REFERENCES

[1] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U2-net: Going deeper with nested u-structure for salient object detection," *Pattern recognition*, vol. 106, p. 107404, 2020.

[2] A. Van Den Oord, O. Vinyals *et al.*, "Neural discrete representation learning," *Advances in neural information processing systems*, vol. 30, 2017.

[3] A. Shahid and M. Mushtaq, "A survey comparing specialized hardware and evolution in tpus for neural networks," *2020 IEEE 23rd International Multitopic Conference (INMIC)*, 2020.

[4] M. Mohammadpour, Z. Heidari, M. Mirghorbani, and H. Hashemi, "Smartphones, tele-ophthalmology, and vision 2020," *International Journal of Ophthalmology*, 2017.

[5] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "Ghostnet: More features from cheap operations," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[6] T. Zhang, G.-J. Qi, B. Xiao, and J. Wang, "Interleaved group convolutions for deep neural networks," *International Conference on Computer Vision*, 2017.

[7] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

[8] D. Li, J. Hu, C. Wang, X. Li, Q. She, L. Zhu, T. Zhang, and Q. Chen, "Involution: Inverting the inherence of convolution for visual recognition," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[9] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv:1502.03167 [cs.LG]*, Mar. 2015.

[10] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," *arXiv:1607.08022 [cs.CV]*, Nov. 2017.

[11] B. S. Chhikara and K. Parang, "Global cancer statistics 2022: The trends projection analysis," *Chemical Biology Letters*, 2023.

[12] J. Lee, "Resection of diminutive and small colorectal polyps: What is the optimal technique?" *Clinical Endoscopy*, 2016.

[13] J. C. van Rijn, J. B. Reitsma, J. Stoker, P. M. Bossuyt, S. J. van Deventer, and E. Dekker, "Polyp miss rate determined by tandem colonoscopy: A systematic review," *The American Journal of Gastroenterology*, 2006.

[14] S. B. Ahn, D. S. Han, J. H. Bae, T. J. Byun, J. P. Kim, and C. S. Eun, "The miss rate for colorectal adenoma determined by quality-adjusted, back-to-back colonoscopies," *Gut and Liver*, 2012.

[15] N. H. Kim, Y. S. Jung, W. S. Jeong, H.-J. Yang, S.-K. Park, K. Choi, and D. I. Park, "Miss rate of colorectal neoplastic polyps and risk factors for missed polyps in consecutive colonoscopies," *Intestinal Research*, 2017.

[16] D. K. Rex, C. S. Cutler, G. T. Lemmel, E. Y. Rahmani, D. W. Clark, D. J. Helper, G. A. Lehman, and D. G. Mark, "Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies," *Gastroenterology*, 1997.

[17] J. L. Teh, J. R. Tan, L. Lau, N. Saxena, A. Salim, A. Tay, A. Shabbir, S. Chung, M. Hartman, and J. B.-Y. So, "Longer examination time improves detection of gastric cancer during diagnostic upper gastrointestinal endoscopy," *Clinical Gastroenterology and Hepatology*, 2015.

[18] P. Wang, P. Liu, J. R. G. Brown, T. M. Berzin, G. Zhou, S. Lei, X. Liu, L. Li, and X. Xiao, "Lower adenoma miss rate of computer-aided detection-assisted colonoscopy vs routine white-light colonoscopy in a prospective tandem study," *Gastroenterology*, 2020.

[19] Y. Mori, S. ei Kudo, T. M. Berzin, M. Misawa, and K. Takeda, "Computer-aided diagnosis for colonoscopy," *Endoscopy*, 2017.

[20] S. Y. Quan, M. T. Wei, J. Lee, R. Mohi-Ud-Din, R. Mostaghim, R. Sachdev, D. Siegel, Y. Friedlander, and S. Friedland, "Clinical evaluation of a real-time artificial intelligence-based polyp detection system: a us multi-center pilot study," *Scientific Reports*, 2022.

