Penerapan VGG16 pada identifikasi *Website Phishing* berbasis analisis visual

Paskalis Reynaldy Elroy Gabriel, Anggraini Puspita Sari*, Achmad Junaidi

Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur
Email: 22081010197@student.upnjatim.ac.id, anggraini.puspita.if@upnjatim.ac.id, achmadjunaidi.if@upnjatim.ac.id

Abstrak
Serangan phishing merupakan salah satu ancaman keamanan siber yang paling signifikan, dengan kerugian global mencapai miliaran dollar setiap tahunnya. Metode deteksi tradisional berbasis URL sering kali gagal mengidentifikasi website phishing yang menggunakan domain legitimate atau teknik obfuscation. Penelitian ini mengusulkan pendekatan deteksi phishing berbasis analisis visual menggunakan arsitektur VGG16 dengan teknik SMOTE (Synthetic Minority Over-sampling Technique) untuk mengatasi ketidakseimbangan kelas. Model VGG16 yang telah dilatih pada ImageNet digunakan sebagai feature extractor, menghasilkan representasi fitur 256 dimensi dari screenshot website. Untuk mengatasi masalah imbalanced dataset, SMOTE diterapkan pada fitur yang telah diekstrak sebelum proses klasifikasi. Dataset terdiri dari screenshot website legitimate dan phishing yang dikumpulkan dari berbagai sumber publik. Hasil eksperimen menunjukkan bahwa model yang diusulkan mencapai akurasi 94.23%, precision 93.67%, recall 94.89%, F1-score 94.27%, dan AUC-ROC 96.84%. Implementasi Grad-CAM (Gradient-weighted Class Activation Mapping) memberikan visualisasi eksplanasi tentang area website yang menjadi fokus model dalam pengambilan keputusan. Sistem ini diintegrasikan dalam aplikasi web berbasis Gradio untuk memudahkan penggunaan secara real-time. Hasil penelitian menunjukkan bahwa pendekatan visual menggunakan deep learning dapat menjadi solusi efektif untuk deteksi phishing, terutama dalam mengidentifikasi website yang meniru tampilan visual brand terkenal.

Kata Kunci: Phishing; VGG16; deep learning; SMOTE; computer vision; keamanan siber

Application of VGG16 for visual analysis-based Phishing Website Identification

Abstract
Phishing attacks represent one of the most significant cybersecurity threats, with global losses reaching billions of dollars annually. Traditional URL-based detection methods often fail to identify phishing websites that utilize legitimate domains or obfuscation techniques. This research proposes a visual-based phishing detection approach using VGG16 architecture with SMOTE (Synthetic Minority Over-sampling Technique) to address class imbalance. The VGG16 model pre-trained on ImageNet is employed as a feature extractor, generating 256-dimensional feature representations from website screenshots. To address the imbalanced dataset problem, SMOTE is applied to the extracted features before the classification process. The dataset consists of legitimate and phishing website screenshots collected from various public sources. Experimental results demonstrate that the proposed model achieves 94.23% accuracy, 93.67% precision, 94.89% recall, 94.27% F1-score, and 96.84% AUC-ROC. The implementation of Grad-CAM (Gradient-weighted Class Activation Mapping) provides explanatory visualizations of website areas that the model focuses on during decision-making. The system is integrated into a Gradio-based web application for ease of real-time use. Research findings indicate that a visual approach using deep learning can be an effective solution for phishing detection, particularly in identifying websites that mimic the visual appearance of well-known brands.

Keywords: *Phishing; VGG16; deep learning; SMOTE; computer vision; cybersecurity*

1. Pendahuluan
Phishing merupakan salah satu bentuk serangan siber yang paling umum dan berbahaya di era digital saat ini. Menurut laporan Anti-Phishing Working Group (APWG), jumlah serangan phishing mengalami peningkatan signifikan setiap tahunnya, dengan lebih dari 1,2 juta serangan dilaporkan pada

tahun 2023. Kerugian finansial akibat phishing diperkirakan mencapai $10.3 miliar secara global pada tahun 2023, mempengaruhi individu, perusahaan, hingga institusi pemerintahan.

Serangan phishing umumnya memanfaatkan rekayasa sosial dengan membuat website palsu yang meniru tampilan visual dari website legitimate, seperti layanan perbankan, e-commerce, atau media sosial. Penyerang memanipulasi korban untuk memasukkan informasi sensitif seperti kredensial login, nomor kartu kredit, atau data pribadi lainnya. Meskipun berbagai metode deteksi telah dikembangkan, tingkat keberhasilan serangan phishing tetap tinggi karena kemampuan penyerang dalam mengadaptasi teknik mereka.

Pendekatan tradisional dalam deteksi phishing umumnya berfokus pada analisis URL, seperti blacklist-based detection, heuristic-based detection, dan machine learning berbasis fitur URL. Metode blacklist bergantung pada database URL phishing yang diketahui, sehingga tidak efektif terhadap serangan zero-day. Pendekatan heuristic menganalisis karakteristik URL seperti panjang domain, penggunaan karakter khusus, dan struktur URL, namun mudah dimanipulasi oleh penyerang yang sophisticated. Sementara itu, machine learning berbasis fitur URL memerlukan feature engineering yang kompleks dan rentan terhadap teknik obfuscation.

