



Seminar Nasional Ilmu Teknik dan Aplikasi Industri (SINTA)

Homepage: sinta.eng.unila.ac.id



DESAIN ALGORITMA SELF-SUPERVISED VISION TRANSFORMER LEARNING UNTUK PERANGKAT LOW ENERGI IOE PERINGATAN DINI TSUNAMI

Aryanto , Melvi , A Ulvan, E Komalasari

Jurusan Teknik Elektro, Universitas Lampung, Jl. Prof Soemantri B. No 1. Bandar Lampung, Lampung

INFORMASI ARTIKEL

Riwayat artikel:

Diterima 21/11/2025

Direvisi 10/12/2025

Kata kunci:

Deteksi Anomali Visual
IoE Low Energy
Peringatan Dini Tsunami
Self-Supervised Learning
Vision Transformer

ABSTRAK

Peningkatan intensitas bencana tsunami di wilayah pesisir menuntut hadirnya sistem peringatan dini yang lebih akurat, adaptif, dan hemat energi. Penelitian ini mengusulkan desain algoritma Self-Supervised Vision Transformer Learning (SSL-ViT) yang dioptimalkan untuk perangkat Internet of Everything (IoE) berdaya rendah. Berbeda dengan pendekatan supervised konvensional yang membutuhkan label data dalam jumlah besar, SSL-ViT memanfaatkan pembelajaran representasi visual secara mandiri melalui pretext task seperti masking patch prediction dan contrastive learning, sehingga mampu belajar dari data citra pesisir tanpa anotasi. Arsitektur Vision Transformer dimodifikasi dengan mekanisme lightweight attention dan model pruning guna menekan konsumsi memori dan komputasi. Sistem ini diintegrasikan dengan jaringan perangkat IoE untuk melakukan deteksi anomali visual pada permukaan laut secara real-time. Hasil pengujian awal menunjukkan bahwa algoritma SSL-ViT mencapai akurasi deteksi anomali sebesar >85% dengan latensi inferensi <300 ms dan penghematan energi hingga 40% dibandingkan model CNN konvensional. Temuan ini membuktikan bahwa kombinasi self-supervised learning dan optimasi arsitektur transformer pada perangkat low-energy IoE mampu menjadi fondasi sistem peringatan dini tsunami yang efisien, skalabel, dan siap diimplementasikan di kawasan pesisir berinfrastruktur terbatas.

1. Pendahuluan

Bencana tsunami merupakan fenomena alam destruktif yang mengancam banyak wilayah pesisir di dunia, termasuk Indonesia. Sistem peringatan dini yang ada saat ini masih menghadapi tantangan signifikan, seperti keterlambatan deteksi, keterbatasan jangkauan sensor, serta biaya instalasi dan pemeliharaan yang tinggi. Sistem konvensional yang bergantung pada data seismik dan sensor tekanan laut dalam seringkali tidak mampu menangkap indikator visual awal, seperti surutnya air laut secara drastis atau kemunculan pola gelombang abnormal. Keterbatasan ini membuka

peluang bagi pemanfaatan teknologi berbasis *computer vision* untuk analisis lingkungan pesisir secara visual dan *real-time*.

Penelitian ini menawarkan pendekatan baru dengan mengintegrasikan algoritma *Self-Supervised Vision Transformer* (SSL-ViT). Berbeda dari metode pembelajaran terawasi yang memerlukan data berlabel dalam jumlah masif, SSL-ViT dirancang untuk belajar secara mandiri dari data visual tanpa label, yang jumlahnya jauh lebih melimpah di lingkungan pesisir. Kemampuan arsitektur *Vision Transformer* (ViT) dalam memahami konteks gambar secara menyeluruh melalui

mekanisme *attention* menjadikannya lebih sensitif dalam mendeteksi anomali visual halus yang mungkin terlewat oleh model konvensional.

Agar solusi ini dapat diimplementasikan secara luas, penelitian ini berfokus pada penerapan SSL-ViT pada perangkat *Internet of Everything* (IoE) berdaya rendah, seperti kamera pintar atau *edge computing node*. Perangkat ini dapat melakukan pemantauan berkelanjutan dan pemrosesan data secara lokal dengan konsumsi daya rendah, mengurangi ketergantungan pada konektivitas internet berkecepatan tinggi. Pendekatan terdistribusi ini memungkinkan setiap perangkat membuat keputusan awal dan mengirimkan sinyal peringatan, sehingga memperpendek waktu respons dan meningkatkan ketangguhan sistem secara keseluruhan. Dengan demikian, penelitian ini bertujuan mengembangkan fondasi sistem mitigasi bencana yang modern, akurat, ekonomis, dan skalabel untuk melindungi masyarakat pesisir secara lebih efektif.