[21] S. Ali, D. Jha, N. Ghatwary, S. Realdon, R. Cannizzaro, O. E. Salem, D. Lamarque, C. Daul, M. A. Riegler, K. V. Anonsen, A. Petlund, P. Halvorsen, J. Rittscher, T. de Lange, and J. E. East, "A multi-centre polyp detection and segmentation dataset for generalisability assessment," *Scientific Data*, 2023.

[22] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, "Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians," *Computerized Medical Imaging and Graphics*, 2015.

[23] J. Bernal, J. Sánchez, and F. Vilariño, "Towards automatic polyp detection with a polyp appearance model," *Pattern Recognition*, 2012.

[24] S. A. Hicks, D. Jha, V. Thambawita, P. Halvorsen, H. L. Hammer, and M. A. Riegler, "The endotect 2020 challenge: Evaluation and comparison of classification, segmentation and inference time for endoscopy," *ICPR International Workshops and Challenges*, 2021.

[25] J. Silva, A. Histace, O. Romain, X. Dray, and B. Granado, "Toward embedded detection of polyps in wce images for early diagnosis of colorectal cancer," *International Journal of Computer Assisted Radiology and Surgery*, 2013.

[26] D. Jha, N. K. Tomar, S. Ali, M. A. Riegler, H. D. Johansen, D. Johansen, T. de Lange, and P. Halvorsen, "Nanonet: Real-time polyp segmentation in video capsule endoscopy and colonoscopy," in *Proceedings of the 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, 2021, pp. 37–43.

[27] D. Jha, P. H. Smedsrud, M. A. Riegler, P. Halvorsen, T. de Lange, D. Johansen, and H. D. Johansen, "Kvasir-seg: A segmented polyp dataset," *MultiMedia Modeling*, 2019.

[28] L. F. Sánchez-Peralta, L. Bote-Curiel, A. Picón, F. M. Sánchez-Margallo, and J. B. Pagador, "Deep learning to find colorectal polyps in colonoscopy: A systematic literature review," *Artificial Intelligence in Medicine*, 2020.

[29] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lecture Notes in Computer Science*, 2015.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[31] D.-P. Fan, G.-P. Ji, T. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "Pranet: Parallel reverse attention network for polyp segmentation," *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, 2020.

[32] N. Tajbakhsh, S. R. Gurudu, and J. Liang, "Automated polyp detection in colonoscopy videos using shape and context information," *IEEE Transactions on Medical Imaging*, 2016.

[33] D. Jha, P. H. Smedsrud, M. A. Riegler, D. Johansen, T. de Lange, P. Halvorsen, and H. D. Johansen, "Resunet++: An advanced architecture for medical image segmentation," *2019 IEEE International Symposium on Multimedia (ISM)*, 2019.

[34] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual u-net," *IEEE Geoscience and Remote Sensing Letters*, 2018.

[35] D. Jha, S. Ali, N. K. Tomar, H. D. Johansen, D. Johansen, J. Rittscher, M. A. Riegler, and P. Halvorsen, "Real-time polyp detection, localization and segmentation in colonoscopy using deep learning," *IEEE Access*, 2021.

[36] R. Karmakar and S. Nooshabadi, "Mobile-polypnet: Lightweight colon polyp segmentation network for low-resource settings," *Journal of Imaging*, 2022.

[37] G. S., G. C., S. H. C., A. Sreenivas, A. M. John, and A. A.S., "Lite-deep : Improved auto encoder-decoder for polyp segmentation," *2021 8th International Conference on Smart Computing and Communications (ICSCC)*, 2021.

[38] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

[39] C. Yang and J. Yang, "Lightweight real-time polyp segmentation network based on factorized convolution unit," *2022 2nd International Symposium on Artificial Intelligence and its Application on Media (ISAIAM)*, 2022.

[40] K. Yamaguchi, K. Sakamoto, T. Akabane, and Y. Fujimoto, "A neural network for speaker-independent isolated word recognition," *First International Conference on Spoken Language Processing (ICSLP 1990)*, 1990.

