

Novas Métricas de Avaliação de Qualidade para Imagens Petrofísicas

Rewbenio Araújo Frota

*PETROBRAS e Departamento de Eng. Elétrica
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) Centro de Tecnologia, Universidade Federal do Ceará (UFC)*
Rio de Janeiro, Brasil

Guilherme de Alencar Barreto

*Programa de Pós-Graduação em Eng. de Teleinformática
Centro de Tecnologia, Universidade Federal do Ceará (UFC)*
Fortaleza, Brasil

Marley Maria Bernardes Rebuzzi Vellasco

*Departamento de Eng. Elétrica
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)*
Rio de Janeiro, Brasil

Candida Menezes de Jesus

PETROBRAS
Rio de Janeiro, Brasil

Abstract—O emprego de métricas tradicionais de visão computacional para avaliação da qualidade de imagem (IQA, do inglês *Image Quality Assessment*), sejam elas clássicas ou mais recentes, baseadas em aprendizado profundo, mostrou-se inconclusivo no que diz respeito à avaliação de perfis de imagem acústica de paredes de poços de petróleo geradas por modelos generativos. A metodologia BERGSOM, por sua vez, utiliza aprendizado regenerativo para obter tais perfis de imagem a partir de perfis básicos, como perfis de raios gama, densidade e porosidade. Este trabalho emprega os modelos treinados para a metodologia BERGSOM com o propósito de propor duas novas métricas: z_{IQA} e b_{IQA} . Ademais, são incluídos limiares de qualidade para as métricas propostas, com base nos dados de treinamento. A aplicação desses limiares permite definir se as entradas empregadas para a geração dos perfis de imagem por meio da metodologia BERGSOM estão significativamente distantes das condições de treinamento do modelo. As novas métricas demonstraram maior coerência em comparação com as métricas tradicionais de visão computacional. Além da coerência apresentada, a métrica b_{IQA} tem a vantagem adicional de poder ser calculada em diferentes situações reais, o que não ocorre com as demais métricas avaliadas, que demandam uma imagem de referência para seu cálculo.

Index Terms—Avaliação de Qualidade de Imagens, Rede SOM, Petrofísica

I. INTRODUÇÃO

Perfis de imagem de parede de poços representam um dos dados mais importantes¹ para auxiliar na caracterização de rochas em subsuperfície e atestar suas propriedades porosas, que são de interesse atualmente para diversas indústrias, tais quais produção de óleo e gás [1], estocagem geológica de carbono [2] e hidrogênio [3], além de energia geotérmica [4]. Tais perfis de imagem de paredes de poços são derivados de

¹Além dos perfis de imagem, a indústria também adquire perfis mais simples, ora chamados perfis básicos, que registram propriedades da rocha como densidade, porosidade e saturação de fluidos ao longo do poço. Os perfis básicos, têm baixo custo de aquisição e são adquiridos em quase todos os poços perfurados. Perfis de imagem acústica contêm maior conteúdo de informação, o que acarreta maior custo de aquisição. Além disso, apresentam maior propensão a falhas de aquisição e têm processamento mais lento, o que pode impactar a tomada de decisões em tempo real.

medições acústicas realizadas nos poços e são tipicamente avaliadas por especialistas humanos que julgam sua qualidade com base em seu potencial para identificar estruturas geológicas e diferenciá-las de artefatos produzidos durante a perfuração.

Modelos generativos baseados em Aprendizado de Máquinas para geração [5] ou melhoria [6] de perfis de imagem, trouxeram a necessidade de avaliar numericamente sua qualidade. Nesse caso, o mais usual é tomar métricas de avaliação de qualidade de imagem (IQA, do inglês *Image Quality Assessment*) já consagradas da visão computacional, como MSE e SSIM [7], [8], ou mais recentes baseadas em aprendizado profundo, como LPIPS [9]. Contudo, trabalhos recentes, e.g. [5] que avaliam perfis de imagem gerados por modelos generativos usando métricas IQA convencionais de visão computacional sugerem que elas podem levar a avaliações divergentes, o que sugere que tais métricas não sejam coerentes na avaliação desse tipo de dado. Um dos problemas apresentados, por exemplo, é que todas as métricas citadas acima são do tipo FR (*full reference*), i.e., precisam sempre de uma imagem de referência para que sejam calculadas, o que não é realista na maior parte dos casos. Outro problema é que diferentes métricas avaliam de forma diferente conjuntos de imagem com variados graus de degradação visual.

Este trabalho propõe duas novas métricas IQA: uma do tipo FR e outra do tipo NR (*no reference*), i.e., que não precisa de uma imagem de referência para seu cálculo. Ambas baseadas numa recente metodologia de aprendizado regenerativo [5] chamada BERGSOM (do inglês *Bidirectional Expanded ReGenerative SOM*), capaz de realizar geração cruzada bidirecional entre dados de diferentes modalidades. Essa metodologia permite estimar perfis sintéticos de imagem de poços a partir de perfis mais simples², que são mais rápidos e menos custosos em se adquirir. A metodologia

²Neste trabalho, os perfis de imagem são estimados a partir de sete perfis básicos de entrada: GR, DEN, NEU, DTC, NMR PhiT, PE, log(*RES*).

