

Bibliography

- [1] Athar, Ali, Mahadevan, Sabarinath, Osep, Aljosa, Leal-Taixé, Laura, and Leibe, Bastian. “Stem-seg: Spatio-temporal embeddings for instance segmentation in videos”. In: *European Conference on Computer Vision*. Springer. 2020, pp. 158–177.
- [2] Ayyadevara, Yeshwanth Reddy; V Kishore. *Modern Computer Vision with PyTorch*. Packt Publishing, 2020.
- [3] Caelles, Sergi, Pont-Tuset, Jordi, Perazzi, Federico, Montes, Alberto, Maninis, Kevis-Kokitsi, and Van Gool, Luc. “The 2019 davis challenge on vos: Unsupervised multi-object segmentation”. In: *arXiv preprint arXiv:1905.00737* (2019).
- [4] Cao, Zheng, Principe, José C, Ouyang, Bing, Dalgleish, Fraser, Vuorenkoski, Anni, Ramos, Brian, and Alsenas, Gabriel. “Marine animal classification using UMSLI in HBOI optical test facility”. In: *Multimedia Tools and Applications* 76.21 (2017), pp. 23117–23138.
- [5] Carion, Nicolas, Massa, Francisco, Synnaeve, Gabriel, Usunier, Nicolas, Kirillov, Alexander, and Zagoruyko, Sergey. “End-to-end object detection with transformers”. In: *European Conference on Computer Vision*. Springer. 2020, pp. 213–229.
- [6] Chan, Sixian, Huang, Cheng, Bai, Cong, Ding, Weilong, and Chen, Shengyong. “Res2-UNeXt: a novel deep learning framework for few-shot cell image segmentation”. In: *Multimedia Tools and Applications* (2021), pp. 1–14.
- [7] Chaudhary, Aayush K, Kothari, Rakshit, Acharya, Manoj, Dangi, Shusil, Nair, Nitinraj, Bailey, Reynold, Kanan, Christopher, Diaz, Gabriel, and Pelz, Jeff B. “RITnet: real-time semantic segmentation of the eye for gaze tracking”. In: *2019*

- IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. IEEE. 2019, pp. 3698–3702.
- [8] Chu, Xiangxiang, Zhang, Bo, Tian, Zhi, Wei, Xiaolin, and Xia, Huaxia. “Do We Really Need Explicit Position Encodings for Vision Transformers?” In: *arXiv preprint arXiv:2102.10882* (2021).
- [9] Deng, Jia, Dong, Wei, Socher, Richard, Li, Li-Jia, Li, Kai, and Fei-Fei, Li. “Imagenet: A large-scale hierarchical image database”. In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee. 2009, pp. 248–255.
- [10] Fan, Zhibo, Yu, Jin-Gang, Liang, Zhihao, Ou, Jiarong, Gao, Changxin, Xia, Gui-Song, and Li, Yuanqing. “Fgn: Fully guided network for few-shot instance segmentation”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020, pp. 9172–9181.
- [11] French, Robert M. “Catastrophic forgetting in connectionist networks”. In: *Trends in cognitive sciences 3.4* (1999), pp. 128–135.
- [12] Geiger, Andreas, Lenz, Philip, and Urtasun, Raquel. “Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite”. In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2012.
- [13] He, Kaiming, Gkioxari, Georgia, Dollár, Piotr, and Girshick, Ross. “Mask r-cnn”. In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2961–2969.
- [14] Hoiem, Derek, Divvala, Santosh K, and Hays, James H. “Pascal VOC 2008 challenge”. In: *World Literature Today* (2009).
- [15] Honda, Hiroto. *Digging into Detectron 2 — part 1*. URL: <https://medium.com/@hirotoschwert/digging-into-detectron-2-47b2e794fabd>.
- [16] Kendall, Alex, Gal, Yarin, and Cipolla, Roberto. “Multi-task learning using uncertainty to weigh losses for scene geometry and semantics”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 7482–7491.

- [17] Kirkpatrick, James, Pascanu, Razvan, Rabinowitz, Neil, Veness, Joel, Desjardins, Guillaume, Rusu, Andrei A, Milan, Kieran, Quan, John, Ramalho, Tiago, Grabska-Barwinska, Agnieszka, et al. “Overcoming catastrophic forgetting in neural networks”. In: *Proceedings of the national academy of sciences* 114.13 (2017), pp. 3521–3526.
- [18] Lin, Chung-Ching, Hung, Ying, Feris, Rogerio, and He, Linglin. “Video instance segmentation tracking with a modified vae architecture”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 13147–13157.
- [19] Lin, Tsung-Yi, Maire, Michael, Belongie, Serge, Hays, James, Perona, Pietro, Ramanan, Deva, Dollár, Piotr, and Zitnick, C Lawrence. “Microsoft coco: Common objects in context”. In: *European conference on computer vision*. Springer. 2014, pp. 740–755.
- [20] Liu, Shu, Qi, Lu, Qin, Haifang, Shi, Jianping, and Jia, Jiaya. “Path aggregation network for instance segmentation”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 8759–8768.
- [21] Maynard-Reid, Margareth. *Hands-On Transfer Learning with TensorFlow 2.0*. 9781789953947. Packt Publishing, 2020.
- [22] Mazur-Milecka, Magdalena, Kocejko, Tomasz, and Ruminski, Jacek. “Deep Instance Segmentation of Laboratory Animals in Thermal Images”. In: *Applied Sciences* 10.17 (2020), p. 5979.
- [23] Mazur-Milecka, Magdalena and Ruminski, Jacek. “Deep learning based thermal image segmentation for laboratory animals tracking”. In: *Quantitative InfraRed Thermography Journal* (2020), pp. 1–18.
- [24] Nesteruk, Sergey, Shadrin, Dmitrii, and Pukalchik, Mariia. “Image Augmentation for Multitask Few-Shot Learning: Agricultural Domain Use-Case”. In: *arXiv preprint arXiv:2102.12295* (2021).
- [25] Pan, Sinno Jialin and Yang, Qiang. “A survey on transfer learning”. In: *IEEE Transactions on knowledge and data engineering* 22.10 (2009), pp. 1345–1359.

