

Spatial Distribution Analysis of Images in GAN Training with Small Datasets

Lucas F. Buzuti, Tatiany M. Heiderich, Carlos E. Thomaz

¹Departamento de Engenharia Elétrica – Laboratório de Processamento de Imagens
Centro Universitário FEI – São Bernardo do Campo – SP – Brasil

lucas.buzuti@outlook.com, tatianyms@gmail.com, cet@fei.edu.br

Abstract. *Generative Adversarial Networks (GANs) are being increasingly used to artificially generate various types of data. Their training requires a sufficiently large dataset and becomes a challenge with small datasets. Recent works have proposed new approaches to training GANs with small samples. In this paper, we analyze the spatial distribution of the real and synthetic data of small datasets, building subspaces randomly, but varying their level of sparsing. To vary the level of sparsing, we propose the algorithm named k -Sparsest Sample. Our results show that small sets with the sparsest spatial distribution are able to generate data with high diversity.*

Resumo. *Redes Adversárias Generativas (GANs) estão sendo cada vez mais usadas para gerar artificialmente vários tipos de dados. O treinamento dessas redes requer um conjunto de dados suficientemente grande e se torna um desafio com pequenos conjuntos. Trabalhos recentes propuseram novas abordagens para o treinamento de GANs com poucas amostras. Este trabalho analisa a distribuição espacial dos dados reais e sintéticos desses conjuntos, construindo subespaços de forma aleatória e variando o nível de espalhamento. Para variar o nível de espalhamento, este trabalho propõe o algoritmo k -Amostras Esparsas. Os resultados mostraram que pequenos conjuntos com uma distribuição espacial mais espalhada tendem a gerar dados com mais diversidade.*

1. Introdução

Estudos envolvendo Redes Adversárias Generativas (*Generative Adversarial Nets*, GANs) [Goodfellow et al. 2014] são cada vez mais impressionantes, sintetizando artificialmente vários tipos de dados como imagens, vídeos e áudios de alta qualidade [Miyato et al. 2018, Zhang et al. 2019, Karras et al. 2020b]. Entretanto, o treinamento dessas redes requer um conjunto suficientemente grande, sendo um desafio para diversas áreas que sofrem da carência de dados [Yi et al. 2019, Lin et al. 2020, Norori et al. 2021, Wu et al. 2021].

Dadas as restrições impostas pela Lei Geral de Proteção de Dados Pessoais (LGPD), Brasil, e no Regulamento Geral sobre a Proteção de Dados (*General Data Protection Regulation*, GDPR), União Europeia, dados de áreas como da saúde estão sujeitos a serem mais limitados diante dessas restrições. Utilizar GANs para gerar dados sintéticos a partir de pequenos conjuntos aumenta a solução de problemas restritos à falta de dados, uma vez que os dados gerados não correspondem a nenhum indivíduo.

O principal problema de treinar GANs com pequenos conjuntos de dados (conjuntos limitados) é o sobre-ajuste do discriminador aos exemplos de treinamento levando ao colapso do modo no gerador. Como alternativa para esse problema, [Wang et al. 2018] propuseram um método para transferir conhecimento prévio de uma GAN para o treinamento de pequenos conjuntos. Trabalhos recentes têm proposto novas abordagens no treinamento de GANs para pequenos conjuntos [Noguchi and Harada 2019, Karras et al. 2020a]. [Noguchi and Harada 2019] abordaram o treinamento de GANs com no máximo 100 imagens e [Karras et al. 2020a] propuseram o Aumento de Dados no treinamento, não especificando o número mínimo ou máximo de imagens.

Devido à restrição do número de imagens de treinamento estabelecido pelo [Noguchi and Harada 2019], este trabalho analisa a distribuição espacial das imagens reais (conjunto de imagens) e sintéticas escolhendo 100 imagens de treinamento de forma aleatória e variando o nível de espalhamento. Para variar o nível de espalhamento este trabalho propõe o algoritmo *k*-Amostras Esparsas (*k-Sparsest Samples*, KSS).