é ainda capaz de prover essas novas métricas IQA que se mostraram coerentes com a degradação visual observada nas imagens, além de serem mais consistentes que as métricas IQA tradicionais para variações da imagem dentro do mesmo poço. Outra vantagem é que uma das métricas é do tipo NR, ou seja, sempre pode ser calculada em casos reais.

A continuação deste artigo apresenta uma breve revisão da metodologia BERGSOM e sua aplicação na estimativa de perfis de imagem acústica em poços a partir de perfis básicos. Em seguida, introduzem-se as novas métricas IQA e seu desempenho é analisado em relação às métricas tradicionais. Por fim, as principais conclusões são apresentadas e discutidas.

II. REVISÃO DA METODOLOGIA BERGSOM PARA REGERAÇÃO DE PERFIS DE IMAGEM

A metodologia BERGSOM [5] e suas variações [10], [11] realizam aprendizado multimodal baseado em memória heteroassociativa, usando a rede SOM (*Self-Organizing Map*) [12]. BERGSOM permite inferência cruzada entre diferentes modalidades de dados, em especial é um método efetivo para aprendizado regenerativo³ [13], que ocorre quando dados da modalidade mais complexa, em termos de densidade de informação, são estimados a partir de entradas de modalidade mais simples (por exemplo, gerar imagens a partir de séries temporais). BERGSOM é baseado na rede SOM expandida, i.e., treinada com uma entrada aumentada \mathbf{x}_{aug} ⁴ composta pela concatenação de dados que representam modalidades diferentes:

$$\mathbf{x}_{aug} = [z_1, \dots, z_m, y_1, \dots, y_n]^T, \quad (1)$$

na qual m e n são as dimensões da representação (\mathbf{z}) da imagem e do vetor de perfis básicos (\mathbf{y}), respectivamente.

Além da rede SOM, BERGSOM usa um autocodificador convolucional variacional (β -CoVAE⁵), cujos objetivos são: (i) aprender uma representação compacta ($\mathbf{z} \in \mathcal{Z}$) dos perfis de imagem; (ii) prover um modelo generativo para perfis de imagem a partir de representações latentes (\mathbf{z}). A Fig. 1 ilustra

³Opta-se pelo termo “aprendizado regenerativo” para traduzir *regeneration learning*, a exemplo da tradução “aprendizado generativo” para *generative learning*. Enquanto se diz que dados são *gerados* por aprendizado generativo, usa-se aqui o termo dados *regerados* por aprendizado regenerativo. Não se usa o verbo regenerar, pois esse termo indica outro tipo de aplicação, por exemplo, recuperação ou regeneração de dados corrompidos por diferentes técnicas.

⁴Este artigo segue a notação a seguir: (i) p é o número de exemplos de treinamento e q é o número de exemplos de teste; (ii) $\mathbf{x} \in \mathbb{R}^{26 \times 180}$, é um segmento de perfil de imagem a uma dada profundidade, tal que $X = [\mathbf{x}_1 \dots \mathbf{x}_p]$; (iii) $\mathbf{y} \in \mathbb{R}^n$, é um vetor com n ($= 7$) perfis básicos correspondentes à profundidade da imagem \mathbf{x} no poço, compondo um conjunto $Y = [\mathbf{y}_1 \dots \mathbf{y}_p]$; (iv) $\mathbf{z} \in \mathbb{R}^m$, é a representação compacta das imagens \mathbf{x} construída por um modelo β -CoVAE, de modo que $Z = [z_1 \dots z_p] Z \in \mathcal{Z}$ (o espaço latente).

⁵O modelo β -CoVAE foi implementado com três camadas convolucionais no codificador (respectivamente 64, 32 e 16 filtros 3×3) e no decodificador (respectivamente 16, 32 e 64 filtros 3×3), que são conectadas por duas camadas densas. A camada mais interna contém as representações latentes. A camada final do modelo, também densa, é responsável pela reconstrução. Todas as camadas convolucionais utilizam uma função de ativação SELU e têm seus pesos inicializados com normalização He [14]. Não se utilizam camadas de *dropout* nem de *pooling*. Todas as camadas densas são lineares.

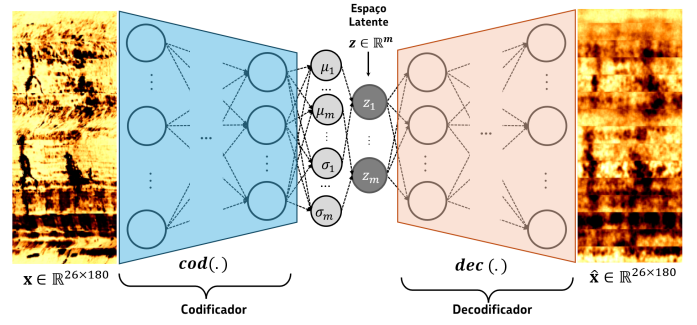


Fig. 1. Esquemático do modelo β -CoVAE utilizado na metodologia BERGSOM para reconstruir perfis de imagem. A dimensão do espaço latente ($\dim(Z)$) é um hiperparâmetro da metodologia BERGSOM.

o treinamento do modelo β -CoVAE, com destaque para os elementos principais desse tipo de rede: codificador $cod(\cdot)$, espaço latente de dimensão $\dim(Z) = m$ e decodificador $dec(\cdot)$.

