

## Chapter 6 - REFERENCES

- [1] “CAVIAR Test Case Scenarios.” [Online]. Available: <https://groups.inf.ed.ac.uk/vision/DATASETS/CAVIAR/CAVIARDATA1/>
- [2] I. Sodagar, “The MPEG-dash standard for multimedia streaming over the internet,” *IEEE Multimedia*, vol. 18, no. 4, pp. 62–67, 2011, doi: 10.1109/MMUL.2011.71.
- [3] P. Soviany and R. T. Ionescu, “Optimizing the Trade-off between Single-Stage and Two-Stage Object Detectors using Image Difficulty Prediction,” *Proceedings - 2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2018*, pp. 209–214, 2018, doi: 10.1109/SYNASC.2018.00041.
- [4] C. Canel *et al.*, “Scaling Video Analytics on Constrained Edge Nodes,” 2019.
- [5] J. Chen and X. Ran, “Deep Learning With Edge Computing: A Review,” *Proceedings of the IEEE*, 2019, doi: 10.1109/JPROC.2019.2921977.
- [6] S. A. Noghabi, L. Cox, S. Agarwal, and G. Ananthanarayanan, “The Emerging Landscape of Edge Computing,” *GetMobile: Mobile Computing and Communications*, vol. 23, no. 4, pp. 11–20, 2020, doi: 10.1145/3400713.3400717.
- [7] A. Jonathan, A. Chandra, and J. Weissman, “WASP: Wide-area Adaptive Stream Processing,” *Middleware 2020 - Proceedings of the 2020 21st International Middleware Conference*, pp. 221–235, 2020, doi: 10.1145/3423211.3425668.
- [8] T. Y.-H. Chen, H. Balakrishnan, L. Ravindranath, and P. Bahl, “GLIMPSE,” *GetMobile: Mobile Computing and Communications*, vol. 20, no. 1, pp. 26–29, 2016, doi: 10.1145/2972413.2972423.

- [9] T. Zhang, A. Chowdhery, P. Bahl, K. Jamieson, and S. Banerjee, “The design and implementation of a wireless video surveillance system,” *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, vol. 2015-Septe, pp. 426–438, 2015, doi: 10.1145/2789168.2790123.
- [10] K. Apicharttrisorn, X. Ran, J. Chen, S. V Krishnamurthy, and A. K. Roy-Chowdhury, “Frugal following: Power thrifty object detection and tracking for mobile augmented reality,” *SenSys 2019 - Proceedings of the 17th Conference on Embedded Networked Sensor Systems*, pp. 96–109, 2019, doi: 10.1145/3356250.3360044.
- [11] X. Ran, H. Chen, X. Zhu, Z. Liu, and J. Chen, “DeepDecision: A Mobile Deep Learning Framework for Edge Video Analytics,” *Proceedings - IEEE INFOCOM*, vol. 2018-April, pp. 1421–1429, 2018, doi: 10.1109/INFOCOM.2018.8485905.
- [12] K. Du *et al.*, “Server-Driven Video Streaming for Deep Learning Inference,” *SIGCOMM 2020 - Proceedings of the 2020 Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication*, pp. 557–570, 2020, doi: 10.1145/3387514.3405887.
- [13] S. Petrocchi, G. Giorgi, and M. G. C. A. Cimino, “A Real-Time Deep Learning Approach for Real-World Video Anomaly Detection,” *ACM International Conference Proceeding Series*, 2021, doi: 10.1145/3465481.3470099.
- [14] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” pp. 248–255, 2010, doi: 10.1109/CVPR.2009.5206848.