2. Fundamentação Teórica

2.1. Topologia

[Noguchi and Harada 2019] propuseram a adaptação de estatísticas de lote (*batch-statistics adaptation*, BSA) transferindo (congelando os pesos) todo o gerador de origem com novos parâmetros treináveis (β e γ) para adaptar as estatísticas de lote ocultas na geração de imagens sintéticas com conjuntos de destino extremamente limitados (máximo 100 imagens). No entanto, o gerador não é treinado de forma adversária, em vez disso, $L1$ e perda perceptiva são utilizados, assim, o modelo generativo pode ser transferido para pequenos conjuntos reduzindo o número de parâmetros treináveis [Noguchi and Harada 2019].

2.2. Métricas de Avaliação

O **Inception Score (IS)** [Salimans et al. 2016] é um índice comumente aplicado para avaliar o desempenho das GANs. Para calcular o IS, as imagens geradas são enviadas para a rede *Inception* [Szegedy et al. 2016] que foi pré-treinada no conjunto de dados *ImageNet*. Do ponto de vista dos três aspectos para GANs ideais, o IS se concentra em medir a Herança e a Diversidade. Especificamente, $\mathbf{x} \in G$ é uma imagem gerada, $y = \text{InceptionNet}(\mathbf{x})$ é o rótulo obtido da rede *Inception* pré-treinada ao inserir a imagem \mathbf{x} . Para todas as imagens geradas, temos o conjunto de rótulos Y . O IS é definido:

$$\exp(\mathbb{E}_{\mathbf{x}}[D_{KL}(p(y | \mathbf{x}) || p(y))]) = \exp(H(y) - \mathbb{E}_{\mathbf{x}}[H(y | \mathbf{x})]), \quad (1)$$

em que D_{KL} é a divergência de Kullback–Leibler (KL) e $H(\cdot)$ representa a entropia. Como a Criatividade não é considerada pelo IS, este não tem capacidade de detectar o sobre-ajuste.

A métrica **Fréchet Inception Distance (FID)** [Heusel et al. 2017] também usa a rede *Inception* pré-treinada. Em vez de rótulos de saída, usa vetores de características das camadas *pooling* finais do *InceptionNet*. Todas as imagens reais e geradas são inseridas na rede para extrair seus vetores de características. Seja $\varphi(\cdot) = \text{InceptionNet_lastPooling}(\cdot)$ o extrator de características e, $F_r = \varphi(R)$ e $F_g = \varphi(G)$

dois grupos de vetores de características extraídos dos conjuntos de imagens reais e sintéticas. Considerando que as distribuições de F_r e F_g são Gaussianas multivariadas, portanto, $F_r \sim \mathcal{N}(\mu_r, \Sigma_r)$ e $F_g \sim \mathcal{N}(\mu_g, \Sigma_g)$, assim sendo,

$$FID(r, g) = \|\mu_r - \mu_g\|_2^2 + Tr \left(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{\frac{1}{2}} \right), \quad (2)$$

em que (μ_r, Σ_r) e (μ_g, Σ_g) são a média e covariância dos vetores de características das imagens reais e sintéticas geradas pelo modelo, respectivamente.

Outra métrica denominada **Índice de Criatividade, Herança e Diversidade (CID)** [Guan and Loew 2019] calcula três índices, denominados Criatividade (C), Herança (I) e Diversidade (D), de forma independente e os combinam para quantificar as imagens sintéticas, medindo a diferença e semelhança entre as imagens reais e sintéticas. Tradicionalmente, o CID utiliza o Índice de Semelhança Estrutural (*Structural Similarity Index*, SSIM) para calcular a Criatividade e a Diversidade, já o cálculo da Herança utiliza a Matriz de Co-ocorrência de Nível de Cinza (*Gray-level Co-occurrence Matrix*, GLCM).

3. k-Amostras Esparsas

Para analisar a distribuição espacial das imagens reais e sintéticas escolhendo N imagens de treinamento, este trabalho propõe o algoritmo k-Amostras Esparsas (*k-Sparsest Samples*, KSS) de modo a variar o nível de espalhamento das amostras. O KSS manipula o espaço original do conjunto de dados para construir um subespaço de modo independente, mas mantendo todas as propriedades originais.

Algoritmo 1: k-Amostras Esparsas (K-Sparsest Samples, KSS)

Entrada: Matriz de dados $\mathbf{X} \in \mathbb{R}^{N \times d}$, em que $\{\mathbf{x}^{(n)} \in \mathbf{X}\}_{n=1}^N$ (N número de amostras e d número de variáveis), k tamanho do subespaço, \mathcal{G} número de agrupamento e p número de componentes.