- [26] Payer, Christian, Štern, Darko, Neff, Thomas, Bischof, Horst, and Urschler, Martin. “Instance segmentation and tracking with cosine embeddings and recurrent hourglass networks”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2018, pp. 3–11.
- [27] Perry, Jonathan and Fernandez, Amanda S. “EyeSeg: Fast and Efficient Few-Shot Semantic Segmentation”. In: *European Conference on Computer Vision*. Springer. 2020, pp. 570–582.
- [28] Pont-Tuset, Jordi, Perazzi, Federico, Caelles, Sergi, Arbeláez, Pablo, Sorkine-Hornung, Alex, and Van Gool, Luc. “The 2017 davis challenge on video object segmentation”. In: *arXiv preprint arXiv:1704.00675* (2017).
- [29] Pont-Tuset, Jordi, Perazzi, Federico, Caelles, Sergi, Arbeláez, Pablo, Sorkine-Hornung, Alex, and Van Gool, Luc. “The 2017 davis challenge on video object segmentation”. In: *arXiv preprint arXiv:1704.00675* (2017).
- [30] Redmon, Joseph, Divvala, Santosh, Girshick, Ross, and Farhadi, Ali. “You only look once: Unified, real-time object detection”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 779–788.
- [31] Ren, Shaoqing, He, Kaiming, Girshick, Ross, and Sun, Jian. “Faster r-cnn: Towards real-time object detection with region proposal networks”. In: *arXiv preprint arXiv:1506.01497* (2015).
- [32] Ruder, Sebastian. *Transfer Learning - Machine Learning’s Next Frontier*. <http://ruder.io/transfer-learning/>. 2017.
- [33] Sweden, AI. *Baltic Seabird dataset*. https://www.ai.se/sites/default/files/content/resource/files/baltic_seabird_dataset_description.pdf. [Online; accessed 22-February-20021]. 2021.
- [34] Taylor, Matthew E and Stone, Peter. “Transfer learning for reinforcement learning domains: A survey.” In: *Journal of Machine Learning Research* 10.7 (2009).
- [35] Vaswani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N, Kaiser, Lukasz, and Polosukhin, Illia. “Attention is all you need”. In: *arXiv preprint arXiv:1706.03762* (2017).

- [36] Voigtlaender, Paul, Krause, Michael, Osep, Aljosa, Luiten, Jonathon, Sekar, Berin Balachandar Gnana, Geiger, Andreas, and Leibe, Bastian. “Mots: Multi-object tracking and segmentation”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, pp. 7942–7951.
- [37] Voigtlaender, Paul, Krause, Michael, Osep, Aljosa, Luiten, Jonathon, Sekar, Berin Balachandar Gnana, Geiger, Andreas, and Leibe, Bastian. “MOTS: Multi-Object Tracking and Segmentation”. In: *CVPR*. 2019.
- [38] Wang, Shuo Hong, Zhao, Jing Wen, and Chen, Yan Qiu. “Robust tracking of fish schools using CNN for head identification”. In: *Multimedia Tools and Applications* 76.22 (2017), pp. 23679–23697.
- [39] Wang, Xinlong, Kong, Tao, Shen, Chunhua, Jiang, Yuning, and Li, Lei. “Solo: Segmenting objects by locations”. In: *European Conference on Computer Vision*. Springer. 2020, pp. 649–665.
- [40] Wang, Yuqing, Xu, Zhaoliang, Wang, Xinlong, Shen, Chunhua, Cheng, Baoshan, Shen, Hao, and Xia, Huaxia. “End-to-End Video Instance Segmentation with Transformers”. In: *arXiv preprint arXiv:2011.14503* (2020).
- [41] Wu, Yuxin, Kirillov, Alexander, Massa, Francisco, Lo, Wan-Yen, and Girshick, Ross. *Detectron2*. 2019.
- [42] Xu, Ning, Yang, Linjie, Fan, Yuchen, Yang, Jianchao, Yue, Dingcheng, Liang, Yuchen, Price, Brian, Cohen, Scott, and Huang, Thomas. “Youtube-vos: Sequence-to-sequence video object segmentation”. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018, pp. 585–601.
- [43] Xu, X. and Kakadiaris, I. A. “Joint Head Pose Estimation and Face Alignment Framework Using Global and Local CNN Features”. In: *2017 12th IEEE International Conference on Automatic Face Gesture Recognition (FG 2017)*. 2017, pp. 642–649. DOI: 10.1109/FG.2017.81.
- [44] Yang, Linjie, Fan, Yuchen, and Xu, Ning. “Video instance segmentation”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, pp. 5188–5197.
- [45] Zhao, Amy, Balakrishnan, Guha, Durand, Fredo, Guttag, John V, and Dalca, Adrian V. “Data augmentation using learned transformations for one-shot medical image segmentation”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, pp. 8543–8553.