Jurnal KomtekInfo
https://jkomtekinfo.org/ojs
2023 Vol. 10 No. 4 Hal: 165-169 p-ISSN: 2356-0010, e-ISSN: 2502-8758

Ilusi Keterjelasan Kuantifikasi Penjajaran Grad-CAM: Studi Kasus Masker Segmentasi Warna Otomatis Citra Penyakit Padi

Agung Ramadhanu[1]✉, Rely Maryani[2]

[1]Magister Teknik Informatika, Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK, Jl. Raya Lubuk Begalung, 25221, Padang, Indonesia

[2]Teknik Informatika, Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK, Jl. Raya Lubuk Begalung, 25221, Padang, Indonesia

✉ agung_ramadhanu@upiyptk.ac.id

Abstract

Saliency maps are widely used to explain convolutional networks in plant disease studies, yet they are usually shown as a few selected images without numbers, and the numbers that do exist depend on reference masks that are rarely validated. This diagnostic case study pairs a visual saliency map with quantitative alignment scores for one image per class and checks whether an automatic reference mask is fit for that purpose. Grad-CAM saliency was computed at the last convolutional block (7 × 7) of an ImageNet-pretrained EfficientNet-B0 that serves as a negative control: it was not fine-tuned on rice images and the target is the highest ImageNet logit. The first image of each class in a deduplicated public Kaggle dataset was used (four classes, four images). Reference masks were obtained by unsupervised CIELAB K-Means (k = 3) inside an HSV/Otsu leaf mask and are intended to mark lesions. Three measures were computed: Lesion Energy Hit Rate (H_E), binary Intersection-over-Union (IoU_B, threshold 0.4), and a Normalized Energy Gain (G), the difference between H_E and the chance level implied by mask area. H_E was 63.4% (Blast), 52.9% (Brown Spot), 45.2% (Bacterial Blight), and 1.9% (Healthy); IoU_B was 42.1%, 33.9%, 10.6%, and 2.9% for Brown Spot, Bacterial Blight, Blast, and Healthy, respectively. G was −13.4 percentage points for Blast and between +0.7 and +12.1 for the other three images (mask areas measured from the figure, hence approximate). Visual inspection showed that in these images the masks largely cover green leaf area and not the visible lesions (for example, the brown spot in the Blast image lies outside the mask). The scores therefore measure alignment with colour-cluster masks, cannot be read as lesion alignment, and illustrate an interpretability illusion when the mask is not validated. With one image per class, they illustrate cases and are not class averages.

Keywords: Explainable AI, Grad-CAM, Attention Alignment, Segmentation Mask, Rice Leaf Disease

Abstrak

Peta saliensi banyak dipakai untuk menjelaskan jaringan konvolusional pada studi penyakit tanaman, tetapi umumnya ditampilkan sebagai beberapa citra pilihan tanpa angka, dan angka yang ada bergantung pada masker rujukan yang jarang divalidasi. Studi kasus diagnostik ini memasangkan peta saliensi dengan skor penjajaran kuantitatif pada satu citra per kelas dan memeriksa apakah masker rujukan otomatis layak dipakai untuk tujuan itu. Grad-CAM dihitung pada blok konvolusi terakhir (7 × 7) EfficientNet-B0 pralatih ImageNet yang dipakai sebagai kontrol negatif: model tidak di-*fine-tune* pada citra padi dan targetnya adalah logit ImageNet tertinggi. Citra pertama tiap kelas pada dataset publik Kaggle yang telah dideduplikasi dipakai (empat kelas, empat citra). Masker rujukan diperoleh dari K-Means CIELAB (k = 3) di dalam masker daun HSV/Otsu dan dimaksudkan menandai lesi. Tiga ukuran dihitung: *Lesion Energy Hit Rate* (H_E), *Intersection-over-Union* biner (IoU_B, ambang 0,4), dan *Normalized Energy Gain* (G), yaitu selisih H_E dengan tingkat peluang yang ditentukan luas masker. H_E adalah 63,4% (Blast), 52,9% (Brown Spot), 45,2% (Bacterial Blight), dan 1,9% (Healthy); IoU_B adalah 42,1%, 33,9%, 10,6%, dan 2,9% berturut-turut untuk Brown Spot, Bacterial Blight, Blast, dan Healthy. Nilai G adalah −13,4 poin persentase pada Blast dan antara +0,7 dan +12,1 pada tiga citra lain (luas masker diukur dari gambar sehingga bersifat perkiraan). Pemeriksaan visual menunjukkan bahwa pada citra-citra ini masker sebagian besar menutupi area daun hijau dan bukan lesi yang tampak (misalnya bercak cokelat pada citra Blast berada di luar masker). Skor ini karenanya mengukur penjajaran dengan masker klaster warna, tidak dapat dibaca sebagai penjajaran lesi, dan menjadi ilusi keterjelasan bila masker tidak divalidasi. Dengan satu citra per kelas, skor ini menggambarkan kasus, bukan rerata kelas.