Saída: S_E subespaço de tamanho k do espaço original \mathbf{X} .

1 $\mathbf{Y} = \text{PCA}(\mathbf{X}, p)$, em que $\{\mathbf{y}^{(n)} \in \mathbf{Y}\}_{n=1}^N$ $\mathbf{Y} \in \mathbb{R}^p$

2 $\{\mu_g\}_{g=1}^{\mathcal{G}}, \{c_g^{(n)}\}_{n=1, g=1}^{N, \mathcal{G}} = k\text{Means}(\mathbf{Y}, \mathcal{G})$

3 $l = k/\mathcal{G}$ (quantidade de amostras por agrupamento).

4 $S := \{\{\mathbf{y}^{(n)}\}_g \mid c_g^{(n)} = g\}_{n=1, g=1}^{N, \mathcal{G}}$

5 **para cada** $g \in \mathcal{G}$ **faça**

6 $\Sigma_S = S^{(g)}(S^{(g)})^T$

7 $D^{(g)} = \sqrt{(S^{(g)} - \mu_g)\Sigma_S^{-1}(S^{(g)} - \mu_g)^T}$

8 ordenar $D^{(g)}$ no maior para o menor em relação à $S^{(g)}$. Selecionar as l primeiras amostras de $S^{(g)}$ e adicionar em S_E .

9 $S_E \leftarrow S_l^{(g)}$

10 **fim**

O arcabouço do KSS é composto por três etapas: (1) aplicação da Análise de Componentes Principais (*Principal Component Analysis*, PCA), Algoritmo 2, a fim de reduzir a dimensionalidade das imagens e impactos de problemas com a memória e o tempo de processamento computacional; (2) utilização do *k-Means*, Algoritmo 3, para

definir \mathcal{G} agrupamentos; (3) em cada agrupamento é calculada a distância de Mahalanobis visando determinar as imagens de maior distância em relação à média do agrupamento. Os detalhes do KSS são mostrados no Algoritmo 1.

Algoritmo 2: Análise de Componentes Principais (PCA)

- Entrada:** Matriz de dados $\mathbf{X} \in \mathbb{R}^{N \times d}$ e p número de componentes.
Saída: projeção da matriz de dados $\mathbf{Y} \in \mathbb{R}^p$.
- 1 $\Sigma_{\mathbf{X}} = (\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T$
 - 2 Calcula-se os autovalores e autovetores de $\Sigma_{\mathbf{X}}$.
 - 3 $\Sigma_{\mathbf{X}}\mathbf{V} = \Lambda\mathbf{V}$ Ordena \mathbf{V} sobre
 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq \dots \geq \lambda_N \quad \lambda_i \in \Lambda \quad \forall i = 1, 2, \dots, N$. Portanto,
 $\mathbf{V} = [\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(p)}]$
 - 4 $\mathbf{Y} = \mathbf{V}^T\mathbf{X}$
-

Algoritmo 3: k-Means

- Entrada:** Matriz de dados $\mathbf{X} \in \mathbb{R}^{N \times d}$ e \mathcal{G} número de agrupamento.
Saída: g centroides $\{\mu_g\}_{g=1}^{\mathcal{G}}$ e rótulos $\{c_g^{(n)}\}_{n=1, g=1}^{N, \mathcal{G}}$ para cada $\mathbf{x}^{(n)}$.
- 1 Inicialização aleatória $\{\mu_g \in \mathbb{R}^p\}_{g=1}^{\mathcal{G}}$.
 - 2 **enquanto não converge faça**
 - 3 $c_g^{(n)} := \arg \min_g \|\mathbf{x}^{(n)} - \mu_g\|^2$
 - 4 $\mu_g := \frac{\sum_{n=1}^N 1_{\{c^{(n)}=g\}} \mathbf{x}^{(n)}}{\sum_{n=1}^N 1_{\{c^{(n)}=g\}}}, \quad \forall g \in \{1, \dots, \mathcal{G}\}$
 - 5 **fim**
-

A construção do subespaço é definido através de k (tamanho do subespaço) e do número de agrupamentos \mathcal{G} . Quanto maior for \mathcal{G} mais espalhados os dados estarão no subespaço e $k \geq \mathcal{G} \quad \forall \{k, \mathcal{G} \in \mathbb{N} \mid k \neq 0, \mathcal{G} \neq 0\}$. Entretanto, o nível de espalhamento do subespaço pode ser limitado ao calcular e inverter a matriz de covariância dos dados de cada agrupamento para calcular a distância de Mahalanobis, ocasionando no erro de singularidade, deste caso $\mathcal{G} \ll k$.