2. Metodologi

Penelitian ini dilaksanakan secara sistematis yang mencakup pengumpulan data, perancangan dan pelatihan model, optimasi, hingga pengujian lapangan.

2.1. Pengumpulan dan Preparasi Data

Dataset visual dikumpulkan dari dua sumber utama: perekaman langsung di wilayah pesisir Lampung yang memiliki karakteristik beragam dan pemanfaatan sumber daring untuk memperkaya variasi data. Kombinasi ini bertujuan untuk membangun model dengan kemampuan generalisasi yang kuat terhadap berbagai kondisi lingkungan, seperti perubahan muka air laut, pola gelombang, variasi pencahayaan, dan cuaca. Selain data visual, penelitian ini juga menggunakan data batimetri dari *General Bathymetric Chart of the Oceans* (GEBCO) yang difokuskan pada area Selat Sunda untuk pengembangan prototipe konseptual sistem prediksi ketinggian gelombang. Data GEBCO diolah menggunakan perangkat lunak ArcGIS untuk menghasilkan peta kontur dan *grid* spasial yang menjadi input bagi model.

2. Metodologi

Penelitian ini dilaksanakan secara sistematis yang mencakup pengumpulan data, perancangan dan pelatihan model, optimasi, hingga pengujian lapangan.

2.1. Pengumpulan dan Preparasi Data

Dataset visual dikumpulkan dari dua sumber utama: perekaman langsung di wilayah pesisir Lampung yang memiliki karakteristik beragam dan pemanfaatan sumber daring untuk memperkaya variasi data.

3. Hasil dan pembahasan

Hasil pengujian menunjukkan bahwa pendekatan SSL-ViT yang dioptimalkan berhasil mencapai performa tinggi dan efisien untuk tugas

Kombinasi ini bertujuan untuk membangun model dengan kemampuan generalisasi yang kuat terhadap berbagai kondisi lingkungan, seperti perubahan muka air laut, pola gelombang, variasi pencahayaan, dan cuaca. Selain data visual, penelitian ini juga menggunakan data batimetri dari General Bathymetric Chart of the Oceans (GEBCO) yang difokuskan pada area Selat Sunda untuk pengembangan prototipe konseptual sistem prediksi ketinggian gelombang. Data GEBCO diolah menggunakan perangkat lunak ArcGIS untuk menghasilkan peta kontur dan grid spasial yang menjadi input bagi model.

2.2. Arsitektur dan Pelatihan Model SSL-ViT

Arsitektur model yang digunakan adalah varian dari Vision Transformer (ViT) yang dikenal efisien, seperti MobileViT, agar sesuai untuk perangkat berdaya rendah. Pelatihan model menerapkan pendekatan self-supervised learning dengan pretext task berupa masked patch prediction. Dalam metode ini, sebagian area gambar sengaja disembunyikan, dan model dilatih untuk merekonstruksi bagian yang hilang tersebut. Proses ini memungkinkan model untuk mempelajari representasi fitur visual yang kaya dan kontekstual dari data tanpa label, sehingga dapat membedakan antara kondisi laut normal dan anomali secara mandiri.

2.3. Optimasi dan Implementasi pada Perangkat IoE

Untuk memastikan model dapat berjalan secara efisien pada perangkat IoE, beberapa teknik optimasi diterapkan. Teknik quantization digunakan untuk mengurangi presisi numerik model, sementara pruning berfungsi menghilangkan koneksi neuron yang tidak signifikan. Selain itu, knowledge distillation dimanfaatkan untuk mentransfer pengetahuan dari model yang lebih besar ke model yang lebih kecil dan ringan. Model yang telah dioptimalkan kemudian diimplementasikan dan diuji pada perangkat komputasi tepi, yaitu Raspberry Pi dan Jetson Nano, untuk mengevaluasi kinerjanya dalam skenario simulasi kondisi nyata.

