

ABSTRAK

Penelitian ini bertujuan untuk mengklasifikasikan berbagai jenis ikan koi dengan menggunakan *Vision Transformer* (ViT). Penulis mencatat bahwa penelitian sebelumnya menggunakan *Support Vector Machine* (SVM) dan *Convolutional Neural Network* (CNN) untuk identifikasi ikan koi dengan tingkat akurasi 84%. Meskipun hasil CNN tinggi, ada ruang untuk peningkatan akurasi. Sebagai metode penelitian, penulis mengusulkan penerapan arsitektur ViT untuk meningkatkan akurasi klasifikasi ikan koi. ViT berasal dari *Transformer*, menggunakan mekanisme perhatian diri dan memiliki kekuatan representasi data lebih baik daripada CNN. Jumlah dan kelas dataset ikan koi dipertahankan dari penelitian sebelumnya, sementara dataset gambar ikan koi yang digunakan pada penelitian ini dikumpulkan dari internet. Penelitian dilakukan dengan menerapkan ViT dan menganalisisnya sebagai *classifier model* pada dataset tersebut. Hasilnya menunjukkan tingkat akurasi rata-rata 88% pada semua kelas data pengujian. Selain itu, penulis melibatkan dua pengujian tambahan untuk mengevaluasi pengaruh augmentasi data terhadap performa model. Hasil pengujian menunjukkan bahwa model dengan augmentasi gambar latih mencapai tingkat akurasi 95%, sedangkan model tanpa augmentasi mencapai akurasi 87%. Analisis ini memberikan wawasan yang signifikan tentang dampak positif augmentasi terhadap kinerja sistem, khususnya pada konteks klasifikasi ikan koi. Dengan demikian, penelitian ini tidak hanya memvalidasi efektivitas ViT dalam meningkatkan akurasi, tetapi juga membuka ruang eksplorasi untuk algoritma dan teknik baru, termasuk penerapan augmentasi data dalam penelitian klasifikasi ikan koi. Hasil ini menjadi kontribusi penting dalam pengembangan lebih lanjut di bidang ini.

Kata kunci: *Deep Learning, Vision Transformer, PyTorch, Citra, Ikan Koi.*

ABSTRACT

This research aims to classify various types of koi fish using Vision Transformer (ViT). The author notes that previous studies utilized Support Vector Machine (SVM) and Convolutional Neural Network (CNN) for koi fish identification with an accuracy rate of 84%. Despite the high accuracy achieved by CNN, there is room for improvement. As the research method, the author proposes the implementation of ViT architecture to enhance the classification accuracy of koi fish. ViT derived from Transformer, employs a self-attention mechanism and exhibits superior data representation capabilities compared to CNN. The number and classes of the koi fish dataset are retained from previous studies, while the image dataset used in this research is collected from the internet. The research is conducted by applying ViT and analyzing it as a classifier model on the dataset. The results show an average accuracy rate of 89% across all testing data classes. In addition, the authors included two additional tests to evaluate the effect of data augmentation on model performance. The test results showed that the model with training image augmentation achieved an accuracy rate of 95%, while the model without augmentation achieved an accuracy of 87%. This analysis provides significant insight into the positive impact of augmentation on system performance, particularly in the context of koi fish classification. Thus, this study not only validates the effectiveness of ViT in improving accuracy, but also opens up exploration space for new algorithms and techniques, including the application of data augmentation in koi fish classification research. This result is an important contribution to further development in this field.

Keywords: *Deep Learning, Vision Transformer, PyTorch, Image, Koi Fish.*