kemampuan model dalam mengenali objek dengan kondisi pencahayaan, orientasi, atau skala yang berbeda Wu & Z, (2024). Augmentasi klasik seperti rotasi, *flipping*, dan perubahan intensitas cahaya dinilai tidak selalu cukup untuk menghasilkan generalisasi optimal.

Teknik **MixUp** merupakan metode *mixed-sample augmentation* yang menggabungkan dua citra beserta labelnya melalui interpolasi linier, sehingga menghasilkan sampel baru yang mampu meningkatkan regularisasi model dan mengurangi *overfitting* Park et al, (2022). **CutMix** bekerja dengan menggantikan sebagian area suatu citra dengan patch dari citra lain, diikuti pencampuran label secara proporsional sehingga mempertahankan konsistensi informasi spasial yang penting bagi model Xu et al, (2021). Studi terbaru menunjukkan bahwa kombinasi kedua teknik ini mampu menghasilkan model yang lebih stabil, lebih tahan terhadap noise, dan lebih kuat dalam menghadapi variasi domain Kim et al, (2023). Pada konteks citra kelautan yang memiliki tingkat variabilitas visual tinggi, pendekatan hybrid *MixUp* dan *CutMix* bahkan terbukti meningkatkan ketahanan model terhadap pergeseran distribusi Tanaka et al, (2024). Selain itu, penelitian teoretis terbaru juga menunjukkan hubungan matematis antara *MixUp* dan *CutMix* melalui formulasi *manifold interpolation*, yang memperkuat justifikasi penggunaan keduanya secara simultan (Li et al, 2024).

Exploratory Data Analysis (EDA) juga memiliki peran penting dalam memahami karakteristik dataset sebelum dimasukkan ke dalam model. EDA memberikan informasi tentang distribusi kelas, kualitas citra, keberadaan noise, data outlier, hingga kemungkinan ketidakseimbangan kelas yang dapat mempengaruhi performa model Emmert-Streib et al, (2020). Dengan demikian, EDA menjadi tahapan awal yang menentukan arah pemilihan metode augmentasi, strategi *fine-tuning*, dan pendekatan evaluasi.

MATERIALS AND METHODS

Pada tahapan ini, alur penelitian dijelaskan secara sistematis untuk memberikan gambaran menyeluruh mengenai proses yang digunakan dalam pengembangan model klasifikasi citra. Pada Gambar 1 menyajikan rangkaian langkah mulai dari akuisisi data, pra-pemrosesan, analisis data awal, optimasi parameter, perancangan arsitektur CNN, tahap evaluasi hingga analisis perbandingan. Penyusunan tahapan ini bertujuan untuk memastikan setiap proses berjalan terstruktur, teratur, dan selaras dengan tujuan penelitian sehingga model yang dihasilkan memiliki performa dan generalisasi yang optimal.



Gambar 1 Tahapan Penelitian

Gambar 1 menampilkan tahapan metodologi penelitian yang disusun secara sistematis untuk menghasilkan model klasifikasi citra ubur-ubur berbasis *Convolutional Neural Network* (CNN) yang optimal. Alur penelitian dimulai dari akuisisi data hingga evaluasi dan analisis perbandingan model, sehingga setiap tahap saling berhubungan dalam meningkatkan kualitas model akhir.

Pada tahap akuisisi data, peneliti mengumpulkan 900 citra ubur-ubur yang terdiri dari enam kelas berbeda dengan ukuran gambar yang bervariasi (179 - 224 piksel). Proses ini dilakukan untuk memastikan tersedianya data yang representatif, dan agar model memiliki variasi sampel yang cukup untuk mempelajari ciri visual masing-masing kelas. Data diambil dari sumber publik untuk menjaga reproduktibilitas dan aksesibilitas penelitian.

Selanjutnya, pada tahap *pra-pemrosesan*, data dipersiapkan melalui proses *resize* ditargetkan menjadi 224x224 piksel, dengan normalisasi 1/255, serta pemberian augmentasi dasar dan lanjutan (*MixUp* dan *CutMix*). Tahap ini dilakukan untuk menyamakan format input serta agar variasi

citra meningkat, sehingga model lebih tahan terhadap *overfitting*. Teknik *MixUp* dan *CutMix* secara khusus diterapkan untuk menciptakan sampel sintesis yang lebih variatif sehingga model dapat belajar fitur yang lebih stabil dan general.

Tahap *Exploratory Data Analysis* (EDA) dilakukan untuk meninjau struktur dataset, meliputi jumlah kelas, distribusi jumlah citra (720 untuk pelatihan dan 180 untuk validasi), serta variasi ukuran berkas citra (179-224 piksel). Kegiatan EDA dilakukan untuk memahami konteks dan karakteristik data, serta agar peneliti dapat menentukan strategi pelatihan yang sesuai, seperti penentuan augmentasi dan kebutuhan balancing dataset.

Pada tahap optimasi parameter, dilakukan penentuan hyperparameter meliputi jumlah epoch (50), ukuran batch (32), laju pembelajaran (1e-5), dan pembukaan 30 layer terakhir pada arsitektur ResNet50. Optimasi ini dilakukan untuk mencari konfigurasi pelatihan yang memberikan performa terbaik dan agar model mampu mengekstraksi fitur visual dengan lebih efektif dari citra ubur-ubur yang memiliki struktur kompleks.