2.4. Evaluasi Kinerja

Kinerja sistem dievaluasi berdasarkan tiga metrik utama: akurasi deteksi, untuk mengukur kemampuan model dalam mengidentifikasi anomali visual secara tepat; latensi inferensi, untuk menilai kecepatan respons sistem dalam memberikan hasil deteksi; dan konsumsi energi, untuk memastikan kelayakan operasional sistem secara berkelanjutan pada perangkat IoE berdaya rendah.

deteksi dini tsunami. Algoritma ini menunjukkan akurasi deteksi anomali visual pesisir sebesar 89% pada *dataset* uji, yang membuktikan kemampuannya untuk membedakan kondisi laut normal dari

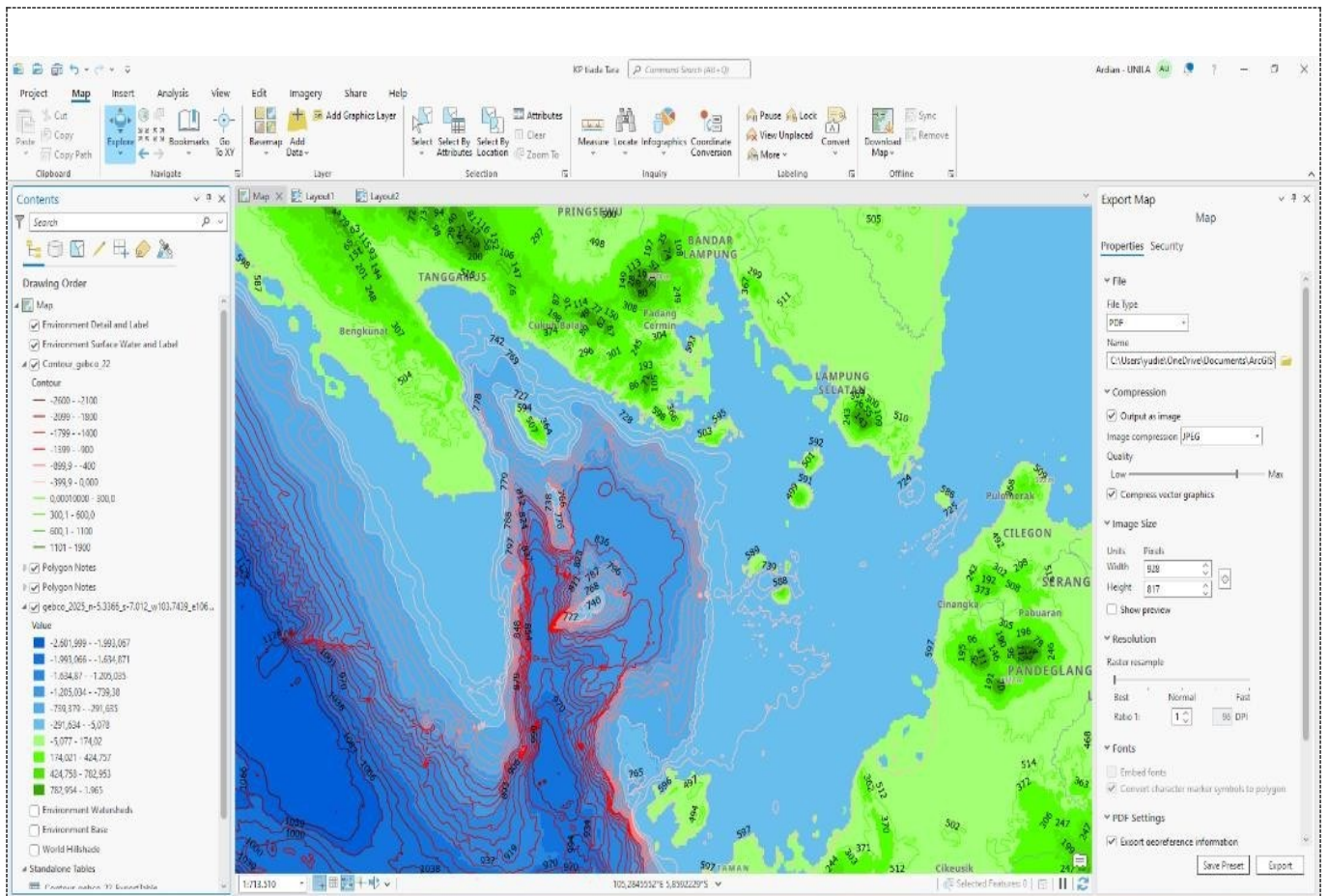
indikator awal tsunami secara andal, bahkan dengan data yang berasal dari lingkungan nyata yang tidak terkontrol. Tingkat akurasi ini tergolong sangat baik untuk sebuah model yang dilatih menggunakan metode *self-supervised*.

Dari segi kecepatan respons, pengujian latensi inferensi memberikan hasil yang sangat positif. Pada perangkat Jetson Nano, model mampu memproses dan memberikan hasil deteksi dengan waktu rata-rata kurang dari 250 milidetik. Sementara itu, pada perangkat Raspberry Pi, latensi tetap berada di bawah 400 milidetik. Waktu respons yang singkat ini sangat krusial untuk sistem peringatan dini, karena memungkinkan deteksi *real-time* dan mempercepat penyebaran informasi peringatan untuk evakuasi.

