

Few Data Diversification in Training Generative Adversarial Networks

Lucas Fontes Buzuti

Department of Electrical Engineering

FEI University Center

São Bernardo do Campo-SP, Brazil

lucas.buzuti@outlook.com

Carlos Eduardo Thomaz

Department of Electrical Engineering

FEI University Center

São Bernardo do Campo-SP, Brazil

cet@fei.edu.br

Abstract—The first GANs have initially produced sharp images in relatively small resolution and with limited variations, and unstable training. Later works proposed new GAN models capable of generating sharp images in high resolution and with a high level of variation. However, these models use unlimited and highly diversified image sets. We discuss here the use of these models with real-world image sets, since they are composed of limited sample size sets.

Index Terms—generative adversarial net, limited data set, synthetic images, limited sample size sets

I. INTRODUÇÃO

Modelos generativos profundos estão possibilitando a geração artificial de dados. A Rede Adversária Generativa (Generative Adversarial Network, GAN) [1] é o modelo generativo mais popular atualmente, produzindo resultados impressionantes em três áreas: geração de dados [2] [3] [4], tradução de imagem para imagem [5] [6] e super-resolução [7] [8].

Embora a GAN tenha demonstrado resultados impressionantes na geração artificial de dados, especialmente em imagens [3], não é capaz de gerar novos dados a partir de conjuntos limitados. Esta incapacidade de não geração de dados também ocorre em conjuntos com baixa diversificação, como mostra este artigo. Por exemplo, um conjunto de dados médicos que contém 50 exames de 30 pacientes totaliza 1500 dados, mas o conjunto conta com apenas 30 indivíduos. Muitos conjuntos de dados necessitam de voluntários e permissão de uso por um comitê de ética, o que torna a coleta dos dados desafiadora, assim resultando em um conjunto limitado. Outras áreas tais como, detecção automatizada de incidentes de tráfego [9] e classificação de variedades de sementes de culturas [10] também sofrem de conjuntos de dados limitados, pois as aquisições desses dados são muito caras e/ou difíceis de se obter.

Os trabalhos sobre geração sintética de imagens utilizando GANs partem do pressuposto que os conjuntos de dados terão um número elevado de imagens e diversidades, também estão considerando a alta resolução nas imagens, por exemplo, 512^2 e 1024^2 pixels [2] [3]. Raramente, no mundo real, há conjuntos ideais em todos os aspectos, principalmente com imagens de alta resolução.

Por definição, um conjunto de dados em visão computacional, aprendizado de máquina e aprendizado profundo é um conjunto de imagens digitais utilizadas para testar, treinar e avaliar o desempenho de algoritmos e modelos. Entretanto, a definição não define o que são conjuntos de dados limitados e é superficialmente abordada nos artigos, sendo: um conjunto limitado é um conjunto com poucos dados [9] [10] [11] [12].

Para definir amplamente o que são os conjuntos de dados limitados, foram utilizados o conceito de amostra da estatística e o conceito de diversidade de dados [13], portanto,

- Uma amostra é um subconjunto de pessoas, itens ou eventos de uma população que é coletada e analisada para fazer inferências. Para representar a população bem, uma amostra deve ser coletada aleatoriamente e ser adequadamente grande, mas finita.
- A diversificação de dados pode fornecer amostras com informações suficientes para treinar um modelo de aprendizado de máquina. A diversidade nos dados de treinamento visa maximizar as informações contidas nos dados.

Logo, se um conjunto de dados não obedecer um desses conceitos este artigo considera que o conjunto de dados é limitado.

O artigo propõe discutir a utilização dos conjuntos limitados para o treinamento de GANs, mas focando na abordagem que não é discutida no treinamento: a baixa diversificação dos dados nos conjuntos limitados.

A estrutura do artigo é composta por 6 seções. A Seção 2 relata os trabalhos sobre modelos generativos profundos. Na Seção 3, é descrita sucintamente a proposta deste artigo. A Seção 4 descreve os modelos utilizados neste artigo. A Seção 5 traz discussões dos trabalhos atuais e os resultados de dois modelos GANs treinados com conjuntos limitados. Por fim, a Seção 6 descreve a conclusão e os trabalhos futuros.

