

Daftar Pustaka

- Abrar Masril M, Putra Caniago D. 2023. Optimasi Teknologi Computer Vision pada Robot Industri Sebagai Pemindah Objek Berdasarkan Warna. 11(1):46–57. <https://doi.org/10.26760/elkomika.v11.i1.46>.
- Chen C, Anjum S, Gurari D. 2022 Feb 4. Grounding Answers for Visual Questions Asked by Visually Impaired People. <http://arxiv.org/abs/2202.01993>.
- Citra EE, Fudholi DH, Dewa CK. 2023. Implementasi Arsitektur EfficientNetV2 Untuk Klasifikasi Gambar Makanan Tradisional Indonesia. *JURNAL MEDIA INFORMATIKA BUDIDARMA*. 7(2):766. <https://doi.org/10.30865/mib.v7i2.5881>.
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, *et al.* 2020 Okt 22. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. <http://arxiv.org/abs/2010.11929>.
- Dr. Visu P, Narayanan K S, Ahmed A SN, S AK. 2024. Visual Questioner using Bert Transformer. *International Journal of Science and Management Studies (IJSMS)*. 7(1):83–89. <https://doi.org/10.51386/25815946/ij sms-v7i1p113>.
- Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, Liu Y, Topol E, Dean J, Socher R. 2021. Deep learning-enabled medical computer vision. *NPJ Digit Med*. 4(1). <https://doi.org/10.1038/s41746-020-00376-2>.
- Géron A. 2023. *Hands-on machine learning with Scikit-Learn, Keras and TensorFlow : concepts, tools, and techniques to build intelligent systems*. 3rd Edition. O'Reilly Media, Inc.
- Guo MH, Xu TX, Liu JJ, Liu ZN, Jiang PT, Mu TJ, Zhang SH, Martin RR, Cheng MM, Hu SM. 2022. Attention mechanisms in computer vision: A

- survey. *Comput Vis Media (Beijing)*. 8(3):331–368. <https://doi.org/10.1007/s41095-022-0271-y>.
- Hakim B. 2021. Analisa Sentimen Data Text Preprocessing Pada Data Mining Dengan Menggunakan Machine Learning. *JBASE - Journal of Business and Audit Information Systems*. 4(2). <https://doi.org/10.30813/jbase.v4i2.3000>.
- Hamad Ali A, Mohsin Absulazeez A. 2024. Transfer Learning In Machine Learning: A Review Of Methods And Applications. *Indonesian Journal of Computer Science*., siap terbit.
- He X, Sun J, Chen H, Li D. 2022. Campo: Cost-Aware Performance Optimization for Mixed-Precision Neural Network Training. Di dalam: *USENIX Annual Technical Conference*. hlm 505–518. <https://www.usenix.org/conference/atc22/presentation/he>.
- Janiesch C, Zschech P, Heinrich K. 2021. Machine learning and deep learning. *Electronic Markets*. 31:685–695. <https://doi.org/10.1007/s12525-021-00475-2/Published>.
- Khurana D, Koli A, Khatter K, Singh S. 2023. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl*. 82(3):3713–3744. <https://doi.org/10.1007/s11042-022-13428-4>.
- Koshti D, Gupta A, Kalla M, Sharma A. 2024. TRANS-VQA: Fully Transformer-Based Image Question-Answering Model Using Question-guided Vision Attention. *Inteligencia Artificial*. 27(73):111–128. <https://doi.org/10.4114/intartif.vol27iss73pp111-128>.
- Li H, Han D, Chen C, Chang CC, Li KC, Li D. 2023. A Visual Question Answering Network Merging High- and Low-Level Semantic Information. *IEICE Trans Inf Syst*. E106.D(5):581–589. <https://doi.org/10.1587/transinf.2022DLP0002>.
- Liu L, Su X, Guo H, Zhu D. 2022. A Transformer-based Medical Visual Question Answering Model. Di dalam: *Proceedings - International*

- Conference on Pattern Recognition*. Volume ke-2022-August. Institute of Electrical and Electronics Engineers Inc. hlm 1712–1718.
- Lu S, Ding Y, Liu M, Yin Z, Yin L, Zheng W. 2023. Multiscale Feature Extraction and Fusion of Image and Text in VQA. *International Journal of Computational Intelligence Systems*. 16(1). <https://doi.org/10.1007/s44196-023-00233-6>.
- Luo Z. 2024. The Application and Development of Image Recognition Technology in the Field of Computer Vision. Volume ke-2023.
- Min B, Ross H, Sulem E, Veyseh AP Ben, Nguyen TH, Sainz O, Agirre E, Heinz I, Roth D. 2021 Nov 1. Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey. <http://arxiv.org/abs/2111.01243>.
- Obe R, Kaufmann B, Baird K, Kadel S, Soltani Y, Cham M, Gobbert MK, Barajas CA, Jiang Z, Sharma VR, *et al.* 2023. Accelerating Real-Time Imaging for Radiotherapy: Leveraging Multi-GPU Training with PyTorch.
- Peng L, An G, Ruan Q. 2022. Transformer-based Sparse Encoder and Answer Decoder for Visual Question Answering. Di dalam: *International Conference on Signal Processing Proceedings, ICSP*. Volume ke-2022-October. Institute of Electrical and Electronics Engineers Inc. hlm 120–123.
- Rezapour M. 2024. Cross-attention Based Text-image Transformer for Visual Question Answering. *Recent Advances in Computer Science and Communications*. 17(4). <https://doi.org/10.2174/0126662558291150240102111855>.
- Rifaldi D, Abdul Fadlil, Herman. 2023. Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet “Mental Health.” *Decode: Jurnal Pendidikan Teknologi Informasi*. 3(2):161–171. <https://doi.org/10.51454/decode.v3i2.131>.

- Saputra AC, Saragih AS, Ronaldo D. 2024. Perbandingan Nilai Akurasi Distilbert dan BERT pada Dataset Analisis Sentimen Lembaga Kursus. *Jurnal Teknologi Informasi*. 18:160–171. <https://doi.org/10.47111/JTI>.
- Sree CJ, Ravi Kumar G, Challari GM, Kumar V. 2024. Evolution of Transfer Learning and Generalization in Machine Learning: Advancements Unveiled. www.ijfmr.com.
- Tan M, Le Q V. 2021. EfficientNetV2: Smaller Models and Faster Training. <https://github.com/google/>.
- Wang W, Bao H, Dong L, Bjorck J, Peng Z, Liu Q, Aggarwal K, Mohammed OK, Singhal S, Som S, *et al.* 2022 Agu 22. Image as a Foreign Language: BEiT Pretraining for All Vision and Vision-Language Tasks. <http://arxiv.org/abs/2208.10442>.
- Wu J, Lu J, Sabharwal A, Mottaghi R. 2022. Multi-Modal Answer Validation for Knowledge-Based VQA. Di dalam: *AAAI-22*. www.aaai.org.
- Yang Z, Gan Z, Wang J, Hu X, Lu Y, Liu Z, Wang L. 2022. An Empirical Study of GPT-3 for Few-Shot Knowledge-Based VQA. Di dalam: *AAAI-22*. hlm 3081–3089. www.aaai.org.
- Zaghoulani W, Vladimir I, Ruiz M. 2023. COVID-Twitter-BERT: A natural language processing model to analyse COVID-19 content on Twitter. <https://github.com/digitalepidemiologylab/covid-twitter-bert>.
- Zhang S, Yang B, Li Z, Ma Z, Liu Y, Bai X. 2024 Mei 9. Exploring the Capabilities of Large Multimodal Models on Dense Text. <http://arxiv.org/abs/2405.06706>.