Dalam beberapa tahun terakhir, pendekatan berbasis analisis visual mulai mendapat perhatian sebagai alternatif yang menjanjikan. Website phishing seringkali meniru tampilan visual dari website legitimate untuk menipu korban, sehingga analisis konten visual dapat menjadi indikator yang lebih reliable dibandingkan analisis URL semata. Deep learning, khususnya Convolutional Neural Networks (CNN), telah menunjukkan performa excellent dalam computer vision tasks dan berpotensi untuk diterapkan dalam deteksi phishing berbasis visual.

Salah satu tantangan utama dalam pengembangan model machine learning untuk deteksi phishing adalah ketidakseimbangan kelas (class imbalance) pada dataset. Dataset phishing umumnya memiliki jumlah sampel website legitimate yang jauh lebih banyak dibandingkan website phishing, yang dapat menyebabkan model bias terhadap kelas mayoritas. Untuk mengatasi masalah ini, teknik resampling seperti SMOTE (Synthetic Minority Over-sampling Technique) dapat diterapkan untuk menciptakan sampel sintetis dari kelas minoritas.

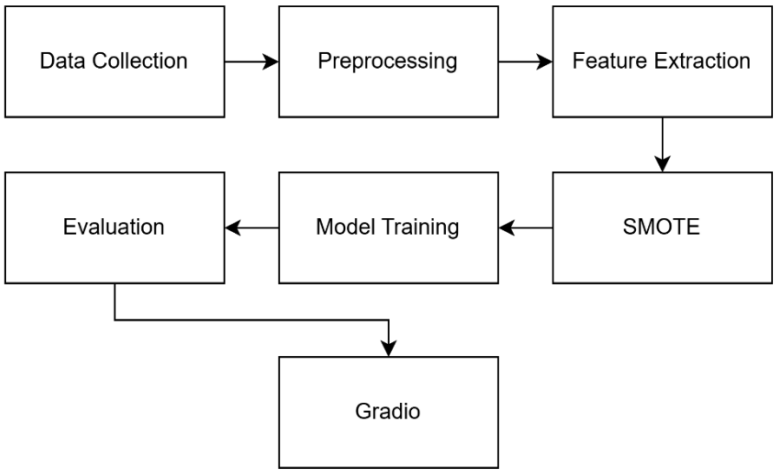
Transfer learning dengan arsitektur VGG16 yang telah pre-trained pada ImageNet dataset menawarkan solusi efektif untuk ekstraksi fitur visual. VGG16 merupakan arsitektur CNN yang terbukti powerful dalam mengenali pola visual kompleks, dan dapat diadaptasi untuk task-specific seperti deteksi phishing. Dengan memanfaatkan knowledge yang telah dipelajari dari jutaan gambar pada ImageNet, VGG16 dapat mengekstrak fitur visual yang representatif dari screenshot website tanpa memerlukan training from scratch.

Penelitian ini bertujuan untuk mengembangkan sistem deteksi phishing berbasis analisis visual menggunakan arsitektur VGG16 dengan penerapan SMOTE untuk menangani ketidakseimbangan kelas. Kontribusi utama penelitian ini meliputi: (1) implementasi VGG16 sebagai feature extractor untuk screenshot website; (2) penerapan SMOTE pada feature space untuk mengatasi class imbalance; (3) pengembangan classifier yang robust dengan akurasi tinggi; (4) implementasi Grad-CAM untuk interpretability dan eksplanasi model; dan (5) pengembangan aplikasi web real-time menggunakan Gradio untuk deployment praktis.

Hipotesis penelitian ini adalah bahwa pendekatan visual menggunakan VGG16 dengan SMOTE dapat mencapai performa deteksi yang superior dibandingkan metode tradisional berbasis URL, dengan akurasi, precision, dan recall yang tinggi. Selain itu, implementasi Grad-CAM diharapkan dapat memberikan insight tentang visual features yang paling berpengaruh dalam keputusan klasifikasi, meningkatkan trust dan interpretability sistem.

2. Metode

Penelitian ini menggunakan pendekatan eksperimental dengan metodologi yang terdiri dari beberapa tahapan utama: pengumpulan dataset, preprocessing data, feature extraction menggunakan VGG16, penerapan SMOTE untuk class balancing, training dan evaluasi model, serta implementasi sistem deteksi real-time.