4. Metodologia Experimental

Este trabalho utilizou 4 conjuntos de imagens contendo faces. O primeiro foi o **Banco de Dados Neonatal da UNIFESP** [Heiderich et al. 2015] com 3290 imagens coloridas com resolução de 450×233 pixels de 34 recém-nascidos. O segundo, o **Banco de Faces da FEI** [Thomaz and Giraldo 2010] contendo 14 imagens para cada um dos 200 indivíduos, mais especificamente 100 homens e 100 mulheres, um total de 2800 imagens com resolução de 640×480 pixels. O terceiro, o conjunto de imagens **Flickr-Face-HQ (FFHQ)** [Karras et al. 2019] contém 70.000 imagens de alta qualidade com resolução de 1024×1024 pixels e variações consideráveis em termos de idade, etnia e plano de fundo da imagem. O quarto, o conjunto de imagens **MetFaces** [Karras et al. 2020a] contém 1336 imagens de alta qualidade com resolução de 1024×1024 pixels de rostos humanos extraídos de obras de arte.

Antes do treinamento do modelo SNGAN-128 por meio da adaptação de estatísticas de lote [Noguchi and Harada 2019], todos os conjuntos de imagens foram redimensionados para 128×128 pixels, na intenção de padronizar todos os experimentos, utilizando a operação bicúbica. Apenas no Banco de Dados Neonatal da UNIFESP foi realizado uma detecção facial, antes da operação de redimensionamento. O CID utiliza o SSIM para calcular a Criatividade e a Diversidade, entretanto, [Zhang et al. 2018] mostraram que o SSIM não leva em conta muitas nuances da percepção humana, assim, propuseram a semelhança de fragmento de imagem perceptiva aprendida (*Learned Perceptual Image Patch Similarity*, LPIPS). Portanto, para calcular a Criatividade e a Diversidade utilizou-se o LPIPS ao invés do SSIM.

Após o redimensionamento, foram construídos 6 subespaços com 100 imagens de cada conjunto, sendo que 4 desses subespaços foram construídos com o KSS (PCA com $\geq 90\%$ da variância total), de forma a variar o nível de espalhamento, portanto, $k = 100$. Para os conjuntos FFHQ, FEI e UNIFESP \mathcal{G} é igual à 1,2,10 e 100. No conjunto MetFaces ocorreu erro de singularidade ao computar o KSS com $\mathcal{G} = 100$, sendo assim, utilizou-se \mathcal{G} igual à 1,2,10 e 80. Os subespaços restantes foram construídos, um de forma aleatória e o outro utilizando as etapas (1) e (3) do processo do KSS (PCA com $\geq 90\%$ da variância total). Na etapa (3) desse processo foi calculado a distância de Mahalanobis adquirindo as imagens de menor distância em relação à média do espaço original, assim, denominando esse subespaço de baixa variância.

5. Resultados e Discussões

De forma a validar igualmente as construções dos 6 subespaços, redimensionou-se o conjunto FEI para \mathbb{R}^2 , $p = 2$. As Figuras 1a-1f mostram o espaço original e os subespaços extraídos do original. Construindo os subespaços com o KSS, fixando k em 100 imagens e variando \mathcal{G} em 1, 2, 10 e 100, observa-se os subespaços se tornando mais espalhados a medida que \mathcal{G} aumenta, Figuras 1a-1d. A Figura 1f mostra o subespaço de 100 imagens construído de forma aleatória e a Figura 1e mostra o subespaço de baixa variância com 100 imagens.