[41] J. J. Weng, N. Ahuja, and T. S. Huang, "Hierarchical models of object recognition in cortex," *1993 (4th) International Conference on Computer Vision*, 1993.

[42] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nature Neuroscience*, 1999.

[43] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "Eca-net: Efficient channel attention for deep convolutional neural networks," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[44] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[45] M. D. Zeiler, G. W. Taylor, and R. Fergus, "Adaptive deconvolutional networks for mid and high level feature learning," *2011 International Conference on Computer Vision*, 2011.

[46] Raspberry Pi Foundation, "Raspberry pi 4 model b specifications," accessed on Nov. 27, 2023. [Online]. Available: https://www.raspberrypi.com/products/raspberry-pi-4-model-b/specifications/

[47] Nvidia Corporation, "Jetson nano developer kit," accessed on Nov. 27, 2023. [Online]. Available: https://developer.nvidia.com/embedded/jetson-nano-developer-kit

[48] Google LLC, "Dev board," accessed on Nov. 27, 2023. [Online]. Available: https://coral.ai/products/dev-board/

[49] Z. DeVito, "Torchscript: Optimized execution of pytorch programs," *Program Transformations for Machine Learning – Workshop at NeurIPS 2019*, Dec. 14, 2019.

[50] Nvidia Corporation, "Nvidia deep learning tensorrt documentation," accessed on Nov. 27, 2023. [Online]. Available: https://docs.nvidia.com/deeplearning/tensorrt/

[51] The Linux Foundation, "Open neural network exchange," accessed on Nov. 27, 2023. [Online]. Available: https://onnx.ai/

[52] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[53] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, 1945.

[54] T. Sørensen, "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons," *Kongelige Danske Videnskabernes Selskab*, 1948.

[55] G. K. Gilbert, "Finley's tornado predictions," *American Meteorological Journal*, 1884.

[56] P. Jaccard, "The distribution of the flora in the alpine zone," *New Phytologist*, 1912.

[57] T. T. Tanimoto, "An elementary mathematical theory of classification and prediction," *Internal IBM Technical Report*, 1958.

[58] J. Yerushalmy, "Statistical problems in assessing methods of medical diagnosis, with special reference to x-ray techniques," *Public Health Reports*, 1947.

[59] D. G. Altman and J. M. Bland, "Statistics notes: Diagnostic tests 2: predictive values," *BMJ*, 1994.

[60] R. Margolin, L. Zelnik-Manor, and A. Tal, "How to evaluate foreground maps?" *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[61] D.-P. Fan, M.-M. Cheng, Y. Liu, T. Li, and A. Borji, "Structure-measure: A new way to evaluate foreground maps," *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017.

[62] D.-P. Fan, C. Gong, Y. Cao, B. Ren, M.-M. Cheng, and A. Borji, "Enhanced-alignment measure for binary foreground map evaluation," *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 2018.

[63] S. S. Shapiro and M. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, 1965.

[64] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, 1945.

[65] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The Annals of Mathematical Statistics*, 1947.

[66] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain." *Psychological Review*, 1958.

[67] J. Watt, R. Borhani, and A. K. Katsaggelos, *Machine Learning Refined: Foundations, Algorithms, and Applications 2nd Edition.* Cambridge University Press, 2020.

[68] J. Schmidhuber, "Annotated history of modern ai and deep learning," *arXiv:2212.11279 [cs.NE]*, Dec. 2022.

[69] B. Ramsundar and R. B. Zadeh, *TensorFlow for deep learning: from linear regression to reinforcement learning.* O'Reilly, 2018.

[70] P. W. Battaglia, J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, C. Gulcehre, F. Song, A. Ballard, J. Gilmer, G. Dahl, A. Vaswani, K. Allen, C. Nash, V. Langston, C. Dyer, N. Heess, D. Wierstra, P. Kohli, M. Botvinick, O. Vinyals, Y. Li, and R. Pascanu, "Relational inductive biases, deep learning, and graph networks," *arXiv:1806.01261 [cs.LG]*, Oct. 2018.

[71] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning.* MIT Press, 2018.

