

Bibliography using BibT_EX

- [1] Rameen Abdal, Yipeng Qin, and Peter Wonka. 2019. Image2StyleGAN: How to embed images into the StyleGAN latent space?. In *ICCV*.
- [2] Rameen Abdal, Yipeng Qin, and Peter Wonka. 2020. Image2StyleGAN++: How to Edit the Embedded Images?. In *CVPR*.
- [3] Aviad Aberdam, Dror Simon, and Michael Elad. 2020. When and How Can Deep Generative Models be Inverted? *arXiv preprint arXiv:2006.15555* (2020).
- [4] Kfir Aberman, Mingyi Shi, Jing Liao, Dani Lischinski, Baoquan Chen, and Daniel Cohen-Or. 2019. Deep Video-Based Performance Cloning. *Computer Graphics Forum* (2019).
- [5] Eirikur Agustsson, Michael Tschannen, Fabian Mentzer, Radu Timofte, and Luc Van Gool. 2019. Generative Adversarial Networks for Extreme Learned Image Compression. In *ICCV*.
- [6] Michael Albright and Scott McCloskey. 2019. Source Generator Attribution via Inversion. In *CVPR Workshops*.
- [7] Kara-Ali Aliiev, Dmitry Ulyanov, and Victor Lempitsky. 2019. Neural point-based graphics. *arXiv preprint arXiv:1906.08240* (2019).
- [8] Amjad Almahairi, Sai Rajeshwar, Alessandro Sordani, Philip Bachman, and Aaron Courville. 2018. Augmented CycleGAN: Learning Many-to-Many Mappings from Unpaired Data. In *ICML*.
- [9] Martin Arjovsky, Soumith Chintala, and Léon Bottou. 2017. Wasserstein GAN. In *ICML*.
- [10] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. 2017. Deep Reinforcement Learning: A Brief Survey. *IEEE Signal Processing Magazine* (2017).
- [11] Georgios Arvanitidis, Lars Kai Hansen, and Søren Hauberg. 2018. Latent Space Oddity: on the Curvature of Deep Generative Models. In *ICLR*.
- [12] Hadar Averbuch-Elor, Daniel Cohen-Or, Johannes Kopf, and Michael F Cohen. 2017. Bringing portraits to life. *TOG* (2017).
- [13] Shai Avidan and Ariel Shamir. 2007. Seam carving for content-aware image resizing. *TOG* (2007).
- [14] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. 2016. Layer normalization. *arXiv preprint arXiv:1607.06450* (2016).
- [15] Kyungjune Baek, Yunje Choi, Youngjung Uh, Jaejun Yoo, and Hyunjeong Shim. 2020. Rethinking the Truly Unsupervised Image-to-Image Translation. *arXiv preprint arXiv:2006.06500* (2020).
- [16] Guha Balakrishnan, Amy Zhao, Adrian V Dalca, Fredo Durand, and John Guttag. 2018. Synthesizing Images of Humans in Unseen Poses. In *CVPR*.
- [17] Mukund Balasubramanian, Eric L Schwartz, Joshua B Tenenbaum, Vin de Silva, and John C Langford. 2002. The isomap algorithm and topological stability. *Science* 295, 5552 (2002), 7–7.
- [18] Shumeet Baluja and Ian Fischer. 2017. Adversarial transformation networks: Learning to generate adversarial examples. *arXiv preprint arXiv:1703.09387* (2017).
- [19] Aayush Bansal, Shugao Ma, Deva Ramanan, and Yaser Sheikh. 2018. RecycleGAN: Unsupervised video retargeting. In *ECCV*.
- [20] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman. 2009. PatchMatch: A randomized correspondence algorithm for structural image editing. *TOG* (2009).
- [21] David Bau, Steven Liu, Tongzhou Wang, Jun-Yan Zhu, and Antonio Torralba. 2020. Rewriting a Deep Generative Model. In *ECCV*.
- [22] David Bau, Hendrik Strobelt, William Peebles, Jonas Wulff, Bolei Zhou, Jun-Yan Zhu, and Antonio Torralba. 2019. Semantic Photo Manipulation with a Generative Image Prior. *TOG* 38, 4 (2019).
- [23] David Bau, Hendrik Strobelt, William Peebles, Jonas Wulff, Bolei Zhou, Jun-Yan Zhu, and Antonio Torralba. 2019. Semantic photo manipulation with a generative image prior. *TOG* 38, 4 (2019), 59.
- [24] David Bau, Jun-Yan Zhu, Hendrik Strobelt, Agata Lapedriza, Bolei Zhou, and Antonio Torralba. 2020. Understanding the role of individual units in a deep neural network. *National Academy of Sciences* (2020).
- [25] David Bau, Jun-Yan Zhu, Hendrik Strobelt, Bolei Zhou, Joshua B. Tenenbaum, William T. Freeman, and Antonio Torralba. 2019. GAN Dissection: Visualizing and Understanding Generative Adversarial Networks. In *ICLR*.
- [26] David Bau, Jun-Yan Zhu, Jonas Wulff, William Peebles, Hendrik Strobelt, Bolei Zhou, and Antonio Torralba. 2019. Inverting layers of a large generator. In *ICLR Workshop*, Vol. 2. 4.
- [27] David Bau, Jun-Yan Zhu, Jonas Wulff, William Peebles, Hendrik Strobelt, Bolei Zhou, and Antonio Torralba. 2019. Seeing what a gan cannot generate. In *ICCV*. 4502–4511.
- [28] Mohamed Ishmael Belghazi, Sai Rajeswar, Olivier Mastropietro, Negar Rostamzadeh, Jovana Mitrovic, and Aaron Courville. 2018. Hierarchical adversarially learned inference. *arXiv preprint arXiv:1802.01071* (2018).
- [29] Mikhail Belkin and Partha Niyogi. 2003. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation* 15, 6 (2003), 1373–1396.
- [30] Sagie Benaim, Ron Mokady, Amit Bermano, Daniel Cohen-Or, and Lior Wolf. 2020. Structural-analogy from a Single Image Pair. *arXiv preprint arXiv:2004.02222* (2020).
- [31] Sagie Benaim and Lior Wolf. 2017. One-sided unsupervised domain mapping. In *NeurIPS*.
- [32] Sagie Benaim and Lior Wolf. 2018. One-Shot Unsupervised Cross Domain Translation. In *NeurIPS*.
- [33] Yoshua Bengio, Aaron Courville, and Pascal Vincent. 2013. Representation learning: A review and new perspectives. *TPAMI* (2013).
- [34] Bharat Lal Bhatnagar, Cristian Sminchisescu, Christian Theobalt, and Gerard Pons-Moll. 2020. Combining implicit function learning and parametric models for 3d human reconstruction. *arXiv preprint arXiv:2007.11432* (2020).
- [35] Siavash Arjomand Bigdeli, Matthias Zwicker, Paolo Favaro, and Meiguang Jin. 2017. Deep mean-shift priors for image restoration. In *NeurIPS*. 763–772.
- [36] Dmitri Bitouk, Neeraj Kumar, Samreen Dhillon, Peter Belhumeur, and Shree K Nayar. 2008. Face swapping: automatically replacing faces in photographs. In *SIGGRAPH*.
- [37] Volker Blanz and Thomas Vetter. 1999. A morphable model for the synthesis of 3D faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*.
- [38] Yochai Blau, Roey Mechrez, Radu Timofte, Tomer Michaeli, and Lihi Zelnik-Manor. 2018. The 2018 PIRM challenge on perceptual image super-resolution. In *ECCV Workshop*.
- [39] Yochai Blau and Tomer Michaeli. 2018. The perception-distortion tradeoff. In *CVPR*.
- [40] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *JMLR* 3, Jan (2003), 993–1022.
- [41] Piotr Bojanowski, Armand Joulin, David Lopez-Paz, and Arthur Szlam. 2017. Optimizing the latent space of generative networks. *arXiv preprint arXiv:1707.05776* (2017).
- [42] Nicolas Bonneel, James Tompkin, Kalyan Sunkavalli, Deqing Sun, Sylvain Paris, and Hanspeter Pfister. 2015. Blind video temporal consistency. *TOG* (2015).
- [43] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G. Dimakis. 2017. Compressed Sensing using Generative Models. In *ICML*, Vol. 70. 537–546.
- [44] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G Dimakis. 2017. Compressed sensing using generative models. *arXiv preprint arXiv:1703.03208* (2017).
- [45] Konstantinos Bousmalis, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan. 2017. Unsupervised pixel-level domain adaptation with generative adversarial networks. In *CVPR*.
- [46] George EP Box, Gwilym M Jenkins, and Gregory C Reinsel. 2011. *Time series analysis: forecasting and control*. Vol. 734. John Wiley & Sons.
- [47] Andrew Brock, Jeff Donahue, and Karen Simonyan. 2019. Large scale GAN training for high fidelity natural image synthesis. In *ICLR*.
- [48] Gabriel J. Brostow, Jamie Shotton, Julien Fauqueur, and Roberto Cipolla. 2008. Segmentation and Recognition Using Structure from Motion Point Clouds. In *ECCV*.
- [49] Matthew Brown, David G Lowe, et al. 2003. Recognising panoramas.. In *ICCV*.
- [50] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language Models are Few-Shot Learners. *arXiv preprint arXiv:2005.14165* (2020).
- [51] Christian Buchta, Martin Kober, Ingo Feinerer, and Kurt Hornik. 2012. Spherical k-means clustering. *Journal of Statistical Software* 50, 10 (2012), 1–22.
- [52] Peter Burt and Edward Adelson. 1983. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications* (1983).
- [53] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyu Zhu. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on scientific computing* 16, 5 (1995), 1190–1208.
- [54] Emmanuel J Candès et al. 2006. Compressive Sampling. In *ICM*, Vol. 3. 1433–1452.
- [55] Emmanuel J Candès, Justin Romberg, and Terence Tao. 2006. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory* 52, 2 (2006), 489–509.
- [56] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2018. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *arXiv preprint arXiv:1812.08008* (2018).

- [57] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. In *CVPR*.
- [58] Joao Carreira and Andrew Zisserman. 2017. Quo vadis, action recognition? a new model and the kinetics dataset. In *CVPR*.
- [59] Caroline Chan, Shiry Ginosar, Tinghui Zhou, and Alexei A Efros. 2019. Everybody dance now. In *ICCV*.
- [60] Olga Chelnokova, Bruno Laeng, Marie Eikemo, Jeppe Riegels, Guro Løseth, Hedda Maurud, Frode Willoch, and Siri Leknes. 2014. Rewards of beauty: the opioid system mediates social motivation in humans. *Molecular psychiatry* 19, 7 (2014), 746–747.
- [61] Anpei Chen, Ruiyang Liu, Ling Xie, and Jingyi Yu. 2020. A Free Viewpoint Portrait Generator with Dynamic Styling. *arXiv preprint arXiv:2007.03780* (2020).
- [62] Bor-Chun Chen, Chu-Song Chen, and Winston H Hsu. 2014. Cross-age reference coding for age-invariant face recognition and retrieval. In *ECCV*. 768–783.
- [63] Dongdong Chen, Jing Liao, Lu Yuan, Nenghai Yu, and Gang Hua. 2017. Coherent online video style transfer. In *ICCV*.
- [64] Jingwen Chen, Jiawei Chen, Hongyang Chao, and Ming Yang. 2018. Image blind denoising with generative adversarial network based noise modeling. In *CVPR*.
- [65] Lele Chen, Zhiheng Li, Ross K Maddox, Zhiyao Duan, and Chenliang Xu. 2018. Lip movements generation at a glance. In *ECCV*.
- [66] Lele Chen, Ross K Maddox, Zhiyao Duan, and Chenliang Xu. 2019. Hierarchical cross-modal talking face generation with dynamic pixel-wise loss. In *CVPR*.
- [67] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. 2017. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* (2017).
- [68] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *ECCV*.
- [69] Mark Chen, Alec Radford, Rewon Child, Jeff Wu, Heewoo Jun, Prafulla Dhariwal, David Luan, and Ilya Sutskever. 2020. Generative Pretraining from Pixels. In *ICML*.
- [70] Nutan Chen, Alexej Klushyn, Richard Kurl, Xueyan Jiang, Justin Bayer, and Patrick van der Smagt. 2018. Metrics for Deep Generative Models. In *AISTATS*, Vol. 84. 1540–1550.
- [71] Qifeng Chen and Vladlen Koltun. 2017. Photographic image synthesis with cascaded refinement networks. In *ICCV*.
- [72] Shu-Yu Chen, Wanchao Su, Lin Gao, Shihong Xia, and Hongbo Fu. 2020. Deep-FaceDrawing: Deep Generation of Face Images from Sketches. *TOG* 39, 4 (2020), 72:1–72:16.
- [73] Tao Chen, Ming-Ming Cheng, Ping Tan, Ariel Shamir, and Shi-Min Hu. 2009. Sketch2photo: Internet image montage. *TOG* (2009).
- [74] Weilun Chen, Zhaoxiang Zhang, Xiaolin Hu, and Baoyuan Wu. 2020. Boosting decision-based black-box adversarial attacks with random sign flip. In *ECCV*.
- [75] Xi Chen, Yan Duan, Rein Houthoofd, John Schulman, Ilya Sutskever, and Pieter Abbeel. 2016. InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets. In *NeurIPS*. 2172–2180.
- [76] Xi Chen, Nikhil Mishra, Mostafa Rohaninejad, and Pieter Abbeel. 2018. Pixel-SNAIL: An improved autoregressive generative model. In *ICML*.
- [77] Xu Chen, Jie Song, and Otmar Hilliges. 2019. Monocular neural image based rendering with continuous view control. In *ICCV*. 4090–4100.
- [78] Xinyuan Chen, Chang Xu, Xiaokang Yang, and Dacheng Tao. 2018. Attention-GAN for object transfiguration in wild images. In *ECCV*.
- [79] Yunjin Chen and Thomas Pock. 2016. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *TPAMI* 39, 6 (2016), 1256–1272.
- [80] Yun-Chun Chen, Chen Gao, Esther Robb, and Jia-Bin Huang. 2020. NAS-DIP: Learning Deep Image Prior with Neural Architecture Search. In *ECCV*.
- [81] Zhiqin Chen and Hao Zhang. 2019. Learning implicit fields for generative shape modeling. In *CVPR*. 5939–5948.
- [82] Anton Cherepkov, Andrey Voynov, and Artem Babenko. 2020. Navigating the GAN Parameter Space for Semantic Image Editing. *arXiv preprint arXiv:2011.13786* (2020).