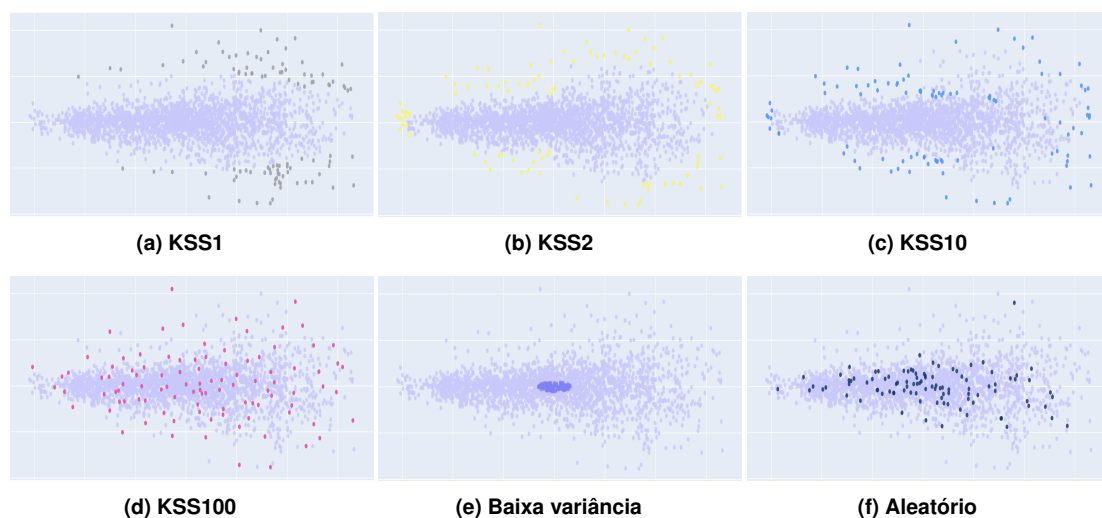


Figura 1. Visualização em \mathbb{R}^2 , $p = 2$, da construção dos subespaços a partir do espaço original.

O subespaço de baixa variância (*Center*) de cada conjunto de imagens obteve pontuações IS e os índices CID menores em relação aos outros subespaços, Tabela 1. As imagens pertencentes ao subespaço de baixa variância estão espacialmente agrupadas, logo os resultados dessas métricas indicam imagens sintéticas geradas com baixa criatividade e diversidade. Em outras palavras, as imagens têm atributos semânticos semelhantes ou iguais, tornando a distância espacial entre elas quase nula. Devido a esse agrupamento, a pontuação FID foi melhor do que nos outros subespaços. Isso ocorreu devido ao FID medir a semelhança/qualidade entre as imagens reais e sintéticas, Tabela 1. Como o subespaço de baixa variância têm imagens agrupadas, o modelo não foi capaz de gerar imagens com alta criatividade e diversidade (Figuras 2a, 2d, 3c e 4b) a partir das imagens reais, ocasionando o problema de sobre-ajuste mensurado pelo índice de Criatividade (C), Tabela 1.

Tabela 1. Análise quantitativa dos subespaços de cada conjunto de imagens.

Conjunto de Dados		IS \uparrow	FID \downarrow	C	I	D \uparrow	CID \uparrow
Flickr Face HQ [Karras et al. 2019]	Random	2,26 \pm 0,02	193,65	1,00	0,60	5,65	3,41
	Center	1,62 \pm 0,01	154,84	0,63	0,51	2,91	0,93
	KSS1	2,60 \pm 0,02	231,16	1,00	0,71	8,40	5,98
	KSS2	2,44 \pm 0,01	236,07	1,00	0,70	8,39	5,86
	KSS10	2,49 \pm 0,01	227,86	1,00	0,73	8,37	6,11
	KSS100	2,50 \pm 0,03	226,53	1,00	0,64	8,35	5,34
MetFaces [Karras et al. 2020a]	Random	2,21 \pm 0,02	153,58	1,00	0,63	5,36	3,37
	Center	1,67 \pm 0,01	141,08	0,81	0,59	3,10	1,47
	KSS1	2,10 \pm 0,02	162,29	1,00	0,66	7,36	4,86
	KSS2	2,15 \pm 0,01	161,15	1,00	0,65	7,41	4,81
	KSS10	2,19 \pm 0,01	164,52	1,00	0,62	7,41	4,59
	KSS80	2,29 \pm 0,02	174,79	1,00	0,68	7,13	4,81
FEI [Thomaz and Giraldo 2010]	Random	2,09 \pm 0,02	47,51	0,11	0,78	2,12	0,18
	Center	1,61 \pm 0,01	64,25	0,00	0,00	0,00	0,00
	KSS1	1,98 \pm 0,02	70,06	0,79	0,84	3,62	2,40
	KSS2	1,86 \pm 0,01	96,18	0,95	0,81	4,11	3,18
	KSS10	1,96 \pm 0,02	82,21	0,82	0,79	3,41	2,22
	KSS100	2,06 \pm 0,01	52,76	0,63	0,80	2,67	1,35
Unifesp [Heiderich et al. 2015]	Random	1,88 \pm 0,01	103,41	0,71	0,69	3,90	1,91
	Center	1,10 \pm 0,00	64,97	0,00	0,42	0,00	0,00
	KSS1	1,92 \pm 0,01	121,90	0,92	0,77	3,97	2,81
	KSS2	2,01 \pm 0,01	122,84	0,92	0,75	4,52	3,11
	KSS10	2,05 \pm 0,02	113,64	0,98	0,70	6,00	4,15
	KSS100	2,04 \pm 0,01	106,01	0,91	0,66	5,66	3,39