O uso da metodologia BERGSOM para regerar perfis de imagem acústica a partir de perfis básicos pode ser resumida nos seguintes passos:

- Passo 1:** aprender representações compactas \mathbf{z} das imagens \mathbf{x} usando um modelo β -CoVAE;
- Passo 2:** aprender um mapeamento conjunto entre modalidades cruzadas \mathbf{z} e \mathbf{y} , por meio do treinamento de uma rede SOM expandida (Eq. 1);
- Passo 3:** inferência cruzada (que, nesse caso, se trata de aprendizado regenerativo):

- Apresentar somente o vetor com perfis básicos, i.e., \mathbf{y} como entrada para a rede SOM expandida treinada no Passo 2 para determinar o neurônio vencedor i^* usando apenas a componente \mathbf{w}_i^y dos pesos da rede:

$$i^* = \arg \min_{i \in \mathcal{A}} \|\mathbf{y} - \mathbf{w}_i^y\|; \quad (2)$$

- Determinado o vencedor i^* , extrair de seu peso \mathbf{w}_{i^*} a componente $\mathbf{w}_{i^*}^z$ relativa à representação da imagem: $\hat{\mathbf{z}} \leftarrow \mathbf{w}_{i^*}^z$;
- Regerar o segmento do perfil de imagem a partir de $\hat{\mathbf{z}}$: $\hat{\mathbf{x}} \leftarrow dec(\hat{\mathbf{z}})$.

A seguir, serão avaliadas as métricas IQA tradicionais para imagens de perfis acústicos gerados por aprendizado generativo segundo dois tipos de teste:

Teste A: aplicar métricas IQA tradicionais em imagens reconstruídas pelo modelo β -CoVAE completo alimentado diretamente com imagens reais. Nesse caso, as imagens reais de entrada são também usadas como referência para cálculo das métricas FR;

Teste B: aplicar métricas IQA tradicionais em imagens regeradas pela metodologia BERGSOM usando o decodificador do modelo β -CoVAE. Nesse caso, as imagens reconstruídas pelo modelo β -CoVAE no teste anterior serão usadas como referência para cálculo das métricas FR.

A. Teste A para Métricas IQA Tradicionais

No primeiro experimento para averiguar a efetividade das métricas IQA FR tradicionais, elas são calculadas para as imagens reconstruídas pelo modelo β -CoVAE diretamente alimentado com imagens reais do poço⁶. Os resultados do experimento são ilustrados na Fig. 2.

As reconstruções mostradas na Fig. 2 são obtidas por diferentes modelos β -CoVAE alimentados diretamente com imagens reais de um trecho do poço A. Cada imagem foi reconstruída por um modelo β -CoVAE correspondente a um valor adotado para a dimensão da representação latente ($8 \leq \dim(\mathcal{Z}) \leq 1024$). Vale destacar que qualquer imagem regenerada a partir dos perfis básicos usando BERGSOM (Teste 2) tem como limite máximo de qualidade a correspondente imagem reconstruída pelo decodificador do modelo β -CoVAE alimentado com a imagem real.

A análise visual da Fig. 2 dá a impressão de haver uma tendência de melhora das imagens reconstruídas com o aumento de $\dim(\mathcal{Z})$, o que é esperado intuitivamente. No entanto, comparar as imagens acima apenas visualmente não permite precisar a relação entre a qualidade das imagens e a dimensão do espaço latente aprendido pelos diferentes modelos β -CoVAE. Além disso, o olho humano tem dificuldade de discernir diferenças a partir de um certo valor de $\dim(\mathcal{Z})$. Isso remete à necessidade do uso de métricas IQA, em especial métricas do tipo FR, que fazem comparação pareada entre uma imagem de referência \mathbf{x}^{ref} e a imagem cuja qualidade deve ser avaliada, $\hat{\mathbf{x}}$. O erro médio quadrático, MSE (Equação 3) é uma métrica inicial natural.

$$\text{MSE}(\mathbf{x}^{\text{ref}}, \hat{\mathbf{x}}) = \|\mathbf{x}^{\text{ref}} - \hat{\mathbf{x}}\|^2 \quad (3)$$

