REFERENCES

[1] X. Chen, H. Ma, J. Wan, B. Li and T. Xia, "Multi-View 3D Object Detection Network for Autonomous Driving," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1907-1915, 2017.

[2] Y. Chen, G. Liu, Y. Xu, P. Pan and Y. Xing, "PointNet++ Network Architecture with Individual Point Level and Global Features on Centroid for ALS Point Cloud Classification," *Remote Sens,* vol. 13, no. 472, pp. 1-17, 2021.

[3] A. Geiger, P. Lenz and R. Urtasun, "Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite," *IEEE Conference on Computer Vision and Pattern Recognition,* pp. 3354-3361, 2012.

[4] D. Yin, Q. Zhang, J. Liu, X. Liang, Y. Wang, J. Maanpaa, H. Ma, J. Hyyppä and R. Chen, "CAE-LO: LiDAR Odometry Leveraging Fully Unsupervised Convolutional Auto-Encoder for Interest Point Detection and Feature Description," *arXiv preprint arXiv:2001.01354,* 2020.

[5] X. Chen, S. Li, B. Mersch, L. Wiesmann, J. Gall, J. Behley and C. Stachniss, "Moving Object Segmentation in 3D LiDAR Data: A Learning-Based Approach Exploiting Sequential Data," *IEEE Robotics and Automation Letters,* vol. 6, no. 4, pp. 6529-6536, 2021.

[6] X. Chen, A. Milioto, E. Palazzolo, P. Giguère, J. Behley and C. Stachniss, "SuMa++: Efficient LiDAR-based Semantic SLAM," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* 2019.

[7] D. J. Yoon, H. Zhang, M. Gridseth, H. Thomas and T. D. Barfoot, "Unsupervised Learning of Lidar Features for Use in a Probabilistic Trajectory Estimator," *IEEE Robotics and Automation Letters,* vol. 6, no. 2, pp. 2130-2138, 2021.

[8] G. Wang, X. Wu, S. Jiang, Z. Liu, Wang and Hesheng, "Efficient 3D Deep LiDAR Odometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 45, no. 5, pp. 5749-5765, 2023.

[9] M. Horn, N. Engel, V. Belagiannis, M. Buchholz and K. Dietmayer, "DeepCLR: Correspondence-Less Architecture for Deep End-to-End Point Cloud Registration," *IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC),* 2020.

[10] P. Adis, N. Horst and M. Wien, "D3dlo: Deep 3d Lidar Odometry," *IEEE International Conference on Image Processing,* pp. 3128-3132, 2021.

[11] Z. Tan, X. Zhang, S. Teng, L. Wang and F. Gao, "A Review of Deep Learning-Based LiDAR and Camera Extrinsic Calibration," *Sensors,* vol. 3878, pp. 1-42, 2024.

[12] B. Zoph, V. Vasudevan, J. Shlens and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 8697-8710, 2018.

[13] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 2117-2125, 2017.

[14] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *ArXiv:1502.03167 [Cs],* pp. 1-11, 2015.

[15] A. V. S. Abhishek and S. Kotni, "Detectron2 Object Detection & Manipulating Images using Cartoonization," *International Journal of Engineering Research & Technology (IJERT),* vol. 10, no. 8, pp. 322-326, 2021.

[16] J. Lee, M. Back, S. S. Hwang and I. Y. Chun, "Improved Real-Time Monocular SLAM Using Semantic Segmentation on Selective Frames," *IEEE Transactions on Intelligent Transportation Systems,* vol. 24, no. 3, pp. 2800 - 2813, 2023.

[17] L. Caltagirone, M. Bellone, L. Svensson and M. Wahde, "LIDAR–camera fusion for road detection using fully convolutional neural networks," *Robotics and Autonomous Systems,* vol. 111, pp. 125-131, 2019.

[18] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan and O. Beijbom, "nuScenes: A multimodal dataset for autonomous driving," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 11621-11631, 2020.

[19] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *International Conference on Learning Representations (ICLR),* 2015.

[20] A. Roy, N. Gale and L. Hong, "Automated traffic surveillance using fusion of Doppler radar and video information," *Mathematical and Computer Modelling,* vol. 54, no. 1-2, pp. 531-543, 2011.

[21] D. J. Yeong, G. V. Hernandez, J. Barry and J. Walsh, "Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review," *Sensors,* vol. 21, no. 2140, pp. 1-37, 2021.

[22] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge: Cambridge university press, 2003.

[23] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 770-778, 2016.

[24] Y. Wu and K. He, "Group Normalization," *Proceedings of the European conference on computer vision (ECCV),* pp. 3-19, 2018.