- [15] A. Milan, L. Leal-Taixé, T. Taixé, I. Reid, S. Roth, and K. Schindler, “MOT16: A Benchmark for Multi-Object Tracking,” 2016.
- [16] Y. Li, A. Padmanabhan, P. Zhao, Y. Wang, G. H. Xu, and R. Netravali, “Reducto: On-Camera Filtering for Resource-Efficient Real-Time Video Analytics,” *SIGCOMM 2020 - Proceedings of the 2020 Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication*, pp. 359–376, 2020, doi: 10.1145/3387514.3405874.
- [17] J. Jiang, G. Ananthanarayanan, P. Bodik, S. Sen, and I. Stoica, “Chameleon: Scalable adaptation of video analytics,” *SIGCOMM 2018 - Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pp. 253–266, 2018, doi: 10.1145/3230543.3230574.
- [18] D. Kang, J. Emmons, F. Abuzaid, P. Bailis, and M. Zaharia, “NoScope,” *Proceedings of the VLDB Endowment*, vol. 10, no. 11, pp. 1586–1597, 2017, doi: 10.14778/3137628.3137664.
- [19] B. Haynes, A. Mazumdar, A. Alaghi, M. Balazinska, L. Ceze, and A. Cheung, “LightDB,” *Proceedings of the VLDB Endowment*, vol. 11, no. 10, pp. 1192–1205, 2018, doi: 10.14778/3231751.3231768.
- [20] G. Ananthanarayanan, V. Bahl, L. Cox, A. Crown, S. Noghahi, and Y. Shu, “Demo: Video analytics - Killer app for edge computing,” *MobiSys 2019 - Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 695–696, 2019, doi: 10.1145/3307334.3328589.
- [21] M. Shan, S. Zhang, M. Xiao, and Y. Zhao, “LENS,” *Computer Networks*, vol. 218, 2022, doi: 10.1016/J.COMNET.2022.109392.
- [22] R. Lee, S. I. Venieris, and N. D. Lane, “Deep Neural Network–based Enhancement for Image and Video Streaming Systems: A Survey and Future

- Directions,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 8, 2021, doi: 10.1145/3469094.
- [23] “Microsoft Stream (Classic) video delivery and network overview - Microsoft Stream | Microsoft Learn.”
- [24] Z. Xiao, Z. Xia, H. Zheng, B. Y. Zhao, and J. Jiang, “Towards Performance Clarity of Edge Video Analytics,” *6th ACM/IEEE Symposium on Edge Computing, SEC 2021*, pp. 148–164, 2021, doi: 10.1145/3453142.3491272.
- [25] “Video Analytics: Technologies and Use Cases.”
- [26] J. Yi, C. Min, and F. Kawsar, “Vision Paper: Towards Software-Defined Video Analytics with Cross-Camera Collaboration,” *SenSys 2021 - Proceedings of the 2021 19th ACM Conference on Embedded Networked Sensor Systems*, pp. 474–477, 2021, doi: 10.1145/3485730.3493453.
- [27] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012, doi: 10.1109/TCSVT.2012.2221191.
- [28] D. Kang, P. Bailis, and M. Zaharia, “Blazelt: Optimizing declarative aggregation and limit queries for neural networkbased video analytics,” *Proceedings of the VLDB Endowment*, vol. 13, no. 4, pp. 533–546, 2019, doi: 10.14778/3372716.3372725.
- [29] D. Y. Fu *et al.*, “Rekall: Specifying Video Events using Compositions of Spatiotemporal Labels,” 2019.
- [30] Y. Lu, A. Chowdhery, S. Kandula, and S. Chaudhuri, “Accelerating machine learning inference with probabilistic predicates,” *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 1493–1508, 2018, doi: 10.1145/3183713.3183751.

- [31] K. Hsieh *et al.*, “Focus: Querying Large Video Datasets with Low Latency and Low Cost,” *Proceedings of the 13th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2018*, pp. 269–286, 2018.
- [32] I. Xarchakos and N. Koudas, “SVQ,” *Proceedings of the 2019 International Conference on Management of Data*, pp. 2013–2016, 2019, doi: 10.1145/3299869.3320230.
- [33] F. Bastani *et al.*, “MIRIS: Fast Object Track Queries in Video,” *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 1907–1921, 2020, doi: 10.1145/3318464.3389692.
- [34] T. Xu, L. M. Botelho, and F. X. Lin, “VStore: A Data Store for Analytics on Large Videos,” *Proceedings of the 14th EuroSys Conference 2019*, p. 17, 2018, doi: 10.1145/3302424.3303971.
- [35] M. Daum, B. Haynes, D. He, A. Mazumdar, and M. Balazinska, “TASM: A Tile-Based Storage Manager for Video Analytics,” *Proc Int Conf Data Eng*, vol. 2021-April, pp. 1775–1786, 2020, doi: 10.1109/ICDE51399.2021.00156.
- [36] X. Zeng, B. Fang, H. Shen, and M. Zhang, “Distream: Scaling live video analytics with workload-adaptive distributed edge intelligence,” *SenSys 2020 - Proceedings of the 2020 18th ACM Conference on Embedded Networked Sensor Systems*, pp. 409–421, 2020, doi: 10.1145/3384419.3430721.
- [37] Y. Guo *et al.*, *Live Video Analytics at Scale with Approximation and Delay-Tolerance*, vol. 40, no. 4. 2017.
- [38] C. C. Hung *et al.*, “VideoEdge: Processing camera streams using hierarchical clusters,” *Proceedings - 2018 3rd ACM/IEEE Symposium on Edge Computing, SEC 2018*, pp. 115–131, 2018, doi: 10.1109/SEC.2018.00016.

