

Redução do Erro de Representação em Sensoriamento Compressivo com Modelos Generativos Usando Ajuste por Pivô

João V. D. Sobrinho¹, Igor D. Alvarenga¹, Miguel Elias M. Campista¹

¹ GTA/DEL-POLI/PEE-COPPE
Universidade Federal do Rio de Janeiro (UFRJ)

{joaosobrinho, alvarenga, miguel}@gta.ufrj.br

Abstract. *Reducing the transmitted data volume is essential for implementing networks with resource-limited and energy-constrained devices. In this sense, compressive sensing becomes a powerful alternative, as it moves the most computationally complex task to the central server node, in contrast to the traditional compression scheme. Recently, a combination of compressive sensing and generative models appeared, giving rise to CSGM (Compressive Sensing using Generative Models). Although CSGM reduces the reconstruction error, it introduces the so-called representation error. This paper proposes a technique based on model retraining in decompression time to reduce the representation error in CSGM. In this way, expanding the scope of the generative model to include the desired signal becomes possible. The results show performance gains in signal reconstruction of up to 30% compared with Deep Image Prior (DIP) and Wavelet Thresholding (WT) techniques, traditionally used in the literature.*

Resumo. *A redução do volume de dados transmitidos é essencial para implementação de redes com dispositivos de baixo poder computacional e limitações de energia. Nesse sentido, o sensoriamento compressivo torna-se uma alternativa poderosa, já que passa ao nó servidor central a tarefa de maior complexidade computacional, em contraste ao esquema tradicional de compressão. Recentemente, uma combinação entre sensoriamento compressivo e modelos generativos foi desenvolvida, dando origem ao CSGM (Compressive Sensing using Generative Models). Apesar do CSGM reduzir o erro de reconstrução, este introduz o chamado erro de representação. Para reduzir o erro de representação em CSGM, este trabalho introduz uma técnica baseada no retreinamento do modelo em tempo de descompressão. Dessa forma, a expansão do alcance do modelo generativo para incluir o sinal desejado torna-se possível. Os resultados mostram ganhos de desempenho na reconstrução do sinal de até 30% em comparação com as técnicas Deep Image Prior (DIP) e Wavelet Thresholding (WT), tradicionalmente utilizadas na literatura.*

1. Introdução

O crescente interesse em aplicações executadas em dispositivos com limitações energéticas e computacionais, como os que compõem as Redes de Sensores Sem Fio (*Wireless Sensor Networks* - WSN) e os sistemas de Internet das Coisas (*Internet of*

Things - IoT), gera desafios quanto a escalabilidade das implementações. Dentre esses desafios, a redução do consumo energético apresenta-se como uma das mais relevantes, já que a tarefa de transmissão é responsável por uma parcela considerável da demanda energética dos dispositivos participantes [Abbasian Dehkordi et al. 2020]. Diferentes abordagens, conseqüentemente, são propostas na literatura, tais como a agregação [Kaur and Munjal 2020] e a compressão [Ketshabetswe et al. 2021] de dados, a colheita de energia (*energy harvesting*) [Singh et al. 2021] e as técnicas de roteamento [Nakas et al. 2020].

Inicialmente proposta por David L. Donoho, a técnica de sensoriamento compressivo (*Compressive Sensing* - CS) apresenta vantagens para cenários com restrições energéticas em nós nas bordas computacionais [Donoho 2006]. Isso ocorre porque o CS possibilita a inversão da relação de complexidade entre as etapas de compressão e descompressão, quando comparados aos métodos tradicionais. Ou seja, diferente de configurações em que a etapa de compressão apresenta elevado custo associado, o sensoriamento compressivo permite que a compressão seja realizada por meio de apenas uma subamostragem do sinal desejado. Dessa forma, os nós das bordas, com restrições energéticas e computacionais, são capazes de comprimir os dados a baixos custos, enquanto os nós centrais assumem a tarefa de reconstrução do sinal subamostrado, com elevado custo. Além da redução do gasto energético associado à transmissão dos dados, o sensoriamento compressivo apresenta também vantagens quanto à própria atividade de sensoriamento, já que os dispositivos utilizados podem realizar menos medições para capturar o sinal desejado.

Apesar da vantajosa relação de custos entre as etapas de compressão e descompressão, o sensoriamento compressivo clássico apresenta limitações que podem dificultar a implementação da técnica em maior escala. Dentre elas, a considerável exigência de que o sinal original seja compressível, ou seja, que haja uma representação esparsa do mesmo em algum domínio de interesse [Wu et al. 2019]. Em contraponto, o sensoriamento compressivo com modelos generativos (*Compressive Sensing using Generative Models* - CSGM) dispensa a exigência de esparsidade, substituindo a restrição pela presença do sinal original no alcance de um modelo generativo [Bora et al. 2017]. Logo, o uso do CSGM torna a recuperação de um sinal subamostrado possível, desde que o próprio sinal possa ser gerado por um modelo generativo, como uma Rede Adversarial Generativa (*Generative Adversarial Network* - GAN). Além de eliminar a exigência de esparsidade, o CSGM promove também ganhos de desempenho, apresentando erros de reconstrução semelhantes aos alcançados com o sensoriamento compressivo tradicional, só que para sinais amostrados a taxas até 10 vezes menores. Apesar das vantagens obtidas pelo uso do CSGM, tanto referente ao desempenho quanto à não exigência de esparsidade, a técnica gera novos desafios. O uso de um modelo generativo no processo de reconstrução tem como risco a incapacidade da geração de um sinal satisfatoriamente similar ao original no espaço de representação. A diferença entre o sinal original e a aproximação obtida pelo modelo é tida na literatura como erro de representação [Asim et al. 2020].