Tahap pemodelan CNN melibatkan penggunaan arsitektur ResNet50 pretrained ImageNet, yang kemudian dimodifikasi dengan penambahan lapisan Global Average Pooling, Dense 128, Dropout 0.5, dan Dense 6 kelas. Pemilihan arsitektur ini dilakukan untuk memanfaatkan kemampuan ekstraksi fitur mendalam yang sudah terbukti kuat, serta agar model dapat disesuaikan dengan kebutuhan klasifikasi enam kelas ubur-ubur pada penelitian ini.

Tahap evaluasi dilakukan untuk mengukur performa model menggunakan metrik akurasi dan loss. Evaluasi berbasis *validation set* dilakukan untuk menguji sejauh mana model mampu menggeneralisasi pola yang dipelajari terhadap data yang tidak digunakan saat pelatihan.

Hasil evaluasi menunjukkan bahwa model dengan augmentasi *MixUp* + *CutMix* memberikan performa terbaik, dengan akurasi validasi mencapai 47.22% dan loss 1.3127, lebih tinggi dibandingkan model baseline yang hanya mencapai 41.11% dengan loss 1.4373. Temuan ini menunjukkan bahwa augmentasi lanjutan secara signifikan meningkatkan generalisasi model, serta agar model mampu mengenali citra yang lebih bervariasi dibandingkan dengan augmentasi dasar saja.

Untuk memperjelas alur pelaksanaan penelitian serta hubungan antarproses yang terlibat, Tabel 3.1 disusun sebagai ringkasan tahapan metodologis yang digunakan dalam pengembangan dan evaluasi model klasifikasi citra ubur-ubur. Tabel ini menyajikan setiap tahap beserta input, output, dan parameter utama yang

digunakan, sehingga memberikan gambaran terstruktur mengenai proses mulai dari pengumpulan data hingga evaluasi model. Penyusunan tabel ini bertujuan mempermudah pemahaman pembaca terhadap langkah-langkah penelitian serta memastikan bahwa keseluruhan proses dapat direplikasi secara sistematis dan konsisten.

Hasil perbandingan menunjukkan bahwa setiap variasi metode memberikan dampak berbeda terhadap performa model. Model baseline dengan augmentasi standar menghasilkan akurasi validasi 41.11% dan loss 1.4373, menunjukkan keterbatasan generalisasi ketika data memiliki variasi visual yang kompleks. Model dengan augmentasi lanjutan *MixUp* dan *CutMix* mencapai akurasi 47.22% dengan loss 1.3127, menunjukkan peningkatan 6.11% dibanding baseline. Peningkatan ini terjadi karena kedua teknik tersebut menghasilkan sampel baru yang lebih bervariasi sehingga model memperoleh representasi fitur yang lebih robust. Dengan demikian, model *MixUp* dan *CutMix* terbukti sebagai konfigurasi paling efektif untuk meningkatkan generalisasi dan mengurangi *overfitting* pada dataset ubur-ubur.

RESULTS AND DISCUSSION

Hasil Augmentasi *MixUp* Dan *CutMix* Pada Dataset Ubur-Ubur Sebelum melangkah pada tahap pelatihan model, tahap eksplorasi data awal (*Exploratory Data Analysis/EDA*) dilakukan untuk mendapatkan pemahaman yang komprehensif mengenai karakteristik dataset. Analisis ini difokuskan pada aspek variasi visual serta proporsi jumlah sampel antar kelas guna memastikan dataset memiliki struktur yang ideal bagi pelatihan *Convolutional Neural Network* (CNN). Pemastian keseimbangan data pada tahap ini sangat krusial untuk meminimalkan potensi bias yang dapat memengaruhi kemampuan model dalam mengenali setiap kelas ubur-ubur secara objektif. Visualisasi sebaran data tersebut disajikan pada Gambar 4.5.

Contoh Gambar Augmentasi Kustom (CutMix/Mixup)



Gambar 2 Contoh Gambar Hasil Augmentasi *CutMix* dan *MixUp*

Berdasarkan Gambar 2 mendemonstrasikan hasil penerapan strategi augmentasi data tingkat lanjut

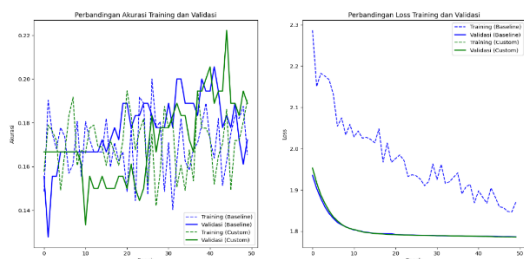
yang mengintegrasikan teknik *CutMix* dan *MixUp* pada dataset ubur-ubur. Secara visual, gambar ini memperlihatkan mekanisme manipulasi spasial di mana bagian (*patch*) dari satu citra kelas tertentu disisipkan secara acak ke dalam citra kelas lain. Hal ini menciptakan variasi data latih yang kompleks, di mana model dipaksa untuk memproses informasi dari dua kelas objek yang berbeda dalam satu representasi visual tunggal.