- [39] Y. Lu, A. Chowdhery, and S. Kandula, “Optasia: A relational platform for efficient large-scale video analytics,” *Proceedings of the 7th ACM Symposium on Cloud Computing, SoCC 2016*, pp. 57–70, 2016, doi: 10.1145/2987550.2987564.
- [40] “Scanner: Efficient Video Analysis at Scale | Request PDF.”
- [41] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2014.
- [42] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386.
- [43] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.
- [44] T. G. Armstrong, V. Ponnekanti, D. Borthakur, and M. Callaghan, “LinkBench: A database benchmark based on the facebook social graph,” *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 1185–1196, 2013, doi: 10.1145/2463676.2465296.
- [45] M. Hanyao, Y. Jin, Z. Qian, S. Zhang, and S. Lu, “Edge-assisted online on-device object detection for real-time video analytics,” *Proceedings - IEEE INFOCOM*, vol. 2021-May, 2021, doi: 10.1109/INFOCOM42981.2021.9488741.
- [46] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Trans Pattern Anal Mach Intell*, vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.

- [47] W. Liu *et al.*, “SSD: Single Shot MultiBox Detector,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9905 LNCS, pp. 21–37, 2015, doi: 10.1007/978-3-319-46448-0\_2.
- [48] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, “End-to-End Object Detection with Transformers,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12346 LNCS, pp. 213–229, 2020, doi: 10.1007/978-3-030-58452-8\_13.
- [49] L. Liu, H. Li, and M. Gruteser, “Edge Assisted Real-time Object Detection for Mobile Augmented Reality,” *The 25th Annual International Conference on Mobile Computing and Networking*, vol. 19, doi: 10.1145/3300061.
- [50] Z. Chen *et al.*, “An empirical study of latency in an emerging class of edge computing applications for wearable cognitive assistance,” *2017 2nd ACM/IEEE Symposium on Edge Computing, SEC 2017*, 2017, doi: 10.1145/3132211.3134458.
- [51] S. Yi, Z. Hao, Q. Zhang, Q. Zhang, W. Shi, and Q. Li, “LAVEA: Latency-Aware video analytics on edge computing platform,” *2017 2nd ACM/IEEE Symposium on Edge Computing, SEC 2017*, 2017, doi: 10.1145/3132211.3134459.
- [52] T. Huang, R. X. Zhang, and L. Sun, “Deep reinforced bitrate ladders for adaptive video streaming,” *NOSSDAV 2021 - Proceedings of the 2021 Workshop on Network and Operating System Support for Digital Audio and Video, Part of MMSys 2021*, pp. 67–73, 2021, doi: 10.1145/3458306.3458873.
- [53] “RTMP Streaming: What Is the Real-Time Messaging Protocol?”
- [54] “MPEG-DASH: Dynamic Adaptive Streaming Over HTTP | Wowza.”

- [55] S. Ritika and D. Rao, “Data Augmentation of Railway Images for Track Inspection,” 2018.
- [56] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” 2018.
- [57] “The many use-cases for emerging technologies all depend on viable edge computing capabilities.”
- [58] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 580–587, 2013, doi: 10.1109/CVPR.2014.81.
- [59] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, “Aggregated Residual Transformations for Deep Neural Networks,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 5987–5995, 2016, doi: 10.1109/CVPR.2017.634.
- [60] S.-H. Gao, M.-M. Cheng, K. Zhao, X.-Y. Zhang, M.-H. Yang, and P. Torr, “Res2Net: A New Multi-scale Backbone Architecture,” *IEEE Trans Pattern Anal Mach Intell*, vol. 43, no. 2, pp. 652–662, 2019, doi: 10.1109/TPAMI.2019.2938758.
- [61] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 6517–6525, 2016, doi: 10.1109/CVPR.2017.690.
- [62] Z. Tian, C. Shen, H. Chen, and T. He, “FCOS: Fully Convolutional One-Stage Object Detection,” *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2019-Octob, pp. 9626–9635, 2019, doi: 10.1109/ICCV.2019.00972.