O presente trabalho propõe uma forma de redução do erro de representação em CSGM por meio de ajuste por pivô (*Pivot Tuning for Latent Based Editing of Real Images* - PTI) [Roich et al. 2022]. A técnica de ajustes por pivô, originalmente utilizada para a tarefa de edição de imagens, é adaptada para o contexto do sensoriamento compres-

sivo, sendo utilizada como uma etapa adicional ao método tradicional. Ou seja, primeiro obtém-se o valor do pivô, a entrada do modelo que gera o sinal mais similar ao desejado, e em seguida são feitos ajustes ao modelo em si por meio do retreinamento da rede neural utilizada. Dessa forma, os ajustes permitem que a capacidade de geração do modelo seja expandida em tempo de descompressão para incluir o sinal desejado, tomando o pivô como uma aproximação inicial.

Os resultados dos experimentos realizados indicam que a técnica proposta é bem sucedida na tarefa de redução do erro de representação, alcançando reduções de até 30% quando comparada ao DIP e 85% quando comparada ao CSGM tradicional. Tais resultados são favoráveis à implementação da técnica em sistemas como o de IoT, uma vez que a redução do volume de amostras que precisa ser transmitido promove ganhos energéticos aos dispositivos envolvidos. Apesar do ganho obtido, resultados evidenciam também que o CSGM-PTI requer um maior número de iterações para que o erro da reconstrução convirja. Além disso, foi realizada também comparações quantitativa e qualitativa entre a técnica proposta e o WT. Apesar de obter menor desempenho de reconstrução pela métrica utilizada, uma análise qualitativa das imagens geradas pelo CSGM-PTI indica que ele é capaz de preservar a estrutura geral de características relevantes da imagem comprimida.

Este trabalho está organizado da seguinte forma. A Seção 2 discute os trabalhos relacionados. A Seção 3 introduz inicialmente conceitos básicos de sensoriamento compressivo e modelos generativos. A seguir, a Seção 4 propõe o emprego de ajuste por pivô como forma de redução do erro de representação de sinais. A Seção 5 apresenta a metodologia utilizada e os resultados dos experimentos. Por fim, na Seção 6 é feita uma discussão quanto as conclusões que podem ser tomadas e as possibilidades de trabalhos futuros.

2. Trabalhos Relacionados

Devido às limitações dos modelos utilizados em CSGM, diversas alternativas para a redução do erro de representação existem na literatura. Uma técnica especialmente relevante para a proposta do presente trabalho é o *Deep Image Prior* (DIP) [Ulyanov et al. 2018], que pode ser considerado como uma generalização do PTI. O DIP é uma técnica de resolução de problemas inversos em processamento de imagens que se baseia no uso de uma rede neural artificial (*Artificial Neural Network* - ANN) com pesos inicializados em valores aleatórios. No contexto de CSGM, a entrada da rede é fixada, enquanto seus pesos são ajustados em tempo de descompressão para que a saída da rede adéque-se ao sinal que se deseja reconstruir. Dessa forma, elimina-se o problema do erro de representação, já que não há mais uma busca pelo sinal em um modelo previamente treinado e, conseqüentemente, o viés associado a tal treinamento. Contudo, tal abordagem não faz proveito dos modelos de alta capacidade de representação disponíveis na literatura, como o StyleGAN-2 [Karras et al. 2020b]. Espera-se que com o uso de tais modelos seja possível obter ganhos de desempenho.

Daras et al. propõem a técnica de Otimização de camadas intermediárias (*Intermediate Layer Optimization* - ILO) [Daras et al. 2021] que, de forma análoga ao DIP e ao PTI, realiza ajustes pontuais à rede geradora com o intuito de expandir sua capacidade de geração. Entretanto, a ILO difere quanto a abordagem tomada para a implementação de tais ajustes. Após uma etapa de inversão da rede geradora G utilizada no processo de des-

compressão, fixa-se a entrada da rede, z , e é realizada uma divisão do gerador em camadas intermediárias. Com tais camadas, é possível executar ajustes e obter representações intermediárias mais adequadas ao sinal desejado, $z_i = G_i(z_{i-1})$. Tais representações são, por sua vez, utilizadas em sucessivas etapas de inversão e ajuste das camadas intermediárias. Esse processo resulta em uma rede geradora dividida em um número de sub-redes com entradas ajustadas especificamente para expandir a capacidade de geração da rede completa, de forma a incluir o sinal que se deseja reconstruir.

Em contraponto, Asim et al. propõem o uso de modelos generativos inversíveis no lugar das GANs. Os autores argumentam que, como esse tipo de modelo é construído a partir de redes matematicamente inversíveis, o uso dessa arquitetura elimina, por definição, o erro de representação [Asim et al. 2020]. Entretanto, o treinamento de modelos generativos inversíveis pode ser desafiador e custoso. Além disso, como tais modelos não apresentam a mesma presença ubíqua na literatura que as GANs, o acesso a modelos pré-treinados pode ser dificultado.

Neste trabalho, é proposto o uso da técnica do ajuste por pivô, ou PTI, no contexto de CSGM. Diferente de outras abordagens da literatura, o CSGM-PTI, além de recorrer à estrutura clássica do CSGM, adapta para o contexto de sensoriamento compressivo a ideia de retreinamento da rede geradora para a redução do erro de representação. Além disso, a limitação dos ajustes por meio de um regularizador local, evita que a técnica cause distorções em outros sinais que o modelo generativo pode gerar. Por fim, esse esquema permite que se tome proveito da disponibilidade de modelos generativos com alta capacidade de geração, como a StyleGAN-2.