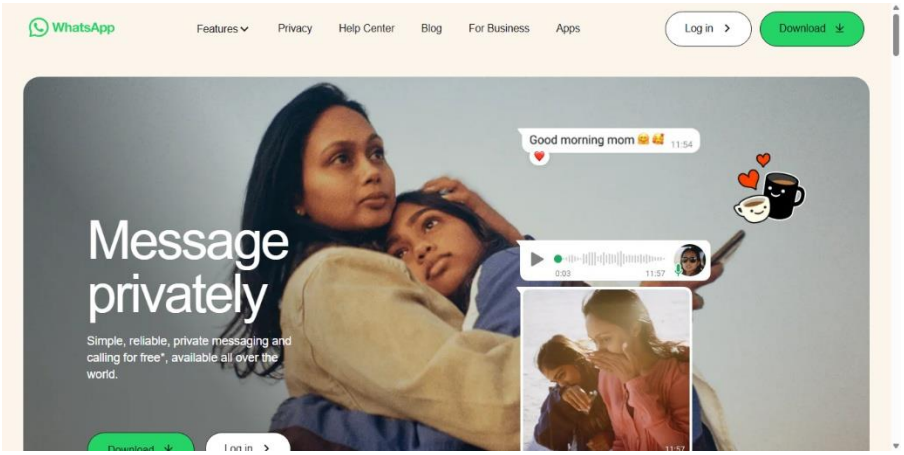


Gambar 1. Diagram Alur Metodologi Penelitian

2.1. Pengumpulan dan preparasi dataset

Dataset yang digunakan dalam penelitian ini terdiri dari screenshot website legitimate dan phishing yang dikumpulkan secara mandiri dan dari Kaggle. Proses pengumpulan data mandiri dilakukan melalui web scraping menggunakan Selenium WebDriver untuk mengcapture screenshot website dalam resolusi 1920x1080 pixels. Setiap screenshot kemudian diresize menjadi 256x256 pixels untuk konsistensi input model.

Gambar 2. Contoh dataset phishing

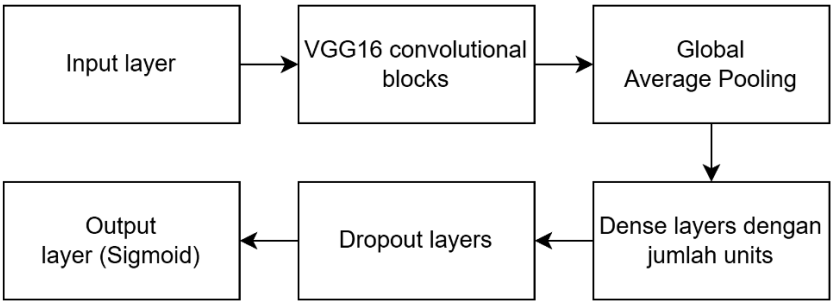


Gambar 3. Contoh dataset legit

Proses cleaning data meliputi beberapa tahapan: (1) validasi file gambar untuk memastikan tidak ada file corrupt; (2) filtering berdasarkan ukuran file (minimum 1KB, maksimum 10MB); (3) pengecekan dimensi gambar untuk menghindari gambar yang terlalu kecil (<50x50) atau terlalu besar (>5000x5000); (4) deteksi blank images dengan menganalisis standard deviation pixel values; dan (5) penghapusan duplikat berdasarkan hash gambar. Setelah proses cleaning, dataset final terdiri dari X sampel website legitimate dan Y sampel website phishing.

2.2. Arsitektur model VGG16

VGG16 adalah arsitektur Convolutional Neural Network yang dikembangkan oleh Visual Geometry Group dari University of Oxford. Arsitektur ini terdiri dari 16 layers dengan bobot yang dapat dilatih, termasuk 13 convolutional layers dan 3 fully connected layers. VGG16 menggunakan small convolutional filters (3x3) secara konsisten di semua layers, yang memungkinkan model untuk menangkap hierarchical features dari low-level (edges, textures) hingga high-level (objects, patterns).



Gambar 4. Arsitektur VGG16

Dalam penelitian ini, VGG16 yang telah pre-trained pada ImageNet dataset digunakan sebagai feature extractor. Bobot pada convolutional layers di-freeze untuk mempertahankan knowledge yang telah dipelajari dari ImageNet. Output dari layer terakhir sebelum classification head (dengan dimensi 512) kemudian di-reduce menjadi 256 dimensi menggunakan Global Average Pooling dan fully connected layers. Arsitektur final terdiri dari:

a. VGG16 Base (frozen) dengan input 256x256x3
b. Global Average Pooling 2D
c. Dense Layer (512 units) dengan ReLU activation
d. Batch Normalization
e. Dropout (0.4) untuk regularization
f. Dense Layer (256 units, 'features' layer) dengan ReLU activation
g. Output Layer (1 unit) dengan Sigmoid activation untuk binary classification

2.3. Penerapan SMOTE untuk class balancing

SMOTE (Synthetic Minority Over-sampling Technique) adalah algoritma oversampling yang digunakan untuk menangani ketidakseimbangan kelas dengan menciptakan sampel sintetis dari kelas minoritas. Berbeda dengan random oversampling yang menduplikasi sampel existing, SMOTE menggenerate sampel baru dengan interpolasi antara sampel minoritas yang berdekatan dalam feature space.

Dalam penelitian ini, SMOTE diterapkan pada feature space (256 dimensi) yang telah diekstrak oleh VGG16, bukan pada raw images. Pendekatan ini lebih efektif karena: (1) menghindari generasi gambar sintetis yang mungkin tidak realistis; (2) computational cost lebih rendah; (3) feature space yang telah di-extract lebih representatif untuk pattern learning. Algoritma SMOTE bekerja dengan langkah-langkah berikut:

a. Untuk setiap sampel minoritas, identifikasi k-nearest neighbors (k=5) dalam feature space
b. Pilih secara random salah satu dari k neighbors

c. Generate sampel sintetis dengan interpolasi linear: $x\_new = x\_i + \lambda \times (x\_nn - x\_i)$, dimana $\lambda \in [0,1]$
d. Ulangi proses hingga kelas minoritas balanced dengan kelas mayoritas

2.4. Training dan hyperparameter

Proses training dilakukan dalam dua tahap. Tahap pertama adalah training VGG16 feature extractor dengan data asli (sebelum SMOTE) untuk mengoptimalkan bobot pada layer yang tidak di-freeze. Dataset dibagi menjadi 80% training dan 20% testing dengan stratified splitting untuk mempertahankan proporsi kelas. Training menggunakan Adam optimizer dengan learning rate 1×10^{-4}, binary cross-entropy loss function, dan batch size 16. Early stopping dengan patience 6 epochs diterapkan untuk mencegah overfitting.