II. TRABALHOS RELACIONADOS

A classificação automatizada de imagens é uma tarefa desafiadora devido à variabilidade de padrões presentes. Desde a adoção das Redes Neurais Profundas (Deep Neural Networks, DNNs), o estado da arte vem se redefinindo rapidamente [20] [21]. Entretanto, para avançar precisa-se resolver a questão prática dos conjuntos de dados limitados, pois as aquisições de dados podem ser burocraticamente complexas e difíceis.

TABLE I

PRINCIPAIS TRABALHOS SOBRE GANS NA ÁREA DE GERAÇÃO DE IMAGENS SINTÉTICAS. A TABELA APRESENTA OS CONJUNTOS DE IMAGENS UTILIZADOS NO TREINAMENTO DE CADA MODELO E ALGUNS ATRIBUTOS DOS CONJUNTOS DE IMAGENS (COMO, RESOLUÇÃO DAS IMAGENS E TAMANHO DO CONJUNTO).

Modelo	Ano	Base de Imagem	Resolução	Classes	Número de Imagens
Generative Adversarial Nets [1]	2014	MNIST, Frey Face	28^2 , 20×28	10, N/A	60k, 2k
Conditional Generative Adversarial Nets [14]	2014	MNIST, TFD, CIFAR-10	28^2 , 32^2 , 32^2	10, 7, 10	60k, 98k, 60k
Laplacian Generative Adversarial Network [15]	2015	CIFAR-10, STL, LSUN	32^2 , 96^2 , 64^2 (downsampled)	10, N/A, 10	100k, 100k, 10M
Deep Convolutional Generative Adversarial Networks [16]	2016	LSUN(Bedrooms), ImageNet, Faces dataset	64^2 , 32^2 , N/A	1, 1k, 1	3M, 1.2M, 350k
Information Maximizing Generative Adversarial Nets [17]	2016	MNIST, CelebA, SVHN, 3D Faces, 3D Chairs	28^2 , 32^2 , 32^2 , 32^2 , 64^2	10, N/A, 10, N/A, N/A	60k, 200k, 73k, N/A, N/A
Wasserstein GAN [18]	2017	LSUN(Bedrooms)	64^2	1	3M
Wasserstein GAN Gradient Penalty [19]	2017	ImageNet, LSUN(Bedrooms), CIFAR-10	32^2 , 128^2 , 32^2	1k, 1, 10	1.2M, 3M, 100k
Progressive Growing of GANs [2]	2018	CelebA-HQ, LSUN(Bedrooms), CIFAR-10	1024^2 , 256^2 , 32^2	1, 1, 10	30k, 3M, 100k
Style-Based Generator Architecture for Generative Adversarial Networks [3]	2019	Flickr-Faces-HQ, LSUN(Bedrooms), LSUN(Cars), LSUN(Cats),	1024^2 , 256^2 , 512×384 , 256^2	1, 1, 1, 1	70k, N/A, 46M, N/A

Para contornar este problema, modelos generativos profundos podem aprender a sintetizar dados. O modelo mais recente e promissor é a Rede Adversária Generativa (GAN) [1].

Muitos modelos foram propostos para gerar imagens sintéticas baseados na GAN [16] [17] [2] [3], na Tabela I mostra alguns trabalhos propostos e os conjuntos de imagens utilizados. Uma desvantagem das GANs é a resolução das imagens de treinamento para sintetizar a distribuição do conjunto [22]. A grande maioria dos trabalhos são avaliados em conjuntos de imagens com baixa resolução, como CIFAR (32^2 pixels) e MNIST (28^2 pixels), e com uma amostra adequadamente grande, Tabela I.

Poucos trabalhos mostraram resultados promissores para a geração de imagens de alta resolução. Por exemplo, o procedimento de treinamento progressivo proposto por Karras *et al.* [2] gera rostos de celebridades com alta resolução (1024^2 pixels). O treinamento progressivo começa alimentando a rede com imagens de baixa resolução. Progressivamente, o modelo recebe imagens de treinamento de resolução cada vez mais altas enquanto amplifica a influência das respectivas camadas para a saída. Em trabalho posterior, Karras *et al.* [3] propuseram um modelo GAN baseado em estilo, StyleGAN, e mantiveram o treinamento progressivo do trabalho anterior. A StyleGAN foi modelada para aprender de forma não supervisionado atributos de alto nível (por exemplo, pose, identidade) e variação estocástica (por exemplo, sardas, estilo do cabelo) nas imagens sintéticas, gerando rostos de pessoas comuns em alta resolução (1024^2 pixels). Karras *et al.* não modificaram o discriminador e a função de perda do modelo. Na mesma direção, Wang *et al.* [22] geraram imagens de alta resolução a partir de mapas semânticos e de instâncias. Propuseram o uso de múltiplos discriminadores e geradores operando

em diferentes resoluções para avaliar detalhes refinados e consistências globais das imagens sintéticas.

