

## DAFTAR PUSTAKA

- Ahmed, F., and M. Yeasin., 2017, Optimization and *Evaluation* of deep architectures for ambient awareness on a sidewalk, in *Neural Networks (IJCNN)*, 2017 International Joint Conference on. IEEE, pp. 2692–2697.
- Ahmed, F., Md Mahmud, S., Al-Fahad, R., Alam, S., and M Yeasin., 2018, *Image captioning* for Ambient Awareness on a Sidewalk, The University of Memphis.
- Ahsan, H., Bhalla, N., Bhatt, D., Shah, K., 2021, Multi-Modal *Image captioning* for the Visually Impaired, University of Massachusetts Amherst
- Al-Malla, M.A., Jafar, A., & Ghneim, N., 2022, *Image captioning* model using attention and object features to mimic human image understanding. *J Big Data* 9, 20. <https://doi.org/10.1186/s40537-022-00571-w>
- Amritkar, C., Jabade, V., 2018, *Image caption* generation using *deep learning* technique. In 2018 fourth international conference on computing communication control and automation (ICCUBEA). IEEE, Pune, pp 1–4
- Artyani, I., 2019, Simulasi metode convolutional neural network dan *Long Short-Term Memory* untuk *generate Image captioning* pada gambar lalu lintas kendaraan berbahasa Indonesia, Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah Jakarta.
- Beddiar, DR., Oussalah, M. & Seppänen, T., 2023, Automatic *captioning* for medical imaging (MIC): a rapid review of literature, *Artif Intell Rev* 56, 4019–4076
- Bourne, R, R, A., Flaxman, S, R., Braithwaite, T., Cicinelli, M, V., Das, A., 2017, Vision *Loss* Expert Group. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis. *Lancet Glob. Health* 2017, 5, e888–e897.
- Chen, L., Zhang, H., Xiao, J., Nie, L., Shao, J., Liu, W., Chua, T., 2017, SCA-CNN: Spatial and channel-wise attention in convolutional networks for *Image captioning*, in: *Proceedings of the IEEE Conference on Computer vision and Pattern Recognition (CVPR)*, pp. 6298–6306.
- Cubuk, E, D., Zoph, B., Mane, D., Vasudevan, V., & Le, Q. V. 2019. Autoaugment: Learning augmentation strategies from data. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 113-123).
- Dittakan, K., Prompitak, K., Thungklang, P., 2023, *Image caption* generation using transformer learning methods: a case study on instagram *Image*. *Multimed Tools Appl*.
- Elman, J, L., 1990, Finding structure in time, *Cognitive science*, vol. 14, no. 02, p. 179–211

- Ferreira, W., and Vlachos, A., 2016, “Emergent: a novel data-set for stance classification,” in Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies, pp. 1163–1168
- Gers, F. A., Schmidhuber, J., & Cummins, F. 2000, Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451–2471
- Habarulema, J. B., 2010, “A contribution to TEC modelling over Southern Africa using GPS data,” no. August.
- Hagan, M. T., Demuth, H., and Beale, M., 1996, *Neural Network Design*. USA: PWS Publishing Co.
- Han, M., Chen, W., & Moges, A.D., 2019, Fast *image captioning* using LSTM. *Cluster Comput* 22 (Suppl 3), 6143–6155. <https://doi.org/10.1007/s10586-018-1885-9>
- He, K., Zhang, X., Ren, S., dan Sun, J., 2016, Deep Residual Learning for *Image Recognition*. IEEE Conference on *Computer vision and Pattern Recognition (CVPR)*, Las Vegas.
- He, X., Yang, Y., Shi, B., Bai, X., 2019, VD-SAN: visual-densely semantic attention network for *Image caption* generation, *Neurocomputing* 328, 48–55.
- Hilal, A. M., Alrowais, F., Al-Wesabi, F. N., Marzouk, R., 2023, Red Deer Optimization with *Artificial intelligence Enabled Image captioning System* for Visually Impaired People, *Computer Systems Science and Engineering*
- Hochreiter, S., and Schmidhuber, J. U, 1997, “*Long Short-Term Memory*,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780.
- Huang, L., Wang, W., Chen, J., Wei, X, Y., 2019, Attention on Attention for *Image captioning*, Macau University of Science and Technology.
- Jiang, Z., Zhao, L., Li, S., Jia, Y., & Liquan, Z. 2020, Real-time object detection method for embedded devices. ArXiv.
- Kaur, M., Kaur, H., 2023, An efficient deep learning based hybrid model for *image caption* generation, *International Journal of Advanced Computer Science and Applications*, 14(3)
- Kesavan, V., Muley, V., dan Kolhekar, M., 2019, *Deep learning Based Automatic Image Caption Generation*, Global Conference for Advancement in Technology (GCAT), Bangalore.
- Kesavan, V., Muley, V., dan Kolhekar, M., 2019, *Deep learning Based Automatic Image Caption Generation*, Global Conference for Advancement in Technology (GCAT), Bangalore.
- Kim, D. J., Choi, J., Oh, T. H., Kweon, I. S., 2019, *Image captioning* with very scarce supervised data: adversarial semi-supervised learning approach arXiv preprint arXiv:1909.02201