- [63] D. Bhatt *et al.*, “CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope,” *Electronics 2021, Vol. 10, Page 2470*, vol. 10, no. 20, p. 2470, 2021, doi: 10.3390/ELECTRONICS10202470.
- [64] X. Zhou, J. Zhuo, and P. Krahenbuhl, “Bottom-up Object Detection by Grouping Extreme and Center Points,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-June, pp. 850–859, 2019, doi: 10.1109/CVPR.2019.00094.
- [65] “YOLOv8 - Ultralytics | Revolutionizing the World of Vision AI.”
- [66] “GitHub - openimages/dataset: The Open Images dataset.”
- [67] “Recognition of human actions.”
- [68] “Action Recognition in the Dark.”
- [69] R. Jayaswal and M. Dixit, “Human Behavior dataset (HBD21),” vol. 1, 2021, doi: 10.17632/XH9PGB3W8C.1.
- [70] “UCSD Anomaly Detection Dataset.”
- [71] “ARID Dataset - V7 Open Datasets.”
- [72] “GitHub - frankzhangrui/Darknet-Yolo: Object Detection End-to-End framework.”
- [73] “Deep Learning Guide: Choosing Your Data Annotation Tool.”
- [74] “GitHub - heartexlabs/labelImg: LabelImg is now part of the Label Studio community. The popular image annotation tool created by Tzutalin is no longer actively being developed, but you can check out Label Studio, the open source data labeling tool for imag.”
- [75] A. Dutta and A. Zisserman, “The VIA annotation software for images, audio and video,” *MM 2019 - Proceedings of the 27th ACM International*

*Conference on Multimedia*, pp. 2276–2279, 2019, doi: 10.1145/3343031.3350535.

- [76] “GitHub - microsoft/VoTT: Visual Object Tagging Tool: An electron app for building end to end Object Detection Models from Images and Videos.”
- [77] “Roboflow: Give your software the power to see objects in images and video.”
- [78] “GitHub - jsbroks/coco-annotator: :pencil2: Web-based image segmentation tool for object detection, localization, and keypoints.”
- [79] “A Gentle Introduction to Model Selection for Machine Learning - MachineLearningMastery.com.”
- [80] “RetinaNet Explained | Papers With Code.”
- [81] B. Wu, F. Iandola, P. H. Jin, and K. Keutzer, “SqueezeDet: Unified, Small, Low Power Fully Convolutional Neural Networks for Real-Time Object Detection for Autonomous Driving,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2017-July, pp. 446–454, 2016, doi: 10.1109/CVPRW.2017.60.
- [82] Y. Chen, C. Han, N. Wang, and Z. Zhang, “Revisiting Feature Alignment for One-stage Object Detection,” 2019.
- [83] “The proposed DetectNet neural network training model with a new layer... | Download Scientific Diagram.”
- [84] S. A. Sanchez, H. J. Romero, and A. D. Morales, “A review: Comparison of performance metrics of pretrained models for object detection using the TensorFlow framework,” *IOP Conf Ser Mater Sci Eng*, vol. 844, no. 1, 2020, doi: 10.1088/1757-899X/844/1/012024.

- [85] “Object Detection YOLO v1 , v2, v3 | by Venkata Krishna Jonnalagadda | Medium.”
- [86] “YOLO v4 or YOLO v5 or PP-YOLO? Which should I use? | Towards Data Science.”
- [87] “YOLOv8 for Object Detection Explained [Practical Example] | by Encord | Encord | Mar, 2023 | Medium.”
- [88] “Which metrics are used to evaluate a multiclass classification model’s performance?”
- [89] “Video, slides and scripts for DDS - Kuntai Du posts.”
- [90] “Object detection: speed and accuracy comparison (Faster R-CNN, R-FCN, SSD, FPN, RetinaNet and YOLOv3) | by Jonathan Hui | Medium.”
- [91] “Binary Classification – LearnDataSci.”
- [92] “sampaththennakoon/dds-v2: Server-driven Video Streaming for Deep Learning Inference.” <https://github.com/sampaththennakoon/dds-v2> (accessed Jul. 13, 2023).