Aspek fundamental yang ditonjolkan dalam visualisasi ini adalah perubahan pada label target (anotasi kelas). Berbeda dengan pendekatan klasifikasi tradisional yang menggunakan label biner (*one-hot encoding*), teknik ini menghasilkan label lunak (*soft labels*) yang proporsional, seperti yang terlihat pada keterangan gambar (contoh: *blue_jellyfish*: 0.55; *compass_jellyfish*: 0.45). Interpolasi label ini bertujuan untuk memitigasi keyakinan berlebih (*overconfidence*) pada model serta memfasilitasi pembelajaran transisi antar-kelas yang lebih halus.

Secara fungsional, penerapan augmentasi ini berperan sebagai regularisasi yang memperkaya pola lokal dan fitur morfologi ubur-ubur. Dengan mengaburkan batasan spasial dan mencampurkan konteks visual, model didorong untuk mengekstraksi fitur yang lebih robust (tangguh) dan tidak hanya bergantung pada hafalan latar belakang, sehingga secara efektif mengurangi risiko *overfitting* sekaligus meningkatkan kemampuan generalisasi pada data validasi.

Hasil Pelatihan Model Baseline Dan *CutMix* + *MixUp* (Tanpa Fine-Tuning)

Gambar 2 menyajikan dinamika akurasi model selama proses pelatihan pada skenario *baseline*. Dalam eksperimen ini, arsitektur ResNet50 difungsikan sebagai pengekstraksi fitur tetap (*fixed feature extractor*) dengan membekukan seluruh bobot *backbone*, sehingga pembelajaran hanya terfokus pada lapisan klasifikasi akhir (*classifier head*). Visualisasi kurva antara data latih dan validasi ini bertujuan untuk memetakan kemampuan dasar fitur *pretrained* ImageNet dalam mengenali karakteristik visual ubur-ubur, sekaligus menjadi tolak ukur komparatif untuk menilai efektivitas strategi *fine-tuning* pada tahap selanjutnya.



Gambar 2 Hasil Pelatihan Model Baseline Tanpa Fine Tuning. (a) Kurva Training-Validation Accuracy. (b) Kurva Training-Validation Loss

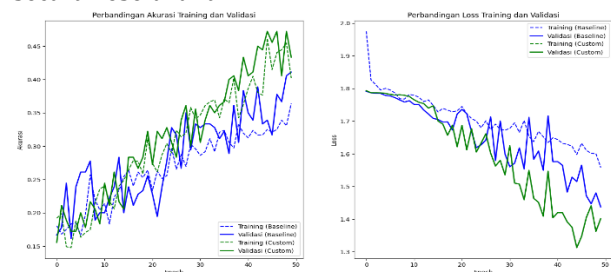
Berdasarkan Gambar 2 (a) menunjukkan dinamika akurasi antara model Baseline (garis biru) dan model dengan augmentasi *CutMix* dan *MixUp* (garis hijau) pada skenario pelatihan tanpa fine-tuning. Garis biru putus-putus (b--) merepresentasikan akurasi *training* Baseline, sedangkan garis biru penuh (b-) menunjukkan akurasi *validation* Baseline. Pada model augmented, garis hijau putus-putus (g--) menunjukkan akurasi *training* *CutMix* dan *MixUp*, sementara garis hijau penuh (g-) merepresentasikan akurasi *validation*-nya. Secara visual, seluruh kurva menunjukkan fluktuasi yang cukup tinggi, mengindikasikan proses pembelajaran yang belum stabil karena hanya lapisan atas (*classification head*) yang dilatih, sementara *backbone* ResNet50 masih dibekukan. Secara kuantitatif, akurasi validasi terbaik masih relatif rendah, yaitu 20,56% (b-) pada *Baseline* dan 22,22% (g-) pada model *CutMix* dan *MixUp*. Peningkatan kecil pada kurva hijau menunjukkan bahwa teknik mixing memberikan regularisasi yang sedikit lebih baik dibandingkan *Baseline*, namun tidak cukup untuk menghasilkan lonjakan performa yang signifikan pada tahap *head-only training*.

Dan berdasarkan Gambar 4.6(b) memperlihatkan pola penurunan loss untuk kedua model. Garis biru putus-putus (b--) merepresentasikan *training loss* *Baseline*, sementara garis biru penuh (b-) menunjukkan *validation loss* *Baseline*. Pada model augmented, garis hijau putus-putus (g--) menggambarkan *training loss* *CutMix* dan *MixUp*, sedangkan garis hijau penuh (g-) menunjukkan *validation loss* yang dihasilkan.

Kurva biru menampilkan penurunan loss yang tidak stabil dengan fluktuasi tajam, mencerminkan proses optimisasi yang kurang efisien. Sebaliknya, kurva hijau (g-- dan g-) menunjukkan penurunan loss yang sedikit lebih mulus dan terkontrol. Namun, perbedaan antar model pada tahap ini masih sangat kecil. Secara kuantitatif, *Best Validation Loss* pada *Baseline* adalah 1,7861 (b-), sementara model *CutMix*-*MixUp* mencapai 1,7855 (g-), menunjukkan perbaikan yang minimal namun konsisten dengan peran *CutMix*/*MixUp* sebagai regularisasi awal.