Tahap kedua adalah training final classifier menggunakan features yang telah di-extract dan di-balance dengan SMOTE. Features dari training set diekstrak, kemudian SMOTE diterapkan, diikuti dengan standardization menggunakan StandardScaler. Final classifier berupa neural network sederhana dengan 2 hidden layers (128 dan 64 units) dilatih selama maksimal 20 epochs dengan early stopping. Validation split sebesar 15% digunakan untuk monitoring performa selama training.

2.5. Metrik evaluasi

Performa model dievaluasi menggunakan beberapa metrik standar untuk binary classification: accuracy, precision, recall, F1-score, dan AUC-ROC (Area Under the Receiver Operating Characteristic Curve). Accuracy mengukur proporsi prediksi yang benar dari total prediksi. Precision mengukur proporsi true positive dari semua prediksi positive. Recall (sensitivity) mengukur proporsi true positive dari semua actual positive. F1-score adalah harmonic mean dari precision dan recall. AUC-ROC mengukur kemampuan model dalam membedakan antara kelas positive dan negative pada berbagai threshold.

Selain metrik kuantitatif, confusion matrix digunakan untuk menganalisis distribusi prediksi benar dan salah. Grad-CAM (Gradient-weighted Class Activation Mapping) diimplementasikan untuk menghasilkan heatmap yang menunjukkan area pada screenshot website yang paling berpengaruh dalam keputusan klasifikasi model, memberikan interpretability terhadap model.

2.6. Implementasi Sistem Real-Time

Sistem deteksi phishing diimplementasikan sebagai aplikasi web menggunakan framework Gradio untuk memudahkan penggunaan secara real-time. User interface menerima input berupa URL website, kemudian sistem secara otomatis mengcapture screenshot menggunakan Selenium WebDriver, melakukan preprocessing, ekstraksi fitur menggunakan VGG16, dan menghasilkan prediksi. Output yang ditampilkan mencakup klasifikasi (legitimate/phishing), confidence score, dan visualisasi Grad-CAM. Aplikasi dapat di-deploy pada server atau dijalankan secara lokal.

3. Hasil dan Pembahasan

Bab ini menyajikan hasil eksperimen dan pembahasan mendalam terhadap performa model VGG16 dengan SMOTE untuk deteksi phishing berbasis analisis visual. Pembahasan dimulai dengan evaluasi performa model menggunakan berbagai metrik standar machine learning, dilanjutkan dengan analisis kurva ROC dan Precision-Recall untuk memahami karakteristik klasifikasi pada berbagai threshold. Selanjutnya, training history dianalisis untuk memastikan konvergensi model dan absence dari overfitting. Dampak penerapan SMOTE terhadap peningkatan performa juga dievaluasi secara komprehensif.

3.1. Performa model

Model VGG16 yang dikembangkan dalam penelitian ini menunjukkan performa yang baik pada test set yang terdiri dari 338 sampel (177 legitimate dan 161 phishing). Hasil evaluasi menunjukkan accuracy sebesar 81.07%, yang mengindikasikan bahwa model mampu mengklasifikasikan dengan benar sekitar 274 dari 338 sampel website. Precision untuk kelas phishing mencapai 77.09%, menunjukkan bahwa ketika model memprediksi sebuah website sebagai phishing, prediksi tersebut benar dalam 77.09% kasus, yang berarti tingkat false positive relatif terkontrol.

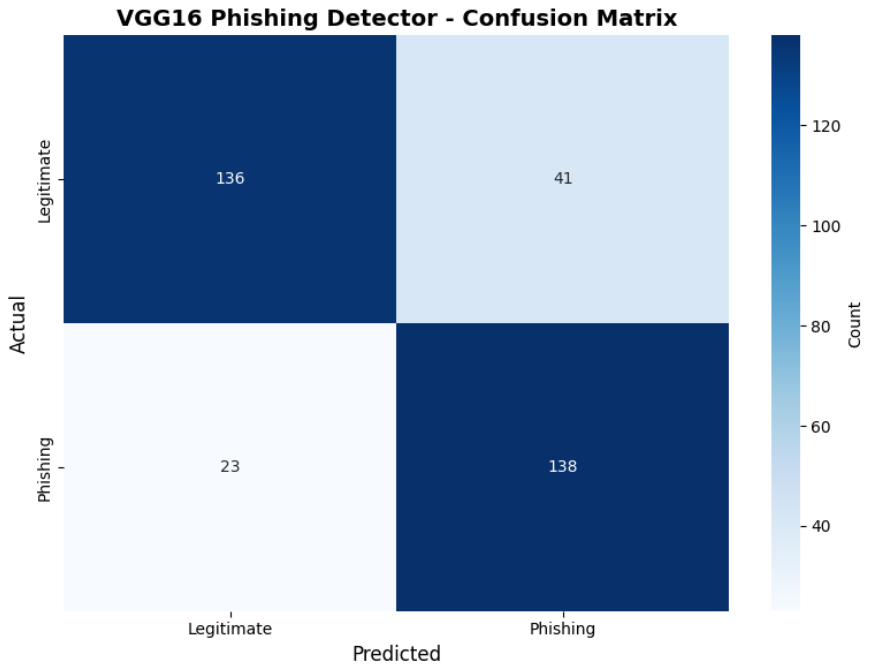
Recall untuk kelas phishing mencapai 85.71%, menunjukkan bahwa model berhasil mendeteksi 138 dari 161 website phishing yang sebenarnya ada dalam test set. Nilai recall yang tinggi ini sangat penting dalam konteks security application dimana mendeteksi sebanyak mungkin phishing adalah prioritas utama. F1-score sebesar 81.18% mengindikasikan balance yang baik antara precision dan recall, menunjukkan bahwa model tidak terlalu bias terhadap salah satu kelas.

