UNIVERSITAS
GADJAH MADA
**Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur**
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Niken Wirasanti, M.Si.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# DAFTAR PUSTAKA

Abdulateef, S. M., & Salman, F. H. (2021). A comprehensive review of image segmentation techniques. Iraqi Journal of Electrical and Electronic Engineering, 17(2), 166–178.

Auliasari, K., & Orisa, M. (2017). Algorithm for identifying objects in the relief image using watershed segmentation. Scientific Journal of Informatics, 4(2), 104–114.

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional *encoder-decoder* architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

Bandyopadhyay, S. (2023). A survey on semantic segmentation techniques. Journal of Visual Communication and Image Representation, 94, 103848.

Bradski, G. (2000). The OpenCV library. https://opencv.org

Cao, J., Bui, T. D., & Bhanu, B. (2019). Cell segmentation based on morphological watersheds. In Proceedings of the IEEE ISBI, 1348-1351.

Cawley, G. C., & Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. Journal of *Machine Learning* Research, 11, 2079–2107. https://www.jmlr.org/papers/v11/cawley10a.html

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic image segmentation with deep convolutional nets, *atrous convolution*, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848. https://doi.org/10.1109/TPAMI.2017.2699184

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). *Encoder-decoder* with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), LNCS Vol. 11211 (pp. 833–851). https://doi.org/10.1007/978-3-030-01234-2_49

Cheng, B., Misra, I., Schwing, A. G., Kirillov, A., & Garg, R. (2022). Masked-attention mask transformer for universal image segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1290–1299). https://doi.org/10.1109/CVPR52688.2022.00135

UNIVERSITAS GADJAH MADA
**Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur**
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Ika Ken Wirasanti, M.Si.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**


Csurka, G., Larlus, D., & Perronnin, F. (2013). What is a good evaluation measure for semantic segmentation? In Proceedings of the British Machine Vision Conference (BMVC). https://doi.org/10.5244/C.27.32

Dong, X., Garratt, M., Anavatti, S. G., & Abbass, H. A. (2022). Towards real-time monocular *depth estimation* for robotics: A survey. IEEE Transactions on Intelligent Transportation Systems, 23(10), 16940–16961. https://doi.org/10.1109/tits.2022.3160741

Frisky, A. Z. K., Harjoko, A., Awaludin, L., Dharmawan, A., Augoestien, N. G., et al. (2021). Registered relief depth (RRD) Borobudur dataset for single-frame depth prediction on one-side artifacts. Data in Brief, 35, 106853. https://doi.org/10.1016/j.dib.2021.106853

Frisky, A. Z. K., Putranto, A., Zambanini, S., & Sablatnig, R. (2021). MCCNet: Multi-Color Cascade Network with Weight Transfer for Single Image Depth Prediction on Outdoor Relief Images. In A. Del Bimbo et al. (Eds.), Pattern Recognition. ICPR International Workshops and Challenges (Lecture Notes in Computer Science, Vol. 12667, pp. 263–278). Springer. https://doi.org/10.1007/978-3-030-68787-8_19

Ge, Y., Guo, B., Zha, P., Jiang, S., Jiang, Z., & Li, D. (2024). 3D reconstruction of ancient buildings using UAV images and neural radiation field with depth supervision. Remote Sensing, 16(3), 473. https://doi.org/10.3390/rs16030473

Geiger, A., Roser, M., & Urtasun, R. (2010). Efficient large-scale stereo matching. In Asian Conference on Computer Vision (ACCV) (pp. 25–38). Springer.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS), JMLR Workshop and Conference Proceedings, 9, 249–256.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., et al. (2014). Recent advances in convolutional neural networks. Pattern Recognition, 77, 354-377.

Gunarto, H. (2007). Documentation of Borobudur reliefs based on digital image processing. Ritsumeikan Journal of Asia Pacific Studies, 22, 53-70.

Harris, C. G., & Stephens, M. (1988). A combined corner and edge detector. In Proceedings of the Alvey Vision Conference (Vol. 15, No. 50, pp. 10–5244).

Hazirbas, C., Ma, L., Domokos, C., & Cremers, D. (2016). FuseNet: Incorporating depth into semantic segmentation via fusion-based CNN architecture. In

UNIVERSITAS
GADJAH MADA
**Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur**
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Ika Ken Wirasanti, M.Si.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**


Proceedings of the Asian Conference on Computer Vision (ACCV) (pp. 213–228). https://doi.org/10.1007/978-3-319-54181-5_14

He, K., Sun, J., & Tang, X. (2013). Guided image filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(6), 1397–1409. https://doi.org/10.1109/TPAMI.2012.213

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 1026–1034).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 770–778). https://doi.org/10.1109/CVPR.2016.90

Hirschmuller, H. (2005). Accurate and efficient stereo processing by semi-global matching and mutual information. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 807–814).