Por considerar somente a intensidade das imagens pixel a pixel e não levar em conta correlações espaciais ao longo da imagem, a métrica MSE pode não corresponder à percepção visual humana da qualidade de uma imagem [7], [16]. O índice de similaridade estrutural (SSIM, do inglês *structural similarity*) [8], compara padrões locais pixel a pixel, mas leva em consideração a informação estrutural da imagem, de modo a ser perceptualmente mais realista. Utiliza-se aqui a definição simplificada conforme [8]

$$\text{SSIM}(\mathbf{x}^{\text{ref}}, \hat{\mathbf{x}}) = \frac{(2\mu_{x^{\text{ref}}} \mu_{\hat{x}} + C_1) (2\sigma_{x^{\text{ref}} \hat{x}} + C_2)}{(\mu_{x^{\text{ref}}}^2 + \mu_{\hat{x}}^2 + C_1) (\sigma_{x^{\text{ref}}}^2 + \sigma_{\hat{x}}^2 + C_2)}, \quad (4)$$

na qual sugerem-se os valores $C_1 = 0,01^2$ e $C_2 = 0,03^2$, que são constantes incluídas para evitar problemas numéricos. O SSIM é baseado em estatísticas locais $\mu_{x^{\text{ref}}}$, $\mu_{\hat{x}}$, $\sigma_{x^{\text{ref}}}$, $\sigma_{\hat{x}}$ e $\sigma_{x^{\text{ref}} \hat{x}}$ são calculados com base numa janela de dimensão 11×11 que desliza pixel a pixel sobre as imagens \mathbf{x}^{ref} e $\hat{\mathbf{x}}$.

⁶Os experimentos usam um conjunto de dados de perfis básicos e de perfis de imagem relativos a cinco poços, nomeados de A a E, perfurados em carbonatos altamente heterogêneos de um reservatório da camada pré-sal da formação Barra Velha [15] no Brasil. Os resultados apresentados no poço A são gerados com modelos (SOM e β -CoVAE) treinados com os dados dos poços B a E. Maiores detalhes sobre a base de dados estão disponíveis no repositório Wellbore Acoustic Image Database (WAID), acessado em 15/08/2025.

Mais recentemente, com a ubiquidade de modelos baseados em aprendizado profundo na área de visão computacional, surgiram métricas IQA baseadas nas representações (ou *embeddings*) aprendidos por modelos profundos em tarefas de classificação de imagens. Uma das mais usadas é a LPIPS, do inglês *Learned Perceptual Image Patch Similarity* [9], que explora o fato de que as representações aprendidas por modelos profundos parecem capturar das imagens informações estruturais e semânticas que são relevantes para a percepção humana. A métrica LPIPS pode ser calculada usando os *embeddings* aprendidos por diferentes modelos de aprendizado profundo, como AlexNet [17] e VGG [18]⁷.

Enquanto MSE e LPIPS são medidas de dissimilaridade entre a imagem avaliada e a imagem de referência, i.e. melhores correspondências devem ocorrer em valores mais baixos da métrica, SSIM é uma medida de similaridade, nesse caso, as melhores imagens devem apresentar valores mais altos.

Aplicando-se as métricas⁸ MSE, LPIPS e SSIM às imagens da Fig. 2, tomando como imagem de referência o perfil de imagem acústica real de um trecho do poço A, obtém-se o comportamento mostrado na Fig. 3. A Fig. 3 indica que as métricas tradicionais parecem adequadas, pois corroboram, de modo geral, a percepção visual de melhoria com o aumento de $\dim(\mathcal{Z})$ para o exemplo dado. Porém, há divergências: enquanto LPIPS e SSIM variam no sentido da melhoria da imagem, o MSE cresce após $\dim(\mathcal{Z}) = 64$, apesar de voltar a cair. Outro ponto de destaque é que esperavam-se valores mais elevados de similaridade estrutural, pois a partir de $\dim(\mathcal{Z}) = 128$ elas são todas muito parecidas com a imagem real, mas o valor máximo de SSIM é de apenas 0,14, i.e., 14% de similaridade. Para fins de comparação, foram registrados valores SSIM de até 0,58 [7] para imagens gravemente distorcidas do rosto de Einstein.

B. Teste B para Métricas IQA Tradicionais

Resta avaliar as métricas IQA tradicionais para o caso de imagens regeneradas pela metodologia BERGSOM a partir de um conjunto de sete perfis básicos. A Fig. 4 mostra imagens regeneradas para o mesmo trecho do mesmo poço que vem sendo usado como exemplo, para diferentes valores da dimensão do espaço das representações $\dim(\mathcal{Z})$ a partir de uma rede SOM expandida com 80×90 unidades⁹. A verificação visual dos resultados mostrados na Fig. 4, indica uma qualidade geral inferior à observada anteriormente, quando se reconstróem as imagens usando os modelos β -CoVAE (Fig. 2). Isso é esperado, dado que o que se propõe é gerar um trecho de imagem de 26×180 pixels usando apenas um vetor de sete componentes como informação de entrada. O nível de

⁷Por uniformidade com o que se tornou praxe na literatura, optou-se aqui pela implementação fornecida pelos proponentes da métrica, disponível em Perceptual Similarity Repository, acessado em 15/08/2025.

⁸Neste trabalho, as imagens são reescaladas para o intervalo $[0, 1]$ tanto para o treinamento das redes neurais utilizadas na metodologia BERGSOM, como para cálculo das métricas IQA FR tradicionais.

⁹A rede SOM usada na metodologia BERGSOM tem as seguintes configurações: vizinhança hexagonal gaussiana, geometria toroidal e inicialização randômica dos pesos.