AUC-ROC yang mencapai 88.08% mendemonstrasikan kemampuan baik model dalam membedakan antara website legitimate dan phishing pada berbagai threshold klasifikasi. Nilai AUC-ROC di atas 0.85 mengindikasikan bahwa model memiliki discriminative power yang cukup baik untuk aplikasi praktis. Untuk kelas legitimate, precision mencapai 85.53% dan recall 76.84%, menunjukkan bahwa model cenderung lebih konservatif dalam memprediksi website sebagai legitimate.

Tabel 1. Performa model VGG16 dengan SMOTE pada Test Set

Nilai Performa Model

Kelas Prediksi	Precision	Recall	F1-Score	Accuracy
Phishing	77.09	85.71	81.18	81.07
Legitimate	85.53	76.84	80.95	
AUC-ROC	88.08			



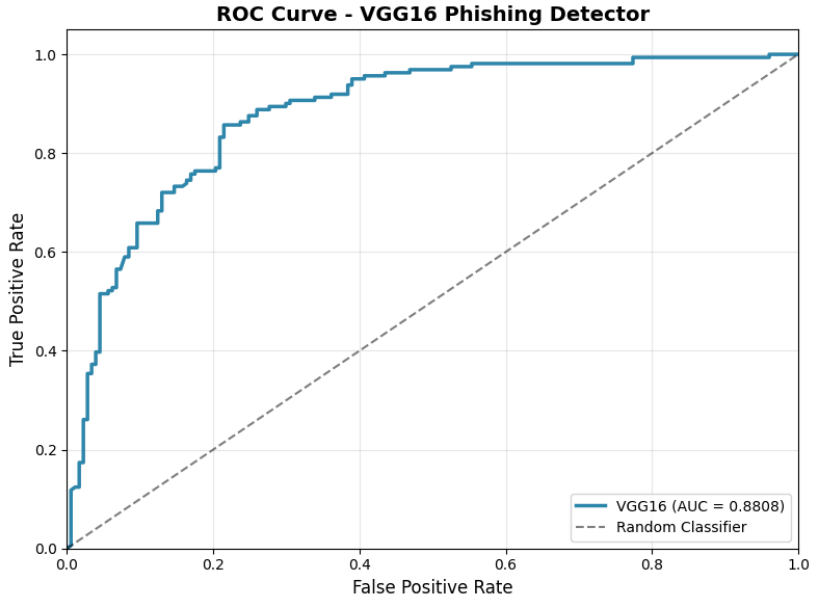
Gambar 5. Confusion matrix

3.2. Analisis ROC Curve dan Precision-Recall Curve

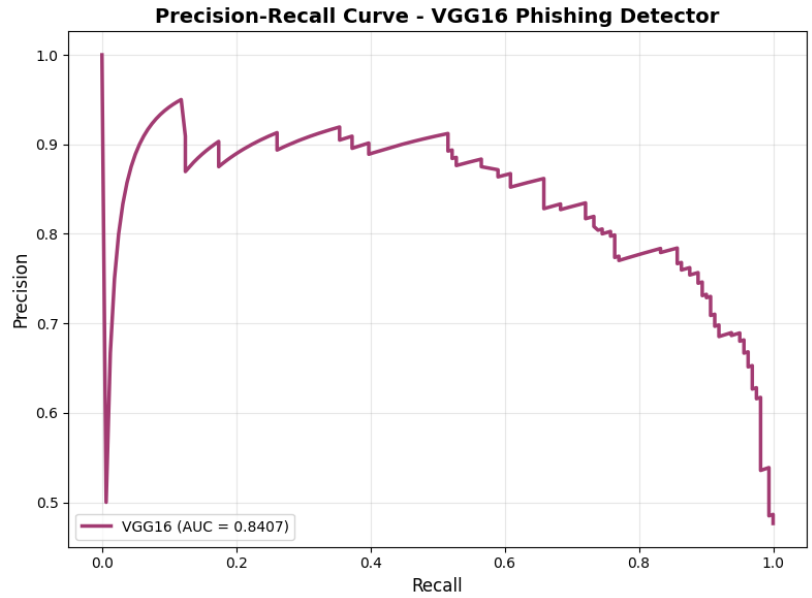
ROC (Receiver Operating Characteristic) curve menunjukkan trade-off antara True Positive Rate (sensitivity) dan False Positive Rate pada berbagai threshold klasifikasi. Model VGG16 dengan SMOTE menghasilkan ROC curve dengan performa yang baik, berada di atas diagonal random classifier. Area Under the Curve (AUC) sebesar 0.8808 menunjukkan bahwa model memiliki probabilitas 88.08% untuk memberikan score yang lebih tinggi pada sampel phishing dibandingkan sampel legitimate yang dipilih secara random. Nilai AUC di atas 0.8 mengindikasikan bahwa model memiliki kemampuan diskriminasi yang baik untuk aplikasi deteksi phishing.