Kata kunci: Explainable AI, Grad-CAM, Penjajaran Atensi, Masker Segmentasi, Penyakit Daun Padi

KomtekInfo is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

1. Pendahuluan

Jaringan saraf konvolusional (CNN) banyak dipakai untuk mengenali penyakit daun padi, tetapi keputusannya sulit ditelusuri. Peta saliensi seperti Grad-CAM [1] dan variannya seperti Grad-CAM++ [2] sering dipakai untuk menunjukkan area citra yang berpengaruh pada prediksi. Pada literatur penyakit padi, peta tersebut kerap dipakai sebagai ilustrasi. Rahman dkk. [3], misalnya, menerapkan Grad-CAM pada model padi dan menilainya secara kualitatif pada satu citra kelas *Brown Spot*, tanpa angka yang mengukur kesesuaian peta dengan lesi. Grad-CAM juga dipakai pada klasifikasi penyakit daun kentang untuk membantu menafsirkan prediksi model [4].

Pemeriksaan semacam itu penting karena akurasi tinggi dapat menutupi kelemahan penjelasan, dan peta yang tampak masuk akal pada beberapa contoh dapat menimbulkan ilusi keterjelasan (*interpretability illusion*): kesan bahwa model memperhatikan hal yang relevan padahal belum ada ukuran yang menguji kesan itu. Keandalan metode saliensi sendiri pun dipertanyakan: pada beberapa metode, atribusi dapat berubah oleh pergeseran konstan pada masukan yang tidak mengubah keluaran model [5]. Pada dataset penyakit tanaman, bias latar belakang telah dilaporkan [6]. Pemilihan contoh yang bersifat subjektif memperlemah bukti visual. Penelitian ini juga bertumpu pada contoh, sehingga batas tersebut berlaku padanya; yang ditambahkan adalah angka yang

Diterima: 22-10-2023 | Revisi: 24-11-2023 | Diterbitkan: 31-12-2023 | doi: 10.35134/komtekinfo.v10i3.718

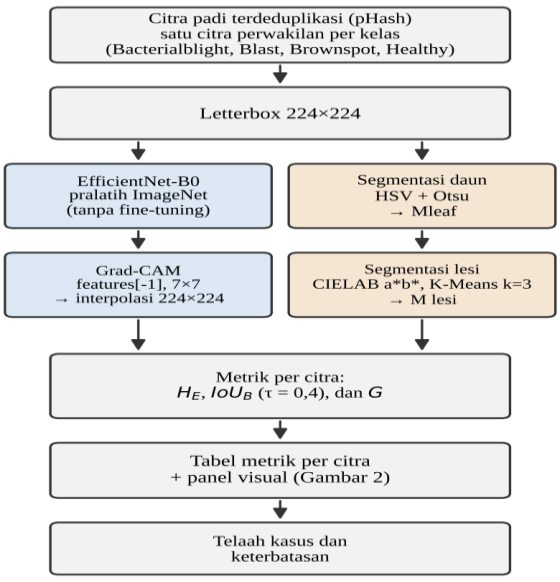
dipasangkan langsung dengan visualisasi dan aturan pemilihan contoh yang tetap.