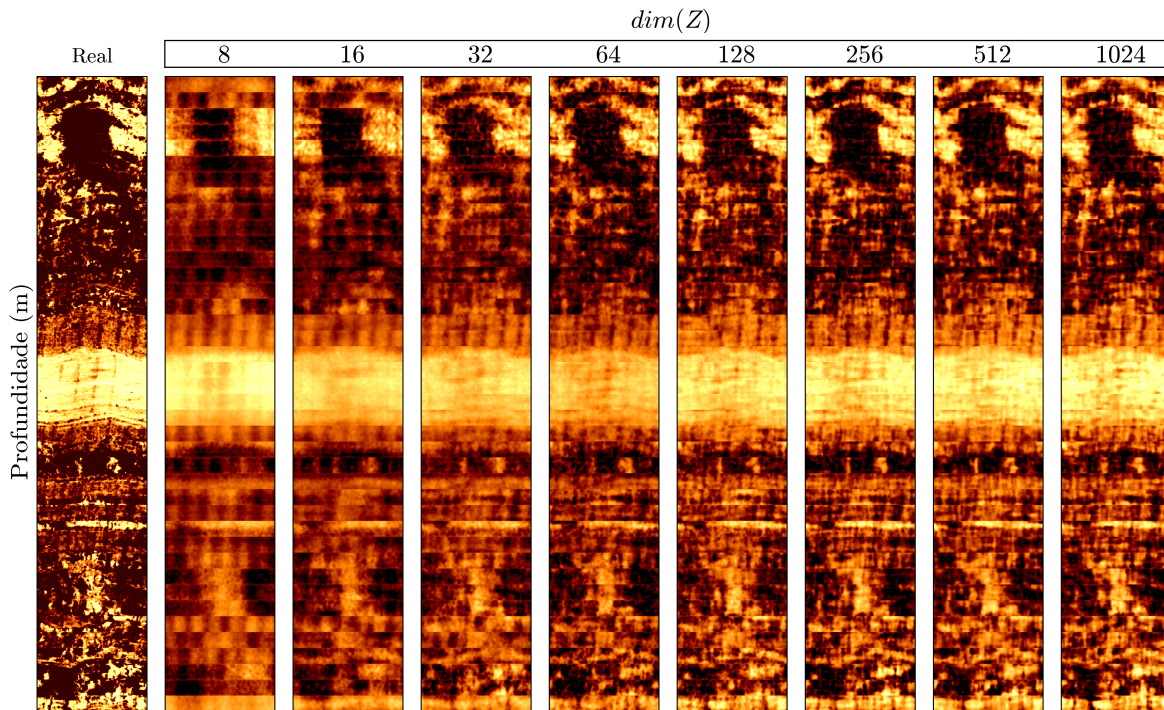


Fig. 2. Imagem real e imagens reconstruídas pelos diversos modelos β -CoVAE variando-se $dim(\mathcal{Z})$. Intervalo 2177,4 - 2184,9 m do poço A.

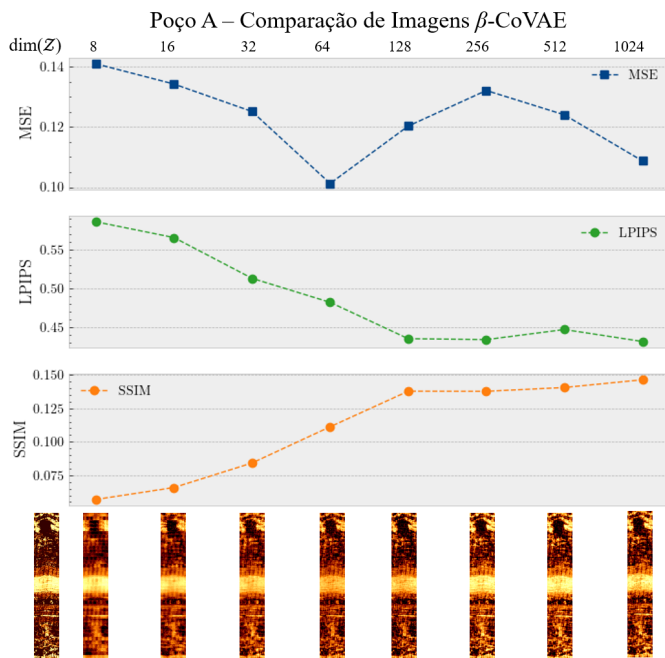


Fig. 3. Métricas MSE, LPIPS e SSIM para as imagens da Fig. 2. As métricas FR são calculadas com referência à imagem real no lado esquerdo.

qualidade apresentado, porém, indicando as grandes estruturas e contrastes entre camadas bem representados, pode ser suficiente para algumas aplicações citadas na Seção I. As imagens acima são avaliadas quanto às métricas IQA, conforme ilustra a Fig. 5.