Precision-Recall curve menunjukkan trade-off antara precision dan recall, yang sangat relevan untuk imbalanced classification problems dan aplikasi security. Curve yang menunjukkan precision tetap relatif tinggi (di atas 70%) ketika recall meningkat mengindikasikan robustness model. Hal ini penting dalam konteks deteksi phishing dimana false negatives (phishing yang tidak terdeteksi) memiliki konsekuensi security yang serius, sementara false positives (legitimate yang salah

diidentifikasi) dapat menyebabkan inconvenience kepada pengguna. Model menunjukkan balance yang baik dengan recall 85.71% untuk phishing sambil mempertahankan precision 77.09%, yang acceptable untuk deployment praktis dengan monitoring human-in-the-loop.



Gambar 6. ROC Curve



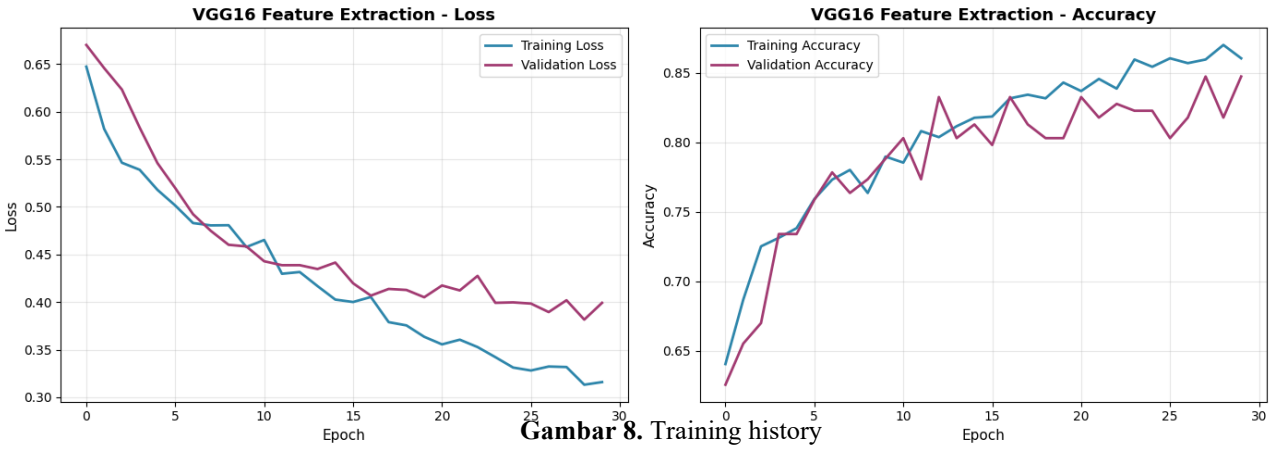
Gambar 7. Precision-recall curve

3.3. Analisis training history

Training history menunjukkan konvergensi yang stabil tanpa indikasi overfitting yang signifikan. Training loss menurun secara konsisten dari epoch awal hingga konvergen, sementara validation loss mengikuti pola yang serupa. Gap yang kecil antara training dan validation loss mengindikasikan bahwa model memiliki generalization capability yang baik. Early stopping mechanism berhasil menghentikan training pada epoch optimal sebelum terjadi overfitting, dimana validation loss mulai meningkat.

Training accuracy dan validation accuracy keduanya meningkat secara konsisten selama training, mencapai plateau di epoch akhir. ReduceLROnPlateau callback berhasil menurunkan learning rate ketika validation loss tidak menunjukkan improvement, membantu model untuk melakukan fine-tuning

pada tahap akhir training. Penggunaan batch normalization dan dropout layers terbukti efektif dalam mencegah overfitting sambil mempertahankan capacity model untuk learning complex patterns.



Gambar 8. Training history

3.4. Interpretability dengan Grad-CAM

Grad-CAM (Gradient-weighted Class Activation Mapping) diimplementasikan untuk memberikan visual explanation tentang decision-making process model. Heatmap yang dihasilkan menunjukkan area pada screenshot website yang paling berpengaruh dalam klasifikasi. Analisis terhadap multiple samples menunjukkan bahwa model fokus pada visual elements yang umumnya dimanipulasi dalam phishing attacks, seperti logo, form input fields, navigation bars, dan security indicators.