[25] S. Santurkar, D. Tsipras, A. Ilyas and A. Madry, "How Does Batch Normalization Help Optimization?," *Advances in neural information processing systems,* vol. 31, 2018.

[26] N. Bjorck, C. P. Gomes, B. Selman and K. Q. Weinberger, "Understanding batch normalization," *Advances in neural information processing systems,* vol. 31, 2018.

[27] P. Luo, J. Ren, Z. Peng, R. Zhang and J. Li, "Differentiable learning-to-normalize via switchable normalization," *arXiv preprint arXiv:1806.10779,* 2018.

[28] D. Mishkin and J. Matas, "All you need is a good init," *arXiv preprint arXiv:1511.06422,* 2015.

[29] J. L. Ba, J. R. Kiros and G. E. Hinton, "Layer normalization," *ArXiv e-prints,* 2016.

[30] P. Dollar, R. Appel, S. Belongie and P. Perona, "Fast feature pyramids for object detection," *IEEE transactions on pattern analysis and machine intelligence,* vol. 36, no. 8, pp. 1532-1545, 2014.

[31] Z. Cai and N. Vasconcelos, "Cascade r-cnn: Delving into high quality object detection," *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 6154 - 6162, 2018.

[32] Z. Huang, L. Huang, Y. Gong, C. Huang and X. Wang, "Mask scoring r-cnn," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,* pp. 6409-6418, 2029.

[33] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg, "Ssd: Single shot multibox detector," *Computer Vision-ECCV 2016,* pp. 21-37, 2016.

[34] Z. Wang, J. Zhu, S. Fu, S. Mao and Y. Ye, "RFPNet: Reorganizing feature pyramid networks for medical image segmentation," *Computers in biology and medicine,* vol. 163, 2023.

[35] K. He, G. Gkioxari, P. Dollar and R. Girshick, "Mask r-cnn," *Proceedings of the IEEE international conference on computer vision,* pp. 2961-2969, 2017.

[36] Q. M. ul Haq, S. J. Ruan, M. A. Haq, S. Karam, J. L. Shieh, P. Chondro and D. Q. Gao, "An incremental learning of yolov3 without catastrophic forgetting for smart city applications," *IEEE Consumer Electronics Magazine,* vol. 11, no. 5, pp. 56-63, 2021.

[37] J. Yosinski and J. a. B. Y. a. L. H. Clune, "How transferable are features in deep neural networks?," *Advances in neural information processing systems,* vol. 27, 2014.

[38] M. Raghu, C. Zhang, J. Kleinberg and S. Bengio, "Transfusion: Understanding transfer learning for medical imaging," *Advances in neural information processing systems,* vol. 32, 2019.

[39] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering,* vol. 22, no. 10, pp. 1345-1356, 2009.

[40] H. Azizpour, A. S. Razavian, J. Sullivan, A. Maki and S. Carlsson, "Factors of transferability for a generic convnet representation," *IEEE transactions on pattern analysis and machine intelligence,* vol. 38, no. 9, pp. 1790-1802, 2015.

[41] M. Huh, P. Agrawal and A. A. Efros, "What makes ImageNet good for transfer learning?," *arXiv preprint arXiv:1608.08614,* 2016.

[42] S. Ruder, M. E. Peters, S. Swayamdipta and T. Wolf, "Transfer learning in natural language processing," *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: Tutorials,* pp. 15-18, 2019.

[43] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," *arXiv preprint arXiv:1801.06146,* 2018.

[44] S. Kornblith, J. Shlens and Q. V. Le, "Do better imagenet models transfer better?," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,* pp. 2661-2671, 2019.

[45] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 580-587, 2014.

[46] D. Mahajan, R. Girshick, V. Ramanathan, K. He, M. Paluri, Y. Li, A. Bharambe and L. Van Der Maaten, "Exploring the limits of weakly

supervised pretraining," *Proceedings of the European conference on computer vision (ECCV),* pp. 181-196, 2018.

[47] K. Karsch, C. Liu and S. B. Kang, "Depth Transfer: Depth Extraction from Video Using Non-Parametric Sampling," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* vol. 36, no. 11, pp. 2144-2158, 2014.

[48] M. Liu, M. Salzmann and X. He, "Discrete-Continuous Depth Estimation from a Single Image," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 716-723, 2014.

[49] A. Saxena, M. Sun and A. Y. Ny, "Make3D: Learning 3D Scene Structure from a Single Still Image," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* vol. 31, no. 5, pp. 824-840, 2009.

[50] D. Eigen, C. Puhrsch and R. Fergus, "Depth Map Prediction from a Single Image using a Multi-Scale Deep Network," *Advances in neural information processing systems,* vol. 27, 2014.