[72] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, 1991.

[73] K. Fukushima, "Visual feature extraction by a multilayered network of analog threshold elements," *IEEE Transactions on Systems Science and Cybernetics*, 1969.

[74] C. von der Malsburg, "Self-organization of orientation sensitive cells in the striate cortex," *Kybernetik*, 1973.

[75] F. Rosenblatt, "Principles of neurodynamics. perceptrons and the theory of brain mechanisms," Cornell Aeronautical Laboratory, Inc., Tech. Rep., 1961.

[76] M. Minsky and S. Papert, *Perceptrons: An introduction to computational geometry*. MIT Press, 1969.

[77] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*. MIT Press, 1987, ch. 8, pp. 318–362.

[78] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems*, 1989.

[79] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, 1991.

[80] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv:1312.6114 [stat.ML]*, Dec. 2022.

[81] M. Connor, G. Canal, and C. Rozell, "Variational autoencoder with learned latent structure," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2021, pp. 2359–2367.

[82] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

[83] L. Cayton, *Algorithms for manifold learning*. eScholarship, University of California, 2008.

[84] C. Fefferman, S. Mitter, and H. Narayanan, "Testing the manifold hypothesis," *Journal of the American Mathematical Society*, 2016.

[85] S. Linnainmaa, *The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors*. Master's Thesis, University of Helsinki, 1970.

[86] P. J. Werbos, "Applications of advances in nonlinear sensitivity analysis," in *System Modeling and Optimization: Proceedings of the 10th IFIP Conference New York City, USA, August 31–September 4, 1981*. Springer, 2005, pp. 762–770.

[87] G. W. Leibniz and J. M. Child, *The early mathematical manuscripts of Leibniz: translated from the Latin texts published by Carl Immanuel Gerhardt with critical and historical notes*. Merchant Books, 2007.

[88] C. Lemaréchal, "Cauchy and the gradient method," *Doc Math Extra*, vol. 251, no. 254, p. 10, 2012.

[89] H. Robbins and S. Monro, "A stochastic approximation method," *The Annals of Mathematical Statistics*, 1951.

[90] D. C. Liu and J. Nocedal, "On the limited memory bfgs method for large scale optimization," *Mathematical Programming*, 1989.

[91] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015.

[92] I. Sutskever, J. Martens, G. Dahl, and G. E. Hinton, "On the importance of initialization and momentum in deep learning," in *International conference on machine learning*. PMLR, 2013, pp. 1139–1147.

[93] G. E. Hinton, N. Srivastava, and K. Swersky, "Lecture 6a overview of mini-batch gradient descent," *Coursera Lecture Slides*, 2012.

[94] A. Graves, "Generating sequences with recurrent neural networks," *arXiv:1308.0850 [cs.NE]*, Jun. 2014.

[95] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, *Dive into Deep Learning*. Cambridge University Press, 2023.

[96] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[97] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv:1312.4400 [cs.NE]*, Mar. 2014.

[98] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[99] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[100] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[101] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10 012–10 022.

[102] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," *arXiv:1607.06450 [stat.ML]*, Jul. 2016.

[103] Y. Wu and K. He, "Group normalization," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 3–19.

[104] S. Santurkar, D. Tsipras, A. Ilyas, and A. Madry, "How does batch normalization help optimization?" *Advances in neural information processing systems*, vol. 31, 2018.

[105] R. Balestriero and R. G. Baraniuk, "Batch normalization explained," *arXiv:2209.14778 [cs.LG]*, Sep. 2022.

[106] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Improved texture networks: Maximizing quality and diversity in feed-forward stylization and texture synthesis," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6924–6932.

[107] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[108] G. Ghiasi, H. Lee, M. Kudlur, V. Dumoulin, and J. Shlens, "Exploring the structure of a real-time, arbitrary neural artistic stylization network," *arXiv:1705.06830 [cs.CV]*, Aug. 2017.

