ABSTRAK

No-Reference Image Quality Assessment (NR-IQA) merupakan salah satu bidang penting dalam pengolahan citra digital yang bertujuan untuk menilai kualitas citra tanpa memerlukan citra referensi. Permasalahan utama pada NR-IQA adalah bagaimana memprediksi kualitas perseptual citra hanya berdasarkan citra terdistorsi. Seiring berkembangnya metode berbasis *deep learning*, *Vision Transformer* (ViT) menunjukkan potensi yang baik dalam tugas analisis citra. Namun, jumlah parameter yang besar pada ViT dapat meningkatkan risiko *overfitting*, terutama ketika digunakan pada dataset yang berukuran relatif kecil. Oleh karena itu, penelitian ini bertujuan untuk menganalisis pengaruh strategi pembekuan *layer* (*freeze layer*) pada arsitektur ViT-Base terhadap performa model dalam tugas NR-IQA serta menentukan konfigurasi pembekuan *layer* yang paling optimal.

Penelitian menggunakan dataset CSIQ yang terdiri atas 866 citra terdistorsi dengan nilai *Differential Mean Opinion Score* (DMOS) sebagai *ground truth*. Tahap praproses data meliputi *resize*, konversi ke tensor, dan normalisasi citra. Model ViT-Base yang telah diprelatih (*pretrained*) pada ImageNet dimodifikasi menjadi model regresi dengan menambahkan *regression head*. Evaluasi dilakukan menggunakan metode *5-Fold Cross Validation* dengan lima skenario pembekuan *layer* yang berbeda. Kinerja model dievaluasi menggunakan metrik *Pearson Linear Correlation Coefficient* (PLCC), *Spearman Rank Correlation Coefficient* (SRCC), dan *Root Mean Square Error* (RMSE).

Hasil penelitian menunjukkan bahwa strategi pembekuan *layer* memberikan pengaruh terhadap performa model. Skenario terbaik diperoleh pada konfigurasi S2, yaitu dengan membekukan *patch embedding* dan melatih seluruh blok *Transformer encoder* serta *regression head*. Skenario tersebut menghasilkan nilai PLCC sebesar 0,9643 dan SRCC sebesar 0,9412. Hasil tersebut menunjukkan bahwa mempertahankan representasi fitur dasar dari model *pretrained* sekaligus melakukan *fine-tuning* pada bagian *encoder* mampu memberikan keseimbangan yang baik antara kemampuan generalisasi dan akurasi prediksi. Selain meningkatkan performa, strategi pembekuan sebagian *layer* juga mampu mengurangi jumlah parameter yang dilatih sehingga proses pelatihan menjadi lebih efisien. Dengan demikian, penelitian ini menunjukkan bahwa pendekatan *partial fine-tuning* melalui strategi *freeze layer* merupakan metode yang efektif untuk meningkatkan kinerja ViT-Base pada tugas NR-IQA dengan dataset berukuran terbatas.

Kata Kunci: *No-Reference Image Quality Assessment, Vision Transformer Base,* Pembekuan Layer, Prediksi Kualitas Citra, Efisiensi Komputasi

ABSTRACT

No-Reference Image Quality Assessment (NR-IQA) has become increasingly important in digital image processing, particularly due to the rapid growth of user-generated visual content distributed through online platforms. Assessing image quality without access to a reference image presents a significant challenge, as the system must accurately predict perceptual quality based solely on distorted inputs. This study focuses on investigating how the implementation of a layer-freezing strategy in the Vision Transformer Base (ViT-Base) architecture affects the performance of the NR-IQA method in generating digital image quality scores. In addition, the research examines the resulting level of prediction accuracy and computational efficiency produced by different freezing configurations. The main problem addressed is determining the optimal layer-freezing strategy that balances predictive performance and computational cost.

The research methodology consists of data acquisition, model development, training configuration, and performance evaluation stages. The dataset used is derived from the CSIQ image quality database, which contains distorted images along with their corresponding subjective quality scores in the form of DMOS. The data undergo preprocessing steps, including image transformation and partitioning into training and testing subsets. The ViT-Base model is adapted for regression tasks to predict image quality scores. Several experimental scenarios are designed by applying different layer-freezing strategies, where certain transformer layers are kept frozen while others are fine-tuned during training. The training process utilizes a regression-based loss function and an adaptive optimization algorithm. Model performance is evaluated by measuring the correlation between predicted scores and subjective scores, as well as analyzing computational efficiency in terms of training stability and resource utilization.

The experimental results demonstrate that the application of appropriate layer-freezing strategies significantly influences model performance. Freezing lower-level layers while fine-tuning higher-level layers provides a balance between maintaining pretrained feature representations and adapting the model to the specific NR-IQA task. Certain freezing configurations achieve higher correlation values between predicted and subjective scores, indicating improved prediction accuracy. Moreover, partial layer freezing reduces computational complexity and training time compared to full fine-tuning, thereby improving computational efficiency without substantially degrading performance. Overall, this study provides empirical evidence regarding the impact of layer-freezing strategies on ViT-Base for NR-IQA and offers practical recommendations for achieving an effective trade-off between accuracy and efficiency in digital image quality prediction.

Keywords: *No-Reference Image Quality Assessment, Vision Transformer Base, Layer Freezing, Image Quality Prediction, Computational Efficiency*