- [83] Dennis Child. 1990. *The essentials of factor analysis*. Cassell Educational.
- [84] Wonwoong Cho, Sungha Choi, David Keetae Park, Inkyu Shin, and Jaegul Choo. 2019. Image-to-image translation via group-wise deep whitening-and-coloring transformation. In *CVPR*.
- [85] Inchang Choi, Orazio Gallo, Alejandro Troccoli, Min H Kim, and Jan Kautz. 2019. Extreme view synthesis. In *ICCV*.
- [86] Yunje Choi, Minje Choi, Munnyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. 2018. StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation. In *CVPR*.
- [87] Tomer Cohen and Lior Wolf. 2019. Bidirectional one-shot unsupervised domain mapping. In *ICCV*.
- [88] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. 2006. Color harmonization. In *ACM SIGGRAPH*. 624–630.
- [89] Edo Collins, Radhakrishna Achanta, and Sabine Susstrunk. 2018. Deep Feature Factorization for Concept Discovery. In *ECCV*, Vol. 14. 352–368.
- [90] Edo Collins and Sabine Susstrunk. 2019. Deep Feature Factorization For Content-Based Image Retrieval And Localization. In *ICIP*. 874–878.
- [91] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. In *CVPR*.
- [92] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine learning* 20, 3 (1995), 273–297.
- [93] Pedro Costa, Adrian Galdran, Maria Ines Meyer, Meindert Niemeijer, Michael Abràmoff, Ana Maria Mendonça, and Aurélio Campilho. 2017. End-to-end Adversarial Retinal Image Synthesis. *IEEE Transactions on Medical Imaging* (2017).
- [94] Rémi Courset, Marine Rougier, Richard Palluel-Germain, Anniq Smeding, Juliette Manto Jonte, Alan Chauvin, and Dominique Muller. 2018. The Caucasian and North African French Faces (CaNAFF): A face database. *International Review of Social Psychology* 31, 1 (2018).
- [95] Antonia Creswell and Anil Anthony Bharath. 2018. Inverting the generator of a generative adversarial network. *TNNLS* (2018).
- [96] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta, and Anil A Bharath. 2018. Generative adversarial networks: An overview. *IEEE Signal Processing Magazine* (2018).
- [97] Ryan Dahl, Mohammad Norouzi, and Jonathon Shlens. 2017. Pixel recursive super resolution. In *ICCV*.
- [98] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. 2017. ScanNet: Richly-annotated 3d reconstructions of indoor scenes. In *CVPR*.
- [99] Bo Dai, Sanja Fidler, Raquel Urtasun, and Dahua Lin. 2017. Towards diverse and natural image descriptions via a conditional GAN. In *ICCV*. 2970–2979.
- [100] Navneet Dalal and Bill Triggs. 2005. Histograms of oriented gradients for human detection. In *CVPR*.
- [101] Giannis Daras, Augustus Odena, Han Zhang, and Alexandros G Dimakis. 2020. Your local GAN: Designing two dimensional local attention mechanisms for generative models. In *CVPR*. 14531–14539.
- [102] Mike E Davies and Christopher J James. 2007. Source separation using single channel ICA. *Signal Processing* 87, 8 (2007), 1819–1832.
- [103] Peter Dayan, Geoffrey E Hinton, Radford M Neal, and Richard S Zemel. 1995. The helmholtz machine. *Neural Computation* (1995).
- [104] Emmanuel de Bézenac, Ibrahim Ayed, and Patrick Gallinari. 2019. Optimal unsupervised domain translation. *arXiv preprint arXiv:1906.01292* (2019).
- [105] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. In *CVPR*. 248–255.
- [106] Emily Denton, Soumith Chintala, Arthur Szlam, and Rob Fergus. 2015. Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks. In *NeurIPS*.
- [107] Emily L Denton and Vighnesh Birodgar. 2017. Unsupervised learning of disentangled representations from video. In *NeurIPS*.
- [108] Sounak Dey, Pau Riba, Anjan Dutta, Josep Lladós, and Yi-Zhe Song. 2019. Doodle to search: Practical zero-shot sketch-based image retrieval. In *CVPR*. 2179–2188.
- [109] Gunet S Dhillon, Kamyar Azizzadenesheli, Zachary C Lipton, Jeremy Bernstein, Jean Kossai, Aran Khanna, and Anima Anandkumar. 2018. Stochastic activation pruning for robust adversarial defense. *arXiv preprint arXiv:1803.01442* (2018).
- [110] Laurent Dinh, David Krueger, and Yoshua Bengio. 2015. NICE: Non-linear independent components estimation. In *ICLR*.
- [111] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. 2017. Density estimation using real NVP. In *ICLR*.
- [112] Carl Doersch. 2016. Tutorial on Variational Autoencoders. *arXiv preprint arXiv:1606.05908* (2016).
- [113] Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. 2016. Adversarial feature learning. *arXiv preprint arXiv:1605.09782* (2016).
- [114] Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. 2017. Adversarial feature learning. In *ICLR*. *ICLR*.
- [115] Jeff Donahue and Karen Simonyan. 2019. Large scale adversarial representation learning. In *NIPS*. 10542–10552.
- [116] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. 2015. Image super-resolution using deep convolutional networks. *TPAMI* (2015).
- [117] Chao Dong, Chen Change Loy, and Xiaoou Tang. 2016. Accelerating the super-resolution convolutional neural network. In *ECCV*.
- [118] Haoye Dong, Xiaodan Liang, Ke Gong, Hanjiang Lai, Jia Zhu, and Jian Yin. 2018. Soft-gated warping-GAN for pose-guided person image synthesis. In *NeurIPS*.
- [119] Hao Dong, Simiao Yu, Chao Wu, and Yike Guo. 2017. Semantic Image Synthesis via Adversarial Learning. In *ICCV*.
- [120] David L Donoho. 2006. Compressed sensing. *IEEE Transactions on Information Theory* 52, 4 (2006), 1289–1306.
- [121] David L Donoho and Carrie Grimes. 2003. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. *Proceedings of the National*

- Academy of Sciences* 100, 10 (2003), 5591–5596.
- [122] Alexey Dosovitskiy and Thomas Brox. 2016. Generating images with perceptual similarity metrics based on deep networks. In *NeurIPS*.
 - [123] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. 2017. CARLA: An Open Urban Driving Simulator. In *CoRL*.
 - [124] Changying Du, Changde Du, Xingyu Xie, Chen Zhang, and Hao Wang. 2018. Multi-view adversarially learned inference for cross-domain joint distribution matching. In *ACM KDD*. 1348–1357.
 - [125] Yilun Du and Igor Mordatch. 2019. Implicit generation and generalization in energy-based models. In *NeurIPS*.
 - [126] Simon Duane, Anthony D Kennedy, Brian J Pendleton, and Duncan Roweth. 1987. Hybrid monte carlo. *Physics letters B* 195, 2 (1987), 216–222.
 - [127] Vincent Dumoulin, Ishmael Belghazi, Ben Poole, Olivier Mastropietro, Alex Lamb, Martin Arjovsky, and Aaron Courville. 2016. Adversarially learned inference. *arXiv preprint arXiv:1606.00704* (2016).
 - [128] Vincent Dumoulin, Ishmael Belghazi, Ben Poole, Olivier Mastropietro, Alex Lamb, Martin Arjovsky, et al. 2020. 3D Morphable Face Models—Past, Present, and Future. *TOG* 39, 5 (2020), 1–38.
 - [129] Collins Edo, Bala Raja, Price Bob, and Süsstrunk Sabine. 2020. Editing in Style: Uncovering the Local Semantics of GANs. In *CVPR*.
 - [130] Bernhard Egger, William AP Smith, Ayush Tewari, Stefanie Wuhner, Michael Zollhoefer, Thabo Beeler, Florian Bernard, Timo Bolkart, Adam Kortylewski, Sami Romdhani, et al. 2020. 3D Morphable Face Models—Past, Present, and Future. *TOG* 39, 5 (2020), 1–38.
 - [131] Mathias Eitz, Kristian Hildebrand, Tamy Boubekeur, and Marc Alexa. 2010. Sketch-based image retrieval: Benchmark and bag-of-features descriptors. *TVCG* 17, 11 (2010), 1624–1636.
 - [132] Robert F Engle. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society* (1982), 987–1007.
 - [133] Härkönen Erik, Hertzmann Aaron, Lehtinen Jaakko, and Paris Sylvain. 2020. GANSpace: Discovering Interpretable GAN Controls. *arXiv preprint arXiv:2004.02546* (2020).
 - [134] Patrick Esser, Robin Rombach, and Björn Ommer. 2020. A Disentangling Invertible Interpretation Network for Explaining Latent Representations. In *CVPR*.
 - [135] Patrick Esser, Ekaterina Sutter, and Björn Ommer. 2018. A Variational U-Net for Conditional Appearance and Shape Generation. In *CVPR*.
 - [136] Yanbo Fan, Baoyuan Wu, Tuanhui Li, Yong Zhang, Mingyang Li, Zhifeng Li, and Yujun Yang. 2020. Sparse adversarial attack via perturbation factorization. In *ECCV*.
 - [137] Yanbo Fan, Baoyuan Wu, Tuanhui Li, Yong Zhang, Mingyang Li, Zhifeng Li, and Yujun Yang. 2020. Sparse Adversarial Attack via Perturbation Factorization. In *ECCV*.
 - [138] Tiantian Fang and Alexander Schwing. 2019. Co-Generation with GANs using AIS based HMC. In *NeurIPS*.
 - [139] Chelsea Finn, Ian Goodfellow, and Sergey Levine. 2016. Unsupervised learning for physical interaction through video prediction. In *NeurIPS*.
 - [140] Asja Fischer and Christian Igel. 2012. An introduction to restricted Boltzmann machines. In *Iberoamerican congress on pattern recognition*.
 - [141] John Flynn, Michael Broxton, Paul Debevec, Matthew DuVall, Graham Fyfe, Ryan Overbeck, Noah Snavely, and Richard Tucker. 2019. DeepView: View synthesis with learned gradient descent. In *CVPR*.
 - [142] John Flynn, Ivan Neulander, James Philbin, and Noah Snavely. 2016. Deepstereo: Learning to predict new views from the world’s imagery. In *CVPR*.
 - [143] Ohad Fried, Ayush Tewari, Michael Zollhöfer, Adam Finkelstein, Eli Shechtman, Dan B Goldman, Kyle Genova, Zeyu Jin, Christian Theobalt, and Maneesh Agrawala. 2019. Text-based editing of talking-head video. *TOG* (2019).
 - [144] Aviv Gabbay and Yedid Hoshen. 2019. Style generator inversion for image enhancement and animation. *arXiv preprint arXiv:1906.11880* (2019).
 - [145] Oran Gafni, Lior Wolf, and Yaniv Taigman. 2020. Vid2Game: Controllable Characters Extracted from Real-World Videos. In *ICLR*.
 - [146] Tomer Galanti, Lior Wolf, and Sagie Benaim. 2018. The role of minimal complexity functions in unsupervised learning of semantic mappings. In *ICLR*.
 - [147] Leonardo Galteri, Lorenzo Seidenari, Marco Bertini, and Alberto Del Bimbo. 2017. Deep generative adversarial compression artifact removal. In *ICCV*.
 - [148] Yaroslav Ganin, Daniil Kononenko, Diana Sungatullina, and Victor Lempitsky. 2016. Deepwarp: Photorealistic image resynthesis for gaze manipulation. In *ECCV*. 311–326.
 - [149] Pablo Garrido, Levi Valgaerts, Ole Rehmsen, Thorsten Thormahlen, Patrick Perez, and Christian Theobalt. 2014. Automatic face reenactment. In *CVPR*.
 - [150] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2016. Image style transfer using convolutional neural networks. In *CVPR*.
 - [151] Yuying Ge, Ruimao Zhang, Lingyun Wu, Xiaogang Wang, Xiaou Tang, and Ping Luo. 2019. A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images. In *CVPR*.
 - [152] Adam Gettey. 2020. GitHub - Face Recognition. https://github.com/ageitgey/face_recognition
 - [153] Stuart Geman and Donald Geman. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *TPAMI* 6 (1984), 721–741.
 - [154] Jiahao Geng, Tianjia Shao, Youyi Zheng, Yanlin Weng, and Kun Zhou. 2018. Warp-guided GANs for single-photo facial animation. *TOG* (2018).
 - [155] Mathieu Germain, Gregor Karol, Murray Iain, and Larochelle Hugo. 2015. MADE: Masked Autoencoder for Distribution Estimation. In *ICML*.
 - [156] Samuel Gershman and Noah Goodman. 2014. Amortized inference in probabilistic reasoning. In *Proceedings of the annual meeting of the cognitive science society*, Vol. 36.
 - [157] Allen Gersho and Robert M Gray. 2012. *Vector quantization and signal compression*. Vol. 159. Springer Science & Business Media.
 - [158] Michaël Gharbi, Gaurav Chaurasia, Sylvain Paris, and Frédo Durand. 2016. Deep joint demosaicking and denoising. *TOG* (2016).
 - [159] Arnab Ghosh, Viveka Kulharia, Vinay Nambodiri, Philip HS Torr, and Puneet K Dokania. 2018. Multi-Agent Diverse Generative Adversarial Networks. In *CVPR*.
 - [160] Arnab Ghosh, Viveka Kulharia, Vinay P Nambodiri, Philip HS Torr, and Puneet K Dokania. 2018. Multi-agent diverse generative adversarial networks. In *CVPR*.
 - [161] Arnab Ghosh, Richard Zhang, Puneet K. Dokania, Oliver Wang, Alexei A. Efros, Philip H. S. Torr, and Eli Shechtman. 2019. Interactive Sketch & Fill: Multiclass Sketch-to-Image Translation. In *ICCV*.
 - [162] Anna C. Gilbert, Yi Zhang, Kibok Lee, Yuting Zhang, and Honglak Lee. 2017. Towards Understanding the Invertibility of Convolutional Neural Networks. In *IJCAI*. 1703–1710.
 - [163] Lore Goetschalckx, Alex Andonian, Aude Oliva, and Phillip Isola. 2019. Ganalyze: Toward visual definitions of cognitive image properties. In *ICCV*.
 - [164] Vladislav Golyanik, Kihwan Kim, Robert Maier, Matthias Nießner, Didier Stricker, and Jan Kautz. 2017. Multiframe scene flow with piecewise rigid motion. In *3DV*.
 - [165] Abel Gonzalez-Garcia, Joost Van De Weijer, and Yoshua Bengio. 2018. Image-to-image translation for cross-domain disentanglement. In *NeurIPS*.