3. Fundamentação teórica

Esta seção introduz conceitos básicos de sensoriamento compressivo e modelos generativos. Em seguida, a aplicação do ajuste por pivô em CSGM é proposta.

3.1. Sensoriamento compressivo

O paradigma tradicional de sensoriamento compressivo afirma que é possível recuperar um sinal x de dimensão N a partir de um conjunto de amostras y de dimensão $M < N$, desde que x seja compressível. O sinal x é compressível se este possuir uma representação k -esparsa em alguma base conhecida, tal que:

$$y = A \cdot x + \eta, \quad (1)$$

onde A é uma matriz de amostragem e η representa o possível ruído associado ao processo de amostragem. Como $M < N$, tal esquema configura um sistema indeterminado de equações, havendo infinitas soluções para x que poderiam gerar y . Porém, como se assume que o sinal original é compressível, pode-se reformular o problema como:

$$y = A \cdot \psi \cdot s + \eta, \quad (2)$$

em que ψ representa a base em que o sinal original é k -esparso e s é a sua representação esparsa. Dessa forma, para recuperar o sinal original, basta impor como limitação à solução do problema que s seja esparsa.

A busca pela solução que garanta a maior esparsidade configura um problema NP-difícil, já que seria necessário minimizar a norma l_0 . Entretanto, Candes et al. demonstram que a minimização da norma l_1 serve como uma aproximação quando a matriz de amostragem utilizada respeita condições como a propriedade da isometria restrita (*Restricted Isometry Property* - RIP) [Candes et al. 2006, Candes 2008]. Essa propriedade define como deve ser construída a matriz A que descreve o esquema de amostragem utilizado. Apesar da restrição imposta, em muitos casos o uso de matrizes aleatórias é suficiente para garantir conformidade com a RIP. Portanto, o sinal x pode ser recuperado por meio do seguinte problema de otimização:

$$\min \|s\|_{l_1} \quad | \quad y = A \cdot \psi \cdot s + \eta, \quad (3)$$

que pode ser resolvido por algoritmos de otimização linear convexa conhecidos na literatura, como o *basis pursuit* [Chen and Donoho 1994], que utiliza técnicas de programação linear para decompor o sinal em sua representação esparsa.

3.2. Modelos generativos

Modelos generativos podem ser compreendidos como técnicas de aprendizado não supervisionado em que, a partir de um conjunto de amostras, obtém-se uma distribuição de probabilidade que aproxima a distribuição referente às amostras utilizadas. Dessa forma, um dos desafios associados à modelagem generativa refere-se a como obter um modelo que, a partir de um conjunto de dados, consiga gerar dados semelhantes aos originais. O processo realizado para obtenção de tal aproximação difere para cada técnica. Entretanto, as diferentes abordagens de modelagem generativa podem ser agrupadas quanto a possibilidade de acesso, explícito ou implícito, à distribuição de probabilidade obtida.

As redes adversárias generativas [Goodfellow et al. 2014] são um exemplo de modelo generativo implícito. Isso é, elas permitem a geração de dados sem que haja a definição explícita da densidade de probabilidade, que pode ser de difícil manipulação. Para tal, a técnica aprende um processo de geração de amostras feito pelo treinamento concomitante de duas redes neurais artificiais (*Artificial Neural Networks* - ANN).

Uma característica central das GANs é o uso de uma estratégia adversarial de treinamento. Ou seja, é construída uma rede geradora G , que deve competir com uma segunda rede discriminadora D . Enquanto G tem como objetivo gerar amostras convincentes da distribuição de probabilidade que se deseja aproximar, D é treinada para classificar suas entradas como provenientes do conjunto de dados original ou geradas por G . Dessa forma, espera-se que à medida que D aprimore a sua capacidade de classificação, G tenha que criar amostras que melhor aproximem o conjunto original. Como resultado, tem-se uma rede geradora que é capaz de mapear um vetor de ruído z como entrada para uma amostra que seja compatível com o conjunto de dados oriundo de uma distribuição p_x .

3.3. Sensoriamento compressivo usando modelos generativos

Em contrapartida ao método tradicional de sensoriamento compressivo, o modelo proposto por Bora et al. não depende mais da esparsidade para garantir a recuperação do sinal [Bora et al. 2017]. Em compensação, a CSGM utiliza como estrutura a presença do sinal desejado no alcance de um modelo generativo modelado como $G : \mathbb{R}^k \rightarrow \mathbb{R}^N$, tal

que o problema pode ser reescrito da seguinte forma:

$$\min \|A \cdot G(z) - y\|_{l_2} \quad | \quad y = A \cdot G(z) + \eta, \quad (4)$$

sendo $G(z)$ a saída do modelo generativo. Essa formulação altera a maneira como é realizada a busca pela solução do problema. A obtenção da resposta que apresente representação esparsa na base desejada torna-se desnecessária. Em contrapartida, a busca se direciona ao vetor $z \in \mathbb{R}^k$ que, quando aplicado em $(A \cdot G(z))$, gera um conjunto de medições condizentes com as amostras inicialmente tomadas em y . Para o caso das GANs, utilizadas no presente trabalho, tal processo de busca é conhecido na literatura como inversão de GAN [Xia et al. 2022].