Entretanto, o principal problema de pequenos conjuntos de dados no treinamento de GANs é o discriminador, uma vez que supera rapidamente os exemplos de treinamento. A função do discriminador é classificar suas entradas como falsas ou reais, mas devido ao sobreajuste, ele rejeita tudo que não seja o conjunto de dados de treinamento como falso. Como resultado, o gerador recebe muito pouco retorno para melhorar suas gerações e o treinamento entra em colapso. Karras *et al.* [11] propuseram um pipeline de aumento (aumento do discriminador adaptativo) para as GANs serem treinadas com pequenos conjuntos de dados e, Noguchi e Harada [12] abordaram um novo método de transferência para um gerador, assim atualizando apenas os parâmetros de escala e deslocamento.

III. PROPOSTA DO TRABALHO

Este artigo apresenta uma discussão sobre pequenos conjuntos de imagens para o treinamento das GANs. Além dessa discussão, que é abordada por trabalhos recentes, este artigo também aborda um tópico que não é discutido nesses trabalhos: a baixa diversificação dos dados para o treinamento das GANs. Por fim, uma análise é feita das GANs projetadas para conjuntos ilimitados e limitados, mas treinadas com conjuntos de poucos dados e baixa diversificação.

IV. MODELOS

A. PG-GAN

A GAN de crescimento progressivo, ou simplesmente PG-GAN, proposta por Karras *et al.* [2] é uma extensão da GAN, principalmente no jeito de treinar o modelo, que permite o gerador gerar imagens sintéticas de alta resolução e alta

UNIFESP

3290 img, 128^2



FEI Face Database

2400 img, 128^2



Fig. 1. Exemplos de imagens do conjunto de dados utilizados para o treinamento dos modelos. As imagens dos conjuntos foram travadas a uma resolução de 128^2 pixels coloridas.

qualidade, como faces fotorrealistas com o tamanho desejado. Em seu trabalho Karras *et al.* geraram rostos de celebridades com resolução de 1024^2 pixels, sendo que o conjunto de imagens utilizado por eles é considerado um conjunto não limitado.

A principal inovação da PG-GAN foi o aumento incremental no tamanho da saída das camadas do gerador, começando com uma saída de 4×4 pixels e dobrando para 8×8 , 16×16 e assim por diante até a resolução de saída desejada. Esse procedimento de treinamento envolve períodos de ajuste fino da camada com uma determinada resolução de saída e períodos de introdução gradual de uma nova camada com uma resolução maior. Todas as camadas permanecem treináveis durante o processo de treinamento, incluindo as camadas existentes quando novas camadas são adicionadas.

A PG-GAN envolve o uso de um modelo gerador e discriminador com a mesma estrutura geral e começando com imagens muito pequenas (exemplo, imagens 4×4 pixels). Durante o treinamento, novos blocos de camadas convolucionais são sistematicamente adicionadas ao modelo do gerador e ao modelo do discriminador. A adição incremental das camadas permite que os modelos aprendam efetivamente detalhes de nível grosseiro e, posteriormente, aprendam detalhes cada vez mais sutis, tanto no lado do gerador quanto do discriminador.

A natureza incremental do treinamento progressivo permitiu que o treinamento descobri-se primeiro a estrutura em grande escala da distribuição do conjunto e, em seguida, desvia-se a atenção para os detalhes de escala cada vez mais finos, em vez de ter que aprender todas as escalas simultaneamente.

B. Adaptação de Estatísticas de Lote

Noguchi e Harada propuseram um novo método de transferência para um gerador e adotaram novos parâmetros, escala e deslocamento, para serem atualizados, assim abordando o problema de pequenos conjuntos de dados no treinamento

de GANs [12]. Ao transferir o conhecimento prévio de um gerador pré-treinado e atualizar apenas os parâmetros de escala e deslocamento, será corrigido os parâmetros do kernel no gerador, podendo reduzir o número de imagens necessárias para treinar um gerador. Os autores realizaram experimentos aplicando este método a um conjunto de dados muito pequeno composto por aproximadamente 100 imagens e mostram que a qualidade é superior aos métodos anteriores à 2019, e que é possível gerar imagens capturando a semântica dos dados. Os modelos utilizados foram a SNGAN com resolução de 128^2 e a BigGAN com resolução de 256^2 .