Ao analisar visualmente os resultados mostrados na Fig. 5, observa-se uma queda da similaridade (SSIM) com o aumento de $dim(\mathcal{Z})$, o que não corrobora a impressão visual obtida ao avaliar a Fig. 4, na qual a imagem parece melhorar até $dim(\mathcal{Z}) = 128$. LPIPS e MSE também não permitem extrair uma clara conclusão quanto à melhor das imagens BERGSOM.

As discrepâncias ilustradas para as métricas IQA nos testes 1 e 2 parecem indicar que essas métricas usuais da literatura não são as mais adequadas para atestar a qualidade de imagens acústicas de paredes de poços obtidas por modelos generativos de aprendizado de máquinas. A seguir, propõem-se novas métricas baseadas na metodologia BERGSOM.

III. MÉTRICAS IQA BASEADAS EM BERGSOM

As métricas utilizadas até aqui para avaliar a qualidade dos perfis de imagem regerados pela metodologia BERGSOM são do tipo FR (“full reference”), ou seja, sempre requerem uma imagem de referência livre de distorções. Em casos reais de aplicação, quando não se tem uma imagem de referência, não será possível calcular quaisquer métricas do tipo FR. Além disso, pelo menos mais duas críticas podem ser feitas para as métricas IQA médias usadas até agora:

- 1) Elas podem divergir da percepção visual sobre as diversas imagens regeradas;
- 2) Numa caso real, um intérprete petrofísico teria maior familiaridade em usar uma medida de qualidade associada à profundidade do perfil analisado, em vez de uma medida única para todo o trecho.

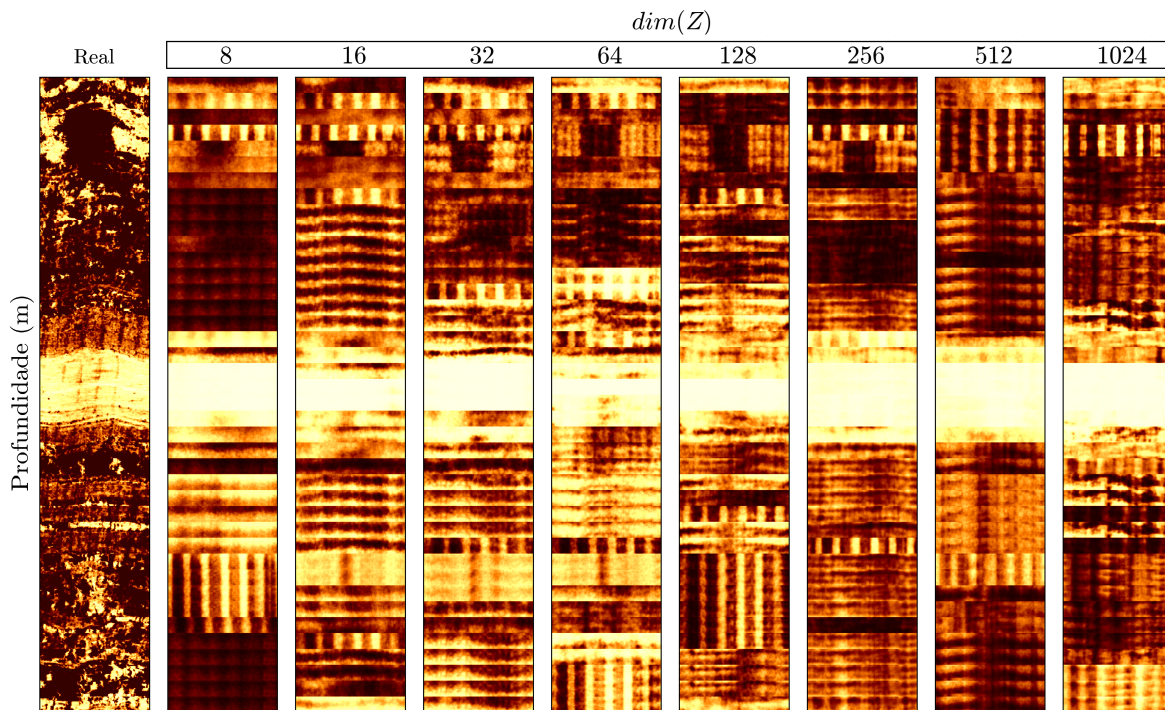


Fig. 4. Perfis de imagem acústica regeradas pela variante BERGSOM usando sete perfis básicos de entrada. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.

Nesta seção, são propostas duas novas métricas IQA que se propõem a solucionar as questões acima. As novas métricas diretamente calculadas a partir da rede SOM expandida treinada para a metodologia BERGSOM (ou sua variante p -BERGSOM [5]). A primeira métrica¹⁰, chamada de z_{IQA} é do tipo FR e a outra, chamada b_{IQA} (ou “blind” IQA) é do tipo NR (“no reference”), i.e., não requer imagens de para seu cálculo. O procedimento para o cálculo de cada uma delas é descrito a seguir. O Algoritmo 1 mostra os passos para cálculo da métrica z_{IQA} .

Vale ressaltar no Algoritmo 1 a necessidade de ter os dados relativos às representações das imagens reais de teste, isso faz com que ela seja classificada como uma métrica FR. O Algoritmo 2 mostra os passos para cálculo da métrica b_{IQA} .