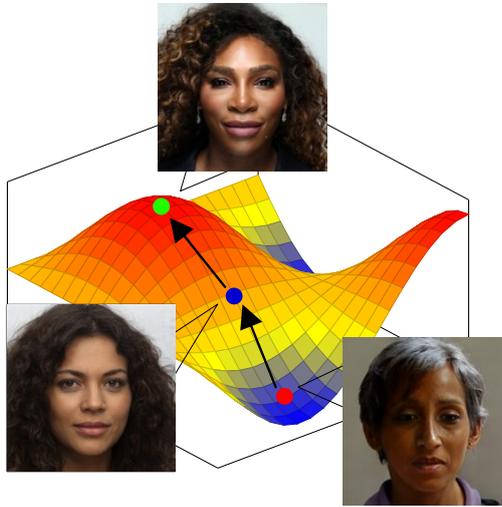


Figura 1. Etapas da inversão de GAN.

A Figura 1 ilustra a inversão de uma GAN utilizando a técnica do gradiente descendente. O ponto vermelho representa uma tentativa inicial de encontrar o sinal desejado, podendo-se utilizar como valor inicial de z um vetor aleatório. Após algumas etapas do algoritmo, espera-se que, ao minimizar o erro de reconstrução entre as medições dos sinais original e gerado, ocorram gradativas aproximações entre os sinais. Tal aproximação é representada na figura pelo ponto intermediário azul. Por fim, caso o sinal original seja bem representado pelo modelo generativo escolhido, aplicando o método do gradiente, torna-se possível encontrar a solução ótima, ilustrada no ponto verde.

4. Ajuste por Pivô em Sensoriamento Compressivo com Modelos Generativos

Originalmente desenvolvido para aplicações em edição de imagens por meio de GANs, a técnica do ajuste por pivô, ou PTI (*Pivotal Tuning for Latent Based Editing of Real Images*), permite que após o processo de inversão de uma GAN, faça-se uso do z' obtido como um ponto de pivô para a execução de ajustes precisos no modelo. Com o retreinamento da rede é possível expandir a sua capacidade de representação e potencialmente incluir nela o sinal original. Para tal, são realizadas mudanças quanto a função de custo referente ao processo de inversão, que deixa de tomar como fixos os pesos

da GAN utilizada.

$$L_{pivo}(G(z'; \theta'), x) = \lambda_{LPIPS} \cdot LPIPS(G(z'; \theta'), x) + \lambda_{l2} \cdot \|G(z'; \theta') - x\|_{l2} \quad (5)$$

A Equação 5 apresenta a configuração referente à parcela central da função de custo utilizada no ajuste de pivô, o L_{pivo} . Construída a partir da associação de duas métricas, a $LPIPS$ [Zhang et al. 2018] e a norma $l2$, a formulação apresentada conta com dois hiperparâmetros escalares λ_{LPIPS} e λ_{l2} para o balanceamento entre as métricas. Tanto a $LPIPS$ quanto a norma $l2$ são calculadas a partir da comparação entre o sinal original x e a saída da rede geradora que tem pesos θ' e como entrada o valor fixado como pivô z' . Adicionalmente, é incluída como uma parcela da função de custo total a L_{local} . Tal inclusão é feita para garantir a manutenção do comportamento do modelo em um escopo global, ou seja, evitar que com o ajuste realizado, seja perdida a capacidade de geração de sinais em outras regiões do alcance do modelo. A Equação 6 descreve a função de custo total que deve ser minimizada durante o ajuste por pivô.

$$L_{total}(G(z'; \theta'), x) = L_{pivo}(G(z'; \theta'), x) + L_{local}(G(z'; \theta'), x) \quad (6)$$

Dessa forma, a etapa do ajuste por pivô é um retreinamento da rede neural utilizada para que ela possa gerar um sinal que corresponda às medições y . O procedimento executado para a obtenção de tal resultado é o problema de otimização descrito pela Equação 7.

$$\min L_{pivo}(G(z'; \theta'), x) + L_{local}(G(z'; \theta'), x) \quad | \quad y = G(z'; \theta') \quad (7)$$

Entretanto, como no contexto de sensoriamento compressivo não se tem acesso ao sinal original, somente um conjunto reduzido de amostras y , não é possível implementar a abordagem original do PTI. O presente trabalho, portanto, propõe uma adaptação da técnica de edição de imagens PTI para que ela possa ser utilizada no contexto de CSGM, fazendo com que o ajuste por pivô baseando-se em sinais subamostrados possa ser realizado. Para esse fim, é necessário realizar alterações na função de custo do ajuste por pivô. Tal alteração é constituída pela compressão por subamostragem dos sinais envolvidos. Além disso, como a métrica $LPIPS$ foi desenvolvida para o uso com imagens não comprimidas, o seu uso pode gerar distorções no processo de treinamento, de forma que opta-se por manter apenas a norma $l2$. Dessa forma, a função de custo pode ser reformulada como:

$$L_{cs}(A \cdot G(z'; \theta'), y) = \lambda_{l2} \cdot \|A \cdot G(z'; \theta') - y\|_{l2}, \quad (8)$$

em que L_{cs} é a nova função de custo associada ao ajuste por pivô em CSGM e y é o conjunto de amostras capturadas diretamente do sinal original pelo mesmo esquema de amostragem descrito pela matriz de amostragem A . O problema de otimização que deve ser solucionado para a reconstrução do sinal passa a ser então descrito como:

$$\min \lambda_{l2} \cdot \|A \cdot G(z'; \theta') - y\|_{l2} + L_{local}(A \cdot G(z'; \theta'), y) \quad | \quad y = A \cdot G(z'; \theta') + \eta, \quad (9)$$

em que se inclui a parcela que garante a coerência global do modelo atualizada com os sinais comprimidos e também o termo η , referente ao possível erro associado ao processo de amostragem.