[109] J. An, H. Xiong, J. Ma, J. Luo, and J. Huan, "Stylenas: An empirical study of neural architecture search to uncover surprisingly fast end-to-end universal style transfer networks," *arXiv:1906.02470 [cs.CV]*, Jun. 2019.

[110] X. Pan, P. Luo, J. Shi, and X. Tang, "Two at once: Enhancing learning and generalization capacities via ibn-net," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 464–479.

[111] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*. Springer, 2016, pp. 630–645.

[112] S. Zagoruyko and N. Komodakis, "Wide residual networks," *arXiv:1605.07146 [cs.CV]*, Jun. 2017.

[113] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[114] P. Chao, C.-Y. Kao, Y.-S. Ruan, C.-H. Huang, and Y.-L. Lin, "Hardnet: A low memory traffic network," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 3552–3561.

[115] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," *arXiv:1505.00387 [cs.LG]*, Nov. 2015.

[116] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[117] S. Hochreiter, *Untersuchungen zu dynamischen neuronalen Netzen*. Diploma thesis, Institut für Informatik, Lehrstuhl Prof. Brauer, Technische Universität München, 1991.

[118] H. Li, Z. Xu, G. Taylor, C. Studer, and T. Goldstein, "Visualizing the loss landscape of neural nets," *Advances in neural information processing systems*, vol. 31, 2018.

[119] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, "Repvgg: Making vgg-style convnets great again," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 13 733–13 742.

[120] O. Weng, G. Marcano, V. Loncar, A. Khodamoradi, N. Sheybani, A. Meza, F. Koushanfar, K. Denolf, J. M. Duarte, and R. Kastner, "Tailor: Altering skip connections for resource-efficient inference," *ACM Transactions on Reconfigurable Technology and Systems*, 2023.

[121] G. Li, J. Zhang, Y. Wang, C. Liu, M. Tan, Y. Lin, W. Zhang, J. Feng, and T. Zhang, "Residual distillation: Towards portable deep neural networks without shortcuts," *Advances in Neural Information Processing Systems*, vol. 33, pp. 8935–8946, 2020.

[122] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan *et al.*, "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1314–1324.

[123] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 116–131.

[124] K. Sun, M. Li, D. Liu, and J. Wang, "Igcv3: Interleaved low-rank group convolutions for efficient deep neural networks," *arXiv:1806.00178 [cs.CV]*, Jul. 2018.

[125] Y. Tang, K. Han, J. Guo, C. Xu, C. Xu, and Y. Wang, "Ghostnetv2: enhance cheap operation with long-range attention," *Advances in Neural Information Processing Systems*, vol. 35, pp. 9969–9982, 2022.

[126] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv:1704.04861 [cs.CV]*, Apr. 2017.

[127] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "Mnasnet: Platform-aware neural architecture search for mobile," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 2820–2828.

[128] T.-J. Yang, A. Howard, B. Chen, X. Zhang, A. Go, M. Sandler, V. Sze, and H. Adam, "Netadapt: Platform-aware neural network adaptation for mobile applications," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 285–300.

[129] G. Xie, J. Wang, T. Zhang, J. Lai, R. Hong, and G.-J. Qi, "Interleaved structured sparse convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8847–8856.

[130] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 33, no. 2, pp. 353–367, 2010.

[131] A. Borji, "What is a salient object? a dataset and a baseline model for salient object detection," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 742–756, 2014.

[132] L. Wang, H. Lu, Y. Wang, M. Feng, D. Wang, B. Yin, and X. Ruan, "Learning to detect salient objects with image-level supervision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 136–145.

[133] D.-P. Fan, G.-P. Ji, G. Sun, M.-M. Cheng, J. Shen, and L. Shao, "Camouflaged object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2777–2787.

[134] S. Krishnan, X. Yang, K. Chan, S. Kumar, and P. Goh, "Intestinal abnormality detection from endoscopic images," in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286)*, vol. 2. IEEE, 1998, pp. 895–898.

[135] S. Hwang, J. Oh, W. Tavanapong, J. Wong, and P. C. De Groen, "Polyp detection in colonoscopy video using elliptical shape feature," in *2007 IEEE International Conference on Image Processing*, vol. 2. IEEE, 2007, pp. II–465.