 - [166] Ian Goodfellow. 2016. NIPS 2016 tutorial: Generative adversarial networks. *arXiv preprint arXiv:1701.00160* (2016).
 - [167] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep learning*. MIT Press.
 - [168] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial networks. In *NeurIPS*.
 - [169] Artur Grigorev, Artem Sevastopolsky, Alexander Vakhitov, and Victor Lempitsky. 2019. Coordinate-based texture inpainting for pose-guided human image generation. In *CVPR*.
 - [170] Jinjin Gu, Yujun Shen, and Bolei Zhou. 2020. Image Processing Using Multi-Code GAN Prior. In *CVPR*.
 - [171] Kuangxiao Gu, Yuqian Zhou, and Thomas S Huang. 2020. FLNet: Landmark Driven Fetching and Learning Network for Faithful Talking Facial Animation Synthesis. In *AAAI*.
 - [172] Shanyan Guan, Ying Tai, Bingbing Ni, Feida Zhu, Feiyue Huang, and Xiaokang Yang. 2020. Collaborative Learning for Faster StyleGAN Embedding. *arXiv preprint arXiv:2007.01758* (2020).
 - [173] Jie Gui, Zhenan Sun, Yonggang Wen, Dacheng Tao, and Jieping Ye. 2020. A review on generative adversarial networks: Algorithms, theory, and applications. *arXiv preprint arXiv:2001.06937* (2020).
 - [174] John T Guibas, Tejal S Virdi, and Peter S Li. 2017. Synthetic Medical Images from Dual Generative Adversarial Networks. *arXiv preprint arXiv:1709.01872* (2017).
 - [175] Riza Alp Güler, Natalia Neverova, and Iasonas Kokkinos. 2018. DensePose: Dense Human Pose Estimation In The Wild. In *CVPR*.
 - [176] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. 2017. Improved training of wasserstein GANs. In *NeurIPS*. 5767–5777.
 - [177] Agrim Gupta, Justin Johnson, Alexandre Alahi, and Li Fei-Fei. 2017. Characterizing and improving stability in neural style transfer. In *ICCV*.
 - [178] David Ha, Andrew Dai, and Quoc V Le. 2016. Hypernetworks. In *ICLR*.
 - [179] David Ha and Jürgen Schmidhuber. 2018. World Models. In *NeurIPS*.
 - [180] Sungjoo Ha, Martin Kersner, Beomsu Kim, Seokjun Seo, and Dongyoung Kim. 2020. MarioNETT: Few-shot Face Reenactment Preserving Identity of Unseen Targets. In *AAAI*.
 - [181] Paul Hand and Vladislav Voroninski. 2020. Global Guarantees for Enforcing Deep Generative Priors by Empirical Risk. *IEEE Transactions on Information Theory* 66, 1 (2020), 401–418.
 - [182] N. Hansen and A. Ostermeier. 2001. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*. *Evolutionary Computation* (2001).
 - [183] Zekun Hao, Xun Huang, and Serge Belongie. 2018. Controllable Video Generation with Sparse Trajectories. In *CVPR*.
 - [184] Harry H Harman. 1976. *Modern factor analysis*. University of Chicago Press.

- [185] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask R-CNN. In *ICCV*.
- [186] Kaiming He, Jian Sun, and Xiaoou Tang. 2010. Single image haze removal using dark channel prior. *TPAMI* 33, 12 (2010), 2341–2353.
- [187] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *CVPR*.
- [188] Tong He, John Collomosse, Hailin Jin, and Stefano Soatto. 2020. Geo-PIFu: Geometry and Pixel Aligned Implicit Functions for Single-view Human Reconstruction. *arXiv preprint arXiv:2006.08072* (2020).
- [189] Zhe He, Adrian Spurr, Xucong Zhang, and Otmar Hilliges. 2019. Photo-realistic Monocular Gaze Redirection using Generative Adversarial Networks. *ICCV* (2019).
- [190] Peter Hedman, Julien Philip, True Price, Jan-Michael Frahm, George Drettakis, and Gabriel Brostow. 2018. Deep blending for free-viewpoint image-based rendering. *TOG* (2018).
- [191] Dan Hendrycks and Kevin Gimpel. 2016. A baseline for detecting misclassified and out-of-distribution examples in neural networks. *arXiv preprint arXiv:1610.02136* (2016).
- [192] Philipp Henzler, Niloy J Mitra, and Tobias Ritschel. 2019. Escaping Plato's Cave: 3D Shape From Adversarial Rendering. In *CVPR*.
- [193] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. 2017. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In *NeurIPS*.
- [194] Irina Higgins, Loïc Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. 2017. beta-VAE: Learning Basic Visual Concepts with a Constrained Variational Framework. In *ICLR*.
- [195] Geoffrey E Hinton, Simon Osindero, and Yee-Whye Teh. 2006. A fast learning algorithm for deep belief nets. *Neural Computation* (2006).
- [196] Geoffrey E Hinton and Ruslan R Salakhutdinov. 2006. Reducing the dimensionality of data with neural networks. *Science* (2006).
- [197] Kary Ho, Andrew Gilbert, Hailin Jin, and John Collomosse. 2020. Neural Architecture Search for Deep Image Prior. *arXiv preprint arXiv:2001.04776* (2020).
- [198] Matthew Hoffman, Francis R Bach, and David M Blei. 2010. Online learning for latent dirichlet allocation. In *NeurIPS*, 856–864.
- [199] Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. 2013. Stochastic variational inference. *JMLR* 14, 1 (2013), 1303–1347.
- [200] Seunghoon Hong, Dingdong Yang, Jongwook Choi, and Honglak Lee. 2018. Inferring Semantic Layout for Hierarchical Text-to-Image Synthesis. In *CVPR*.
- [201] Qiyang Hu, Adrian Waelchli, Tiziano Portenier, Matthias Zwicker, and Paolo Favaro. 2018. Video Synthesis from a Single Image and Motion Stroke. *arXiv preprint arXiv:1812.01874* (2018).
- [202] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. 2017. Densely connected convolutional networks. In *CVPR*.
- [203] Haozhi Huang, Hao Wang, Wenhao Luo, Lin Ma, Wenhao Jiang, Xiaolong Zhu, Zhifeng Li, and Wei Liu. 2017. Real-time neural style transfer for videos. In *CVPR*.
- [204] Wen Huang, Paul Hand, Reinhard Heckel, and Vladislav Voroninski. 2018. A provably convergent scheme for compressive sensing under random generative priors. *arXiv preprint arXiv:1812.04176* (2018).
- [205] Xun Huang and Serge Belongie. 2017. Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. In *ICCV*.
- [206] Xinyu Huang, Xinjing Cheng, Qichuan Geng, Binbin Cao, Dingfu Zhou, Peng Wang, Yuanqing Lin, and Ruigang Yang. 2018. The ApolloScape Dataset for Autonomous Driving. In *CVPR*.
- [207] Xun Huang, Yixuan Li, Omid Poursaeed, John Hopcroft, and Serge Belongie. 2017. Stacked Generative Adversarial Networks. In *CVPR*.
- [208] Xun Huang, Yixuan Li, Omid Poursaeed, John E Hopcroft, and Serge J Belongie. 2017. Stacked Generative Adversarial Networks.. In *CVPR*.
- [209] Xun Huang, Ming-Yu Liu, Serge J. Belongie, and Jan Kautz. 2018. Multimodal Unsupervised Image-to-Image Translation. In *ECCV*, 179–196.
- [210] Xun Huang, Ming-Yu Liu, Serge Belongie, and Jan Kautz. 2018. Multimodal Unsupervised Image-to-Image Translation. *ECCV* (2018).
- [211] Zeng Huang, Yuanlu Xu, Christoph Lassner, Hao Li, and Tony Tung. 2020. ARCH: Animatable Reconstruction of Clothed Humans. In *CVPR*.
- [212] Minyoung Huh, Richard Zhang, Jun-Yan Zhu, Sylvain Paris, and Aaron Hertzmann. 2020. Transforming and Projecting Images into Class-conditional Generative Networks. In *ECCV*.
- [213] Shady Abu Hussein, Tom Tirer, and Raja Giryes. 2019. Image-adaptive GAN based reconstruction. *arXiv preprint arXiv:1906.05284* (2019).
- [214] Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. 2017. Globally and locally consistent image completion. *TOG* (2017).
- [215] Eddy Ilg, Nikolaus Mayer, Tomoy Saikia, Margret Keuper, Alexey Dosovitskiy, and Thomas Brox. 2017. FlowNet 2.0: Evolution of optical flow estimation with deep networks. In *CVPR*.
- [216] Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *ICML*.
- [217] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *CVPR*.
- [218] Ali Jahani, Lucy Chai, and Phillip Isola. 2020. On the "steerability" of generative adversarial networks. In *ICLR*.
- [219] Amir Jamaludin, Joon Son Chung, and Andrew Zisserman. 2019. You said that?: Synthesising talking faces from audio. *IJCV* (2019).
- [220] Xu Jia, Bert De Brabandere, Tinne Tuytelaars, and Luc V Gool. 2016. Dynamic filter networks. In *NeurIPS*.
- [221] Ye Jia, Yu Zhang, Ron J. Weiss, Quan Wang, Jonathan Shen, Fei Ren, Zhifeng Chen, Patrick Nguyen, Ruoming Pang, Ignacio Lopez-Moreno, and Yonghui Wu. 2018. Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis. In *NeurIPS*, 4485–4495.
- [222] Zi-Hang Jiang, Qianyi Wu, Keyu Chen, and Juyong Zhang. 2019. Disentangled Representation Learning for 3D Face Shape. In *CVPR*.
- [223] Y Jin, J Zhang, M Li, Y Tian, and H Zhu. 2017. Towards the high-quality anime characters generation with generative adversarial networks. In *Proceedings of the Machine Learning for Creativity and Design Workshop at NeurIPS*.
- [224] Youngjo Jo and Jongyoul Park. 2019. SC-FEGAN: Face Editing Generative Adversarial Network with User's Sketch and Color. In *ICCV*.
- [225] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. 2016. Perceptual losses for real-time style transfer and super-resolution. In *ECCV*, 694–711.
- [226] Justin Johnson, Agrim Gupta, and Li Fei-Fei. 2018. Image generation from scene graphs. In *CVPR*.
- [227] Matthew Johnson, Gabriel J Brostow, Jamie Shotton, Ognjen Arandjelovic, Vivek Kwatra, and Roberto Cipolla. 2006. Semantic photo synthesis. In *Computer Graphics Forum*.
- [228] Alexia Jolicoeur-Martineau. 2019. The relativistic discriminator: a key element missing from standard GAN. In *ICLR*.
- [229] Alexia Jolicoeur-Martineau. 2019. The relativistic discriminator: a key element missing from standard GAN. In *ICLR*. <https://openreview.net/forum?id=S1erHoR5t7>
- [230] Alexia Jolicoeur-Martineau. 2020. On Relativistic f-Divergences. In *ICML*.
- [231] Donggyu Joo, Doyeon Kim, and Junmo Kim. 2018. Generating a fusion image: One's identity and another's shape. In *CVPR*.
- [232] James T Kajiya. 1986. The rendering equation. In *SIGGRAPH*.
- [233] Nima Khademi Kalantari, Ting-Chun Wang, and Ravi Ramamoorthi. 2016. Learning-based view synthesis for light field cameras. *TOG* (2016).
- [234] Nal Kalchbrenner, Aaron van den Oord, Karen Simonyan, Ivo Danihelka, Oriol Vinyals, Alex Graves, and Koray Kavukcuoglu. 2017. Video pixel networks. *ICML* (2017).
- [235] Takuhiro Kaneko, Kaoru Hiramatsu, and Kunio Kashino. 2017. Generative Attribute Controller with Conditional Filtered Generative Adversarial Networks. In *CVPR*.
- [236] Anton S Kaplanyan, Anton Sochenov, Thomas Leimkühler, Mikhail Okunev, Todd Goodall, and Gizem Rufo. 2019. DeepFovea: neural reconstruction for foveated rendering and video compression using learned statistics of natural videos. *TOG* (2019).
- [237] Levent Karacan, Zeynep Akata, Aykut Erdem, and Erkut Erdem. 2016. Learning to generate images of outdoor scenes from attributes and semantic layouts. *arXiv preprint arXiv:1612.00215* (2016).
- [238] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. 2018. Progressive growing of GANs for improved quality, stability, and variation. In *ICLR*.
- [239] Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In *CVPR*, 4401–4410.
- [240] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and improving the image quality of StyleGAN. In *CVPR*.
- [241] Hiroharu Kato and Tatsuya Harada. 2019. Self-supervised Learning of 3D Objects from Natural Images. *arXiv preprint arXiv:1911.08850* (2019).
- [242] Hyeonwoo Kim, Pablo Garrido, Ayush Tewari, Weipeng Xu, Justus Thies, Matthias Niessner, Patrick Pérez, Christian Richardt, Michael Zollhöfer, and Christian Theobalt. 2018. Deep video portraits. *TOG* (2018).
- [243] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. 2016. Accurate image super-resolution using very deep convolutional networks. In *CVPR*.
- [244] Minsu Kim, Sunghun Joung, Kihong Park, Seunghyong Kim, and Kwanghoon Sohn. 2019. Unpaired Cross-Spectral Pedestrian Detection Via Adversarial Feature Learning. In *ICIP*, IEEE, 1650–1654.
- [245] Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jung Kwon Lee, and Jiwon Kim. 2017. Learning to discover cross-domain relations with generative adversarial networks. In *ICML*.
- [246] Davis E King. 2009. Dlib-ml: A machine learning toolkit. *JMLR* 10 (2009), 1755–1758.
- [247] Davis E. King. 2009. Dlib-ml: A Machine Learning Toolkit. *JMLR* (2009).

- [248] Diederik Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *ICLR*.
- [249] Durk P Kingma and Prafulla Dhariwal. 2018. Glow: Generative flow with invertible 1x1 convolutions. In *NeurIPS*. 10215–10224.
- [250] Durk P Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, and Max Welling. 2016. IAF: Improved variational inference with inverse autoregressive flow. In *NeurIPS*. 4743–4751.
- [251] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational Bayes. In *ICLR*.
- [252] Diederik P Kingma and Max Welling. 2019. An introduction to variational autoencoders. *arXiv preprint arXiv:1906.02691* (2019).
- [253] Teuvo Kohonen and Matti Ruohonen. 1973. Representation of Associated Data by Matrix Operators. *IEEE Trans. Comput.* 22, 7 (1973), 701–702.