A abordagem proposta é utilizada com o objetivo de melhorar o desempenho do CSGM por meio da redução do erro de representação. Tal erro é observado no contexto do CSGM quando o modelo generativo utilizado não é capaz de gerar o sinal que se deseja reconstruir. Nesse caso, o processo de inversão tradicionalmente utilizado na técnica resulta na melhor aproximação que o modelo é capaz de fornecer, o que é exemplificado na Figura 2 pelo ponto verde.

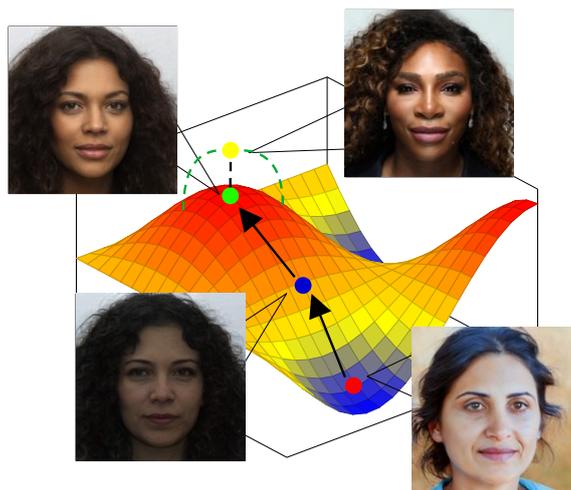


Figura 2. Erro de representação e ajuste por pivô.

No caso de deficiência da capacidade de representação do modelo generativo, a técnica do ajuste por pivô tem como função realizar o retreinamento da rede geradora para possibilitar que ela gere o sinal desejado. Na Figura 2, esse processo é ilustrado por um deslocamento pontual da superfície que representa o modelo para passar a incluir o ponto amarelo. Note que o CSGM-PTI é uma composição de duas etapas: a inversão de GANs tradicionalmente utilizada no contexto de CSGM e o ajuste por pivô. Dessa forma, a técnica se posiciona de forma promissora no escopo de propostas de mitigação de erro de representação. Isso porque o CSGM-PTI se enriquece com avanços de dois populares temas da literatura acadêmica: inversão de GANs e treinamento de redes neurais.

5. Experimentos

Experimentos comparativos entre a qualidade da reconstrução utilizando o CSGM-PTI e outras técnicas tradicionais da literatura foram realizados para avaliar a capacidade da técnica proposta em promover ganhos de desempenho. Uma das técnicas comparadas foi o *Deep Image Prior* (DIP) [Ulyanov et al. 2018], apresentada na Seção 2. Essa escolha foi feita devido à construção análoga das técnicas. Pode-se entender que o CSGM-PTI funciona como o DIP, porém tomando proveito de modelos pré-treinados disponibilizados na literatura para gerar uma aproximação inicial. Devido à sua ampla utilização na literatura e sua vasta capacidade de geração de imagens, o modelo escolhido para todos os experimentos realizados foi a implementação do StyleGAN2-ADA [Karras et al. 2020a] com o uso da base de dados de figuras de faces humanas FFHQ (*Flickr-Faces-HQ*) [Karras et al. 2018]. Esse é um conjunto de dados com 70000 figuras de alta qualidade amplamente utilizado literatura, especialmente na avaliação GANs. A

seguir são apresentados os resultados de experimentos realizados com o intuito de avaliar o desempenho da técnica proposta por meio da comparação com as abordagens tradicionais DIP e WT.

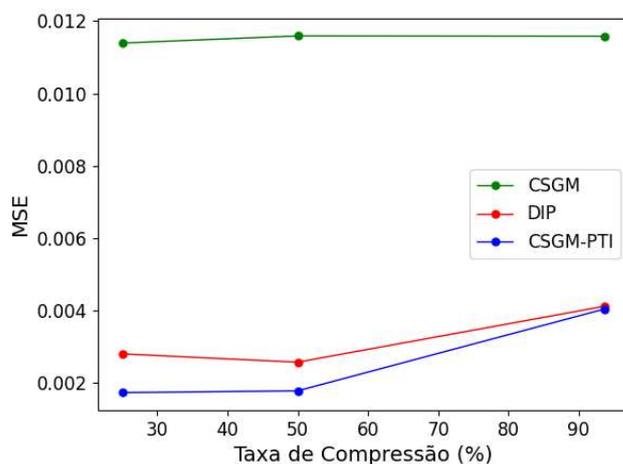


Figura 3. Comparação entre os erros de reconstrução referentes ao uso das técnicas CSGM, DIP e CSGM-PTI para diferentes taxas de compressão.

A Figura 3 apresenta os resultados dos experimentos que comparam a técnica proposta com o DIP e o CSGM tradicional, ou seja, CSGM sem o ajuste do pivô (PTI). O gráfico apresenta a relação do erro de reconstrução medido pelo erro médio quadrático (*Mean Square Error* - MSE) com a taxa de compressão utilizada em cada experimento. Os erros são comparados para imagens comprimidas às taxas de 25%, 50% e 93,75%. É possível observar ganhos de desempenho para todas as taxas de compressão com relação tanto ao CSGM quanto ao DIP, alcançando reduções de até 30% quando comparada ao DIP e 85% quando comparada ao CSGM. Além disso, os resultados demonstram as limitações do CSGM sem o PTI, já que é possível observar que o valor do erro mantém-se praticamente constante para todas as taxas de compressão. Isso ocorre devido à limitação inerente ao gerador utilizado, que não é capaz de representar de forma adequada o sinal que se deseja reconstruir, gerando o erro de representação. Nota-se também que o ganho de desempenho diminui à medida que se aumenta a taxa de compressão, tanto para a técnica proposta quanto para o DIP.