Hasil Pelatihan Model Baseline Dan CutMix + MixUp (Fine-Tuning)

Eksperimen kedua melakukan fine-tuning dengan membuka sekitar 30 lapisan terakhir dari backbone ResNet50, sehingga proses pembelajaran tidak hanya berlangsung pada classifier head, tetapi juga pada representasi fitur tingkat menengah hingga tinggi. Pendekatan ini memungkinkan model menyesuaikan abstraksi visual dengan karakteristik citra ubur-ubur yang berbeda dari ImageNet. Hasilnya, sebagaimana ditunjukkan pada Gambar 3 terjadi peningkatan stabilitas pelatihan dan kenaikan akurasi validasi, disertai semakin kecilnya selisih antara akurasi training dan validasi. Pola ini menunjukkan bahwa proses fine-tuning berhasil memperbaiki kualitas representasi fitur dan meningkatkan kemampuan generalisasi model secara keseluruhan.



Gambar 3 (a) Kurva Training Validation Accuracy (Fine Tuning). (b) Kurva Training Validation Loss

Berdasarkan Gambar 3. (a) memperlihatkan perkembangan akurasi selama tahap *fine-tuning* untuk dua konfigurasi model. Garis biru putus-putus (b--) merepresentasikan akurasi *training* model *Baseline*, sedangkan biru penuh (b-) menunjukkan akurasi *validation*. Pada model dengan augmentasi *CutMix* dan *MixUp*, hijau putus-putus (g--) menggambarkan akurasi *training* dan hijau penuh (g-) merupakan akurasi *validation*. Kedua model menunjukkan pola peningkatan akurasi yang lebih stabil dibandingkan ketika backbone masih dibekukan, menandakan bahwa *fine-tuning* berhasil meningkatkan kemampuan model dalam mempelajari representasi fitur. Namun, kurva hijau tampak secara konsisten berada di atas kurva biru pada hampir seluruh epoch, menunjukkan bahwa mekanisme mixing memperkuat proses pembelajaran setelah backbone dilatih ulang. Secara kuantitatif, model *CutMix* dan *MixUp* mencapai *Best Validation Accuracy* sebesar 47,22%, lebih tinggi dibandingkan *Baseline* yang mencatat 41,11%, sehingga menegaskan keunggulan kombinasi antara *fine-tuning* dan augmentasi mixing.

Pada Gambar 3 (b) memperlihatkan perbandingan loss selama fine-tuning, di mana model *CutMix* dan *MixUp* menampilkan pola

penurunan loss yang lebih stabil dan lebih rendah dibandingkan *Baseline*. Kurva hijau, baik pada training maupun validasi cenderung menunjukkan lintasan yang lebih terkontrol, sedangkan kurva biru milik *Baseline* tampak lebih berfluktuasi. Kondisi ini menunjukkan bahwa mekanisme mixing bertindak sebagai regularisasi yang efektif setelah backbone diperbarui, sehingga model menjadi lebih tangguh terhadap variasi data validasi. Secara kuantitatif, model *CutMix* dan *MixUp* menghasilkan *Best Validation Loss* sebesar 1,3127, lebih rendah daripada *Baseline* yang mencatat 1,4373, menegaskan bahwa konfigurasi ini memberikan proses optimisasi yang lebih efisien serta generalisasi yang lebih baik.

Tabel 1 menyajikan hasil akhir berupa *best validation accuracy* dan *best validation loss* setelah proses fine-tuning. Peningkatan signifikan pada kedua metrik dibandingkan *baseline* menunjukkan bahwa penyesuaian bobot pretrained merupakan komponen kunci dalam meningkatkan performa model pada domain bawah air.

Tabel 1 Ringkasan Hasil Fine-Tuning

Model	Best Accuracy	Val Best Loss
Baseline	41.11%	1.4373
MixUp CutMix	+ 47.22%	1.3127

Berdasarkan Tabel 1 menunjukkan perbandingan kinerja model antara skenario *baseline* dan skenario dengan penerapan augmentasi *MixUp* + *CutMix*. Hasil yang ditampilkan mencakup dua metrik utama, yaitu *best validation accuracy* dan *best validation loss*, yang merefleksikan kemampuan model dalam melakukan generalisasi terhadap data yang tidak terlihat selama pelatihan.

Model *baseline* memperoleh akurasi validasi terbaik sebesar 41.11% dengan nilai loss sebesar 1.4373. Penerapan *MixUp* dan *CutMix* menghasilkan peningkatan akurasi validasi menjadi 47.22%, disertai penurunan loss validasi menjadi 1.3127.