- [254] Alex Krizhevsky, Geoffrey Hinton, et al. 2009. Learning multiple layers of features from tiny images. *Technical Report* (2009).
- [255] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *NeurIPS*.
- [256] Line Kuhnelt, Tom Fletcher, Sarang Joshi, and Stefan Sommer. 2018. Latent space non-linear statistics. *arXiv preprint arXiv:1805.07632* (2018).
- [257] Line Kuhnelt, Tom Fletcher, Sarang C. Joshi, and Stefan Sommer. 2017. Latent Space Non-Linear Statistics. *arXiv preprint arXiv:1706.05587* (2017).
- [258] Orest Kuppyn, Volodymyr Budzan, Mykola Mykhailchuk, Dmytro Mishkin, and Jiri Matas. 2018. DeblurGAN: Blind motion deblurring using conditional adversarial networks. In *CVPR*.
- [259] Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang. 2018. Fast and accurate image super-resolution with deep laplacian pyramid networks. *TPAMI* (2018).
- [260] Wei-Sheng Lai, Jia-Bin Huang, Oliver Wang, Eli Shechtman, Ersin Yumer, and Ming-Hsuan Yang. 2018. Learning Blind Video Temporal Consistency. In *ECCV*.
- [261] Jean-François Lalonde, Derek Hoiem, Alexei A Efros, Carsten Rother, John Winn, and Antonio Criminisi. 2007. Photo clip art. *TOG* (2007).
- [262] Oliver Langner, Ron Dotsch, Gijsbert Bijlstra, Daniel HJ Wigboldus, Skyler T Hawk, and AD Van Knippenberg. 2010. Presentation and validation of the Radboud Faces Database. *Cognition and emotion* 24, 8 (2010), 1377–1388.
- [263] Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, and Ole Winther. 2016. Autoencoding beyond pixels using a learned similarity metric. In *ICML*.
- [264] Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, and Ole Winther. 2016. Autoencoding beyond pixels using a learned similarity metric. In *ICML*.
- [265] Gustav Larsson, Michael Maire, and Gregory Shakhnarovich. 2016. Learning representations for automatic colorization. In *ECCV*. 577–593.
- [266] Vuong Le, Jonathan Brandt, Zhe Lin, Lubomir Bourdev, and Thomas S Huang. 2012. Interactive facial feature localization. In *ECCV*.
- [267] Yann LeCun. 1998. The MNIST database of handwritten digits. <http://yann.lecun.com/exdb/mnist/> (1998).
- [268] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. 2015. Deep learning. *Nature* (2015).
- [269] Yann LeCun, Sumit Chopra, Raia Hadsell, M Ranzato, and F Huang. 2006. A tutorial on energy-based learning. *Predicting structured data* (2006).
- [270] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In *CVPR*.
- [271] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In *CVPR*.
- [272] Alex X Lee, Richard Zhang, Frederik Ebert, Pieter Abbeel, Chelsea Finn, and Sergey Levine. 2018. Stochastic Adversarial Video Prediction. *arXiv preprint arXiv:1804.01523* (2018).
- [273] Cheng-Han Lee, Ziwei Liu, Lingyun Wu, and Ping Luo. 2020. MaskGAN: Towards Diverse and Interactive Facial Image Manipulation. In *CVPR*.
- [274] Cheng-Han Lee, Ziwei Liu, Lingyun Wu, and Ping Luo. 2020. MaskGAN: Towards diverse and interactive facial image manipulation. In *CVPR*.
- [275] D.D. Lee and H.S. Seung. 2016. Learning the parts of objects by non-negative matrix factorization. *Nature* (2016).
- [276] Daniel D Lee and H Sebastian Seung. 2001. Algorithms for non-negative matrix factorization. In *NeurIPS*. 556–562.
- [277] Hsin-Ying Lee, Hung-Yu Tseng, Jia-Bin Huang, Maneesh Singh, and Ming-Hsuan Yang. 2018. Diverse image-to-image translation via disentangled representations. In *ECCV*.
- [278] Kimin Lee, Kibok Lee, Honglak Lee, and Jinwoo Shin. 2018. A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In *NeurIPS*. 7167–7177.
- [279] Soochan Lee, Junsoo Ha, and Gunhee Kim. 2019. Harmonizing maximum likelihood with GANs for multimodal conditional generation. In *ICLR*.
- [280] Qi Lei, Ajil Jalal, Inderjit S Dhillon, and Alexandros G Dimakis. 2019. Inverting Deep Generative models, One layer at a time. In *NeurIPS*. 13910–13919.
- [281] Anat Levin, Dani Lischinski, and Yair Weiss. 2004. Colorization using optimization. *TOG* (2004).
- [282] Bowen Li, Xiaojuan Qi, Thomas Lukasiewicz, and Philip H. S. Torr. 2019. Controllable text-to-image generation. In *NeurIPS*.
- [283] Bowen Li, Xiaojuan Qi, Thomas Lukasiewicz, and Philip HS Torr. 2020. ManiGAN: Text-Guided Image Manipulation. In *CVPR*.
- [284] Chunyun Li, Hao Liu, Changyou Chen, Yuchen Pu, Liqun Chen, Ricardo Henao, and Lawrence Carin. 2017. Alice: Towards understanding adversarial learning for joint distribution matching. In *NeurIPS*. 5495–5503.
- [285] Haoliang Li, Sinno Jialin Pan, Shiqi Wang, and Alex C Kot. 2018. Domain generalization with adversarial feature learning. In *CVPR*. 5400–5409.
- [286] Jianan Li, Xiaodan Liang, Yunchao Wei, Tingfa Xu, Jiashi Feng, and Shuicheng Yan. 2017. Perceptual generative adversarial networks for small object detection. In *CVPR*.
- [287] Ke Li, Tianhao Zhang, and Jitendra Malik. 2019. Diverse image synthesis from semantic layouts via conditional IMLE. In *ICCV*.
- [288] Muyang Li, Ji Lin, Yaoyao Ding, Zhijian Liu, Jun-Yan Zhu, and Song Han. 2020. GAN compression: Efficient architectures for interactive conditional GANs. In *CVPR*.
- [289] Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, and Ming-Hsuan Yang. 2018. Flow-grounded spatial-temporal video prediction from still images. In *ECCV*.
- [290] Yining Li, Chen Huang, and Chen Change Loy. 2019. Dense intrinsic appearance flow for human pose transfer. In *CVPR*.
- [291] Zhengqi Li, Tali Dekel, Forrester Cole, Richard Tucker, Noah Snavely, Ce Liu, and William T Freeman. 2019. Learning the depths of moving people by watching frozen people. In *CVPR*.
- [292] Zhengqi Li and Noah Snavely. 2018. MegaDepth: Learning single-view depth prediction from internet photos. In *CVPR*.
- [293] Ziqiang Li, Rentuo Tao, Hongjing Niu, and Bin Li. 2020. Interpreting the Latent Space of GANs via Correlation Analysis for Controllable Concept Manipulation. *arXiv preprint arXiv:2006.10132* (2020).
- [294] Zhengqi Li, Wenqi Xian, Abe Davis, and Noah Snavely. 2020. Crowdsampling the Plenoptic Function. In *ECCV*.
- [295] Xiaodan Liang, Lisa Lee, Wei Dai, and Eric P Xing. 2017. Dual motion GAN for future-flow embedded video prediction. In *NeurIPS*.
- [296] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. 2017. Enhanced deep residual networks for single image super-resolution. In *CVPR Workshop*.
- [297] Jae Hyun Lim and Jong Chul Ye. 2017. Geometric GAN. *arXiv preprint arXiv:1705.02894* (2017).
- [298] Chih-Jen Lin. 2007. Projected gradient methods for nonnegative matrix factorization. *Neural Computation* 19, 10 (2007), 2756–2779.
- [299] Jianxin Lin, Yingxue Pang, Yingce Xia, Zhibo Chen, and Jiebo Luo. 2020. TuiGAN: Learning Versatile Image-to-Image Translation with Two Unpaired Images. In *ECCV*.
- [300] Zachary C Lipton and Subarna Tripathi. 2017. Precise recovery of latent vectors from generative adversarial networks. *arXiv preprint arXiv:1702.04782* (2017).
- [301] Alexander H. Liu, Yen-Cheng Liu, Yu-Ying Yeh, and Yu-Chiang Frank Wang. 2018. A Unified Feature Disentangler for Multi-Domain Image Translation and Manipulation. In *NeurIPS*. 2595–2604.
- [302] Dong C Liu and Jorge Nocedal. 1989. On the limited memory BFGS method for large scale optimization. *Mathematical programming* 45, 1-3 (1989), 503–528.
- [303] Guilin Liu, Fitsum A Reda, Kevin J Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. 2018. Image inpainting for irregular holes using partial convolutions. In *ECCV*.
- [304] Lingjie Liu, Weipeng Xu, Marc Habermann, Michael Zollhoefer, Florian Bernard, Hyeonwoo Kim, Wenping Wang, and Christian Theobalt. 2020. Neural Human Video Rendering by Learning Dynamic Textures and Rendering-to-Video Translation. *TVCG* (2020).
- [305] Lingjie Liu, Weipeng Xu, Michael Zollhoefer, Hyeonwoo Kim, Florian Bernard, Marc Habermann, Wenping Wang, and Christian Theobalt. 2019. Neural rendering and reenactment of human actor videos. *TOG* (2019).
- [306] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. 2017. Unsupervised Image-to-Image Translation Networks. In *NeurIPS*.
- [307] Ming-Yu Liu, Xun Huang, Arun Mallya, Tero Karras, Timo Aila, Jaakko Lehtinen, and Jan Kautz. 2019. Few-shot Unsupervised Image-to-Image Translation. In *ICCV*.
- [308] Ming-Yu Liu, Xun Huang, Arun Mallya, Tero Karras, Timo Aila, Jaakko Lehtinen, and Jan Kautz. 2019. Few-shot unsupervised image-to-image translation. In *ICCV*.

- [309] Ming-Yu Liu and Oncel Tuzel. 2016. Coupled Generative Adversarial Networks. In *NeurIPS*.
- [310] Shaohui Liu, Xiao Zhang, Jianqiao Wangni, and Jianbo Shi. 2019. Normalized diversification. In *CVPR*.
- [311] Wen Liu, Zhixin Piao, Jie Min, Wenhan Luo, Lin Ma, and Shenghua Gao. 2019. Liquid Warping GAN: A unified framework for human motion imitation, appearance transfer and novel view synthesis. In *ICCV*.
- [312] Xialei Liu, Joost van de Weijer, and Andrew D. Bagdanov. 2017. RankIQ: Learning from Rankings for No-Reference Image Quality Assessment. In *Proc. IEEE Int. Conf. Comput. Vis.* 1040–1049.
- [313] Xihui Liu, Guojun Yin, Jing Shao, Xiaogang Wang, et al. 2019. Learning to predict layout-to-image conditional convolutions for semantic image synthesis. In *NeurIPS*.
- [314] Yunfan Liu, Qi Li, Zhenan Sun, and Tieniu Tan. 2020. Style Intervention: How to Achieve Spatial Disentanglement with Style-based Generators? *arXiv preprint arXiv:2011.09699* (2020).
- [315] Ziwei Liu, Ping Luo, Shi Qiu, Xiaogang Wang, and Xiaoou Tang. 2016. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In *CVPR*. 1096–1104.
- [316] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2015. Deep Learning Face Attributes in the Wild. In *ICCV*.
- [317] Ziwei Liu, Sijie Yan, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2016. Fashion landmark detection in the wild. In *ECCV*. 229–245.
- [318] Stephen Lombardi, Tomas Simon, Jason Saragih, Gabriel Schwartz, Andreas Lehrmann, and Yaser Sheikh. 2019. Neural volumes: Learning dynamic renderable volumes from images. *arXiv preprint arXiv:1906.07751* (2019).
- [319] Jonathan Long, Evan Shelhamer, and Trevor Darrell. 2015. Fully convolutional networks for semantic segmentation. In *CVPR*.
- [320] H Christopher Longuet-Higgins. 1981. A computer algorithm for reconstructing a scene from two projections. *Nature* (1981).
- [321] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. 2015. SMPL: A skinned multi-person linear model. *TOG* (2015).
- [322] Dominik Lorenz, Leonard Bereska, Timo Milbich, and Bjorn Ommer. 2019. Unsupervised part-based disentangling of object shape and appearance. In *CVPR*.
- [323] William Lotter, Gabriel Kreiman, and David Cox. 2017. Deep predictive coding networks for video prediction and unsupervised learning. In *ICLR*.
- [324] Boyu Lu, Jun-Cheng Chen, and Rama Chellappa. 2019. Unsupervised Domain-Specific Deblurring via Disentangled Representations. *CVPR* (2019).
- [325] Yu-Ding Lu, Hsin-Ying Lee, Hung-Yu Tseng, and Ming-Hsuan Yang. 2020. Unsupervised Discovery of Disentangled Manifolds in GANs. *arXiv preprint arXiv:2011.11842* (2020).
- [326] Bruce D Lucas, Takeo Kanade, et al. 1981. An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence (IJCAI)* (1981).
- [327] Andreas Lugmayr, Martin Danelljan, Luc Van Gool, and Radu Timofte. 2020. SRFlow: Learning the Super-Resolution Space with Normalizing Flow. In *ECCV*.
- [328] Junyu Luo, Yong Xu, Chenwei Tang, and Jiancheng Lv. 2017. Learning inverse mapping by autoencoder based generative adversarial nets. In *International Conference on Neural Information Processing*. 207–216.
- [329] Chao Ma, Chih-Yuan Yang, Xiaokang Yang, and Ming-Hsuan Yang. 2017. Learning a no-reference quality metric for single-image super-resolution. *CVIU* (2017).
- [330] Fangchang Ma, Ulas Ayaz, and Sertac Karaman. 2018. Invertibility of convolutional generative networks from partial measurements. In *NeurIPS*. 9651–9660.
- [331] Liqian Ma, Xu Jia, Stamatios Georgoulis, Tinne Tuytelaars, and Luc Van Gool. 2019. Exemplar guided unsupervised image-to-image translation with semantic consistency. In *ICLR*.
- [332] Liqian Ma, Xu Jia, Qianru Sun, Bernt Schiele, Tinne Tuytelaars, and Luc Van Gool. 2017. Pose guided person image generation. In *NeurIPS*.
- [333] Liqian Ma, Qianru Sun, Stamatios Georgoulis, Luc Van Gool, Bernt Schiele, and Mario Fritz. 2018. Disentangled person image generation. In *CVPR*.