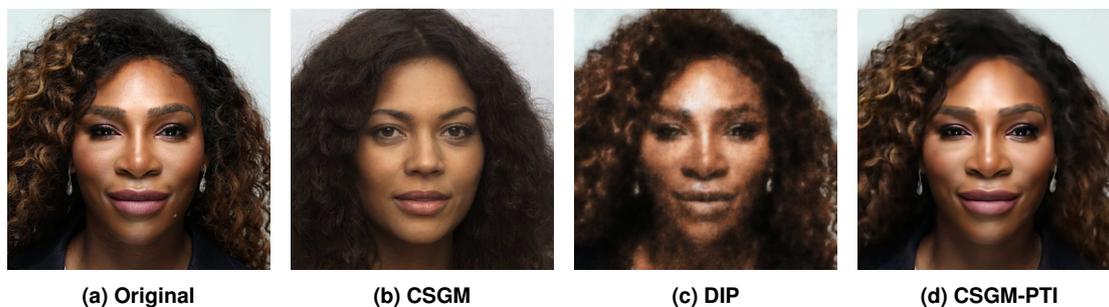


Figura 4. Imagens resultantes do processo de compressão e descompressão utilizando as técnicas CSGM, DIP e CSGM-PTI. A taxa de compressão utilizada foi de 25%.

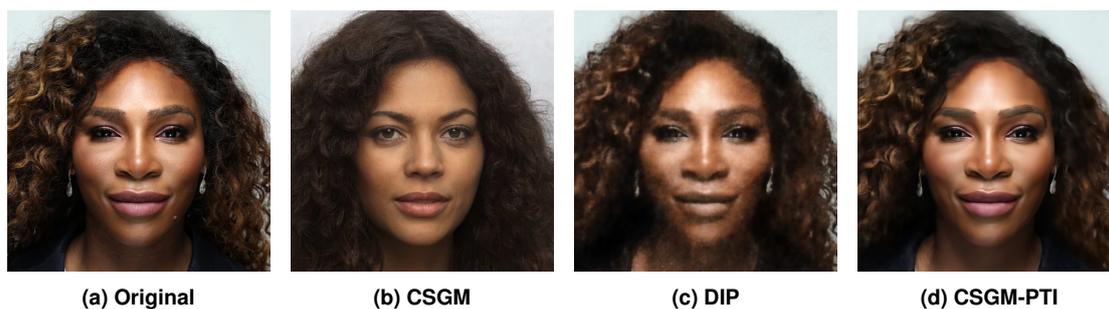


Figura 5. Imagens resultantes do processo de compressão e descompressão utilizando as técnicas CSGM, DIP e CSGM-PTI. A taxa de compressão utilizada foi de 50%.

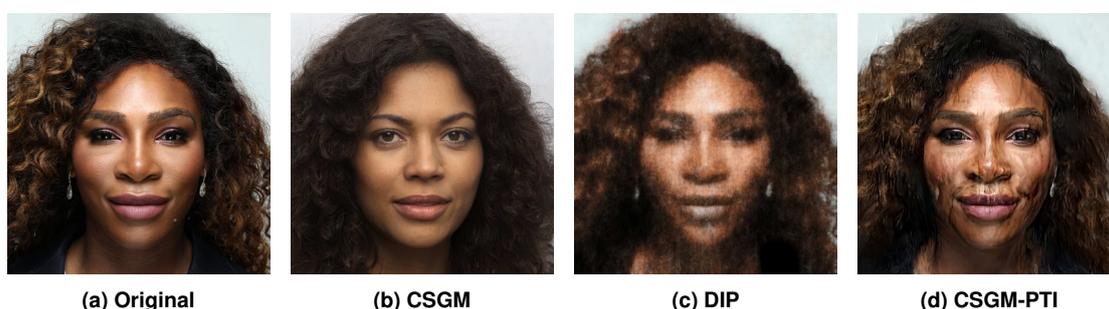


Figura 6. Imagens resultantes do processo de compressão e descompressão utilizando as técnicas CSGM, DIP e CSGM-PTI. A taxa de compressão utilizada foi de 93,75%.

Nas Figuras 4, 5 e 6 são apresentados as imagens que cada técnica gerou após seus processos de compressão e descompressão para cada taxa de compressão empregada. Note o erro de representação evidente entre as imagens geradas pelo CSGM em comparação às outras técnicas. Note ainda que apenas para taxas muito elevadas de compressão o desempenho do DIP se aproxima do desempenho do CSGM-PTI, confirmando os resultados da Figura 3.

Outra análise realizada compara a quantidade de iterações em que o erro de reconstrução leva para convergir no DIP e no CSGM-PTI. Para ambos os procedimentos foram realizadas compressões às taxas de 50% e 93,75%. Além disso, o processo de descompressão foi limitado em 500 iterações, sendo que, no caso do CSGM-PTI, as 200 iterações iniciais referentes ao processo de inversão de GAN e as 300 finais referentes ao processo de ajuste por pivô.

A Figura 7 apresenta os resultados obtidos. É possível observar que, apesar das vantagens quanto ao erro de reconstrução final quando comparada ao DIP, como visto na Figura 3, o CSGM-PTI precisa de um número de iterações superior para alcançar a convergência. Outra observação importante obtida a partir do gráfico é que a taxa de redução do MSE com relação ao número de iterações realizadas permanece semelhante quando se altera a taxa de compressão. Por fim, os resultados mostram a importância do uso do ajuste por pivô, uma vez que é possível verificar que somente a partir da iteração 200, quando se inicia o ajuste, a técnica consegue superar o DIP.