Peningkatan akurasi sebesar 6.11% dan penurunan loss sebesar 0.1246 mengindikasikan bahwa strategi augmentasi berbasis pencampuran sampel mampu memperkaya variasi data pelatihan secara efektif. Dengan demikian, model menjadi lebih robust dalam mengenali pola visual yang bervariasi, sekaligus mengurangi risiko overfitting. Secara keseluruhan, perbandingan ini menunjukkan bahwa integrasi *MixUp* dan *CutMix* memberikan kontribusi positif terhadap stabilitas dan kemampuan generalisasi model pada tugas klasifikasi ubur-ubur.

CONCLUSION

Penelitian ini menunjukkan bahwa peningkatan robustness model ResNet50 untuk klasifikasi ubur-ubur dapat dicapai melalui penerapan teknik augmentasi *MixUp* dan *CutMix* (Hu et al, 2024). Tujuan penelitian untuk meningkatkan akurasi dan ketahanan model terhadap variasi citra bawah laut terpenuhi melalui integrasi kedua teknik ini. Peningkatan akurasi validasi dari 41,11% menjadi 47,22% dan penurunan *validation loss* dari 1,4373 menjadi 1,3127 menunjukkan bahwa model menjadi lebih stabil dan kurang sensitif terhadap *noise*, degradasi warna, dan kondisi pencahayaan yang tidak merata pada citra bawah laut. Temuan ini menegaskan bahwa strategi *mixed-sample augmentation* mampu memperluas distribusi data sintesis sehingga membantu model mempelajari fitur visual secara lebih konsisten.

Penelitian ini memberikan kontribusi terhadap pengembangan metode klasifikasi citra bawah laut dengan menunjukkan bahwa kombinasi *MixUp* dan *CutMix* mampu memberikan regularisasi kuat pada dataset kecil (Park et al, 2022c). Kebaruan penelitian terletak pada penerapan kedua teknik ini secara terintegrasi pada domain ubur-ubur, yang memiliki karakteristik citra kompleks dan rentan hilangnya detail. Pendekatan ini memperkuat pemahaman bahwa robust learning dapat tercapai tidak hanya melalui peningkatan kapasitas model, tetapi juga melalui desain data pelatihan yang lebih kaya dan bervariasi. Implikasi ilmiah dari temuan ini adalah bahwa peningkatan robustness model dapat dicapai secara sistematis melalui strategi augmentasi yang tepat, terutama pada domain dengan keterbatasan data dan kualitas citra.

Penelitian ini memiliki keterbatasan pada ukuran dataset yang relatif kecil dan variasi kelas yang terbatas, kondisi yang diketahui dapat menurunkan generalisasi model CNN (Xu et al, 2023). Keterbatasan ini memengaruhi kemampuan model dalam membangun representasi fitur yang lebih umum untuk kondisi bawah laut yang lebih luas. Selain itu, penelitian hanya mengevaluasi performa model menggunakan metrik akurasi dan loss tanpa menguji ketahanan model terhadap skenario ekstrem seperti blur berat, perubahan warna ekstrem, atau backscatter intens. Keterbatasan ini memberikan ruang untuk penelitian lanjutan yang dapat mengeksplorasi kondisi pengujian lebih komprehensif dan meningkatkan generalisasi model.