Além do cálculo das métricas propostas, a metodologia permite gerar limiares de qualidade para referência. Esses limiares são baseados nos erros de quantização dos dados de treinamento da rede SOM, de modo a construir estatísticas intervalares para esses valores. Os limiares são calculados segundo os passos a seguir, ilustrados na Fig. 7:

- **Limiar z_{IQA} :** aplicar o Algoritmo 1 para os dados de treinamento. O limiar será um percentil arbitrário (P_α) dos valores de z_{IQA} calculados em dados de treino;
- **Limiar b_{IQA} :** aplicar o Algoritmo 2 para os dados de treinamento. O limiar será um percentil arbitrário (P_α) dos valores de b_{IQA} calculados em dados de treino.

¹⁰Optou-se por usar o acrônimo IQA, no inglês *image quality assessment* para as siglas que compõem as métricas propostas.

Modelos: tomar modelos β -CoVAE e SOM treinados para regeração de perfis imagens;

Dados: tomar conjunto de teste com dados de perfis básicos $Y^{teste} = \{\mathbf{y}_1^{teste}, \dots, \mathbf{y}_q^{teste}\}$ e respectivas representações de teste geradas com o modelo β -CoVAE, $Z^{teste} = \{\mathbf{z}_1^{teste}, \dots, \mathbf{z}_q^{teste}\}$;

for $t \leftarrow 1$ to q **do**

Selecionar \mathbf{y}_t^{teste} e \mathbf{z}_t^{teste} ;

Apresentar \mathbf{y}_t^{teste} à rede SOM expandida para definir a unidade vencedora:

$$\hat{i}_t^* = \arg \min_{i \in \mathcal{A}} \|\mathbf{y}_t^{teste} - \mathbf{w}_i^y\|;$$

Extrair dos pesos da unidade vencedora a componente $\mathbf{w}_{\hat{i}_t^*}^z$ referente à representação das imagens;

Atribuir a distância quadrática entre $\mathbf{w}_{\hat{i}_t^*}^z$ e \mathbf{z}_t^{teste} à métrica: $z_{IQA}(\mathbf{z}_t^{teste}, \hat{\mathbf{z}}_t) = \|\mathbf{z}_t^{teste} - \mathbf{w}_{\hat{i}_t^*}^z\|^2$.

end

Algorithm 1: Cálculo da métrica z_{IQA} .

IV. EXPERIMENTOS E RESULTADOS

A Fig. 6 ilustra as métricas z_{IQA} e b_{IQA} (linhas contínuas azuis) para o mesmo intervalo do Poço A usado nos demais exemplos. As linhas verticais verdes são limiares de qualidade para referência calculados usando o percentil P_{95} , ou seja $\alpha = 0,95$ (conforme a Fig. 7). Verifica-se coerência entre as métricas, com destaque para b_{IQA} que mostra erros menores, por exemplo, na região ao redor de 2181 m, no qual a imagem BERGSOM tem a melhor concordância com a imagem

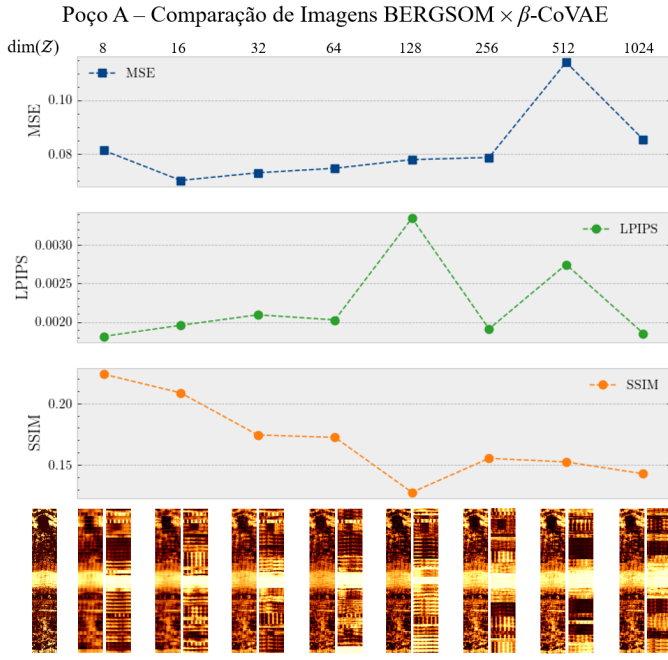


Fig. 5. Imagens regeradas BERGSOM \times imagens β -CoVAE correspondentes para diferentes valores de $dim(\mathcal{Z})$ no intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades. As métricas FR tradicionais são calculadas tomando a respectiva imagem reconstruída pelo modelo β -CoVAE como referência (ao lado esquerdo de cada imagem BERGSOM) usando o mesmo valor para $dim(\mathcal{Z})$.