- Lavie, A., dan Agarwal, A., 2007, METEOR: An Automatic Metric for MT *Evaluation* with High Levels of Correlation with Human Judgments. Proceedings of the Second Workshop on Statistical Machine Translation, Morristown.
- Lin, C.-Y., 2004, ROUGE: A Package for Automatic *Evaluation* of Summaries, Proceedings of the ACL Workshop: Text Summarization Braches Out 2004, Barcelona.
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., dan Zitnick, C.L., 2014, Microsoft COCO: Common Objects in Context, Springer International Publishing, 740–755.
- Liu, Q., Fan, X., Xi, Z., Yin, Z., & Yang, Z. 2022. Object detection based on YOLOv4-Tiny and Improved Bidirectional feature pyramid network. Journal of Physics: Conference Series, 2209(1). <https://doi.org/10.1088/1742-6596/2209/1/012023>
- Liu, W., Chen, S., Guo, L., Zhu, X., dan Liu, J., 2021, CPTR: Full Transformer Network for *Image captioning*, <https://arxiv.org/abs/2101.10804>, diakses tanggal 14 Agustus 2022.
- Liu, L., Wang, Y., and Zhao, H., 2019, "An *Image* Segmentation Method for the blind sidewalks recognition by using the convolutional neural network U-net," 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP), Chongqing, China, pp. 1-4, doi: 10.1109/ICSIDP47821.2019.9172970.
- Lynn, H. M., Pan, S. B., and Kim, P., 2019. A Deep Bidirectional GRU Network Model for Biometric Electrocardiogram Classification Based on Recurrent *Neural Networks*. IEEE Access, 7, 145395–145405.
- Ma, C., Hu Y, Liu, H., Huang, P., Zhu, Y., Dai, D., 2023, Generating *Image* Descriptions of Rice Diseases and Pests Based on DeiT *Feature Encoder*, Applied Sciences., 13(18).
- Ma, Y., Shi, Y., Zhang, M., Li, W., Ma, C., Guo, Y., 2022, Design and Implementation of an Intelligent Assistive Cane for Visually Impaired People Based on an Edge-*Cloud* Collaboration Scheme. Electronics, 11, 2266.
- Makay, B., and Kılıç, V., 2019 A new *Image captioning* approach for visually impaired people. In 2019 11th international conference on electrical and electronics engineering (ELECO), IEEE, 945-949
- Manduchi, R., PhD\* Kurniawan, S. 2011, Mobility-Related Accidents Experienced by People with Visual Impairment, University of Santa Cruz Santa Cruz, CA, pp.10
- Noorviani, Y., “Penerapan Elman Recurrent *Neural Network* untuk Diagnosis Gangguan Autis Pada Anak,” Universitas Gadjah Mada, Yogyakarta, 2011.

- Nursikuwagus, A., Munir, R., dan Khodra, M.L., 2020, *Image captioning* menurut Scientific Revolution Kuhn dan Popper, *Jurnal Manajemen Informatika (JAMIKA)*, 2, 10, 110–121.
- Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J., 2001, BLEU, *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia.
- Pradana, A. W., & Hayaty, M. 2019. The Effect of *Stemming* and Removal of *Stopwords* on the Accuracy of Sentiment Analysis on Indonesian-language Texts. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 4(4), 375-380.
- Rajendran, G., Chitturi, B., and Poornachandran, P., 2018, “Stance-InDepth Deep Neural Approach to Stance Classification,” *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 1646–1653.
- Rane, C., Lashkare, A., Karande, A., Rao, Y, S., 2021, *Image captioning* based Smart Navigation System for Visually Impaired. *Bharatiya Vidya Bhavan’s Sardar Patel Institute of Technology Mumbai, India*
- Reddy, S, M., Suman, C., Saha, S., and Bhattacharyya, P., 2020, “A GRU based fake news prediction system: Working notes for UrduFake - FIRE 2020,” *CEUR Workshop Proc.*, vol. 2826, pp. 464 –468.
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on *Computer vision* and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 779-788.
- Sáez, J.M., Escolano, F., Lozano, M.A. Aerial Obstacle Detection With 3-D Mobile Devices. *IEEE J. Biomed. Health Inform.* 2015, 19, 74–80.
- Safiya, K, M., Pandian, R., 2023, A Real-Time *Image captioning* Framework Using *Computer vision* to Help the Visually Impaired, *Multimedia Tools and Applications*.
- Schuster, M., and Paliwal, K, K., Bidirectional recurrent neural networks, *IEEE Trans. Signal Process.* 45 (11) (1997) 2673–2681.
- Sehgal, S., Sharma, J., Chaudhary, N., 2020, Generating *Image captions* based on *deep learning* and *natural language processing*. In 2020 8th international conference on reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), IEEE, 165–169.
- Selivanov, A., Rogov, O.Y., Chesakov, D., 2023, Medical *Image captioning* via generative pretrained transformers. *Sci Rep* 13, 4171
- Shang, W., Sohn, K., Almeida, D., & Lee, H. 2016. Understanding and improving convolutional neural networks via *concatenated* rectified linear units. In international conference on *machine learning*. pp. 2217-2225
- Sharma, D., Dhiman, C., and Kumar, D., 2022, "Automated *Image Caption* Generation Framework using Adaptive Attention and Bi-LSTM," 2022 IEEE