REFERENCE

- Abbasian Ardakani, G., & al., et. (2024). Interpretation of Artificial Intelligence Models in Healthcare: Evaluation metrics, pitfalls and recommendations. *Journal of Ultrasound in Medicine*.
<https://doi.org/10.1002/jum.16524>
- Ahmed, B., Rau, G., E., Mudunuru, K., M., Karra, S., Tempelman, R., J., Wachtor, J., A., Forien, J.-B., & al., et. (2024). Deep learning with mixup augmentation for improved pore detection during additive manufacturing. *Scientific Reports*.
- Ahmed, B., Rau, G., E., Mudunuru, K., M., Karra, S., Tempelman, R., J., Wachtor, J., A., Forien, J.-B., Guss, M., G., Calta, P., N., DePond, ... M., J. (2024). Deep learning with mixup augmentation for improved pore detection during additive manufacturing. *Scientific Reports*.
- Arafin, P., & al., et. (2022). Performance comparison of multiple convolutional neural network architectures for visual defect classification. *Sensors*, 22(22).
- Avberšek, K., L., & al., et. (2022). *Deep learning in neuroimaging data analysis: Applications, pitfalls and recommendations*.
- Benavoli, A., Corani, G., Mangili, F., Zaffalon, M., Ruggeri, & F. (2020). Bayesian signed-rank test for pairwise comparison of classification algorithms. *Journal of Machine Learning Research*, 1–30.
- Benerradi, J., Clos, J., Landowska, A., Valstar, F., M., Wilson, & M., L. (2023). *Benchmarking framework for machine learning classification from fNIRS data (BenchNIRS): Improving experimental rigor with nested cross-validation and open artifacts*.
- Bouthillier, X., & al., et. (2021). Accounting for variance in machine learning benchmarks. *Proceedings of the National Academy of Sciences*, 118(15).
- Cao, C., Zhou, F., Dai, Y., Wang, J., Zhang, & K. (2023). A survey of mix-based data augmentation: taxonomy, methods, applications, and explainability. *Journal of Big Data*, 11(1).
- Chen, Y., Xu, & L. (2022). Robustness evaluation of convolutional neural networks under input perturbations: The role of validation loss. *Journal of Computer Science*, 145–158.
- Dablain, D., Jacobson, N., K., Bellinger, C., Roberts, M., Chawla, & N., V. (2023). Understanding CNN fragility when learning with imbalanced data. *Machine Learning*, 4785–4810.
- Ditria, M., E., Lopez-Marcano, S., Sievers, K., M., Jinks, L., E., Brown, J., C., Connolly, & R., M. (2020). Automating the analysis of fish abundance using object detection: Optimizing animal ecology with deep learning. *Frontiers in Marine Science*.
- Dong, Y., & al., et. (2023). Competition on robust deep learning. *National Science Review*, 10(6).
- Dumoulin, V., & Visin, F. (2021). A Guide to Convolution Arithmetic for Deep Learning. *ArXiv*.
<https://doi.org/10.48550/arXiv.1603.07285>
- Eghbal-Zadeh, H., & al., et. (2024). *Rethinking data augmentation for adversarial robustness*.
- Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., Dehmer, & M. (2020). An introductory review of deep learning for prediction models with big data. *Frontiers in Artificial Intelligence*.
- Furusho, Y., Ikeda, & K. (2020). *Theoretical analysis of skip connections and batch normalization from generalization and optimization perspectives*.
- Gao, M., Li, S., Wang, K., Bai, Y., Ding, Y., Zhang, B., Guan, N., Wang, & P. (2023). Real-time jellyfish classification and detection algorithm based on improved YOLOv4-tiny and improved underwater image enhancement algorithm. *Scientific Reports*.
- Garcea, F., Gatta, C., Caputo, & B. (2023). Data augmentation for medical imaging: A systematic review. *Computerized Medical Imaging and Graphics*.
- Guo, H., Somayajula, A., S., Hosseini, R., & al., et. (2024). Improving image classification of gastrointestinal endoscopy using curriculum self-supervised learning (C-Mixup). *Scientific Reports*.
- Hao, X., Liu, L., Yang, R., Yin, L., Zhang, L., Li, & X. (2023). A review of data augmentation methods of remote sensing image target recognition. *Remote Sensing*, 15(3).
- Hellín, J., C., Olmedo, A., A., Villedor, A., Gómez, J., López-Benítez, M., Tayebi, & A. (2024). Unraveling the impact of class imbalance on deep-learning models for medical image classification. *Applied Sciences*, 14(8).
- Ho, M., T., Lee, & K., Y. (2023). Comprehensive analysis of performance metrics for deep convolutional neural networks in image recognition tasks. *Computers*, 12(8).
- Hossain, B., M., Iqbal, H., S. S., Islam, M., M., Akhtar, N., M., Sarker, & I., H. (2022). Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images. *Informatics in Medicine Unlocked*.
- Houssein, H., E., & al., et. (2024). *An effective multiclass skin-cancer classification approach using deep convolutional neural networks: accuracy, recall, precision, F1-score, specificity and AUC*.
- Hu, X., Zhang, & L. (2024). Theoretical analysis of CutMix: Patch-based Bayesian priors for robust image classification. *Neural Networks*.
- Imaduddin, H., Kumar, & P. (2024). *Fine-tuning ResNet-50 for the classification of visual impairments from retinal fundus images*.
- Islam, T., Hafiz, S., M., Rahman, J., Kabir, M., M., Mridha, & M., F. (2024). A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions. *Healthcare Analytics*.
- Jiang, X., Yu, H., Zhang, Y., Pan, M., Li, Z., Liu, J., Lv, & S. (2023). An Underwater Image Enhancement Method for a Preprocessing Framework Based on Generative Adversarial Network. *Sensors*, 23(13).
- Jin, G., & al., et. (2024). Performance evaluation of convolutional neural network models for highway distress detection: precision, recall, F1-score and confusion matrix. *Applied Sciences*, 14(10).
- Kandimalla, V., Richard, M., Smith, F., Quirion, J., Torgo, L., Whidden, & C. (2022). Automated detection, classification and counting of fish in fish passages with deep learning. *Frontiers in Marine Science*.
- Kang, J., & al., [et. (2024). Diagnosing oral and maxillofacial diseases using deep learning: data-augmentation strategies including Mixup. *Scientific Reports*.
- Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, & S. (2021). Review on convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 24–49.
- Kaur, B., A., Kaur, M., Singh, D., Roy, S., Amoon, & M. (2023). Efficient skip connections-based residual network (ESRNet) for brain tumor classification. *Diagnostics*, 13(20).
- Khan, H.-M., M., Makoonlall, A., Nazurally, N., Mungloo-Dilmohamud, & Z. (2023). Identification of Crown of Thorns Starfish (COTS) using Convolutional Neural Network (CNN) and attention model. *PLOS ONE*, 18(4).
<https://doi.org/10.1371/journal.pone.0283121>
- Kim, E., H., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, E., M., Ganslandt, & T. (2022). Transfer learning for medical image classification: a literature review. *BMC Medical Imaging*.
- Kim, H., Finn, & C. (2022). *C-Mixup: Improving Generalization in Regression*.
- Kim, T., Cho, H., Oh, & S. (2023). Understanding Robustness of Mixup and CutMix Data Augmentation. *IEEE Access*, 82354–82367.
- Klomp, R., S., Wijnhoven, J., R. G., With, de, & N, P. H. (2023). Performance-Efficiency Comparisons of Channel Attention Modules for ResNets. *Neural Processing Letters*, 55(10), 6797–6813.
- Kotsilieris, T., & al., et. (2022). Regularization techniques for machine learning and their theoretical and practical aspects. *Electronics*, 11(4).