Em contraste com os resultados do subespaço de baixa variância, as pontuações IS e os índices CID para os subespaços KSS e aleatório (*Random*) indicaram geração de imagens sintéticas com alta criatividade e diversidade (Tabela 1), logo a não ocorrência



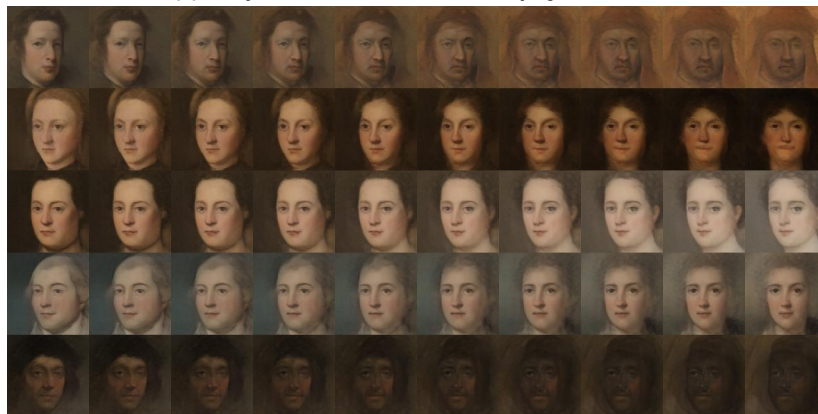
(a) Conjunto Flickr Face HQ subespaço de baixa variância (*Center*).



(b) Conjunto Flickr Face HQ subespaço aleatório (*Random*).



(c) Conjunto Flickr Face HQ subespaço KSS100.



(d) Conjunto MetFaces subespaço de baixa variância (*Center*).

Figura 2. Interpolação no espaço do gerador.



(a) Conjunto MetFaces subespaço aleatório (*Random*).



(b) Conjunto MetFaces subespaço KSS80.



(c) Conjunto FEI subespaço de baixa variância (*Center*).



(d) Conjunto FEI subespaço aleatório (*Random*).

Figura 3. Interpolação no espaço do gerador.



(a) Conjunto FEI subespaço KSS100.



(b) Conjunto Unifesp subespaço de baixa variância (*Center*).



(c) Conjunto Unifesp subespaço aleatório (*Random*).



(d) Conjunto Unifesp subespaço KSS100.

Figura 4. Interpolação no espaço do gerador.

de sobre-ajuste. Devido às imagens contidas nesses subespaços estarem espalhadas espacialmente, a pontuação FID foi ruim. Portanto, as imagens sintéticas geradas nesses subespaços poderão não estar na mesma qualidade das imagens reais. Nas Figuras 2b, 2c, 3a, 3b, 3d, 4a, 4c e 4d pode-se visualizar as imagens geradas nos subespaços KSS e aleatório.