Hirschmuller, H. (2008). Stereo processing by semiglobal matching and mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2), 328–341. https://doi.org/10.1109/TPAMI.2007.1166

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4700–4708). https://doi.org/10.1109/CVPR.2017.243

Intel Corporation. (2026). Intel RealSense D455 depth camera. Retrieved from https://www.intelrealsense.com/depth-camera-d455/

Ji, S., Pan, J., Li, L., Hasegawa, K., Yamaguchi, H., Thufail, F. I., et al. (2023). Semantic segmentation for digital archives of Borobudur reliefs based on soft-edge enhanced *deep learning*. Remote Sensing, 15(4), 956. https://doi.org/10.3390/rs15040956

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations (ICLR). https://doi.org/10.48550/arXiv.1412.6980

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., et al. (2023). Segment anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 4015–4026). https://doi.org/10.1109/ICCV51070.2023.00371

Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief
Candi Borobudur
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Ika Nurlaken Wirasanti, M.Si.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/


Kittler, J., Hatef, M., Duin, R. P. W., & Matas, J. (1998). On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(3), 226–239.

LeCun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional networks and applications in vision. In Proceedings of the IEEE ISCAS, 253-256.

Lei, Y., Gu, L., & Shen, M. (2015). An improved ORB algorithm of extracting and matching features. International Journal of Signal Processing, Image Processing and Pattern Recognition, 8(12), 117–126. https://doi.org/10.14257/ijsip.2015.8.12.11

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2980–2988). https://doi.org/10.1109/ICCV.2017.324

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollar, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 740–755). https://doi.org/10.1007/978-3-319-10602-1_48

Lindeberg, T. (1998). Feature detection with automatic scale selection. International Journal of Computer Vision, 30(2), 79-116.

Lipson, L., Teed, Z., & Deng, J. (2021). RAFT-Stereo: Multilevel recurrent field transforms for stereo matching. In Proceedings of the International Conference on 3D Vision (3DV) (pp. 218–227). https://doi.org/10.1109/3DV53792.2021.00032

Liu, C., Yuen, J., & Torralba, A. (2007). SIFT flow: Dense correspondence across scenes. IEEE TPAMI, 33(5), 978-994.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin Transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 10012–10022). https://doi.org/10.1109/ICCV48922.2021.00986

Lu, J., Sarkis, M., & Lu, G. (2020). Multi-task learning for dense prediction tasks: A survey. IEEE TPAMI, 44(7), 3614-3633.

Ma, H., Zheng, S., Li, Y., Gui, L., Huang, R., & Wei, H. (2018). Confidence-based iterative efficient large-scale stereo matching. Cogent Engineering, 5(1), 1427676. https://doi.org/10.1080/23311916.2018.1427676

Marr, D., & Hildreth, E. (1980). Theory of edge detection. Proceedings of the Royal Society of London. Series B. Biological Sciences, 207(1167), 187–217.

UNIVERSITAS
GADJAH MADA
**Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur**
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Ika Ken Wirasanti, M.Si.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**


Mei, X., Sun, X., Zhou, M., Jiao, S., Wang, H., & Zhang, X. (2011). On building an accurate stereo matching system on graphics hardware. In IEEE International Conference on Computer Vision Workshops (ICCVW) (pp. 467–474).

Mutlag, A. A., Abdul Ghani, M. K., Arunkumar, N., Mohammed, M. A., & Mohd, O. (2020). Enabling technologies for fog computing in healthcare IoT systems. Future Generation Computer Systems, 90, 62-78.

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., et al. (1987). Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing, 39(3), 355–368. https://doi.org/10.1016/S0734-189X(87)80186-X

Remondino, F. (2011). Heritage recording and 3D modeling with photogrammetry and 3D scanning. Remote Sensing, 3(6), 1104–1138. https://doi.org/10.3390/rs3061104

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Rosten, E., & Drummond, T. (2006). *Machine learning* for high-speed corner detection. In Proceedings of the European Conference on Computer Vision (ECCV), LNCS Vol. 3951 (pp. 430-443). https://doi.org/10.1007/11744023_34

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2564–2571). https://doi.org/10.1109/iccv.2011.6126544

Scharstein, D., & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, 47(1–3), 7–42. https://doi.org/10.1023/A:1014573219977

Soekmono, R. (1976). Chandi Borobudur: A monument of mankind. UNESCO Press.