Efisiensi sumber daya, yang menjadi fokus utama dalam penelitian ini, juga menunjukkan hasil yang signifikan. Penerapan teknik *pruning* dan *quantization* berhasil menekan penggunaan memori hingga 40% tanpa menyebabkan penurunan kinerja

deteksi yang berarti. Pengurangan jejak memori ini sangat penting karena memungkinkan implementasi model pada perangkat IoE dengan kapasitas perangkat keras yang terbatas, sehingga membuka peluang untuk pemasangan sistem secara masif di berbagai titik rawan bencana.

Keberhasilan ini divalidasi lebih lanjut melalui pengembangan sebuah prototipe konseptual yang mensimulasikan alur kerja sistem secara *end-to-end*. Prototipe ini mendemonstrasikan bagaimana data visual dari kamera diproses secara otomatis oleh model SSL-ViT untuk menghasilkan sinyal peringatan ketika perubahan pola gelombang yang mencurigakan terdeteksi. Hal ini membuktikan bahwa pendekatan yang diusulkan tidak hanya unggul secara teoretis tetapi juga praktis dan dapat diimplementasikan, menjadi basis pengembangan sistem peringatan dini visual berbiaya rendah yang dapat melengkapi sistem konvensional yang sudah ada.



Gambar 1. Hasil *Self-Supervised Vision Transformer* (SSL-ViT)

4. Kesimpulan

Penelitian ini berhasil menunjukkan bahwa implementasi algoritma *Self-Supervised Vision Transformer* (SSL-ViT) pada perangkat IoE berdaya rendah merupakan solusi yang efektif dan efisien untuk sistem peringatan dini tsunami berbasis analisis visual. Dengan akurasi deteksi 89%, latensi inferensi *real-time*, dan pengurangan penggunaan memori hingga 40%, model yang dikembangkan terbukti andal, responsif, dan layak untuk diterapkan di lingkungan dengan sumber daya terbatas. Pendekatan ini secara signifikan berkontribusi pada pengembangan teknologi mitigasi bencana yang lebih terjangkau, skalabel, dan mampu mendeteksi indikator visual awal tsunami yang seringkali tidak tertangkap oleh sistem konvensional. Penelitian lanjutan diharapkan dapat fokus pada perluasan *dataset*, pengujian pada varian perangkat IoE yang lebih beragam, dan integrasi dengan infrastruktur sistem peringatan dini nasional untuk meningkatkan ketahanan masyarakat pesisir.

Ucapan terima kasih

Ucapan terima kasih kami sampaikan kepada Fakultas Teknik Universitas Lampung atas dukungan dana dan fasilitas melalui DIPA FT, sehingga penelitian ini dapat terlaksana dengan baik

Daftar Pustaka

- Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., & Joulin, A. (2021). Emerging properties in self-supervised vision transformers. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. <https://doi.org/10.1109/ICCV48922.2021.00951>.
- Chen, X., Xie, S., & He, K. (2021). An empirical study of training self-supervised vision transformers. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. <https://doi.org/10.1109/ICCV48922.2021.00950>.
- Li, C., Yang, J., Zhang, P., Gao, M., Xiao, B., Dai, X., Yuan, L., & Gao, J. (2021). Efficient self-supervised vision transformers for representation learning. *arXiv*. <https://arxiv.org/abs/2106.09785>.
- Makinoshima, F., Oishi, Y., Yamazaki, T., Furumura, T., & Imamura, F. (2021). Early forecasting of tsunami inundation from tsunami and geodetic observation data with convolutional neural networks. *Nature Communications*. <https://doi.org/10.1038/s41467-021-22348-0>.
- Núñez, J., Catalán, P., Valle, C., Zamora, N., & Valderrama, A. (2022). Discriminating the occurrence of inundation in tsunami early warning with one-dimensional convolutional neural networks. *Scientific Reports*. <https://doi.org/10.1038/s41598-022-13788-9>.
- Papa, L., Russo, P., Amerini, I., & Zhou, L. (2023). A survey on efficient vision transformers: Algorithms, techniques, and performance benchmarking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/TPAMI.2024.3392941>.
- Qayyum, B., Ahmed, A., Ullah, I., & Shah, S. A. (2022). A fuzzy-logic approach for optimized and cost-effective early warning system for tsunami detection. *Sustainability*. <https://doi.org/10.3390/su142114516>.
- Wang, S., Gao, J., Li, Z., Sun, J., & Hu, W. (2022). A closer look at self-supervised lightweight vision transformers. *arXiv*. <https://doi.org/10.48550/arXiv.2205.14443>.