As pontuações IS e os índices CID dos subespaços, Tabela 1, validam a hipótese de que “conjunto com dados mais distantes espacialmente entre si possibilita a geração de dados sintéticos mais diversos”. Logo, uma distribuição espacial espalhada tende a cobrir mais combinações dos atributos semânticos para sintetizar novos dados. Entretanto, pode não garantir a geração de dados com a mesma qualidade dos dados reais, sendo validado pela pontuação FID.

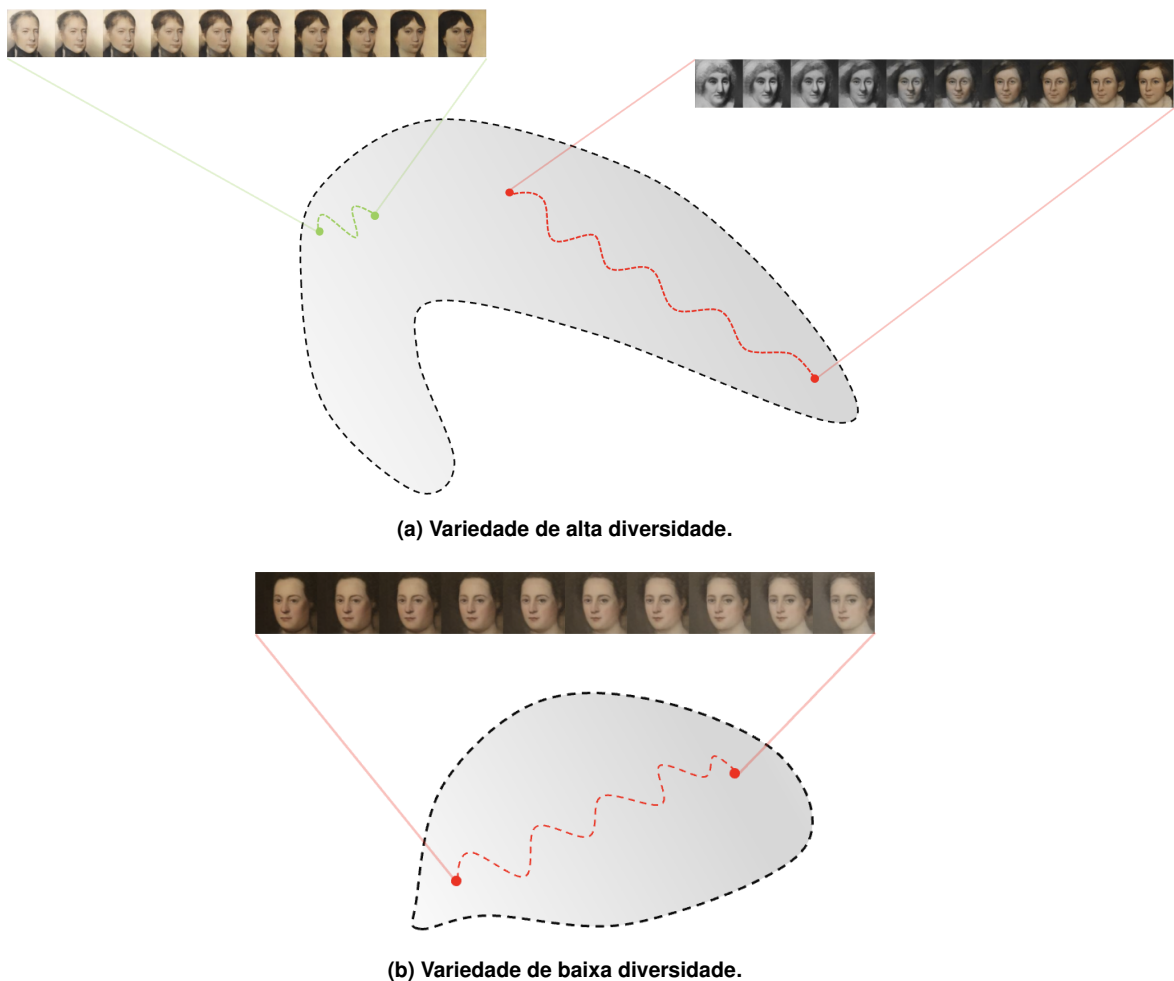


Figura 5. Ideia geral da distribuições espacial de alta e baixa diversidade.