Penilaian kuantitatif atas peta atensi telah dikenal, misalnya *pointing game* yang menguji apakah titik tertinggi peta jatuh pada objek [7]. Ukuran yang menghitung proporsi atribusi di dalam masker rujukan dikenal sebagai *relevance mass accuracy* [8], dan bersama ukuran lokalisasi lain tersedia dalam perangkat Quantus [9]. Di sisi lain, hasil saliensi sendiri perlu dicurigai: uji kewarasan menunjukkan bahwa sebagian metode saliensi dapat menghasilkan peta yang mirip meskipun bobot model diacak [10]. Pada citra radiograf dada, Saporta dkk. [11] membandingkan peta saliensi dengan segmentasi radiolog dan melaporkan bahwa Grad-CAM umumnya melokalisasi patologi lebih baik daripada metode lain yang diuji, tetapi ketujuh metode secara signifikan lebih buruk daripada pembanding manusia. Karena itu, angka penjajaran atensi perlu dibaca bersama batasan model dan masker rujukannya, dan pada penyakit tanaman masker rujukan yang andal menjadi prasyarat yang diperiksa dalam penelitian ini.

Penelitian ini adalah studi kasus diagnostik: penjajaran Grad-CAM [1] terhadap masker hasil segmentasi warna yang dimaksudkan sebagai masker lesi diukur pada satu citra per kelas, memakai EfficientNet-B0 [12] berbobot ImageNet [13] yang tidak di-*fine-tune* pada citra padi. Dua pertanyaan diajukan: (RQ1) seberapa besar energi saliensi pada citra sampel jatuh di dalam masker tersebut dibandingkan jaringan daun lainnya, dan (RQ2) apakah masker tersebut sesuai dengan lesi yang tampak, sehingga skor dapat dibaca sebagai penjajaran lesi. Kontribusinya adalah (a) tiga ukuran berbasis masker (termasuk pembanding peluang berbasis luas masker) yang dilaporkan bersama visualisasinya dan (b) temuan bahwa pada citra sampel masker otomatis tidak sesuai dengan lesi yang tampak, sehingga skor yang tampak masuk akal tidak boleh dibaca sebagai penjajaran lesi.

2. Metodologi Penelitian

Alur penelitian ditunjukkan pada Gambar 1. Tiap citra diproses melalui dua jalur yang bertemu pada perhitungan metrik: jalur saliensi dan jalur masker rujukan.



Gambar 1. Alur Penelitian

2.1 Data

Data berasal dari *Rice Leaf Disease Detection Dataset* yang dideduplikasi berdasarkan *perceptual hash* identik sehingga menghasilkan 4.410 citra unik, mengikuti penelitian pendamping [14] [15]. Audit ini memakai empat kelas pertama menurut urutan abjad, yaitu *Bacterial Blight*, *Blast*, *Brown Spot*, dan *Healthy*; kelas *Tungro* tidak dianalisis. Dari tiap kelas diambil satu citra perwakilan, yaitu citra pertama pada daftar kelas tersebut setelah deduplikasi (indeks 0). Aturan ini tetap dan tidak memilih citra berdasarkan kemiripan visual dengan lesi ideal, tetapi satu citra tidak mewakili keragaman dalam kelas. Citra diubah ke 224 × 224 piksel dengan *letterbox* agar rasio aspek terjaga.

2.2 Model dan Grad-CAM

Model yang diaudit adalah EfficientNet-B0 [12] dengan bobot ImageNet [13] dari pustaka torchvision, dipakai sebagai tulang punggung beku (*frozen backbone*) tanpa dilatih ulang pada citra padi. Kepala klasifikasinya masih berupa kepala 1.000 kelas ImageNet. Skor target diambil dari logit tertinggi, $Y^c = \max_i Z_i$, sehingga yang dijelaskan adalah kelas ImageNet teratas, bukan kelas penyakit padi. Konfigurasi ini sengaja dipakai sebagai kontrol negatif: karena bobot dan kepala belum mengenal penyakit padi, peta yang dihasilkan mencerminkan atensi representasi ImageNet umum dan bukan atensi diagnostik. Perbandingan langsung dengan model yang di-*fine-tune*, dengan Grad-CAM dihitung terhadap logit kelas penyakit sebenarnya, belum dilakukan. Peta saliensi diambil dari lapisan konvolusi terakhir (*features[-1]*), yaitu tensor aktivasi 1.280 × 7 × 7.