- [334] Andrew L Maas, Awni Y Hannun, and Andrew Y Ng. 2013. Rectifier nonlinearities improve neural network acoustic models. In *ICML*.
- [335] Shunta Maeda. 2020. Unpaired Image Super-Resolution using Pseudo-Supervision. In *CVPR*.
- [336] Alireza Makhzani, Jonathon Shlens, Navdeep Jaitly, and Ian J. Goodfellow. 2015. Adversarial Autoencoders. *arXiv preprint arXiv:1511.05644* (2015).
- [337] Arun Mallya, Ting-Chun Wang, Karan Sapra, and Ming-Yu Liu. 2020. World-Consistent Video-to-Video Synthesis. In *ECCV*.
- [338] Qi Mao, Hsin-Ying Lee, Hung-Yu Tseng, Siwei Ma, and Ming-Hsuan Yang. 2019. Mode seeking generative adversarial networks for diverse image synthesis. In *CVPR*.
- [339] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, and Stephen Paul Smolley. 2017. Least squares generative adversarial networks. In *ICCV*.
- [340] Ricardo Martin-Brualla, Rohit Pandey, Shuoran Yang, Pavel Pidlypenskiy, Jonathan Taylor, Julien Valentin, Sameh Khamis, Philip Davidson, Anastasia Tkach, Peter Lincoln, Adarsh Kowdle, Christoph Rhemann, Dan B Goldman, Cem Keskin, Steve Seitz, Shahram Izadi, and Sean Fanello. 2018. LookinGood: Enhancing Performance Capture with Real-Time Neural Re-Rendering. *TOG* (2018).
- [341] Ricardo Martin-Brualla, Noha Radwan, Mehdi SM Sajjadi, Jonathan T Barron, Alexey Dosovitskiy, and Daniel Duckworth. 2020. NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. *arXiv preprint arXiv:2008.02268* (2020).
- [342] Michael Mathieu, Camille Couprie, and Yann LeCun. 2016. Deep multi-scale video prediction beyond mean square error. In *ICLR*.
- [343] Youssef Alami Mejjati, Christian Richardt, James Tompkin, Darren Cosker, and Kwang In Kim. 2018. Unsupervised attention-guided image-to-image translation. In *NeurIPS*.
- [344] Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. 2018. Which training methods for GANs do actually converge?. In *ICML*.
- [345] Moustafa Meshry, Dan B Goldman, Sameh Khamis, Hugues Hoppe, Rohit Pandey, Noah Snaveley, and Ricardo Martin-Brualla. 2019. Neural rendering in the wild. In *CVPR*.
- [346] Ben Mildenhall, Pratul P Srinivasan, Rodrigo Ortiz-Cayon, Nima Khademi Kalantari, Ravi Ramamoorthi, Ren Ng, and Abhishek Kar. 2019. Local light field fusion: Practical view synthesis with prescriptive sampling guidelines. *TOG* (2019).
- [347] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2020. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*.
- [348] Mehdi Mirza and Simon Osindero. 2014. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784* (2014).
- [349] Anish Mittal, Anush K Moorthy, and Alan C Bovik. 2011. Blind/referenceless image spatial quality evaluator. In *Asilomar Conference on Signals, Systems and Computers*. 723–727.
- [350] A. Mittal, R. Soundararajan, and A. C. Bovik. 2013. Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Processing Letters* 20, 3 (2013), 209–212.
- [351] Takeru Miyato, Toshiaki Kataoka, Masanori Koyama, and Yuichi Yoshida. 2018. Spectral Normalization for Generative Adversarial Networks. In *ICLR*.
- [352] Takeru Miyato and Masanori Koyama. 2018. cGANs with projection discriminator. In *ICLR*.
- [353] Sangwoo Mo, Minsu Cho, and Jinwoo Shin. 2019. InstaGAN: Instance-aware image-to-image translation. In *ICLR*.
- [354] Grégoire Montavon and Klaus-Robert Müller. 2012. Deep Boltzmann machines and the centering trick. In *Neural Networks: Tricks of the Trade*. Springer, 621–637.
- [355] Anush Krishna Moorthy and Alan Conrad Bovik. 2011. Blind image quality assessment: From natural scene statistics to perceptual quality. *TIP* 20, 12 (2011), 3350–3364.
- [356] Koki Nagano, Jaewoo Seo, Jun Xing, Lingyu Wei, Zimo Li, Shunsuke Saito, Aviral Agarwal, Jens Fursund, and Hao Li. 2018. paGAN: real-time avatars using dynamic textures. *TOG* (2018).
- [357] Nikhil Naik, Jade Philipoom, Ramesh Raskar, and César Hidalgo. 2014. Streetscore-predicting the perceived safety of one million streetscapes. In *CVPR Workshops*. 779–785.
- [358] Vinod Nair and Geoffrey E Hinton. 2010. Rectified linear units improve restricted boltzmann machines. In *ICML*.
- [359] Pushmeet Kohli Nathan Silberman, Derek Hoiem and Rob Fergus. 2012. Indoor Segmentation and Support Inference from RGBD Images. In *ECCV*.
- [360] Kamyar Nazeri, Eric Ng, Tony Joseph, Faisal Z Qureshi, and Mehran Ebrahimi. 2019. EdgeConnect: Generative image inpainting with adversarial edge learning. *arXiv preprint arXiv:1901.00212* (2019).
- [361] Vladimir Nekrasov, Thanuja Dharmasiri, Andrew Spek, Tom Drummond, Chunhua Shen, and Ian D. Reid. 2019. Real-Time Joint Semantic Segmentation and Depth Estimation Using Asymmetric Annotations. In *ICRA. IEEE*, 7101–7107.
- [362] Natalia Neverova, Riza Alp Guler, and Iasonas Kokkinos. 2018. Dense pose transfer. In *ECCV*.
- [363] Anh Nguyen, Alexey Dosovitskiy, Jason Yosinski, Thomas Brox, and Jeff Clune. 2016. Synthesizing the preferred inputs for neurons in neural networks via deep generator networks. In *NeurIPS*. 3387–3395.
- [364] Anh Nguyen, Jason Yosinski, and Jeff Clune. 2015. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *CVPR*. 427–436.
- [365] Thu Nguyen-Phuoc, Chuan Li, Lucas Theis, Christian Richardt, and Yong-Liang Yang. 2019. HoloGAN: Unsupervised learning of 3d representations from natural images. In *CVPR*.
- [366] Thu Nguyen-Phuoc, Christian Richardt, Long Mai, Yong-Liang Yang, and Niloy Mitra. 2020. BlockGAN: Learning 3D Object-aware Scene Representations from Unlabelled Images. *arXiv preprint arXiv:2002.08988* (2020).
- [367] Simon Niklaus, Long Mai, and Feng Liu. 2017. Video frame interpolation via adaptive convolution. In *CVPR*.
- [368] Simon Niklaus, Long Mai, and Feng Liu. 2017. Video frame interpolation via adaptive separable convolution. In *ICCV*.

- [369] Yuval Nirkin, Yosi Keller, and Tal Hassner. 2019. FSGAN: Subject agnostic face swapping and reenactment. In *ICCV*.
- [370] Sebastian Nowozin, Botond Cseke, and Ryota Tomioka. 2016. f-GAN: Training generative neural samplers using variational divergence minimization. In *NeurIPS*.
- [371] Spingarn-Eliezer Nurit, Banner Ron, and Michaeli Tomer. 2020. GAN Steerability without optimization. *arXiv preprint arXiv:2012.05328* (2020).
- [372] Augustus Odena, Christopher Olah, and Jonathon Shlens. 2017. Conditional image synthesis with auxiliary classifier GANs. In *ICML*.
- [373] Katsunori Ohnishi, Shohei Yamamoto, Yoshitaka Ushiku, and Tatsuya Harada. 2018. Hierarchical Video Generation from Orthogonal Information: Optical Flow and Texture. In *AAAI*.
- [374] Erkki Oja and Zhijian Yuan. 2006. The FastICA algorithm revisited: Convergence analysis. *IEEE Transactions on Neural Networks* 17, 6 (2006), 1370–1381.
- [375] Kyle Olszewski, Zimo Li, Chao Yang, Yi Zhou, Ronald Yu, Zeng Huang, Sitao Xiang, Shunsuke Saito, Pushmeet Kohli, and Hao Li. 2017. Realistic dynamic facial textures from a single image using GANs. In *ICCV*.
- [376] Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. 2016. WaveNet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499* (2016).
- [377] Aaron van den Oord, Nal Kalchbrenner, and Koray Kavukcuoglu. 2016. Pixel recurrent neural networks. *arXiv preprint arXiv:1601.06759* (2016).
- [378] Roy Or-El, Soumyadip Sengupta, Ohad Fried, Eli Shechtman, and Ira Kemelmacher-Shlizerman. 2020. Lifespan Age Transformation Synthesis. In *ECCV*.
- [379] Junting Pan, Chengyu Wang, Xu Jia, Jing Shao, Lu Sheng, Junjie Yan, and Xiaogang Wang. 2019. Video Generation from Single Semantic Label Map. In *CVPR*.
- [380] Xingang Pan, Xiaohang Zhan, Bo Dai, Dahua Lin, Chen Change Loy, and Ping Luo. 2020. Exploiting Deep Generative Prior for Versatile Image Restoration and Manipulation. In *ECCV*.
- [381] Zhaoqing Pan, Weijie Yu, Xiaokai Yi, Asifullah Khan, Feng Yuan, and Yuhui Zheng. 2019. Recent progress on generative adversarial networks (GANs): A survey. *IEEE Access* (2019).
- [382] Lu Pang, Yaowei Wang, Yi-Zhe Song, Tiejun Huang, and Yonghong Tian. 2018. Cross-Domain Adversarial Feature Learning for Sketch Re-identification. In *ACM MM*. 609–617.
- [383] George Papamakarios, Theo Pavlakou, and Iain Murray. 2017. MAF: Masked Autoregressive Flow for Density Estimation. In *NeurIPS*. 2338–2347.
- [384] Taesung Park, Alexei A Efros, Richard Zhang, and Jun-Yan Zhu. 2020. Contrastive Learning for Unpaired Image-to-Image Translation. In *ECCV*.
- [385] Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. 2019. Semantic Image Synthesis with Spatially-Adaptive Normalization. In *CVPR*.
- [386] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. 2017. Automatic differentiation in PyTorch. (2017).
- [387] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A Efros. 2016. Context encoders: Feature learning by inpainting. In *CVPR*.
- [388] Guim Perarnau, Joost Van De Weijer, Bogdan Raducanu, and Jose M Álvarez. 2016. Invertible conditional gans for image editing. *arXiv preprint arXiv:1611.06355* (2016).
- [389] Francesco Pittaluga, Sanjeev J Koppal, Sing Bing Kang, and Sudipta N Sinha. 2019. Revealing scenes by inverting structure from motion reconstructions. In *CVPR*.
- [390] Antoine Plummerault, Hervé Le Borgne, and Céline Hudelot. 2020. Controlling generative models with continuous factors of variations. *ICLR* (2020).
- [391] K R Prajwal, Rudrabha Mukhopadhyay, Vinay P. Nambodiri, and C.V. Jawahar. 2020. Learning Individual Speaking Styles for Accurate Lip to Speech Synthesis. In *CVPR*.
- [392] Albert Pumarola, Antonio Agudo, Aleix M Martinez, Alberto Sanfeliu, and Francesc Moreno-Noguer. 2018. GANimation: Anatomically-aware facial animation from a single image. In *ECCV*.
- [393] Albert Pumarola, Antonio Agudo, Aleix M Martinez, Alberto Sanfeliu, and Francesc Moreno-Noguer. 2020. GANimation: One-shot anatomically consistent facial animation. *IJCV* (2020).
- [394] Albert Pumarola, Antonio Agudo, Alberto Sanfeliu, and Francesc Moreno-Noguer. 2019. Unsupervised person image synthesis in arbitrary poses. In *CVPR*.
- [395] Xiaojuan Qi, Qifeng Chen, Jiaya Jia, and Vladlen Koltun. 2018. Semi-parametric image synthesis. In *CVPR*.
- [396] Shengju Qian, Kwan-Yee Lin, Wayne Wu, Yangxiaokang Liu, Quan Wang, Fumin Shen, Chen Qian, and Ran He. 2019. Make a face: Towards arbitrary high fidelity face manipulation. In *ICCV*.
- [397] Julien Rabin, Gabriel Peyré, Julie Delon, and Marc Bernot. 2011. Wasserstein barycenter and its application to texture mixing. In *International Conference on Scale Space and Variational Methods in Computer Vision*. 435–446.
- [398] Alec Radford, Luke Metz, and Soumith Chintala. 2016. Unsupervised representation learning with deep convolutional generative adversarial networks. In *ICLR*.
- [399] Ankit Raj, Yuqi Li, and Yoram Bresler. 2019. Gan-based projector for faster recovery with convergence guarantees in linear inverse problems. In *ICCV*. 5602–5611.
- [400] Amit Raj, Patsorn Sangkloy, Huiwen Chang, James Hays, Duygu Ceylan, and Jingwan Lu. 2018. SwapNet: Image based garment transfer. In *ECCV*.
- [401] Sai Rajeswar, Fahim Mannan, Florian Golemo, Jérôme Parent-Lévesque, David Vazquez, Derek Nowrouzezahrai, and Aaron Courville. 2020. Pix2Shape: Towards Unsupervised Learning of 3D Scenes from Images Using a View-Based Representation. *IJCV* (2020), 1–16.
- [402] Abdal Rameen, Zhu Peihao, Mitra Niloy, and Wonka Peter. 2020. StyleFlow: Attribute-conditioned Exploration of StyleGAN-Generated Images using Conditional Continuous Normalizing Flows. *arXiv preprint arXiv:2008.02401* (2020).
- [403] Aditya Ramesh, Youngduck Choi, and Yann LeCun. 2018. A Spectral Regularizer for Unsupervised Disentanglement. *arXiv preprint arXiv:1812.01161* (2018).
- [404] MarcAurelio Ranzato, Arthur Szlam, Joan Bruna, Michael Mathieu, Ronan Collobert, and Sumit Chopra. 2014. Video (language) modeling: a baseline for generative models of natural videos. *arXiv preprint arXiv:1412.6604* (2014).
- [405] J. Rapin and O. Teytaud. 2018. Nevergrad - A gradient-free optimization platform. <https://GitHub.com/FacebookResearch/Nevergrad>.
- [406] Sashank J Reddi, Satyen Kale, and Sanjiv Kumar. 2018. On the convergence of adam and beyond. In *ICLR*.