C. StyleGAN2-ADA

Karras *et al.* [11] abordaram o problema de pequenos conjuntos de dados no treinamento de GANs por meio do Aumento do Discriminador Adaptativo, assim propondo a StyleGAN2 com Aumento do Discriminador Adaptativo (StyleGAN2-ADA).

A StyleGAN2-ADA propõe o uso do aumento de dados para o problema do sobreajuste. No entanto, isso não funciona diretamente para o treinamento de GANs, pois o gerador aprenderia a produzir a distribuição aumentada. Este “vazamento” de aumento para os dados gerados, como é descrito pelos autores, é altamente indesejável. Então, os autores propuseram uma técnica de aumento chamada de aumento do discriminador estocástico, em que os aumentos podem ser projetados para não haver “vazamentos”, na condição de serem ignorados com uma probabilidade p diferente de zero. Portanto, se a maioria dos dados alimentados no discriminador não sofrer aumento, o gerador aprenderá a não criar dados com aumento.

Além da utilização do aumento de dados, Karras *et al.* propuseram duas heurísticas de sobreajuste plausíveis para medir o sobreajuste. A primeira é a heurística r_v , Equação 1, que expressa a saída de um conjunto de validação relativo ao conjunto de treinamento e os dados gerados.

$$r_v = \frac{\mathbb{E}[D_{train}] - \mathbb{E}[D_{validation}]}{\mathbb{E}[D_{train}] - \mathbb{E}[D_{generated}]} \quad (1)$$

Uma vez que assume a existência de um conjunto de validação separado de um conjunto de dados já pequeno, não é viável calcular a heurística r_v . Portanto, a segunda heurística proposta por Karras *et al.* foi a r_t , em que estima a parte do conjunto de treinamento que obtém saídas positivas no discriminador para identificar o sobreajuste e adaptar dinamicamente a probabilidade de aumento p conforme o treinamento avança, r_t é definido como:

$$r_t = \mathbb{E}[\text{sign}(D_{train})] \quad (2)$$

assim, quando r_t é alto o valor de p aumenta e quando r_t é baixo o valor de p diminuí.

V. DISCUSSÃO

Os primeiros modelos GANs propostos [1] [14] geravam imagens com baixa resolução e utilizavam conjuntos de imagens ilimitados. A ideia de gerar imagens de alta resolução é difícil, pois uma resolução mais alta torna mais fácil distinguir as imagens geradas das imagens de treinamento [23], assim ampliando drasticamente o problema do gradiente. Grandes resoluções requerem o uso de mini lotes menores, uma vez que restrições de memória e o comprometimento da estabilidade do treinamento ocorrem. As GANs propostas inicialmente (2014-2017) produziam imagens nítidas em resoluções relativamente pequenas e o treinamento continuava instável. Embora os estudos sobre GANs tenham avançados, ainda partem do pressuposto de utilizar conjuntos de dados ilimitados. Na Tabela I têm os atributos dos conjuntos de imagens (como resolução da imagem e tamanho do conjunto) utilizados em cada estudo e é evidente que os conjuntos são ilimitado em sua maioria.

As Figuras 2 - 7 mostram as imagens produzidas pelos modelos PG-GAN (treinamento do zero) [2], StyleGAN2-ADA (treinamento do zero) [11] e SNGAN-128 (treinamento por transferência) [12]. Os modelos foram treinados (os treinamentos foram realizados conforme descritos nos artigos) com dois conjuntos de imagens limitados, não obedecendo os dois conceitos descritos na Seção I, com resolução de 128^2 pixels. O primeiro conjunto (Banco de Imagens da UNIFESP) consiste em 3290 imagens de face de 34 recém-nascidos [24] e o segundo conjunto (FEI Face Database) consiste em 2400 imagens de face de 200 adultos [25], a Figura 1 mostra algumas imagens de ambos os conjuntos. Com esses atributos dos conjuntos, número de imagens n e quantidade de indivíduos i , foi quantificado a diversificação i/n de cada conjunto: o banco de imagens da UNIFESP ficou com uma diversificação de $34/3290 = 0,0103$ e o FEI Face Database ficou com a diversificação de $200/2400 = 0,083$. Quanto mais próximo de 1 maior será a diversificação do conjunto. Em outras palavras, quanto mais dados distintos a amostra tiver, a diversificação fornecida terá informações suficientes

TABLE II
FRÉCHET INCEPTION DISTANCE (FID) PARA OS PROJETOS DOS GERADORES PGAN, STYLEGAN2-ADA E SNGAN-128 (QUANTO MENOR, MELHOR).