- Krichen, M., & al, et. (2023). Convolutional Neural Networks: A Survey. *Computers*, 12(8).
- Lee, H., Lee, H., Kim, & J. (2023). *The effects of mixed sample data augmentation are class dependent*.
- Li, F., Zhao, M., Yu, & X. (2024). Unifying CutMix and Mixup via interpolation manifold theory. *Entropy*, 26(5).
- Li, S., Li, P., He, S., Kuai, Z., Gu, Y., Liu, H., Liu, T., & Lin, Y. (2024). An automatic detection and statistical method for underwater fish based on Foreground Region Convolution Network (FR-CNN). *Journal of Marine Science and Engineering*, 12(8), 1343. <https://doi.org/10.3390/jmse12081343>
- Li, X., Yang, J., Zhang, & Y. (2023). Comparison of classifiers: A robust bootstrap-based statistical framework. *PLOS ONE*, 18(3).
- Li, Y., Wang, H., Guo, J., Zhao, & L. (2023). Efficient fine-tuning of deep CNNs with discriminative learning-rate strategies for domain-specific adaptation. *Sensors*, 23(14).
- Li, Z., Wang, Y., Chen, & X. (2024). *An empirical study of geometric and photometric augmentation magnitudes on CNN robustness*.
- Liao, T., Yang, R., Zhao, P., Zhou, W., He, M., Li, & L. (2022). MDAM-DRNet: Dual channel residual network with multi-directional attention mechanism in strawberry leaf diseases detection. *Frontiers in Plant Science*.
- Liu, W., Gao, S., Han, & J. (2022). The effect of learning rate scheduling and batch normalization on the convergence of deep neural networks. *Applied Sciences*, 12(19).
- Liu, Y., T., Mirzasoleiman, & B. (2022). *Data-Efficient Augmentation for Training Neural Networks*.
- Luo, H., Zhang, Y., Sun, & J. (2023). A probabilistic interpretation of Mixup for regularizing neural networks. *Pattern Recognition Letters*, 1–10.
- Mai, J., Gao, C., Bao, & J. (2025). Domain generalization through data augmentation: A survey of methods, applications, and challenges. *Mathematics*, 13(5).
- Martin-Abadal, M., Ruiz-Frau, A., Hinz, H., González-Cid, & Y. (2020). Jellytoring: Real-time jellyfish monitoring based on deep learning object detection. *Sensors*, 20(6).
- Mienye, D., I., Swart, & T., G. (2024). A comprehensive review of deep learning: Architectures, recent advances, and applications. *Information*, 15(12).
- Miseta, T., Fodor, A., Vathy-Fogarassy, & Á. (2024). Surpassing early stopping: A novel correlation-based stopping criterion for neural networks. *Neurocomputing*.
- Moodley, C., Sephton, B., Rodríguez-Fajardo, V., Forbes, & A. (2021). Deep learning early stopping for non-degenerate ghost imaging. *Scientific Reports*.
- Mumuni, A., Mumuni, & F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*.
- Najeeb, M., Alariyibi, & A. (2024). Imbalanced dataset effect on CNN-based classifier performance for face recognition. *International Journal of Artificial Intelligence & Applications*, 15(1), 25–41.
- Noh, & S.-H. (2021). Performance Comparison of CNN Models Using Gradient Flow Analysis. *Informatics*, 8(3).
- Park, C., Yun, S., Chun, & S. (2022a). *A unified analysis of mixed sample data augmentation: A loss-function perspective*.
- Park, C., Yun, S., Chun, & S. (2022b). A Unified Analysis of Mixed Sample Data Augmentation: A Loss Function Perspective. *Advances in Neural Information Processing Systems*, 17457–17470.
- Park, C., Yun, S., Chun, & S. (2022c). *A Unified Analysis of Mixed Sample Data Augmentation: A Loss Function Perspective*.
- Park, J., Kim, J., Lee, & J. (2024). *Impact of regularization on calibration and robustness: From the representation space perspective*.
- Pineau, J., Vincent-Lamarre, P., Sinha, K., Larivière, V., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Larochelle, & H. (2021). Improving reproducibility in machine learning research: A report from the NeurIPS 2019 reproducibility program. *Journal of Machine Learning Research*, 1–20.
- Rainio, O., Teuho, J., Klén, & R. (2024). Evaluation metrics and statistical tests for machine learning: Common metrics for CNN classification, segmentation and object detection. *Scientific Reports*, 14(1).
- Razavi, M., & al, et. (2024). *ResNet deep models and transfer learning technique for classification of structural content*.
- Rezvani, S., Wang, & X. (2023). A broad review on class imbalance learning techniques. *Applied Soft Computing*.
- Robillard, J., A., Trizna, G., M., Ruiz-Tafur, M., Panduro, D., L., E., Santana, de, D., C., White, E., A., Dikow, B., R., Deichmann, & J., L. (2023). Application of a deep learning image classifier for identification of Amazonian fishes. *Ecology and Evolution*, 13(5).
- Shafiq, M., Gu, & Z. (2022). Deep residual learning for image recognition: A survey. *Applied Sciences*, 12(18).
- Shantharam, M., R., Hossain, Z., M., Pandey, & R. (2024). ML-based pain recognition model using Mixup data augmentation. *Electronics*, 13(6).
- Shin, W., Kahng, H., Kim, & S., B. (2022). Mixup-based classification of mixed-type defect patterns in wafer bin maps. *Computers & Industrial Engineering*.
- Shorten, C., & Khoshgoftaar, T. M. (2020). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 7(1), 1–48. <https://doi.org/10.1186/s40537-019-0197-0>
- Smucny, J., Shi, G., Lesh, A., T., Carter, S., C., Davidson, I., & al, et. (2022). *Data augmentation with Mixup: Enhancing performance of a functional neuroimaging-based prognostic deep learning classifier in recent onset psychosis*.
- Sousa, J. V. M., Almeida, V. R. de, Saraiva, A. A., Santos, D. B. S., Pimentel, P. M. C., & Sousa, L. L. de. (2020). Classification of Pneumonia Images on Mobile Devices with Quantized Neural Network. *Research, Society and Development*, 9(10), e889108382. <https://doi.org/10.33448/rsd-v9i10.8382>
- Sun, S., Chen, J.-N., He, R., Yuille, A., Bai, & S. (2022). *LUMix: Improving Mixup by better modelling label uncertainty*.
- Tanaka, K., Kobayashi, S., Iwamoto, & T. (2024). Hybrid Mixup-CutMix augmentation for domain-robust image recognition in marine visual datasets. *Sensors*, 24(8).
- Tarling, P., Cantor, M., Clapés, A., Escalera, & S. (2022). Deep learning with self-supervision and uncertainty regularization to count fish in underwater images. *PLOS ONE*, 17(5).
- Usmani, A., I., Qadri, T., M., Zia, R., Alrayes, S., F., Saidani, O., Dashtipour, & K. (2023). Interactive effect of learning rate and batch size to implement transfer learning for brain tumour classification. *Electronics*, 12(4).
- Wang, X., & al, ... et. (2023). On the effectiveness of mixed-sample data augmentation for improving model robustness. *Machine Learning with Applications*.
- Wang, Y., Chen, T., Wang, & Z. (2023). Understanding mixed-sample data augmentation through risk variance minimization. *IEEE Transactions on Neural Networks and Learning Systems*, 34(9), 5764–5778.
- Won, S., Bae, S.-H., Kim, & S., T. (2023a). *Analyzing effects of mixed sample data augmentation on model interpretability*.
- Won, S., Bae, S.-H., Kim, & S., T. (2023b). *Effects of mixed sample data augmentation on model interpretability*.
- Wu, P. (2022). A Survey of Few-Shot Learning Research Based on Deep Neural Network. *Frontiers in Computing and Intelligent Systems*, 2(1), 110–115. <https://doi.org/10.54097/fcis.v2i1.3177>
- Wu, & Z. (2024). Image data augmentation techniques based on deep learning: A survey. *Mathematical Biosciences and Engineering*, 21(6), 6190–6224.
- Xu, C., Coen-Pirani, P., Jiang, & X. (2023). Empirical Study of Overfitting in Deep Learning for Predicting Breast Cancer Metastasis. *Cancers*, 15(7).
- Xu, Z., Choi, S., Park, S., Yun, & S. (2021). Improved CutMix training for deep neural networks. *IEEE Access*, 60462–60473.