[136] C. Van Wijk, V. F. Van Ravesteijn, F. M. Vos, and L. J. Van Vliet, "Detection and segmentation of colonic polyps on implicit isosurfaces by second principal curvature flow," *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 688–698, 2010.

[137] H. Zhu, Y. Fan, and Z. Liang, "Improved curvature estimation for shape analysis in computer-aided detection of colonic polyps," *Beijing, China*, p. 19, 2010.

[138] Y. Iwahori, H. Hagi, H. Usami, R. J. Woodham, A. Wang, M. K. Bhuyan, and K. Kasugai, "Automatic polyp detection from endoscope image using likelihood map based on edge information," in *International Conference on Pattern Recognition Applications and Methods*, vol. 2. SciTePress, 2017, pp. 402–409.

[139] S. A. Karkanis, D. K. Iakovidis, D. E. Maroulis, D. A. Karras, and M. Tzivras, "Computer-aided tumor detection in endoscopic video using color wavelet features," *IEEE Transactions on Information Technology in Biomedicine*, 2003.

[140] M. P. Tjoa and S. M. Krishnan, "Feature extraction for the analysis of colon status from the endoscopic images," *BioMedical Engineering OnLine*, 2003.

[141] P. Li, K. L. Chan, and S. M. Krishnan, "Learning a multi-size patch-based hybrid kernel machine ensemble for abnormal region detection in colonoscopic images," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. IEEE, 2005, pp. 670–675.

[142] S. Ameling, S. Wirth, D. Paulus, G. Lacey, and F. Vilariño, "Texture-based polyp detection in colonoscopy," *Bildverarbeitung für die Medizin 2009*, 2009.

[143] M. T. Coimbra and J. P. S. Cunha, "Mpeg-7 visual descriptors—contributions for automated feature extraction in capsule endoscopy," *IEEE Transactions on Circuits and Systems for Video Technology*, 2006.

[144] J. J. C. Fu, Y.-W. Yu, H.-M. Lin, J.-W. Chai, and C. C.-C. Chen, "Feature extraction and pattern classification of colorectal polyps in colonoscopic imaging," *Computerized Medical Imaging and Graphics*, 2014.

[145] A. Lou, S. Guan, H. Ko, and M. H. Loew, "Caranet: Context axial reverse attention network for segmentation of small medical objects," in *Medical Imaging 2022: Image Processing*, vol. 12032. SPIE, 2022, pp. 81–92.

[146] C.-H. Huang, H.-Y. Wu, and Y.-L. Lin, "Hardnet-mseg: A simple encoder-decoder polyp segmentation neural network that achieves over 0.9 mean dice and 86 fps," *arXiv:2101.07172 [cs.CV]*, Jan. 2021.

[147] Q.-H. Trinh, "Meta-polyp: a baseline for efficient polyp segmentation," *arXiv:2305.07848 [eess.IV]*, Jul. 2023.

[148] R.-G. Dumitru, D. Peteleaza, and C. Craciun, "Using duck-net for polyp image segmentation," *Scientific Reports*, vol. 13, no. 1, p. 9803, 2023.

[149] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *arXiv:1606.00915 [cs.CV]*, May 2017.

[150] H. Li, P. Xiong, J. An, and L. Wang, "Pyramid attention network for semantic segmentation," *arXiv:1805.10180 [cs.CV]*, Nov. 2018.

[151] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 3–19.

[152] M. Courbariaux, Y. Bengio, and J.-P. David, "Binaryconnect: Training deep neural networks with binary weights during propagations," *Advances in neural information processing systems*, vol. 28, 2015.

[153] D. Han, J. Kim, and J. Kim, "Deep pyramidal residual networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5927–5935.

[154] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," *arXiv:1710.05941 [cs.NE]*, Oct. 2017.

[155] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T.-S. Chua, "Scacnn: Spatial and channel-wise attention in convolutional networks for image captioning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5659–5667.