- [407] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *CVPR*.
- [408] Scott Reed, Zeynep Akata, Xinchun Yan, Lajanugen Logeswaran, Bernt Schiele, and Honglak Lee. 2016. Generative adversarial text to image synthesis. In *ICML*.
- [409] Scott Reed, Aäron van den Oord, Nal Kalchbrenner, Sergio Gómez Colmenarejo, Ziyu Wang, Dan Belov, and Nando de Freitas. 2017. Parallel multiscale autoregressive density estimation. In *ICML*.
- [410] Erik Reinhard, Michael Adhikhmin, Bruce Gooch, and Peter Shirley. 2001. Color transfer between images. *IEEE Computer Graphics and Applications* (2001).
- [411] Hongyu Ren, Diqi Chen, and Yizhou Wang. 2018. RAN4QA: Restorative Adversarial Nets for No-Reference Image Quality Assessment. In *AAAI*. 7308–7314.
- [412] Jie Ren, Peter J Liu, Emily Fertig, Jasper Snoek, Ryan Poplin, Mark Depristo, Joshua Dillon, and Balaji Lakshminarayanan. 2019. Likelihood ratios for out-of-distribution detection. In *NeurIPS*. 14707–14718.
- [413] Yurui Ren, Xiaoming Yu, Junming Chen, Thomas H Li, and Ge Li. 2020. Deep image spatial transformation for person image generation. In *CVPR*.
- [414] Danilo Jimenez Rezende and Shakir Mohamed. 2015. Variational inference with normalizing flows. In *ICML*.
- [415] Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. 2014. Stochastic backpropagation and approximate inference in deep generative models. In *ICML*.
- [416] Elad Richardson, Yuval Alaluf, Or Patashnik, Yotam Nitzan, Yaniv Azar, Stav Shapiro, and Daniel Cohen-Or. 2020. Encoding in Style: a StyleGAN Encoder for Image-to-Image Translation. *arXiv preprint arXiv:2008.00951* (2020).
- [417] JH Rick Chang, Chun-Liang Li, Barnabas Poczos, BVK Vijaya Kumar, and Aswin C Sankaranarayanan. 2017. One Network to Solve Them All—Solving Linear Inverse Problems Using Deep Projection Models. In *ICCV*. 5888–5897.
- [418] Salah Rifai, Pascal Vincent, Xavier Muller, Xavier Glorot, and Yoshua Bengio. 2011. Contractive auto-encoders: Explicit invariance during feature extraction. In *ICML*.
- [419] Robin Rombach, Patrick Esser, and Bjorn Ommer. 2020. Network-to-Network Translation with Conditional Invertible Neural Networks. In *NeurIPS*, Vol. 33.
- [420] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*.
- [421] Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, and Matthias Nießner. 2018. FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces. *arXiv preprint arXiv:1803.09179* (2018).
- [422] Kevin Roth, Aurelien Lucchi, Sebastian Nowozin, and Thomas Hofmann. 2017. Stabilizing training of generative adversarial networks through regularization. In *NeurIPS*.
- [423] Stefan Roth and Michael J Black. 2005. Fields of experts: A framework for learning image priors. In *CVPR*. 860–867.
- [424] Sam T Roweis and Lawrence K Saul. 2000. Nonlinear dimensionality reduction by locally linear embedding. *science* 290, 5500 (2000), 2323–2326.
- [425] Manuel Ruder, Alexey Dosovitskiy, and Thomas Brox. 2016. Artistic style transfer for videos. In *German Conference on Pattern Recognition*.
- [426] Leonid I Rudin, Stanley Osher, and Emad Fatemi. 1992. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena* (1992).
- [427] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. 2015. ImageNet Large Scale Visual Recognition Challenge. *IJCV* 115, 3 (2015), 211–252.

- [428] Michele A Saad, Alan C Bovik, and Christophe Charrier. 2010. A DCT statistics-based blind image quality index. *IEEE Signal Processing Letters* 17, 6 (2010), 583–586.
- [429] Michele A Saad, Alan C Bovik, and Christophe Charrier. 2012. Blind image quality assessment: A natural scene statistics approach in the DCT domain. *TIP* 21, 8 (2012), 3339–3352.
- [430] Kuniaki Saito, Kate Saenko, and Ming-Yu Liu. 2020. COCO-FUNIT: Few-Shot Unsupervised Image Translation with a Content Conditioned Style Encoder. In *ECCV*.
- [431] Masaki Saito, Eiichi Matsumoto, and Shunta Saito. 2017. Temporal generative adversarial nets with singular value clipping. In *ICCV*.
- [432] Shunsuke Saito, Zeng Huang, Ryota Natsume, Shigeo Morishima, Angjoo Kanazawa, and Hao Li. 2019. PIFu: Pixel-aligned implicit function for high-resolution clothed human digitization. In *ICCV*.
- [433] Shunsuke Saito, Tomas Simon, Jason Saragih, and Hanbyul Joo. 2020. PIFuHD: Multi-Level Pixel-Aligned Implicit Function for High-Resolution 3D Human Digitization. In *CVPR*, 84–93.
- [434] Mehdi SM Sajjadi, Bernhard Schölkopf, and Michael Hirsch. 2017. Enhancenet: Single image super-resolution through automated texture synthesis. In *ICCV*.
- [435] Ruslan Salakhutdinov and Geoffrey Hinton. 2009. Deep boltzmann machines. In *AISTATS*. 448–455.
- [436] Ruslan Salakhutdinov and Hugo Larochelle. 2010. Efficient learning of deep Boltzmann machines. In *AISTATS*. 693–700.
- [437] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. 2016. Improved techniques for training GANs. In *NeurIPS*.
- [438] Tim Salimans, Andrej Karpathy, Xi Chen, and Diederik P Kingma. 2017. Pixel-CNN++: Improving the PixelCNN with discretized logistic mixture likelihood and other modifications. In *ICLR*.
- [439] Pouya Samangouei, Maya Kabkab, and Rama Chellappa. 2018. Defense-GAN: Protecting Classifiers Against Adversarial Attacks Using Generative Models. In *ICLR*.
- [440] Patson Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. 2016. Scribbler: Controlling deep image synthesis with sketch and color. *arXiv preprint arXiv:1612.00835* (2016).
- [441] Katja Schwarz, Yiyi Liao, Michael Niemeyer, and Andreas Geiger. 2020. GRAF: Generative Radiance Fields for 3D-Aware Image Synthesis. *arXiv preprint arXiv:2007.02442* (2020).
- [442] Viraj Shah and Chinmay Hegde. 2018. Solving linear inverse problems using gan priors: An algorithm with provable guarantees. In *ICASSP*. 4609–4613.
- [443] Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli. 2019. Singan: Learning a generative model from a single natural image. In *ICCV*. 4570–4580.
- [444] Wenling Shang, Kihyuk Sohn, Diogo Almeida, and Honglak Lee. 2016. Understanding and improving convolutional neural networks via concatenated rectified linear units. In *ICML*. 2217–2225.
- [445] Eli Shechtman, Yaron Caspi, and Michal Irani. 2005. Space-time super-resolution. *TPAMI* (2005).
- [446] Hamid R Sheikh and Alan C Bovik. 2004. Image information and visual quality. In *ICASSP*.
- [447] Hamid R Sheikh, Alan C Bovik, and Gustavo De Veciana. 2005. An information fidelity criterion for image quality assessment using natural scene statistics. *TIP* 14 (2005).
- [448] Hamid R Sheikh, Muhammad F Sabir, and Alan C Bovik. 2006. A statistical evaluation of recent full reference image quality assessment algorithms. *TIP* 15 (2006).
- [449] Yujun Shen, Jinjin Gu, Xiaou Tang, and Bolei Zhou. 2020. Interpreting the latent space of GANs for semantic face editing. In *CVPR*.
- [450] Yujun Shen and Bolei Zhou. 2020. Closed-Form Factorization of Latent Semantics in GANs. *arXiv preprint arXiv:2007.06600* (2020).
- [451] Ziyi Shen, Wei-Sheng Lai, Tingfa Xu, Jan Kautz, and Ming-Hsuan Yang. 2018. Deep semantic face deblurring. In *CVPR*.
- [452] Wenzhe Shi, Jose Caballero, Ferenc Huszar, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *CVPR*.
- [453] Ken Shoemake. 1985. Animating rotation with quaternion curves. In *SIGGRAPH*. 245–254.
- [454] Ashish Shrivastava, Tomas Pfister, Oncel Tuzel, Josh Susskind, Wenda Wang, and Russ Webb. 2017. Learning from simulated and unsupervised images through adversarial training. In *CVPR*.
- [455] Lin Shuyu and Clark Ronald. 2020. LaDDer: Latent Data Distribution Modelling with a Generative Prior. In *BMVC*.
- [456] Aliaksandra Shysheya, Egor Zakharov, Kara-Ali Aliev, Renat Bashirov, Egor Burkov, Karim Isakov, Aleksei Ivakhnenko, Yuriy Malkov, Igor Pasechnik, Dmitry Ulyanov, et al. 2019. Textured neural avatars. In *CVPR*.
- [457] Aliaksandr Siarohin, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, and Nicu Sebe. 2019. First order motion model for image animation. In *NeurIPS*.
- [458] Aliaksandr Siarohin, Stéphane Lathuilière, Sergey Tulyakov, Elisa Ricci, and Nicu Sebe. 2019. Animating Arbitrary Objects via Deep Motion Transfer. In *CVPR*.
- [459] Aliaksandr Siarohin, Enver Sangineto, Stéphane Lathuilière, and Nicu Sebe. 2018. Deformable GANs for pose-based human image generation. In *CVPR*.
- [460] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. 2013. Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv preprint arXiv:1312.6034* (2013).
- [461] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).
- [462] Vincent Sitzmann, Justus Thies, Felix Heide, Matthias Nießner, Gordon Wetzstein, and Michael Zollhofer. 2019. DeepVoxels: Learning persistent 3d feature embeddings. In *CVPR*.
- [463] Brandon M Smith, Li Zhang, Jonathan Brandt, Zhe Lin, and Jianchao Yang. 2013. Exemplar-based face parsing. In *CVPR*.
- [464] Casper Kaae Sønderby, Tapani Raiko, Lars Maaløe, Søren Kaae Sønderby, and Ole Winther. 2016. Ladder variational autoencoders. In *NeurIPS*.
- [465] Sijie Song, Wei Zhang, Jiaying Liu, and Tao Mei. 2019. Unsupervised person image generation with semantic parsing transformation. In *CVPR*.
- [466] Yang Song, Jingwen Zhu, Dawei Li, Andy Wang, and Hairong Qi. 2019. Talking Face Generation by Conditional Recurrent Adversarial Network. In *IJCAI*.
- [467] Pratul P Srinivasan, Richard Tucker, Jonathan T Barron, Ravi Ramamoorthi, Ren Ng, and Noah Snavely. 2019. Pushing the boundaries of view extrapolation with multiplane images. In *CVPR*.
- [468] Pratul P Srinivasan, Tongzhou Wang, Ashwin Sreelal, Ravi Ramamoorthi, and Ren Ng. 2017. Learning to synthesize a 4D RGBD light field from a single image. In *ICCV*.
- [469] Nitish Srivastava, Elman Mansimov, and Ruslan Salakhutdinov. 2015. Unsupervised learning of video representations using LSTMs. In *ICML*.
- [470] Nitish Srivastava and Russ R Salakhutdinov. 2012. Multimodal learning with deep boltzmann machines. In *NeurIPS*. 2222–2230.
- [471] Hang Su, Varun Jampani, Deqing Sun, Orazio Gallo, Erik Learned-Miller, and Jan Kautz. 2019. Pixel-Adaptive Convolutional Neural Networks. In *CVPR*.
- [472] Tiancheng Sun, Jonathan T Barron, Yun-Ta Tsai, Zexiang Xu, Xueming Yu, Graham Fyfe, Christoph Rhemann, Jay Busch, Paul E Debevec, and Ravi Ramamoorthi. 2019. Single image portrait relighting. *ACM Trans. Graph.* 38, 4 (2019), 79–1.
- [473] Kalyan Sunkavalli, Micah K Johnson, Wojciech Matusik, and Hanspeter Pfister. 2010. Multi-scale image harmonization. *TOG* 29, 4 (2010), 1–10.
- [474] Ilya Sutskever, Geoffrey E Hinton, and Graham W Taylor. 2009. The recurrent temporal restricted boltzmann machine. In *NeurIPS*.
- [475] Supasorn Suwajanakorn, Steven M Seitz, and Ira Kemelmacher-Shlizerman. 2017. Synthesizing Obama: learning lip sync from audio. *TOG* (2017).
- [476] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the inception architecture for computer vision. In *CVPR*. 2818–2826.
- [477] Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. 2017. MemNet: A persistent memory network for image restoration. In *ICCV*.
- [478] Yaniv Taigman, Adam Polyak, and Lior Wolf. 2017. Unsupervised cross-domain image generation. In *ICLR*.
- [479] Yaniv Taigman, Adam Polyak, and Lior Wolf. 2017. Unsupervised Cross-Domain Image Generation. In *ICLR*.
- [480] Hossein Talebi and Peyman Milanfar. 2018. Nima: Neural image assessment. *TIP* 27 (2018).
- [481] Hao Tang, Xiaojuan Qi, Dan Xu, Philip HS Torr, and Nicu Sebe. 2020. Edge Guided GANs with Semantic Preserving for Semantic Image Synthesis. *arXiv preprint arXiv:2003.13898* (2020).
- [482] Xu Tao, Zhang Pengchuan, Huang Qiuyuan, Zhang Han, Gan Zhe, Xiaolei Huang, and He Xiaodong. 2018. AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks. In *CVPR*.
- [483] Ayush Tewari, Mohamed Elgharib, Gaurav Bharaj, Florian Bernard, Hans-Peter Seidel, Patrick Pérez, Michael Zöllhofer, and Christian Theobalt. 2020. StyleRig: Rigging StyleGAN for 3D Control over Portrait Images. In *CVPR*.
- [484] Ayush Tewari, Mohamed Elgharib, Mallikarjun BR, Florian Bernard, Hans-Peter Seidel, Patrick Pérez, Michael Zöllhofer, and Christian Theobalt. 2020. PIE: Portrait Image Embedding for Semantic Control. *TOG* 39, 6 (December 2020).