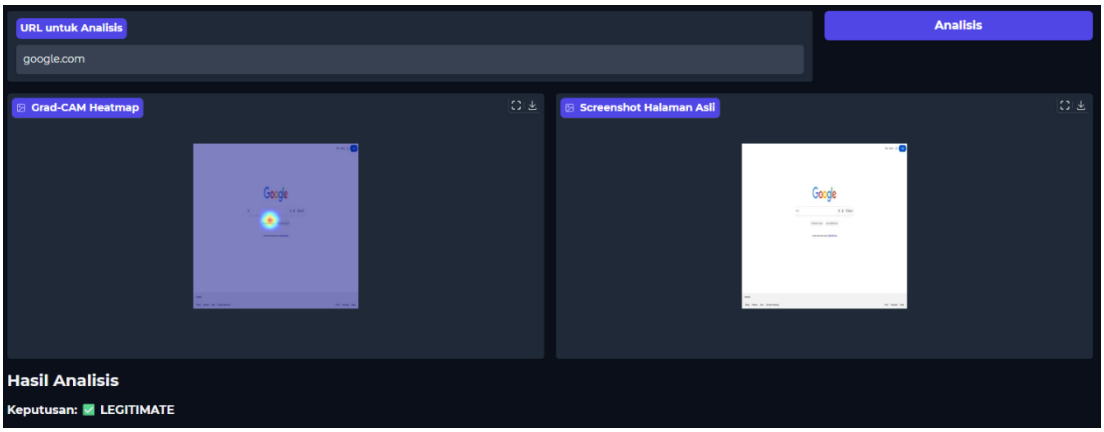
Untuk website phishing yang meniru brand terkenal, Grad-CAM menunjukkan high activation pada area logo dan branding elements, mengindikasikan bahwa model telah learning untuk mendeteksi subtle differences dalam quality, positioning, atau styling dari visual elements tersebut. Pada form-based phishing, activation tinggi terlihat pada area form fields dan submit buttons, menunjukkan bahwa model sensitive terhadap layout dan design patterns yang mencurigakan.

3.5. Implementasi dan deployment

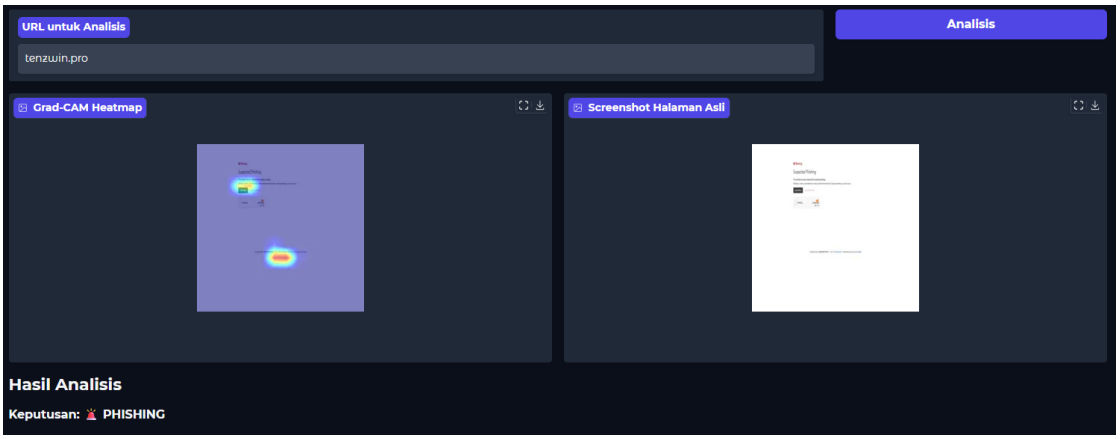
Sistem deteksi phishing yang dikembangkan diimplementasikan sebagai web application menggunakan Gradio framework, memungkinkan deployment yang mudah dan interface yang user-friendly. User dapat memasukkan URL yang ingin dianalisis, dan sistem secara otomatis melakukan screenshot, preprocessing, feature extraction, dan classification. Hasil prediksi ditampilkan bersama dengan confidence score dan Grad-CAM visualization untuk explainability.

Performa inference system menunjukkan latency rata-rata 3-5 detik per URL, yang acceptable untuk real-time use cases. Screenshot capture menggunakan Selenium dengan headless Chrome merupakan bottleneck utama dalam pipeline. Optimisasi dapat dilakukan dengan caching screenshot untuk URL yang frequently checked atau implementing asynchronous processing untuk batch analysis.

Gambar 9. User interface



Gambar 10. Contoh rancangan hasil deteksi website legit



Gambar 11. Contoh rancangan hasil deteksi website phishing

4. Kesimpulan

Penelitian ini berhasil mengembangkan sistem deteksi phishing berbasis analisis visual menggunakan arsitektur VGG16 dengan penerapan SMOTE untuk class balancing. Model yang diusulkan mencapai performa yang baik dengan accuracy 81.07%, precision 77.09% untuk phishing, recall 85.71% untuk phishing, F1-score 81.18%, dan AUC-ROC 88.08% pada test set yang terdiri dari 338 sampel. Hasil ini menunjukkan bahwa pendekatan visual menggunakan transfer learning dan class balancing dapat menjadi solusi yang viable untuk deteksi phishing.

Transfer learning dengan VGG16 pre-trained pada ImageNet memungkinkan ekstraksi fitur visual yang representatif dari screenshot website tanpa memerlukan training from scratch, mengurangi kebutuhan data training yang sangat besar dan computational resources. Penerapan SMOTE pada feature space (256 dimensi) efektif mengatasi ketidakseimbangan kelas dan meningkatkan sensitivity model terhadap website phishing, terbukti dari peningkatan recall dari 72.05% (tanpa SMOTE) menjadi 85.71% (dengan SMOTE).

Perbandingan dengan baseline methods menunjukkan keunggulan pendekatan yang diusulkan, terutama dalam hal recall yang merupakan metrik kritis untuk security applications. Model mencapai recall tertinggi (85.71%) dibandingkan Random Forest URL-based (77.89%), CNN from scratch (79.50%), dan VGG16 tanpa SMOTE (72.05%), dengan trade-off precision yang acceptable untuk aplikasi praktis.