A Figura 8 apresenta os resultados da comparação entre a técnica proposta,

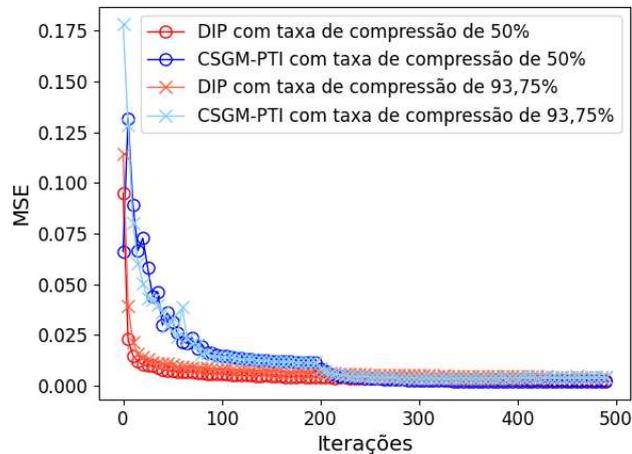


Figura 7. Comparação entre os erros de reconstrução referentes ao uso das técnicas DIP e CSGM-PTI para realizar compressões de 50% e 93,75%. Os erros são apresentados em função do número de iterações realizadas.

CSGM-PTI, e o WT (*Wavelet Thresholding*) que segue o paradigma clássico de compressão. Em contraste às estratégias de sensoriamento compressivo, com o WT não é feita a compressão por subamostragem. Nessa abordagem, o sinal completo é submetido a uma transformada do tipo Wavelet, sendo os coeficientes resultantes filtrados, de forma que apenas os de maior intensidade são mantidos. A compressão é feita portanto pela redução do número de coeficientes da transformada do sinal, e não da subamostragem do mesmo. Tal esquema de compressão apresenta semelhanças a partes da ubíqua técnica de compressão de imagens JPEG2000.

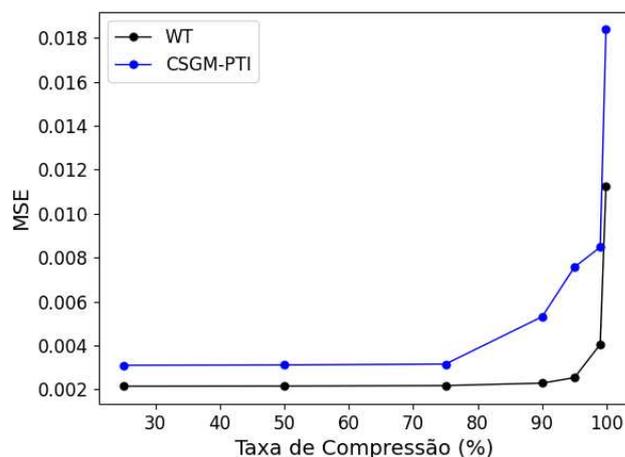


Figura 8. Comparação entre os erros de reconstrução referentes ao uso das técnicas WT e CSGM-PTI para diferentes taxas de compressão.

Os resultados apresentados na Figura 8 indicam que, para todas as taxas de compressão analisadas, o WT obteve menor MSE do que a técnica proposta. Apesar de apresentar melhor desempenho na métrica utilizada, quando são aplicadas altas taxas de com-

pressão, o WT pode eliminar elementos da imagem relevantes para tarefas fundamentais para classificação.



Figura 9. Resultado do processo de compressão e descompressão utilizando o CSGM-PTI a taxas de compressão de 99% e 99,9%.

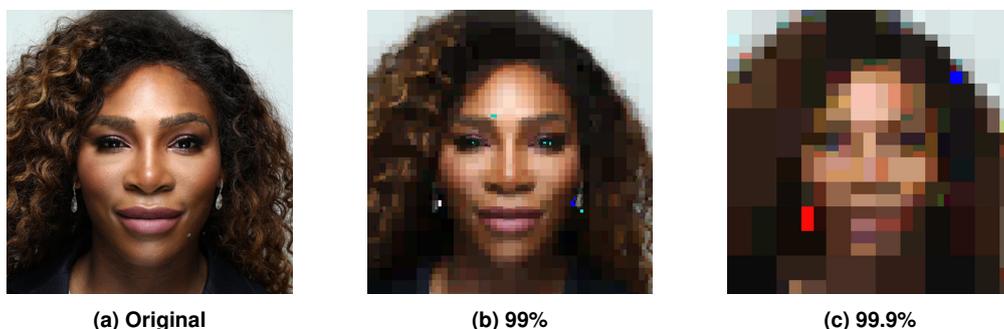


Figura 10. Resultado do processo de compressão e descompressão utilizando o WT a taxas de compressão de 99% e 99,9%.

As Figuras 9 e 10 apresentam uma comparação qualitativa entre a distorção causada pela compressão utilizando o CSGM-PTI e o WT, respectivamente. É possível observar que, à medida que a taxa de compressão se aproxima de 100%, o uso do WT pode dificultar o reconhecimento de alguns dos componentes mais relevantes da imagem original. Os artefatos resultantes dessa compressão mascaram características como o nariz, boca e os olhos. Apesar de também causar distorção, a compressão por meio do CSGM-PTI mantém reconhecíveis algumas dessas características. A manutenção da estrutura de características fundamentais da imagem original pode ser relevante para a tarefa de classificação da figura como uma face humana, por exemplo.