Song, S., Yu, F., Zeng, A., Chang, A. X., Savva, M., & Funkhouser, T. (2017). Semantic scene completion from a single depth image. In Proceedings of the IEEE CVPR, 1746-1754.

Tsintotas, K. A., Bampis, L., & Gasteratos, A. (2022). The revisiting problem in simultaneous localization and mapping: A survey on visual loop closure

UNIVERSITAS GADJAH MADA
**Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur**
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Ika Wirasanti, M.Si.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**


detection. IEEE Transactions on Intelligent Transportation Systems, 23(11), 19929–19953. https://doi.org/10.1109/TITS.2022.3175618

UNESCO World Heritage Centre. (2026). Borobudur Temple Compounds. Retrieved from https://whc.unesco.org/en/list/592

Wan, J., Zhang, J., Xu, L., & Wang, H. (2022). Edge-guided image segmentation with boundary-aware network. Pattern Recognition, 129, 108746.

Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., & Cottrell, G. (2014). Understanding convolution for semantic segmentation. In Proceedings of the IEEE WACV, 1451-1460.

Wirasanti, N., & Murwanto, H. (2020). The reconstruction of a Javanese civilization cultural landscape in 8 AD based on Canggal Inscription in Gendol Hill Complex, Magelang Regency, Indonesia.

Wirasanti, N., Murwanto, H., & Prasetyo, B. (2026). Ikonografi dan narasi visual relief Candi Borobudur: Perspektif multidisiplin. Gadjah Mada University Press.

Wirasanti, N. (2023). Candi dan lingkungan: Abad IX-X Masehi di wilayah Jawa bagian tengah (D. Surani, Penyunt.). Gadjah Mada University Press. ISBN 978-623-359-121-8.

Xie, S., Wang, W., Zhu, X., Hu, H., Wang, X., Yao, T., & Tian, Y. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. Advances in Neural Information Processing Systems (NeurIPS), 34, 12077–12090. https://doi.org/10.48550/arXiv.2105.15203

Yang, L., Zhuo, W., Qi, L., Shi, Y., & Gao, Y. (2024). ST++: Make self-training work better for semi-supervised semantic segmentation. In Proceedings of the IEEE/CVF CVPR, 4268-4277.

Yao, C., Bai, X., & Liu, W. (2016). A unified framework for multi-oriented text detection and recognition. IEEE TIP, 23(11), 4737-4749.

Yeung, H. W., Hou, J., Chen, J., Chung, Y. Y., & Chen, X. (2022). Deep learning for *depth estimation*: A survey. Pattern Recognition, 125, 108497.

Yin, B., Zhang, Z., Du, C., Zhou, Q., Lv, Z., & Fu, C. (2024). DFormer: Rethinking RGBD representation learning for semantic segmentation. In Proceedings of the International Conference on Learning Representations (ICLR). https://doi.org/10.48550/arXiv.2309.09668

Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., & Yoo, Y. (2019). CutMix: Regularization strategy to train strong classifiers with localizable features.

UNIVERSITAS
GADJAH MADA
**Pengembangan Metode Complementary-Stream Sum Fusion Untuk Segmentasi Semantik Pada Citra Relief Candi Borobudur**
Lukman Awaludin, Prof. Drs. Agus Harjoko, M.Sc., Ph.D.; Wahyono, S.Kom., Ph.D.; Dr. Dra. Niken Wirasanti, M.Si.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**


In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 6023–6032).

Yunmar, R. A., & Harjoko, A. (2014). Sistem identifikasi relief pada situs bersejarah menggunakan perangkat mobile Android (studi kasus Candi Borobudur). Indonesian Journal of Computing and Cybernetics Systems (IJCCS), 8(2), 133–144.

Zabih, R., & Woodfill, J. (1994). Non-parametric local transforms for computing visual correspondence. In European Conference on Computer Vision (ECCV) (pp. 151–158). Springer.

Zhang, H., Zhao, J., Shen, D., Chen, Y., & Yang, J. (2023). CMX: Cross-modal fusion for RGB-X semantic segmentation with transformers. IEEE Transactions on Intelligent Transportation Systems, 24(12), 14679–14694. https://doi.org/10.1109/TITS.2023.3300537

Zhang, Z., Zhang, X., Peng, C., Xue, X., & Sun, J. (2018). ExFuse: Enhancing feature fusion for semantic segmentation. In Proceedings of the ECCV, 269-284.

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2881–2890). https://doi.org/10.1109/CVPR.2017.660