Grad-CAM [1] memberi bobot pada tiap kanal k dengan merata-ratakan gradien skor target terhadap aktivasi (*global average pooling*), lalu menjumlahkan kanal secara terbobot:

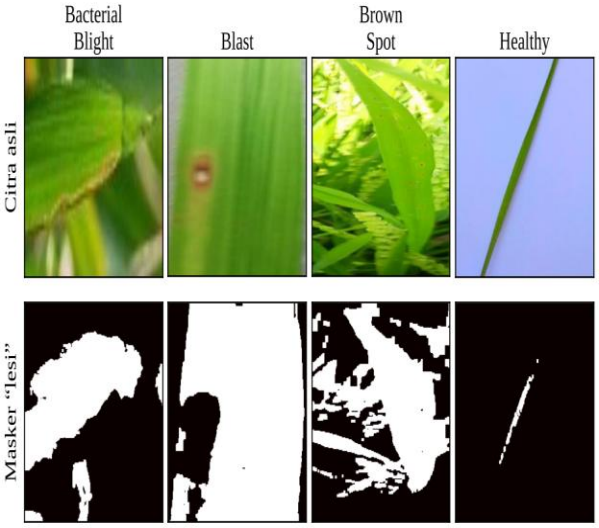
$$\alpha_k^c = (1/Z) \Sigma_{ij} \partial Y^c / \partial A_{ij}^k \quad (1)$$

$$L^c = \mathrm{ReLU}(\Sigma_k \alpha_k^c A^k) \quad (2)$$

dengan Z jumlah posisi spasial (49). Peta L^c dinormalisasi min–maks ke [0, 1] lalu diinterpolasi bilinear dari 7 × 7 ke 224 × 224. Satu sel pada peta 7 × 7 setara area 32 × 32 piksel pada citra masukan, atau sekitar 2% luas citra.

2.3 Masker Rujukan Lesi

Masker daun M_{leaf} diperoleh dengan ambang Otsu pada kanal saturasi ruang HSV, dilanjutkan operasi morfologi. Di dalam daun, piksel dikelompokkan pada kanal a^*b^* ruang CIELAB dengan K-Means (k = 3), dan satu klaster dipilih berdasarkan kroma, dirancang untuk warna nekrotik atau klorotik, sebagai masker lesi M_{lesion}. Masker ini dihasilkan otomatis dari warna dan tidak divalidasi oleh pakar patologi tanaman atau terhadap anotasi, sehingga disebut pseudo-*ground-truth*. Gambar 2 menampilkan citra sampel dan masker tersebut; masker ini diperiksa secara visual pada Bagian 3.1.



Gambar 2. Citra sampel dan masker "lesi" hasil K-Means CIELAB

2.4 Metrik Penjajaran

Lesion Energy Hit Rate mengukur fraksi energi saliensi di dalam lesi terhadap seluruh energi pada peta saliensi, tanpa membatasi pada masker daun:

$$H_E = \Sigma_{\text{lesi}}\, L^c / \Sigma_{\text{seluruh peta}}\, L^c \times 100\% \quad (3)$$

Untuk IoU biner, peta dibinerkan dengan ambang $\tau = 0{,}4$, yaitu $L_\tau = \{(x, y) \mid L^c(x, y) \geq \tau\}$, lalu

$$IoU_B = |L_\tau \cap M_{\text{lesion}}| / |L_\tau \cup M_{\text{lesion}}| \times 100\% \quad (4)$$

Karena H_E bergantung pada luas masker, dipakai pula *Normalized Energy Gain*, yaitu selisih H_E dengan nilai yang diharapkan bila energi tersebar merata:

$$G = H_E - (A_L / A) \times 100\% \quad (5)$$

dengan A_L luas masker lesi dan A luas citra (224 × 224 piksel). Bila energi tersebar merata, H_E sama dengan $(A_L / A) \times 100\%$, sehingga $G \leq 0$ berarti energi di dalam masker tidak lebih terkonsentrasi daripada sebaran merata. Luas diukur dari jumlah piksel putih pada masker di Gambar 3, sehingga G bersifat perkiraan (galat sekitar 1 poin persentase).

Nilai ketiga ukuran dihitung pada tiap citra sampel dan dilaporkan apa adanya. Karena tiap kelas hanya diwakili satu citra, tidak ada rerata, simpangan baku, maupun uji statistik.