A Figura 5 representa uma ideia geral da distribuição espacial à nível de alta e baixa diversidade. A área da linha (preta) tracejada representa a distribuição das imagens reais. Os pontos vermelho e verde representam imagens sintéticas geradas e a linha tracejada que conectam eles representa a interpolação. Na Figura 5a se tem uma distribuição mais diversificada com combinações dos atributos semânticos para sintetizar uma nova imagem pertencente à distribuição real (linha tracejada vermelha), mas nada impede de

gerar imagens semelhantes (linha tracejada verde). Na Figura 5b se tem uma distribuição menos diversificada e não importa a distância entre os pontos, a combinações dos atributos semânticos não terá muitas variações (linha tracejada vermelha).

6. Conclusão

Este trabalho revela a importância da distribuição espacial das imagens para pequenos conjuntos de dados, visto que, conforme as imagens estão distribuídas espacialmente pode-se afetar a criatividade e a diversidade das imagens sintéticas geradas. Para mostrar isso, uma GAN via adaptação de estatísticas de lote [Noguchi and Harada 2019] foi treinada por meio dos subespaços aleatórios variando o nível de espalhamento a partir do espaço original dos conjuntos de imagens. O algoritmo *k*-Amostras Esparsas (*k-Sparsest Samples*, KSS), proposto por este trabalho, foi utilizado para construir os subespaços variando o nível de espalhamento das amostras. Os resultados apresentados aqui evidenciaram que conjuntos com dados mais distantes espacialmente entre si possibilitam a geração de dados sintéticos mais diversos, mas, podem não garantir a geração de dados com a mesma qualidade dos dados reais. Portanto, ao propor modelos de geração de dados sintéticos a partir de conjuntos limitados precisa levar em consideração outras limitações além do número de amostras, como a falta de (ou baixa) diversidade.

Como trabalhos subsequentes busca-se estender essas análises para modelos que trabalham com conjuntos pequenos que não delimitam um número máximo de imagens para o treinamento, tal como a StyleGAN2-ADA [Karras et al. 2020a] e incluir nas análises outras variáveis que não foram abordadas aqui, tal como *outliers*.

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

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Guan, S. and Loew, M. (2019). Evaluation of generative adversarial network performance based on direct analysis of generated images. In *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–5. IEEE.
- Heiderich, T. M., Leslie, A. T. F. S., and Guinsburg, R. (2015). Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements. *Acta Paediatrica*, 104(2):e63–e69.
- Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30.
- Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., and Aila, T. (2020a). Training generative adversarial networks with limited data. In *Proc. NeurIPS*.

- Karras, T., Laine, S., and Aila, T. (2019). A style-based generator architecture for generative adversarial networks. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4396–4405.
- Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. (2020b). Analyzing and improving the image quality of stylegan. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8107–8116.
- Lin, Y., Li, L., Jing, H., Ran, B., and Sun, D. (2020). Automated traffic incident detection with a smaller dataset based on generative adversarial networks. *Accident; analysis and prevention*, 144:105628.
- Miyato, T., Kataoka, T., Koyama, M., and Yoshida, Y. (2018). Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*.
- Noguchi, A. and Harada, T. (2019). Image generation from small datasets via batch statistics adaptation. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2750–2758.
- Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D., and Tzovara, A. (2021). Addressing bias in big data and ai for health care: A call for open science. *Patterns*, 2(10):100347.
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. (2016). Improved techniques for training gans. *Advances in neural information processing systems*, 29:2234–2242.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826.
- Thomaz, C. E. and Giraldi, G. A. (2010). A new ranking method for principal components analysis and its application to face image analysis. *Image and vision computing*, 28(6):902–913.
- Wang, Y., Wu, C., Herranz, L., van de Weijer, J., Gonzalez-Garcia, A., and Raducanu, B. (2018). Transferring gans: generating images from limited data. In *ECCV*.
- Wu, N., Liu, F., Meng, F., Li, M., Zhang, C., and He, Y. (2021). 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*, 9.
- Yi, X., Walia, E., and Babyn, P. (2019). Generative adversarial network in medical imaging: A review. *Medical image analysis*, 58:101552.
- Zhang, H., Sindagi, V., and Patel, V. M. (2019). Image de-raining using a conditional generative adversarial network. *IEEE transactions on circuits and systems for video technology*, 30(11):3943–3956.
- Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*.