[156] A. G. Roy, N. Navab, and C. Wachinger, "Recalibrating fully convolutional networks with spatial and channel "squeeze and excitation" blocks," *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 540–549, 2018.

[157] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[158] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size," *arXiv:1602.07360 [cs.CV]*, Nov. 2016.

[159] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.

[160] H. Cai, L. Zhu, and S. Han, "Proxylessnas: Direct neural architecture search on target task and hardware," *arXiv:1812.00332 [cs.LG]*, Feb. 2019.

[161] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[162] C. Guo, M. Szemenyei, Y. Yi, W. Wang, B. Chen, and C. Fan, "Sa-unet: Spatial attention u-net for retinal vessel segmentation," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 1236–1242.

[163] R. D. H. Al-Fahsi, A. Aqthobirrobbany, I. Ardiyanto, and H. A. Nugroho, "Msgnet: Modified mobilenet-shufflenet-ghostnet network for lightweight retinal vessel segmentation," in *2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*. IEEE, 2023, pp. 94–99.

[164] Y. Guo, H. He, Y. Zhu, and Y. Yu, "Domain generalization person re-identification on attention-aware multi-operation strategery," *arXiv:2210.10409 [cs.CV]*, Oct. 2022.

[165] F. Zhu, Y. Tian, C. Han, Y. Li, J. Nan, N. Yao, and W. Zhou, "Mlabin: Model-level attention and batch-instance style normalization for domain generalization of federated learning on medical image segmentation," *arXiv:2306.17008 [eess.IV]*, Jun. 2023.

[166] X.-Y. Zhou, J.-Q. Zheng, P. Li, and G.-Z. Yang, "Acnn: a full resolution dcnn for medical image segmentation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 8455–8461.

[167] D. Vázquez, J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, A. M. López, A. Romero, M. Drozdzal, A. Courville *et al.*, "A benchmark for endoluminal scene segmentation of colonoscopy images," *Journal of healthcare engineering*, vol. 2017, 2017.

[168] P. H. Smedsrud, V. Thambawita, S. A. Hicks, H. Gjestang, O. O. Nedrejord, E. Næss, H. Borgli, D. Jha, T. J. D. Berstad, S. L. Eskeland, M. Lux, H. Espeland, A. Petlund, D. T. D. Nguyen, E. Garcia-Ceja, D. Johansen, P. T. Schmidt, E. Toth, H. L. Hammer, T. de Lange, M. A. Riegler, and P. Halvorsen, "Kvasir-Capsule, a video capsule endoscopy dataset," *Scientific Data*, vol. 8, no. 1, p. 142, 2021.

[169] V. Sovrasov, "ptflops: a flops counting tool for neural networks in pytorch framework," 2018-2023. [Online]. Available: https://github.com/sovrasov/flops-counter.pytorch

[170] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.

[171] W. Falcon and The PyTorch Lightning team, "PyTorch Lightning," Mar. 2019. [Online]. Available: https://github.com/Lightning-AI/lightning

[172] N. S. Detlefsen, J. Borovec, J. Schock, A. Harsh, T. Koker, L. D. Liello, D. Stancl, C. Quan, M. Grechkin, and W. Falcon, "TorchMetrics - Measuring Reproducibility in PyTorch," Feb. 2022. [Online]. Available: https://github.com/Lightning-AI/torchmetrics

[173] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, "Scipy 1.0: fundamental algorithms for scientific computing in python," *Nature methods*, vol. 17, no. 3, pp. 261–272, 2020.

[174] K. Wickstrøm, M. Kampffmeyer, and R. Jenssen, "Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps," *Medical image analysis*, vol. 60, p. 101619, 2020.

[175] Z. Zhang, Y. Xie, F. Xing, M. McGough, and L. Yang, "Mdnet: A semantically and visually interpretable medical image diagnosis network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6428–6436.

[176] N. K. Tomar, D. Jha, U. Bagci, and S. Ali, "Tganet: Text-guided attention for improved polyp segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 151–160.