Modelo	UNIFESP	FEI Face Database
PG-GAN [2]	18,52	28,69
SNGAN-128 [12]	53,65	62,57
StyleGAN2-ADA [11]	15,73	28,32

para treinar melhor os modelos. A Tabela II mostra a métrica Fréchet Inception Distance (FID) para cada treinamento.

Os modelos PG-GAN e StyleGAN2-ADA treinados com o conjunto de imagens da UNIFESP geraram imagens sintéticas produzindo atributos de alto nível tais como, pose e tons de pele, Figuras 2 e 4. Entretanto, ambos os modelos apresentaram o problema de sobreajuste, quando o modelo gera basicamente as mesmas imagens do conjunto de treinamento. Esse problema ocorreu devido da limitação do conjunto da UNIFESP, ocasionando na superação dos exemplos de treinamento no discriminador e seu retorno para o gerador tornou-se sem sentido, fazendo com que o treinamento começa-se a divergir.



Fig. 2. Imagens produzidas pelo modelo PG-GAN com o conjunto de imagens UNIFESP.

O treinamento de ambos os modelos PG-GAN e StyleGAN2-ADA com o conjunto FEI Face Database também apresentou o problema de sobreajuste. Entretanto, como a diversificação do conjunto FEI Face Database é 8 vezes maior do que o conjunto da UNIFESP as gerações das imagens sintéticas tiveram mais diversidade, produzindo atributos de alto nível tais como, pose, tons de pele e tipos de olhos. As Figuras 3 e 5 mostram essas imagens sintéticas.

Os resultados do modelo SNGAN-128 [12] não foram o esperado, mesmo sendo treinado com um método proposto para



Fig. 3. Imagens produzidas pelo modelo PG-GAN com o conjunto de imagens FEI Face Database.



Fig. 5. Imagens produzidas pelo modelo StyleGAN2-ADA com o conjunto de imagens FEI Face Database.



Fig. 4. Imagens produzidas pelo modelo StyleGAN2-ADA com o conjunto de imagens UNIFESP.



Fig. 6. Imagens produzidas pelo modelo SNGAN-128 treinado via adaptação de estatísticas de lote com o conjunto de imagens UNIFESP.

conjuntos limitados. As imagens sintéticas geradas tiveram a capacidade de produzir o atributo de alto nível pose, entretanto, não foram capazes de sintetizar atributos centrais de ambos os conjuntos, atributos inerentes da face por exemplo, olhos, nariz e boca, Figuras 6 e 7. Na perspectiva dos conjuntos utilizados no treinamento o modelo é limitado para conjuntos com baixa diversificação, ou seja, a falta de informações nos dados, proporcionada pela baixa diversificação, faz com que o modelo não aprenda a distribuição real dos conjuntos.

O problema de sobreajuste que o modelo PG-GAN teve na geração das imagens sintéticas de ambos os conjuntos, Figuras 1, 2 e 3, já era esperado. O modelo PG-GAN foi

projetado para conjuntos de dados ilimitado, logo treiná-lo com conjuntos limitados ocasionou na superação dos exemplos de treinamento no discriminador e o treinamento divergiu-se.

O sobreajuste ocasionado pelo modelo StyleGAN2-ADA não era esperado, Figuras 1, 4 e 5, uma vez que foi projetado para conjuntos limitados e utiliza a técnica de aumento de dados. Olhando na perspectiva dos conjuntos limitados, descritos na Seção I, o modelo StyleGAN2-ADA foi projetado para conjuntos com poucos dados e alta diversidade, logo, a baixa diversificação dos dados de treinamento não forneceu amostra com informações suficientes para o treinamento, assim oca-



Fig. 7. Imagens produzidas pelo modelo SNGAN-128 treinado via adaptação de estatísticas de lote com o conjunto FEI Face Database.

sionando também na superação dos exemplos de treinamento no discriminador.