3. Hasil dan Pembahasan

3.1 Hasil dan Pemeriksaan Masker

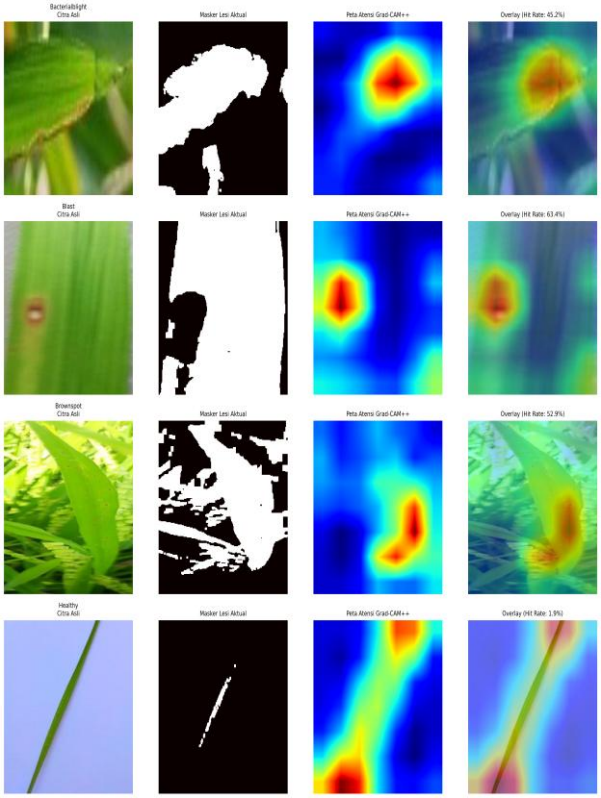
Tabel 1 memuat skor tiap citra sampel dan Gambar 3 menampilkan visualisasi penjajarannya: citra asli, masker, peta Grad-CAM, dan *overlay*. Pada tiga citra berpenyakit, H_E berkisar 45,2–63,4% dan IoU_B berkisar 10,6–42,1%. Kedua metrik mengurutkan citra secara berbeda: menurut H_E urutannya *Blast*, *Brown Spot*, *Bacterial Blight*, sedangkan menurut IoU_B urutannya *Brown Spot*, *Bacterial Blight*, *Blast*. Angka ini berlaku untuk keempat citra tersebut dan tidak boleh dibaca sebagai rerata kelas.

Nilai G berkisar −13,4 hingga +12,1 poin. Hanya *Blast* yang berada di bawah sebaran merata: masker mencakup 76,8% citra, sehingga H_E 63,4% yang tampak tinggi lebih rendah daripada yang diharapkan bila energi tersebar merata (G = −13,4). Pada *Bacterial Blight* dan *Brown Spot*, H_E melampaui luas masker sebesar 12,1 dan 7,9 poin, dan pada *Healthy* hampir sama (+0,7). Selisih positif yang kecil ini tidak menunjukkan penjajaran lesi, karena masker bukan lesi (Gambar 2).

Tabel 1. Skor Penjajaran Saliensi pada Citra Sampel (satu citra per kelas)

Citra sampel	*n*	H_E (%)	IoU_B (%)	A_L/A (%)	G (poin)
Bacterial Blight	1	45,2	33,9	33,1	+12,1
Blast	1	63,4	10,6	76,8	−13,4
Brown Spot	1	52,9	42,1	45,0	+7,9
Healthy	1	1,9	2,9	1,2	+0,7

Keterangan: nilai per citra, $\tau = 0{,}4$; tanpa rerata atau simpangan baku. A_L/A dan G dihitung dari masker pada Gambar 3 (perkiraan).



Gambar 3. Visualisasi penjajaran atensi pada satu citra per kelas

Pemeriksaan visual pada Gambar 2 menunjukkan bahwa masker tidak sesuai dengan lesi yang tampak. Pada citra *Blast*, masker mencakup sebagian besar daun hijau, sedangkan bercak cokelat berinti pucat justru berada di luar masker. Pada citra *Bacterial Blight*, masker berupa area daun hijau. Pada citra *Brown Spot*, masker mencakup daun hijau terang dan sebagian vegetasi sekitarnya, dan bercak kecil tidak terisolasi. Pada citra *Healthy*, masker hanya berupa potongan sempit dari daun. Dengan demikian skor pada Tabel 1 mengukur penjajaran saliensi dengan klaster warna dominan, bukan dengan lesi.