6. Conclusões

Este trabalho apresentou uma proposta para a redução do erro de representação em sensoriamento compressivo com modelos generativos. O CSGM-PTI (*Compressive Sensing using Generative Models - Pivotal Tuning for Latent-based Editing of Real Images*) é fruto da adaptação de uma técnica de edição de imagens que torna possível a sua utilização no contexto de sensoriamento compressivo, onde não há acesso ao sinal completo. Com a técnica proposta é possível reduzir o volume de dados que dispositivos de baixa disponibilidade energética e computacional transmitem em sistemas como redes de sensores. Além disso, o procedimento de compressão em si, é facilitado, sendo transferidas para nós com maior disponibilidade de recursos as etapas mais custosas. A

composição do CSGM-PTI por dois procedimentos distintos apresenta-se como uma vantagem à técnica, já que ambos os seus componentes são de relevante presença na literatura e seus respectivos avanços podem garantir maiores ganhos de desempenho à técnica.

Para a avaliação do CSGM-PTI, foram realizados experimentos comparando-a com técnicas tradicionalmente utilizadas na literatura. Foi possível observar que, para as diferentes taxas de compressão analisadas, a técnica proposta apresentou menor erro de reconstrução do que o DIP (*Deep Image Prior*) e o CSGM tradicional, ou seja, sem o PTI. Apesar do ganho de desempenho, a técnica proposta requer mais iterações para que o seu erro de reconstrução alcance a convergência quando comparada ao DIP. Outro ponto analisado foi a comparação do CSGM-PTI com uma técnica do paradigma tradicional de compressão, a WT (*Wavelet Thresholding*). Em uma análise quantitativa, a abordagem tradicional apresentou menor erro de reconstrução para todas as taxas de compressão analisadas. Entretanto, uma análise qualitativa das imagens resultantes do processo de compressão pelos dois métodos indica que a técnica proposta promove a manutenção da estrutura de algumas das características centrais da imagem original.

As análises realizadas indicam alguns caminhos que podem ser seguidos para promover futuros desenvolvimentos à técnica proposta. Dentre eles, deve-se considerar o estudo dos efeitos da distorção causada pelo CSGM-PTI em tarefas como a de classificação, comparando-os com os efeitos da compressão de paradigma clássico. Além disso, o estudo do impacto do tempo necessário para a reconstrução dos sinais comprimidos no desempenho de comunicação em redes é essencial para considerar a viabilidade do uso da técnica em aplicações reais.

Agradecimentos

O presente trabalho foi realizado com apoio do CNPq, da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, FAPERJ (E-26/211.144/2019), FAPESP (2015/24494-8, 2014/50937-1, 2015/24485-9) e RNP (Programa de Monitoramento (PMon) 2022).

Referências

- Abbasian Dehkordi, S., Farajzadeh, K., Rezazadeh, J., Farahbakhsh, R., Sandrasegaran, K., and Abbasian Dehkordi, M. (2020). A survey on data aggregation techniques in iot sensor networks. *Wireless Networks*, 26(2):1243–1263.
- Asim, M., Daniels, M., Leong, O., Ahmed, A., and Hand, P. (2020). Invertible generative models for inverse problems: mitigating representation error and dataset bias. In *International Conference on Machine Learning*, pages 399–409. PMLR.
- Bora, A., Jalal, A., Price, E., and Dimakis, A. G. (2017). Compressed sensing using generative models. In *International Conference on Machine Learning*, pages 537–546. PMLR.
- Candes, E., Romberg, J., and Tao, T. (2006). Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509.
- Candes, E. J. (2008). The restricted isometry property and its implications for compressed sensing. *Comptes rendus mathématique*, 346(9-10):589–592.

- Chen, S. and Donoho, D. (1994). Basis pursuit. In *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 41–44 vol.1.
- Daras, G., Dean, J., Jalal, A., and Dimakis, A. G. (2021). Intermediate layer optimization for inverse problems using deep generative models.
- Donoho, D. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial networks.
- 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. (2018). A style-based generator architecture for generative adversarial networks.
- Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. (2020b). Analyzing and improving the image quality of stylegan. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8107–8116.
- Kaur, M. and Munjal, A. (2020). Data aggregation algorithms for wireless sensor network: a review. *Ad Hoc Networks*, 100:102083.
- Ketshabetswe, K. L., Zungeru, A. M., Mtengi, B., Lebekwe, C. K., and Prabakaran, S. (2021). Data compression algorithms for wireless sensor networks: A review and comparison. *IEEE Access*, 9:136872–136891.
- Nakas, C., Kandris, D., and Visvardis, G. (2020). Energy efficient routing in wireless sensor networks: a comprehensive survey. *Algorithms*, 13(3):72.
- Roich, D., Mokady, R., Bermano, A. H., and Cohen-Or, D. (2022). Pivotal tuning for latent-based editing of real images. *ACM Transactions on Graphics (TOG)*, 42(1):1–13.
- Singh, J., Kaur, R., and Singh, D. (2021). Energy harvesting in wireless sensor networks: A taxonomic survey. *International Journal of Energy Research*, 45(1):118–140.
- Ulyanov, D., Vedaldi, A., and Lempitsky, V. (2018). Deep image prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Wu, Y., Rosca, M., and Lillcrap, T. (2019). Deep compressed sensing. In Chaudhuri, K. and Salakhutdinov, R., editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 6850–6860. PMLR.
- Xia, W., Zhang, Y., Yang, Y., Xue, J.-H., Zhou, B., and Yang, M.-H. (2022). Gan inversion: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. *CoRR*, abs/1801.03924.