VI. CONCLUSÃO

O artigo abordou a questão prática da utilização de conjuntos limitados no treinamento de Redes Adversárias Gerativas, visto que os conjuntos do mundo real tem essa característica. Os modelos abordados foram os mais recentes no tema de geração artificial de imagens, sendo PG-GAN, StyleGAN2-ADA e SNGAN-128 via Adaptação de Estatísticas de Lote. O artigo discutiu sobre a maioria das GANs serem projetadas para conjuntos ilimitados. A principal contribuição do artigo foi discutir o treinamento dessas GANs em conjuntos com poucos dados e baixa diversidade e os resultados mostraram que as GANs projetadas para conjuntos ilimitados e limitados não consideram a baixa diversidade na amostra.

Vislumbra-se como continuação do trabalho aprofundar os estudos no tema, considerar outras métricas para quantificar as imagens sintéticas geradas e seus desempenhos com conjuntos limitados e, propor um treinamento no que concerne aos conjuntos limitados definido neste artigo.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) processo nº 142560/2020-0 e do Centro Universitário FEI.

REFERÊNCIAS

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, and Y. Bengio, "Generative adversarial nets," in *NIPS*, 2014.

[2] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *ArXiv*, vol. abs/1710.10196, 2018.

[3] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4396–4405, 2019.

[4] M. Huzaifah and L. Wyse, "Deep generative models for musical audio synthesis," *ArXiv*, vol. abs/2006.06426, 2020.

[5] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[6] S. Engelhardt, L. Sharan, M. Karck, R. D. Simone, and I. Wolf, "Cross-domain conditional generative adversarial networks for stereoscopic hyperrealism in surgical training," in *MICCAI*, 2019.

[7] X. Wang, K. Yu, S. Wu, J. Gu, Y.-H. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang, "Esrgan: Enhanced super-resolution generative adversarial networks," in *ECCV Workshops*, 2018.

[8] L. Yang, C. Liu, P. Wang, S. Wang, P. Ren, S. Ma, and W. Gao, "Hifacegan: Face renovation via collaborative suppression and replenishment," *Proceedings of the 28th ACM International Conference on Multimedia*, 2020.

[9] Y. Lin, L. Li, H. Jing, B. Ran, and D. Sun, "Automated traffic incident detection with a smaller dataset based on generative adversarial networks," *Accident; analysis and prevention*, vol. 144, p. 105628, 2020.

[10] N. Wu, F. Liu, F. Meng, M. Li, C. Zhang, and Y. He, "Rapid and accurate varieties classification of different crop seeds under sample-limited condition based on hyperspectral imaging and deep transfer learning," *Frontiers in Bioengineering and Biotechnology*, vol. 9, 2021.

[11] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," in *Proc. NeurIPS*, 2020.

[12] A. Noguchi and T. Harada, "Image generation from small datasets via batch statistics adaptation," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2750–2758, 2019.

[13] Z. Gong, P. Zhong, and W. Hu, "Diversity in machine learning," *IEEE Access*, vol. 7, pp. 64323–64350, 2019.

[14] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[15] E. L. Denton, S. Chintala, A. D. Szlam, and R. Fergus, "Deep generative image models using a laplacian pyramid of adversarial networks," in *NIPS*, 2015.

[16] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *CoRR*, vol. abs/1511.06434, 2016.

[17] X. Chen, Y. Duan, R. Houthoofd, J. Schulman, I. Sutskever, and P. Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," in *NIPS*, 2016.

[18] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," *ArXiv*, vol. abs/1701.07875, 2017.

[19] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," in *NIPS*, 2017.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.

[21] G. Zamzami, R. Paul, D. Goldgof, R. Kasturi, and Y. Sun, "Pain assessment from facial expression: Neonatal convolutional neural network (n-cnn)," *2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, 2019.

[22] T. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8798–8807, 2018.

[23] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in *International conference on machine learning*. PMLR, 2017, pp. 2642–2651.

[24] T. M. Heiderich, A. T. F. S. Leslie, and R. Guinsburg, "Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements," *Acta Paediatrica*, vol. 104, no. 2, pp. e63–e69, 2015.

[25] C. E. Thomaz and G. A. Giraldi, "A new ranking method for principal components analysis and its application to face image analysis," *Image and vision computing*, vol. 28, no. 6, pp. 902–913, 2010.