- [485] A. Tewari, O. Fried, J. Thies, V. Sitzmann, S. Lombardi, K. Sunkavalli, R. Martin-Brualla, T. Simon, J. Saragih, M. Nießner, R. Pandey, S. Fanello, G. Wetzstein, J.-Y. Zhu, C. Theobalt, M. Agrabala, E. Shechtman, D. B Goldman, and M. Zollhofer. 2020. State of the Art on Neural Rendering. *Computer Graphics Forum (EG STAR)* (2020).
- [486] Justus Thies, Michael Zollhofer, and Matthias Nießner. 2019. Deferred neural rendering: Image synthesis using neural textures. *TOG* (2019).
- [487] Justus Thies, Michael Zollhofer, Matthias Nießner, Levi Valgaerts, Marc Stamminger, and Christian Theobalt. 2015. Real-time expression transfer for facial reenactment. *TOG* (2015).

- [488] Justus Thies, Michael Zollhofer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. 2016. Face2Face: Real-time face capture and reenactment of rgb videos. In *CVPR*.
- [489] Justus Thies, Michael Zollhöfer, Christian Theobalt, Marc Stamminger, and Matthias Niessner. 2018. HeadOn: Real-time reenactment of human portrait videos. *TOG* (2018).
- [490] Justus Thies, Michael Zollhöfer, Christian Theobalt, Marc Stamminger, and Matthias Nießner. 2019. Image-guided neural object rendering. In *ICLR*.
- [491] Ilya Tolstikhin, Olivier Bousquet, Sylvain Gelly, and Bernhard Schölkopf. 2018. Wasserstein Auto-Encoders. In *ICLR*. <https://openreview.net/forum?id=HkL7n1-0b>
- [492] Carlo Tomasi and Takeo Kanade. 1992. Shape and motion from image streams under orthography: a factorization method. *IJCV* (1992).
- [493] Tong Tong, Gen Li, Xiejie Liu, and Qinquan Gao. 2017. Image super-resolution using dense skip connections. In *ICCV*.
- [494] Hugo Touvron, Matthijs Douze, Matthieu Cord, and Hervé Jégou. 2020. Powers of layers for image-to-image translation. *arXiv preprint arXiv:2008.05763* (2020).
- [495] Yi-Hsuan Tsai, Xiaohui Shen, Zhe Lin, Kalyan Sunkavalli, Xin Lu, and Ming-Hsuan Yang. 2017. Deep image harmonization. In *CVPR*. 3789–3797.
- [496] Michael Tschannen, Eirikur Agustsson, and Mario Lucic. 2018. Deep generative models for distribution-preserving lossy compression. In *NeurIPS*.
- [497] Richard Tucker and Noah Snavely. 2020. Single-View View Synthesis with Multiplane Images. In *CVPR*. 551–560.
- [498] Sergey Tulyakov, Ming-Yu Liu, Xiaodong Yang, and Jan Kautz. 2018. MoCoGAN: Decomposing motion and content for video generation. In *CVPR*.
- [499] Hsiao-Yu Fish Tung, Adam W Harley, William Seto, and Katerina Fragkiadaki. 2017. Adversarial Inverse Graphics Networks: Learning 2D-to-3D Lifting and Image-to-Image Translation from Unpaired Supervision. In *ICCV*.
- [500] Matthew Turk and Alex Pentland. 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3, 1 (1991), 71–86.
- [501] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. 2016. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022* (2016).
- [502] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. 2018. Deep image prior. In *CVPR*. 9446–9454.
- [503] Aaron Van den Oord, Nal Kalchbrenner, Lasse Espeholt, Oriol Vinyals, Alex Graves, et al. 2016. Conditional image generation with PixelCNN decoders. In *NeurIPS*.
- [504] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing Data using t-SNE. *JMLR* 9 (2008), 2579–2605.
- [505] Tycho FA van der Ouderaa and Daniel E Worrall. 2019. Reversible GANs for memory-efficient image-to-image translation. In *CVPR*.
- [506] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *NeurIPS*. 5998–6008.
- [507] Sundar Vedula, Simon Baker, Peter Rander, Robert Collins, and Takeo Kanade. 1999. Three-dimensional scene flow. In *CVPR*.
- [508] Ruben Villegas, Jimei Yang, Seunghoon Hong, Xunyu Lin, and Honglak Lee. 2017. Decomposing motion and content for natural video sequence prediction. In *ICLR*.
- [509] Ruben Villegas, Jimei Yang, Yuliang Zou, Sungryull Sohn, Xunyu Lin, and Honglak Lee. 2017. Learning to generate long-term future via hierarchical prediction. In *ICML*.
- [510] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. 2008. Extracting and composing robust features with denoising autoencoders. In *ICML*. 1096–1103.
- [511] Daniel Vlasic, Matthew Brand, Hanspeter Pfister, and Jovan Popovic. 2005. Face transfer with multilinear models. *TOG* (2005).
- [512] Carl Vondrick, Hamed Pirsiavash, and Antonio Torralba. 2016. Generating videos with scene dynamics. In *NeurIPS*.
- [513] Carl Vondrick and Antonio Torralba. 2017. Generating the future with adversarial transformers. In *CVPR*.
- [514] Konstantinos Vougioukas, Stavros Petridis, and Maja Pantic. 2019. Realistic speech-driven facial animation with GANs. *IJCV* (2019).
- [515] Andrey Vovnov and Artem Babenko. 2020. Unsupervised Discovery of Interpretable Directions in the GAN Latent Space. *ICML* (2020).
- [516] Jacob Walker, Carl Doersch, Abhinav Gupta, and Martial Hebert. 2016. An uncertain future: Forecasting from static images using variational autoencoders. In *ECCV*.
- [517] Jacob Walker, Kenneth Marino, Abhinav Gupta, and Martial Hebert. 2017. The pose knows: Video forecasting by generating pose futures. In *ICCV*.
- [518] Kevin Wampller and Zoran Popović. 2009. Optimal Gait and Form for Animal Locomotion. *TOG* (2009).
- [519] Jiashun Wang, Chao Wen, Yanwei Fu, Haitao Lin, Tianyun Zou, Xiangyang Xue, and Yinda Zhang. 2020. Neural Pose Transfer by Spatially Adaptive Instance Normalization. In *CVPR*.
- [520] Miao Wang, Guo-Ye Yang, Ruilong Li, Run-Ze Liang, Song-Hai Zhang, Peter. M Hall, and Shi-Min Hu. 2019. Example-Guided Style-Consistent Image Synthesis from Semantic Labeling. In *CVPR*.
- [521] Sheng-Yu Wang, Oliver Wang, Richard Zhang, Andrew Owens, and Alexei A Efros. 2020. CNN-generated images are surprisingly easy to spot... for now. In *CVPR*.
- [522] Ting-Chun Wang, Ming-Yu Liu, Andrew Tao, Guilin Liu, Jan Kautz, and Bryan Catanzaro. 2019. Few-shot video-to-video synthesis. In *NeurIPS*.
- [523] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Guilin Liu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. 2018. Video-to-video Synthesis. In *NeurIPS*.
- [524] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. 2018. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In *CVPR*.
- [525] Xiaolong Wang and Abhinav Gupta. 2016. Generative image modeling using style and structure adversarial networks. In *ECCV*.
- [526] Xiaolong Wang, Abhinav Shrivastava, and Abhinav Gupta. 2017. A-Fast-RCNN: Hard positive generation via adversary for object detection. In *CVPR*.
- [527] Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. 2018. Recovering realistic texture in image super-resolution by deep spatial feature transform. In *CVPR*.
- [528] Xintao Wang, Ke Yu, Chao Dong, Xiaoou Tang, and Chen Change Loy. 2019. Deep network interpolation for continuous imagery effect transition. In *CVPR*. 1692–1701.
- [529] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. 2018. ESRGAN: Enhanced super-resolution generative adversarial networks. In *ECCV*.
- [530] Yaxing Wang, Salman Khan, Abel Gonzalez-Garcia, Joost van de Weijer, and Fahad Shahbaz Khan. 2020. Semi-supervised Learning for Few-shot Image-to-Image Translation. In *CVPR*.
- [531] Yi Wang, Xin Tao, Xiaojuan Qi, Xiaoyong Shen, and Jiaya Jia. 2018. Image inpainting via generative multi-column convolutional neural networks. In *NeurIPS*.
- [532] Zhou Wang and Alan C Bovik. 2009. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine* 26 (2009).
- [533] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *TIP* 13 (2004).
- [534] Zhaowen Wang, Ding Liu, Jianchao Yang, Wei Han, and Thomas Huang. 2015. Deep networks for image super-resolution with sparse prior. In *ICCV*.
- [535] Zhengwei Wang, Qi She, and Tomas E Ward. 2019. Generative adversarial networks in computer vision: A survey and taxonomy. *arXiv preprint arXiv:1906.01529* (2019).
- [536] Zhou Wang, Eero Simoncelli, Alan Bovik, et al. 2003. Multi-scale structural similarity for image quality assessment. In *Asilomar Conference on Signals, Systems, and Computers*.
- [537] Chung-Yi Weng, Brian Curless, and Ira Kemelmacher-Shlizerman. 2019. Photo wake-up: 3d character animation from a single photo. In *CVPR*.
- [538] Yonatan Wexler, Eli Shechtman, and Michal Irani. 2004. Space-time video completion. In *CVPR*.
- [539] Yonatan Wexler, Eli Shechtman, and Michal Irani. 2007. Space-time completion of video. *TPAMI* 29, 3 (2007).
- [540] Tom White. 2016. Sampling generative networks: Notes on a few effective techniques. *arXiv preprint arXiv:1609.04468* (2016).
- [541] Olivia Wiles, Georgia Gkioxari, Richard Szeliski, and Justin Johnson. 2020. SynSin: End-to-end View Synthesis from a Single Image. In *CVPR*.
- [542] Olivia Wiles, A Sophia Koepke, and Andrew Zisserman. 2018. X2Face: A network for controlling face generation using images, audio, and pose codes. In *ECCV*.
- [543] Svante Wold, Kim Esbensen, and Paul Geladi. 1987. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 2, 1-3 (1987), 37–52.
- [544] Erroll Wood, Tadas Baltrušaitis, Louis-Philippe Morency, Peter Robinson, and Andreas Bulling. 2018. GazeDirector: Fully Articulated Eye Gaze Redirection in Video. *Comput. Graph. Forum* 37, 2 (2018), 217–225.
- [545] Jialin Wu, Dai Li, Yu Yang, Chandrajit Bajaj, and Xiangyang Ji. 2018. Dynamic sampling convolutional neural networks. In *ECCV*.
- [546] Shangzhe Wu, Christian Rupprecht, and Andrea Vedaldi. 2020. Unsupervised Learning of Probably Symmetric Deformable 3D Objects from Images in the Wild. In *CVPR*. 1–10.
- [547] Tianyi Wu, Sheng Tang, Rui Zhang, and Yongdong Zhang. 2018. CGNet: A Lightweight Context Guided Network for Semantic Segmentation. *arXiv preprint arXiv:1811.08201* (2018).
- [548] Wayne Wu, Yunxuan Zhang, Cheng Li, Chen Qian, and Chen Change Loy. 2018. ReenactGAN: Learning to reenact faces via boundary transfer. In *ECCV*.
- [549] Yuxin Wu and Kaiming He. 2018. Group normalization. In *ECCV*.
- [550] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. 2019. Detectron2. <https://github.com/facebookresearch/detectron2>.
- [551] Zongze Wu, Dani Lischinski, and Eli Shechtman. 2020. StyleSpace Analysis: Disentangled Controls for StyleGAN Image Generation. *arXiv preprint arXiv:2011.12799* (2020).

- [552] Zifeng Wu, Chunhua Shen, and Anton Van Den Hengel. 2019. Wider or deeper: Revisiting the resnet model for visual recognition. *Pattern Recognition* (2019).
- [553] Weihao Xia, Zhanglin Cheng, Yujiu Yang, and Jing-Hao Xue. 2019. Cooperative Semantic Segmentation and Image Restoration in Adverse Environmental Conditions. *arXiv preprint arXiv:1911.00679* (2019).
- [554] Weihao Xia, Zhanglin Cheng, Yujiu Yang, Jing-Hao Xue, and Wensen Feng. 2020. Controllable Continuous Gaze Redirection. In *ACM MM*.
- [555] Weihao Xia, Yujiu Yang, and Jing-Hao Xue. 2019. Cali-Sketch: Stroke Calibration and Completion for High-Quality Face Image Generation from Poorly-Drawn Sketches. *arXiv preprint arXiv:1911.00426* (2019).
- [556] Weihao Xia, Yujiu Yang, and Jing-Hao Xue. 2020. Unsupervised multi-domain multimodal image-to-image translation with explicit domain-constrained disentanglement. *Neural Networks* (2020).
- [557] Weihao Xia, Yujiu Yang, Jing-Hao Xue, and Baoyuan Wu. 2020. TediGAN: Text-Guided Diverse Image Generation and Manipulation. *arXiv preprint arXiv:2012.03308* (2020).
- [558] Weihao Xia, Yujiu Yang, Jing-Hao Xue, and Jing Xiao. 2020. Domain Fingerprints for No-reference Image Quality Assessment. *IEEE Transactions on Circuits and Systems for Video Technology* (2020).
- [559] Yuanbo Xiangli*, Yubin Deng*, Bo Dai*, Chen Change Loy, and Dahua Lin. 2020. Real or Not Real, that is the Question. In *ICLR*.
- [560] Han Xiao, Kashif Rasul, and Roland Vollgraf. 2017. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747* (2017).
- [561] Taihong Xiao, Jiapeng Hong, and Jinwen Ma. 2018. ELEGANT: Exchanging Latent Encodings with GAN for Transferring Multiple Face Attributes. In *ECCV*. 172–187.
- [562] Junyuan Xie, Ross Girshick, and Ali Farhadi. 2016. Deep3D: Fully automatic 2D-to-3D video conversion with deep convolutional neural networks. In *ECCV*.
- [563] Qizhe Xie, Zihang Dai, Yulun Du, Eduard Hovy, and Graham Neubig. 2017. Controllable invariance through adversarial feature learning. In *NeurIPS*. 585–596.
- [564] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. 2017. Aggregated residual transformations for deep neural networks. In *CVPR*.
- [565] Saining Xie and Zhuowen Tu. 2015. Holistically-nested edge detection. In *ICCV*.
- [566] Wei Xiong, Jiahui Yu, Zhe Lin, Jimei Yang, Xin Lu, Connelly Barnes, and Jiebo Luo. 2019. Foreground-aware image inpainting. In *CVPR*.
- [567] Jingwei Xu, Huazhe Xu, Bingbing Ni, Xiaokang Yang, Xiaolong Wang, and Trevor Darrell. 2020. Hierarchical Style-based Networks for Motion Synthesis. In *ECCV*.