3.2 Pembahasan per Citra

Blast. Citra ini memiliki H_E tertinggi (63,4%) dan IoU_B terendah (10,6%). Karena masker pada citra ini menutupi daun hijau dan tidak mencakup bercak, energi saliensi di dalam masker berada pada jaringan hijau, bukan pada lesi. Nilai H_E yang tinggi karenanya tidak menunjukkan fokus pada lesi, dan seberapa besar energi saliensi menyentuh bercak itu sendiri tidak dapat dibaca dari angka ini.

Brown Spot. Citra ini memberi IoU_B tertinggi (42,1%) dengan H_E 52,9%. Masker pada citra ini besar dan mencakup daun hijau terang, sehingga tumpang tindih dengan area saliensi relatif mudah diperoleh. Nilai tersebut tidak menunjukkan penjajaran dengan bercak cokelat kecil.

Bacterial Blight. Sebanyak 54,8% energi saliensi (100% − 45,2%) berada di luar masker, dengan IoU_B 33,9%. Masker berupa area daun hijau, sehingga energi di luar masker dapat berada pada tepi, area non-hijau, atau latar. Naskah ini tidak menyimpulkan bahwa penyebabnya adalah kontur tepi daun; analisis tepi-versus-bagian-dalam tidak dilakukan.

Healthy. Nilai H_E 1,9% dan IoU_B 2,9% rendah, tetapi masker pada citra ini hanya potongan sempit daun, dan daun sehat tidak memiliki lesi. Nilai ini tidak dapat dibaca sebagai bukti bahwa model tidak mengaktifkan lesi palsu, dan citra ini lebih tepat dipandang sebagai pembanding daripada kontrol negatif yang tervalidasi.

3.3 Implikasi dan Keterbatasan

Yang dapat dipertanggungjawabkan dari studi kasus ini ada dua. Pertama, pada empat citra sampel, skor penjajaran terhadap masker berbeda, dan satu metrik tidak cukup karena H_E dan IoU_B memberi urutan yang berbeda, sedangkan G menunjukkan bahwa H_E yang tampak tinggi pada *Blast* (63,4%) tidak melebihi sebaran merata dan pada tiga citra lain hanya melampauinya 0,7–12,1 poin. Kedua, dan lebih penting, masker klaster warna otomatis dapat menghasilkan skor yang tampak bermakna sementara masker itu sendiri tidak menandai lesi, sehingga masker harus diperiksa atau divalidasi sebelum skor dibaca sebagai penjajaran lesi. Angka yang tampak kuantitatif tetap dapat menimbulkan ilusi keterjelasan bila rujukannya tidak valid. Temuan ini juga tidak dibaca sebagai bukti tentang ketergantungan model pada artefak; klaim itu memerlukan masker tervalidasi, model yang dilatih pada tugas penyakit padi, dan uji yang mengubah isi citra, misalnya menutup lesi atau mengganti latar.

Keterbatasan penelitian ini adalah sebagai berikut. (1) Masker rujukan tidak sesuai dengan lesi yang tampak pada citra sampel dan tidak divalidasi terhadap anotasi, sehingga skor tidak mengukur penjajaran lesi. (2) Satu citra per kelas: tidak ada rerata, simpangan baku, atau uji statistik, dan hasil tidak dapat digeneralisasi ke kelas. (3) Model tidak di-*fine-tune* pada citra padi dan targetnya logit ImageNet tertinggi; ia dipakai sebagai kontrol negatif sehingga peta bukan atensi diagnostik model padi, dan tidak ada perbandingan langsung dengan model yang di-*fine-tune*. (4) Pembanding peluang hanya berupa G berbasis luas masker yang diukur dari gambar (perkiraan); pembanding berupa peta acak belum dilakukan. (5) Ambang $\tau = 0{,}4$ tidak diuji sensitivitasnya. (6) Hanya satu metode saliensi (Grad-CAM orde pertama) dan satu lapisan yang dipakai, dan uji kewarasan [10] belum dilakukan. (7) Kelas *Tungro* tidak dianalisis, dan seluruh citra bersifat *in-distribution*.

