

BIBLIOGRAPHY

- [1] E. I. Kuznetsova, "Media Reality Of The Digital Age: Challenges Of Visuality," Mar. 2022, pp. 555–562. doi: 10.15405/epsbs.2022.03.66.
- [2] R. Pi, J. Zhang, J. Zhang, R. Pan, Z. Chen, and T. Zhang, "Image Textualization: An Automatic Framework for Creating Accurate and Detailed Image Descriptions," Jun. 2024, [Online]. Available: <http://arxiv.org/abs/2406.07502>
- [3] S. S. Patil and P. J. Patel, "Image Captioning using Deep Learning Model for Visually Impaired People," 2024. [Online]. Available: www.ijfmr.com
- [4] B. Richard, "Key Issues Affecting the Inclusion of Alt Text in Scholarly PDF Publications," *Logos (Netherlands)*, vol. 34, no. 1, pp. 44–60, 2023, doi: 10.1163/18784712-03104058.
- [5] C. Ardito *et al.*, "'Honestly I Never Really Thought About Adding a Description': Why Highly Engaged Tweets Are Inaccessible," in *Proceedings*, Springer, Aug. 2021, pp. 373–395. doi: 10.1007/978-3-030-85623-6_23.
- [6] M. Stefanini, M. Cornia, L. Baraldi, S. Cascianelli, G. Fiameni, and R. Cucchiara, "From Show to Tell: A Survey on Deep Learning-based Image Captioning," Jul. 2021, [Online]. Available: <http://arxiv.org/abs/2107.06912>
- [7] Y. Li, "Image Caption using VGG model and LSTM," *Applied and Computational Engineering*, vol. 48, no. 1, pp. 68–77, Mar. 2024, doi: 10.54254/2755-2721/48/20241175.
- [8] A. Nursikuwagus, R. Munir, and M. L. Khodra, "Image Captioning menurut Scientific Revolution Kuhn dan Popper," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 10, Oct. 2020, doi: 10.34010/jamika.v10i2.
- [9] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network," *Physica D*, vol. 404, Mar. 2020, doi: 10.1016/j.physd.2019.132306.
- [10] T. Lin, Y. Wang, X. Liu, and X. Qiu, "A survey of transformers," *AI Open*, vol. 3, pp. 111–132, Jan. 2022, doi: 10.1016/j.aiopen.2022.10.001.
- [11] M. J. Panicker, V. Upadhayay, G. Sethi, and V. Mathur, "Image Caption Generator," *International Journal of Innovative Technology and Exploring Engineering*, vol. 10, no. 3, pp. 87–92, Jan. 2021, [Online]. Available: <https://www.ijitee.org/portfolio-item/C83830110321/>

- [12] C. Wang, Z. Zhou, and L. Xu, "An Integrative Review of Image Captioning Research," in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Jan. 2021. doi: 10.1088/1742-6596/1748/4/042060.
- [13] A. Nurhopipah, J. Suhaman, and A. Widiyanto, "Exploring Pre-Trained Model and Language Model for Translating Image to Bahasa," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 17, no. 4, p. 347, Oct. 2023, doi: 10.22146/ijccs.76389.
- [14] C. Wu, Y. Li, L. Li, L. Wang, and Y. Liu, "Caption Generation from Road Images for Traffic Scene Construction," in *IEEE Intelligent Vehicles Symposium, Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2020, pp. 1271–1276. doi: 10.1109/IV47402.2020.9304746.
- [15] A. Verma, A. K. Yadav, M. Kumar, and D. Yadav, "Automatic Image Caption Generation Using Deep Learning," 2022, doi: 10.21203/rs.3.rs-1282936/v1.
- [16] S. Amirian, K. Rasheed, T. R. Taha, and H. R. Arabnia, "Automatic Image and Video Caption Generation with Deep Learning: A Concise Review and Algorithmic Overlap," 2020, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2020.3042484.
- [17] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," Oct. 2020, [Online]. Available: <http://arxiv.org/abs/2010.11929>
- [18] B. Wu *et al.*, "Visual Transformers: Token-based Image Representation and Processing for Computer Vision," 2020.
- [19] A. Gholami, S. Kim, Z. Dong, Z. Yao, M. W. Mahoney, and K. Keutzer, "A Survey of Quantization Methods for Efficient Neural Network Inference," Mar. 2021, [Online]. Available: <http://arxiv.org/abs/2103.13630>
- [20] F. Petersen, H. Kuehne, C. Borgelt, and O. Deussen, "Differentiable Top-k Classification Learning," Jun. 2022, [Online]. Available: <http://arxiv.org/abs/2206.07290>
- [21] M. Barbella and G. Tortora, "ROUGE metric evaluation for Text Summarization techniques," Fisciano, Mar. 2022. [Online]. Available: <https://ssrn.com/abstract=4120317>
- [22] F. Chiusano, "Two minutes NLP: Learn BLEU metric by example." Accessed: Jun. 04, 2025. [Online]. Available: <https://medium.com/nlplanet/two-minutes-nlp-learn-the-bleu-metric-by-examples-df015ca73a86>
- [23] D. Romadinur and J. Pardede, "Image Captioning Menggunakan Transformer Dan Area Attention," in *2023: Prosiding Diseminasi FTI Genap 2022/2023*, Sep. 2023. Accessed: Jun. 11, 2025. [Online].

Available:

<https://eproceeding.itenas.ac.id/index.php/fti/article/view/3277>

- [24] D. H. Fudholi, A. Zahra, and R. A. N. Nayoan, "A Study on Visual Understanding Image Captioning using Different Word Embeddings and CNN-Based Feature Extractions," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, Feb. 2022, doi: 10.22219/kinetik.v7i1.1394.
- [25] A. Vaswani *et al.*, "Attention Is All You Need," Aug. 2023, [Online]. Available: <http://arxiv.org/abs/1706.03762>