- Delhi Section Conference (DELCON), New Delhi, India, pp. 1-5, doi: 10.1109/DELCON54057.2022.9752859.
- Shorten, C., & Khoshgoftaar, T. M. 2019. A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1). doi: 10.1186/s40537-019-01970.
- Sumbul, G., Nayak, S., Demir, B., 2020, SD-RSIC: Summarization Driven Deep Remote Sensing *Image captioning*, in IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 8, pp. 6922-6934
- Sun, W., Chunshan W., Jingqiu G., Xiang S., Jiuxi L., and Fangfang L., 2023, "Veg-DenseCap: Dense *Captioning* Model for Vegetable Leaf Disease *Images*" *Agronomy* 13, no. 7: 1700.
- Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. 2018. A survey on deep *transfer learning*. In *Artificial Neural Networks and Machine learning—ICANN 2018: 27th International Conference on Artificial Neural Networks*, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27 (pp. 270-279). Springer International Publishing.
- Tan, M., dan Le, Q. V., 2019, EfficientNet: Rethinking Model Scaling for Convolutional *Neural Networks*, Proceedings of the 36th International Conference on *Machine learning*, California.
- Tiwary, T., Mahapatra, R.P., 2023, An accurate generation of *Image captions* for blind people using extended convolutional atom neural network, *Multimed Tools Appl* 82, 3801–3830
- Vedantam, R., Zitnick, C.L., dan Parikh, D., 2015, CIDEr: Consensus-based *Image Description Evaluation*, 2015 IEEE Conference on *Computer vision* and Pattern Recognition (CVPR), Boston.
- Wang, Q., Huang, W., Zhang, X., Li, X., 2021, Word–Sentence Framework for Remote Sensing *Image captioning*, in IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 12, pp. 10532-10543, Dec. 2021, doi: 10.1109/TGRS.2020.3044054.
- Wang, S., dan Zhu, Y., 2021, A Novel *Image Caption* Model Based on Transformer Structure, 2021 IEEE International Conference on Information Communication and Software Engineering (ICICSE), Chengdu.
- Wang, Y., Xu, J., dan Sun, Y., 2022, *End-to-End* Transformer Based Model for *Image captioning*, <https://arxiv.org/abs/2203.15350>, diakses tanggal 14 Agustus 2022.
- Wonohadidjojo, D, M., 2021, “Perbandingan Convolutional *Neural Network* pada *Transfer learning* Method untuk Mengklasifikasikan Sel Darah Putih,” *Ultim. J. Tek. Inform.*, vol. 13, no. 1, pp. 51–57
- Xiao, X., Wang, L., Ding, K., Xiang, S., Pan C., 2019, Deep Hierarchical *Encoder–Decoder* Network for *Image captioning*, in IEEE Transactions on Multimedia, vol. 21, no. 11, pp. 2942-2956

- Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R., Bengio, Y., 2015, Show, attend and tell: Neural *Image caption* generation with visual attention, in: Proceedings of the International conference on *machine learning*, pp. 2048–2057.
- Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. 2014. How Transferable Are Features in Deep Neural Networks? *Advances In Neural Information Processing Systems*, 3320–3328.
- You, Q., Jin, H., Wang, Z., Fang, C., Luo, J 2016 *Image captioning* with semantic attention. In proceedings of the IEEE conference on *computer vision and pattern recognition*, 4651–4659
- Yu, L., Nikandrou, M., Jin, J., Rieser, V., 2023, Quality-agnostic *Image captioning* to Safely Assist People with Vision Impairment. Tianjin University of Technology, Tian jin, China & Heriot-Watt University, Edinburgh, United Kingdom & Alana AI
- Zaman, S., Abrar M. A., Hassan, M, M., Naiful Islam, A, N, M., 2019, A Recurrent Neural Network Approach to *Image captioning* in Braille for Blind-Deaf People, IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), Dhaka, Bangladesh, pp. 49-53
- Zhang, W., Nie, W., Li, X., dan Yu, Y., 2019, *Image Caption* Generation with Adaptive Transformer, 2019 34rd Youth Academic Annual Conference of Chinese Association of Automation (YAC), Jinzhou.
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. 2020, A comprehensive survey on *transfer learning*. *Proceedings of the IEEE*, 109(1), 43-76.