4. Kesimpulan

Studi kasus diagnostik terhadap Grad-CAM dari EfficientNet-B0 pralatih (kontrol negatif) menghasilkan skor penjajaran terhadap masker klaster warna pada satu citra per kelas: pada tiga citra berpenyakit H_E berkisar 45,2–63,4% dan IoU_B berkisar 10,6–42,1%, dengan G antara −13,4 dan +12,1 poin. Pemeriksaan visual menunjukkan bahwa masker tersebut tidak sesuai dengan lesi yang tampak, sehingga skor tidak dapat dibaca sebagai penjajaran lesi dan dapat menimbulkan ilusi keterjelasan. Angkanya juga bukan rerata kelas.

Penelitian lanjutan sebaiknya (1) memakai masker lesi berupa anotasi poligon manual yang divalidasi kriteria agronomis pada subset citra uji (misalnya dengan Labelme atau CVAT), dengan masker otomatis tetap dipakai untuk mengukur galatnya, (2) memperluas evaluasi menjadi minimal 30–50 citra per kelas pada kelima kelas, termasuk *Tungro*, dengan rerata, simpangan baku, dan selang kepercayaan 95%, (3) melatih atau me-*fine-tune* EfficientNet-B0 pada lima kelas penyakit padi dan menghitung Grad-CAM terhadap logit kelas sebenarnya, lalu membandingkannya langsung dengan konfigurasi pralatih ini, dan (4) menambahkan pembanding peta acak, uji sensitivitas τ, dan uji kewarasan.

Daftar Rujukan

[1] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, 2020, doi: 10.1007/s11263-019-01228-7.

[2] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, 2018, pp. 839–847, doi: 10.1109/WACV.2018.00097.

[3] M. M. Rahman *et al.*, "Explainable deep transfer learning framework for rice leaf disease diagnosis and classification," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 15, no. 12, pp. 862–884, 2024.

[4] S. M. Alhammad *et al.*, "Deep learning and explainable AI for classification of potato leaf diseases," *Front. Artif. Intell.*, vol. 7, Art. no. 1449329, 2025, doi: 10.3389/frai.2024.1449329.

[5] P.-J. Kindermans *et al.*, "The (un)reliability of saliency methods," arXiv:1711.00867, 2017.

[6] M. A. Noyan, "Uncovering bias in the PlantVillage dataset," arXiv:2206.04374, 2022.

[7] J. Zhang, S. A. Bargal, Z. Lin, J. Brandt, X. Shen, and S. Sclaroff, "Top-down neural attention by excitation backprop," *Int. J. Comput. Vis.*, vol. 126, no. 10, pp. 1084–1102, 2018, doi: 10.1007/s11263-017-1059-x.

[8] L. Arras, A. Osman, and W. Samek, "CLEVR-XAI: A benchmark dataset for the ground truth evaluation of neural network explanations," *Inf. Fusion*, vol. 81, pp. 14–40, 2022, doi: 10.1016/j.inffus.2021.11.008.

[9] A. Hedström *et al.*, "Quantus: An explainable AI toolkit for responsible evaluation of neural network explanations and beyond," *J. Mach. Learn. Res.*, vol. 24, no. 34, pp. 1–11, 2023.

[10] J. Adebayo, J. Gilmer, M. Muelly, I. Goodfellow, M. Hardt, and B. Kim, "Sanity checks for saliency maps," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 31, 2018, arXiv:1810.03292.

[11] A. Saporta *et al.*, "Benchmarking saliency methods for chest X-ray interpretation," *Nat. Mach. Intell.*, vol. 4, no. 10, pp. 867–878, 2022, doi: 10.1038/s42256-022-00536-x.

[12] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Mach. Learn. (ICML)*, PMLR 97, 2019, pp. 6105–6114.

[13] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015, doi: 10.1007/s11263-015-0816-y.

[14] "Rice Leaf Disease Detection Dataset," Kaggle, pengguna shams069. [Online]. Available: https://www.kaggle.com/datasets/shams069/rice-leaf-disease-detection-dataset

[15] Muhammad, G. U., & Astuti, E. Z. (2025). Implementation of Convolutional Neural Network and Support Vector Machine Classification for Disease Detection in Rice Plants. *Journal of Innovation and Technology Polbeng Series on Informatics (INOVTEK Polbeng-Seri Informatika)*, *10*(3), 1762-1771..