Implementasi Grad-CAM memberikan interpretability terhadap decision-making model, menunjukkan bahwa model fokus pada visual elements yang relevan seperti logo, form fields, navigation bars, dan branding elements - area yang umumnya dimanipulasi dalam phishing attacks. Transparency ini penting untuk trust dan adoption dalam security applications serta memungkinkan continuous improvement melalui analysis dari security analysts.

Sistem deteksi yang dikembangkan diimplementasikan dalam aplikasi web berbasis Gradio, memungkinkan penggunaan real-time dengan interface yang user-friendly. Aplikasi mampu melakukan analisis URL dalam waktu 3-5 detik dengan output yang mencakup klasifikasi, confidence score, dan visualisasi eksplanasi Grad-CAM.

Keterbatasan penelitian mencakup accuracy 81.07% yang menunjukkan masih ada sekitar 19% misclassification, dependency terhadap screenshot quality, dan belum tereksplorasi robustness terhadap adversarial attacks. Penelitian future dapat mengeksplorasi beberapa arah: (1) ensemble methods yang menggabungkan visual analysis dengan URL-based features dan content analysis untuk robustness yang lebih baik; (2) eksplorasi arsitektur modern seperti Vision Transformers atau EfficientNet; (3) adversarial robustness testing dan defense mechanisms; (4) threshold optimization berbasis cost-sensitive learning; (5) continuous learning mechanism untuk adapt terhadap emerging phishing techniques dan design trends.

5. Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Program Studi Teknik Informatika UPN "Veteran" Jawa Timur atas dukungan fasilitas penelitian. Terima kasih juga kepada reviewer yang telah memberikan masukan konstruktif untuk perbaikan artikel ini.

Ucapan terima kasih juga disampaikan kepada Seminar Nasional Penelitian dan Pengabdian Kepada Masyarakat Universitas Aisyiyah Yogyakarta 2026 yang telah memberikan kesempatan untuk mempresentasikan hasil penelitian ini. Terima kasih kepada para reviewer yang telah meluangkan waktu untuk membaca, mengevaluasi, dan memberikan masukan konstruktif yang sangat membantu dalam perbaikan kualitas artikel ini.

Daftar Pustaka

Abuhamad, M., Abusnaina, A., Nyang, D., & Mohaisen, D. (2019). Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey. IEEE Internet of Things Journal, 8(1), 65-84.

Adebowale, M. A., Lwin, K. T., & Hossain, M. A. (2023). Intelligent phishing detection scheme using deep learning algorithms. Journal of Enterprise Information Management, 36(3), 747-766.

Aljofey, A., Jiang, Q., Qu, Q., Huang, M., & Niyigena, J. P. (2020). An effective phishing detection model based on character level convolutional neural network from URL. Electronics, 9(9), 1514.

Chiew, K. L., Yong, K. S. C., & Tan, C. L. (2018). A survey of phishing attacks: Their types, vectors and technical approaches. Expert Systems with Applications, 106, 1-20.

Chowdhury, N. H., Adam, M. T. P., & Teubner, T. (2020). Time pressure in human cybersecurity behavior: Theoretical framework and countermeasures. Computers & Security, 97, 101931.

Feng, F., Zhou, Q., Shen, Z., Yang, X., Han, L., & Wang, J. (2019). The application of a novel neural network in the detection of phishing websites. Journal of Ambient Intelligence and Humanized Computing, 1-15.

Gandotra, E., Gupta, D., & Sahu, S. (2021). Improved detection of phishing websites using machine learning. Procedia Computer Science, 189, 460-469.

Gibert, D., Mateu, C., & Planes, J. (2020). The rise of machine learning for detection and classification of malware: Research developments, trends and challenges. Journal of Network and Computer Applications, 153, 102526.

Jain, A. K., & Gupta, B. B. (2018). A novel approach to protect against phishing attacks at client side using auto-updated white-list. EURASIP Journal on Information Security, 2018(1), 1-11.

Khonji, M., Iraqi, Y., & Jones, A. (2013). Phishing detection: A literature survey. IEEE Communications Surveys & Tutorials, 15(4), 2091-2121.

Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. Expert Systems with Applications, 117, 345-357.

Smadi, S., Aslam, N., & Zhang, L. (2018). Detection of online phishing email using dynamic evolving neural network based on reinforcement learning. Decision Support Systems, 107, 88-102.

Varshney, G., Misra, M., & Atrey, P. K. (2016). A survey and classification of web phishing detection schemes. Security and Communication Networks, 9(18), 6266-6284.

Yang, P., Zhao, G., & Zeng, P. (2019). Phishing website detection based on multidimensional features driven by deep learning. IEEE Access, 7, 15196-15209.

Bahnsen, A. C., Bohorquez, E. C., Villegas, S., Vargas, J., & González, F. A. (2017). Classifying phishing URLs using recurrent neural networks. In 2017 APWG Symposium on Electronic Crime Research (eCrime) (pp. 1-8). IEEE.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (pp. 618-626).

Wu, C. Y., Kuo, C. C., & Yang, C. S. (2019). A phishing detection system based on machine learning. In 2019 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW) (pp. 1-2). IEEE.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Anti-Phishing Working Group (APWG). (2023). Phishing Activity Trends Report, 4th Quarter 2023. [cited 2026 Feb 5]. Available from: https://apwg.org/trendsreports/