- Xu, Z., Liu, C., He, & J. (2022). Geometric interpretation of Mixup in manifold space. *Information Sciences*, 985–1001.
- Zhang, H., Gong, L., Li, X., Liu, F., Yin, & J. (2023). An Underwater Imaging Method of Enhancement via Multi-Scale Weighted Fusion. *Frontiers in Marine Science*.
- Zhang, L., Bian, Y., Jiang, P., Zhang, & F. (2023). A transfer residual neural network based on ResNet-50 for detection of steel surface defects. *Applied Sciences*, 13(9).
- Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., Parmar, & M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*.
- Zhao, Y., Li, R., Xu, & C. (2022). Improving CNN generalization with adversarial and stochastic augmentation. *IEEE Access*, 86421–86433.
- Zhou, X., Li, Y., Wang, H., & Chen, Z. (2024). Cutout-based augmentation for medical imaging: Improving robustness and generalization in deep CNN models. *Biomedical Signal Processing and Control*, 93, 106059. <https://doi.org/10.1016/j.bspc.2024.106059>
- Zhou, Z., & al, et. (2023). Improving the classification accuracy of fishes and invertebrates in underwater images by using pre-trained ResNet50 neural network and error-minimized random vector functional link classifier. *ICES Journal of Marine Science*, 80(5), 1256–1268.