- [568] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhutdinov, Richard Zemel, and Yoshua Bengio. 2015. Show, attend and tell: Neural image caption generation with visual attention. *arXiv preprint arXiv:1502.03044* (2015).
- [569] Tao Xu, Pengchuan Zhang, Qiuyuan Huang, Han Zhang, Zhe Gan, Xiaolei Huang, and Xiaodong He. 2018. AttnGAN: Fine-grained text to image generation with attentional generative adversarial networks. In *CVPR*.
- [570] Xiangyu Xu, Deqing Sun, Jinshan Pan, Yujin Zhang, Hanspeter Pfister, and Ming-Hsuan Yang. 2017. Learning to super-resolve blurry face and text images. In *ICCV*.
- [571] Yinghao Xu, Yujun Shen, Jiapeng Zhu, Ceyuan Yang, and Bolei Zhou. 2020. Generative Hierarchical Features from Synthesizing Images. *arXiv preprint arXiv:2007.10379* (2020).
- [572] Tianfan Xue, Jiajun Wu, Katherine Bouman, and Bill Freeman. 2016. Visual dynamics: Probabilistic future frame synthesis via cross convolutional networks. In *NeurIPS*.
- [573] Chao Yang, Taehwan Kim, Ruizhe Wang, Hao Peng, and C-C Jay Kuo. 2018. ESTHER: Extremely Simple Image Translation Through Self-Regularization.. In *BMVC*.
- [574] Chao Yang, Taehwan Kim, Ruizhe Wang, Hao Peng, and C-C Jay Kuo. 2019. Show, Attend, and Translate: Unsupervised Image Translation With Self-Regularization and Attention. *TIP* (2019).
- [575] Ceyuan Yang, Yujun Shen, and Bolei Zhou. 2019. Semantic hierarchy emerges in deep generative representations for scene synthesis. *arXiv preprint arXiv:1911.09267* (2019).
- [576] Ceyuan Yang, Zhe Wang, Xinge Zhu, Chen Huang, Jianping Shi, and Dahua Lin. 2018. Pose Guided Human Video Generation. In *ECCV*.
- [577] Dingdong Yang, Seunghoon Hong, Yunseok Jang, Tianchen Zhao, and Honglak Lee. 2019. Diversity-sensitive conditional generative adversarial networks. In *ICLR*.
- [578] Shan Yang, Tanya Ambert, Zherong Pan, Ke Wang, Licheng Yu, Tamara Berg, and Ming C Lin. 2016. Detailed garment recovery from a single-view image. *arXiv preprint arXiv:1608.01250* (2016).
- [579] Chun-Han Yao, Chia-Yang Chang, and Shao-Yi Chien. 2017. Occlusion-aware video temporal consistency. In *MM*.
- [580] Joost van de Weijer Yaxing Wang, Lu Yu. 2020. DeepI2I: Enabling Deep Hierarchical Image-to-Image Translation by Transferring from GANs. In *NeurIPS*.
- [581] Peng Ye, Jayant Kumar, Le Kang, and David Doermann. 2012. Unsupervised feature learning framework for no-reference image quality assessment. In *CVPR*. IEEE, 1098–1105.
- [582] Peng Ye, Jayant Kumar, Le Kang, and David Doermann. 2013. Real-time no-reference image quality assessment based on filter learning. In *CVPR*. 987–994.
- [583] Raymond A Yeh, Chen Chen, Teck Yian Lim, Alexander G Schwing, Mark Hasegawa-Johnson, and Minh N Do. 2017. Semantic image inpainting with deep generative models. In *CVPR*. 5485–5493.
- [584] Ran Yi, Yong-Jin Liu, Yu-Kun Lai, and Paul L Rosin. 2019. ApdrawingGAN: Generating artistic portrait drawings from face photos with hierarchical GANs. In *CVPR*. 10743–10752.
- [585] Zili Yi, Hao Zhang, Ping Tan Gong, et al. 2017. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation. In *ICCV*.
- [586] Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs, and Hod Lipson. 2015. Understanding Neural Networks Through Deep Visualization. *arXiv preprint arXiv:1506.06579* (2015).
- [587] Nitzan Yotam, Bermanno Amit, Li Yangyan, and Daniel Cohen-Or. 2020. Disentangling in Latent Space by Harnessing a Pretrained Generator. *arXiv preprint arXiv:2005.07728* (2020).
- [588] Aron Yu and Kristen Grauman. 2014. Fine-Grained Visual Comparisons with Local Learning. In *CVPR*. 192–199.
- [589] Aron Yu and Kristen Grauman. 2017. Semantic Jitter: Dense Supervision for Visual Comparisons via Synthetic Images. In *ICCV*. 5571–5580.
- [590] Fisher Yu, Yinda Zhang, Shuran Song, Ari Seff, and Jianxiong Xiao. 2015. LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop. *arXiv preprint arXiv:1506.03365* (2015).
- [591] Jiahui Yu, Yuchen Fan, Jianchao Yang, Ning Xu, Zhaowen Wang, Xinchao Wang, and Thomas Huang. 2019. Wide activation for efficient and accurate image super-resolution. In *BMVC*.
- [592] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang. 2018. Generative image inpainting with contextual attention. In *CVPR*.
- [593] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang. 2019. Free-form image inpainting with gated convolution. In *ICCV*.
- [594] Viazovetskiy Yuri and Kashin Vladimir, Ivashkin ang Evgeny. 2020. StyleGAN2 Distillation for Feed-forward Image Manipulation. In *ECCV*.
- [595] Sergey Zagoruyko and Nikos Komodakis. 2016. Wide residual networks. *arXiv preprint arXiv:1605.07146* (2016).
- [596] Egor Zakharov, Aliaksandra Shysheya, Egor Burkov, and Victor Lempitsky. 2019. Few-shot adversarial learning of realistic neural talking head models. In *ICCV*.
- [597] Mihai Zanfir, Alin-Ionut Popa, Andrei Zanfir, and Cristian Sminchisescu. 2018. Human appearance transfer. In *CVPR*.
- [598] Matthew D. Zeiler and Rob Fergus. 2014. Visualizing and Understanding Convolutional Networks. *Lecture Notes in Computer Science* (2014), 818–833.
- [599] Yanhong Zeng, Jianlong Fu, Hongyang Chao, and Baining Guo. 2019. Learning pyramid-context encoder network for high-quality image inpainting. In *CVPR*.
- [600] Wujing Zhan, Xinqi Ou, Yunyi Yang, and Long Chen. 2019. DSNet: Joint Learning for Scene Segmentation and Disparity Estimation. In *ICRA*. 2946–2952.
- [601] Chris Zhang, Mengye Ren, and Raquel Urtasun. 2019. Graph hypernetworks for neural architecture search. In *ICLR*.
- [602] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. 2019. Self-Attention Generative Adversarial Networks. In *ICML*. 7354–7363.
- [603] Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaolei Huang, Xiaogang Wang, and Dimitris Metaxas. 2017. StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. In *ICCV*.
- [604] Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaolei Huang, Xiaogang Wang, and Dimitris Metaxas. 2017. StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. In *CVPR*.
- [605] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. 2017. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *TIP* (2017).
- [606] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. 2017. Learning deep CNN denoiser prior for image restoration. In *CVPR*. 3929–3938.
- [607] Pan Zhang, Bo Zhang, Dong Chen, Lu Yuan, and Fang Wen. 2020. Cross-domain Correspondence Learning for Exemplar-based Image Translation. In *CVPR*.
- [608] Quanshi Zhang, Ying Nian Wu, and Song-Chun Zhu. 2018. Interpretable Convolutional Neural Networks. In *CVPR*.
- [609] Richard Zhang, Phillip Isola, and Alexei A Efros. 2016. Colorful image colorization. In *ECCV*.
- [610] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*. 586–595.
- [611] Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S Lin, Tianhe Yu, and Alexei A Efros. 2017. Real-time user-guided image colorization with learned deep priors. In *SIGGRAPH*.
- [612] Xu Zhang, Svebor Karaman, and Shih-Fu Chang. 2019. Detecting and simulating artifacts in GAN fake images. In *WIFD*.

- [613] Xuaner Cecilia Zhang, Yun-Ta Tsai, Rohit Pandey, Xiuming Zhang, Ren Ng, David E Jacobs, et al. 2020. Portrait Shadow Manipulation. *arXiv preprint arXiv:2005.08925* (2020).
- [614] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. 2018. Image super-resolution using very deep residual channel attention networks. In *ECCV*.
- [615] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. 2018. Residual dense network for image super-resolution. In *CVPR*.
- [616] Yuanhan Zhang, Zhenfei Yin, Yidong Li, Guojun Yin, Junjie Yan, Jing Shao, and Ziwei Liu. 2020. CelebA-Spoof: Large-Scale Face Anti-Spoofing Dataset with Rich Annotations. In *ECCV*.
- [617] Zhifei Zhang, Yang Song, and Hairong Qi. 2017. Age Progression/Regression by Conditional Adversarial Autoencoder. In *CVPR*.
- [618] Zhenyue Zhang and Hongyuan Zha. 2004. Principal manifolds and nonlinear dimensionality reduction via tangent space alignment. *SIAM journal on scientific computing* 26, 1 (2004), 313–338.
- [619] Bo Zhao, Lili Meng, Weidong Yin, and Leonid Sigal. 2019. Image generation from layout. In *CVPR*.
- [620] Bo Zhao, Xiao Wu, Zhi-Qi Cheng, Hao Liu, Zequn Jie, and Jiashi Feng. 2018. Multi-view image generation from a single-view. In *MM*.
- [621] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. 2017. Pyramid scene parsing network. In *CVPR*.
- [622] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. 2017. Pyramid Scene Parsing Network. In *CVPR*.
- [623] Junbo Zhao, Michael Mathieu, and Yann LeCun. 2017. Energy-based generative adversarial network. In *ICLR*.
- [624] Chuanxia Zheng, Tat-Jen Cham, and Jianfei Cai. 2019. Pluralistic image completion. In *CVPR*.
- [625] Haitian Zheng, Haofu Liao, Lele Chen, Wei Xiong, Tianlang Chen, and Jiebo Luo. 2020. Example-Guided Image Synthesis across Arbitrary Scenes using Masked Spatial-Channel Attention and Self-Supervision. *arXiv preprint arXiv:2004.10024* (2020).
- [626] Zerong Zheng, Tao Yu, Yebin Liu, and Qionghai Dai. 2020. PaMIR: Parametric Model-Conditioned Implicit Representation for Image-based Human Reconstruction. *arXiv preprint arXiv:2007.03858* (2020).
- [627] Zhao-Heng Zheng, Hao-Tian Zhang, Fang-Lue Zhang, and Tai-Jiang Mu. 2017. Image-based clothes changing system. *Computational Visual Media* (2017).
- [628] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. 2016. Learning Deep Features for Discriminative Localization. In *CVPR*.
- [629] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. 2016. Learning Deep Features for Discriminative Localization. In *CVPR*. IEEE Computer Society, 2921–2929.
- [630] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. 2017. Places: A 10 million image database for scene recognition. *TPAMI* 40 (2017), 1452–1464.
- [631] Bolei Zhou, Agata Lapedriza, Jianxiong Xiao, Antonio Torralba, and Aude Oliva. 2014. Learning Deep Features for Scene Recognition using Places Database. In *NeurIPS*. 487–495.
- [632] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. 2017. Scene Parsing through ADE20K Dataset. In *CVPR*.
- [633] Hao Zhou, Sunil Hadap, Kalyan Sunkavalli, and David W Jacobs. 2019. Deep single-image portrait relighting. In *ICCV*. 7194–7202.
- [634] Hang Zhou, Yu Liu, Ziwei Liu, Ping Luo, and Xiaogang Wang. 2019. Talking face generation by adversarially disentangled audio-visual representation. In *AAAI*.
- [635] Tinghui Zhou, Richard Tucker, John Flynn, Graham Fyfe, and Noah Snavely. 2018. Stereo Magnification: Learning View Synthesis using Multiplane Images. In *SIGGRAPH*.
- [636] Yipin Zhou, Zhaowen Wang, Chen Fang, Trung Bui, and Tamara L Berg. 2019. Dance Dance Generation: Motion Transfer for Internet Videos. *arXiv preprint arXiv:1904.00129* (2019).
- [637] Jiapeng Zhu, Yujun Shen, Deli Zhao, and Bolei Zhou. 2020. In-domain gan inversion for real image editing. In *ECCV*.
- [638] Jiapeng Zhu, Deli Zhao, Bo Zhang, and Bolei Zhou. 2019. Disentangled Inference for GANs with Latently Invertible Autoencoder. *arXiv preprint arXiv:1906.08090* (2019).
- [639] Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alexei A Efros. 2016. Generative visual manipulation on the natural image manifold. In *ECCV*.
- [640] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*.
- [641] Jun-Yan Zhu, Richard Zhang, Deepak Pathak, Trevor Darrell, Alexei A Efros, Oliver Wang, and Eli Shechtman. 2017. Toward multimodal image-to-image translation. In *NeurIPS*.
- [642] Jun-Yan Zhu, Zhoutong Zhang, Chengkai Zhang, Jiajun Wu, Antonio Torralba, Josh Tenenbaum, and Bill Freeman. 2018. Visual Object Networks: Image generation with disentangled 3D representations. In *NeurIPS*.
- [643] Peihao Zhu, Rameen Abdal, Yipeng Qin, and Peter Wonka. 2020. SEAN: Image Synthesis with Semantic Region-Adaptive Normalization. In *CVPR*. 5104–5113.
- [644] Song Chun Zhu and David Mumford. 1997. Prior learning and Gibbs reaction-diffusion. *TPAMI* 19, 11 (1997), 1236–1250.
- [645] Yi Zhu, Karan Sapra, Fitsum A Reda, Kevin J Shih, Shawn Newsam, Andrew Tao, and Bryan Catanzaro. 2019. Improving semantic segmentation via video propagation and label relaxation. In *CVPR*.
- [646] Zhen Zhu, Tengting Huang, Baoguang Shi, Miao Yu, Bofei Wang, and Xiang Bai. 2019. Progressive pose attention transfer for person image generation. In *CVPR*.
- [647] Zhen Zhu, Zhiliang Xu, Ansheng You, and Xiang Bai. 2020. Semantically Multimodal Image Synthesis. In *CVPR*. 5467–5476.
- [648] C Lawrence Zitnick, Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski. 2004. High-quality video view interpolation using a layered representation. *TOG* (2004).