Modelos: tomar modelos β -CoVAE e SOM treinados para reeração de perfis imagens;

Dados: tomar conjunto de teste com dados de perfis básicos $Y^{teste} = \{y_1^{teste}, \dots, y_q^{teste}\}$;

for $t \leftarrow 1$ to q **do**

Selecionar y_t^{teste} ;

Apresentar y_t^{teste} à rede SOM expandida para definir a unidade vencedora:

$$i_t^* = \arg \min_{i \in \mathcal{A}} \|y_t^{teste} - w_i^y\|;$$

Atribuir a distância quadrática entre $w_{i_t^*}^y$ e y_t^{teste} à métrica: $b_{IQA}(y_t^{teste}, w_{i_t^*}^y) = \|y_t^{teste} - w_{i_t^*}^y\|^2$.

end

Algorithm 2: Cálculo da métrica b_{IQA} .

reconstruída pelo modelo β -CoVAE. O mesmo não ocorreu com a métrica z_{IQA} , que aumentou em torno da profundidade 2181 m. Isso demonstra consistência da métrica b_{IQA} frente a variações ocorridas na imagem.

A Fig. 8 mostra todas as métricas IQA discutidas até aqui, inclusive as métricas tradicionais do tipo FR, porém mostradas como séries de valores indexados em profundidade no intervalo 2177,4 - 2184,9 m do poço A. Para as métricas MSE, SSIM e LPIPS, os valores foram calculados comparando-se, a cada passo de reeração das imagens de 26×180 pixels, a imagem BERGSOM regerada com a imagem correspondente gerada pelo modelo β -CoVAE (no exemplo mostrado, foram regeradas 40 imagens para a composição

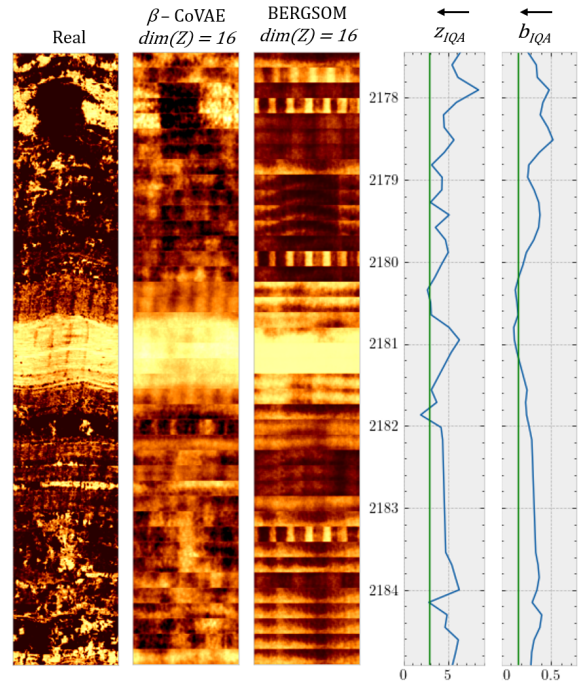


Fig. 6. Métricas z_{IQA} e b_{IQA} no intervalo 2177,4 - 2184,9 m do poço A. A métrica z_{IQA} é do tipo FR, enquanto a métrica b_{IQA} é do tipo NR, i.e., não depende de imagens de referência para seu cálculo. As setas apontam a direção em que cada métrica indica melhora na imagem.

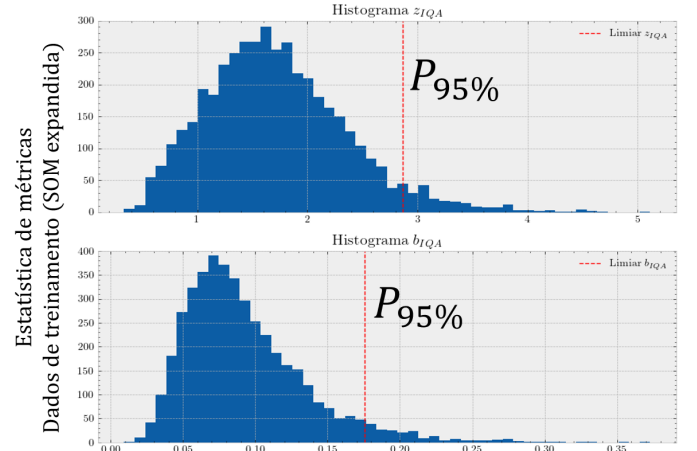


Fig. 7. Ilustração do cálculo de limiares de qualidade para as métricas z_{IQA} e b_{IQA} . Percentis são usados para definir os limiares a partir dos quais podem-se desprezar os valores mais altos das métricas calculadas com dados de teste.

de todo o trecho), o que implica um custo computacional muito superior ao necessário para o cálculo das métricas propostas z_{IQA} e b_{IQA} . Quanto à capacidade de verificar as melhores imagens, porém, observa-se que as métricas MSE, SSIM e LPIPS não são superiores às novas métricas ora propostas. Por exemplo, na região ao redor de 2181 m na qual as imagens têm boa concordância, apenas a métrica b_{IQA} apresentou melhora, quando as demais apresentaram piora em sua indicação. Vale destacar o comportamento da métrica SSIM, cujo limiar calculado com dados de treino, que chegou

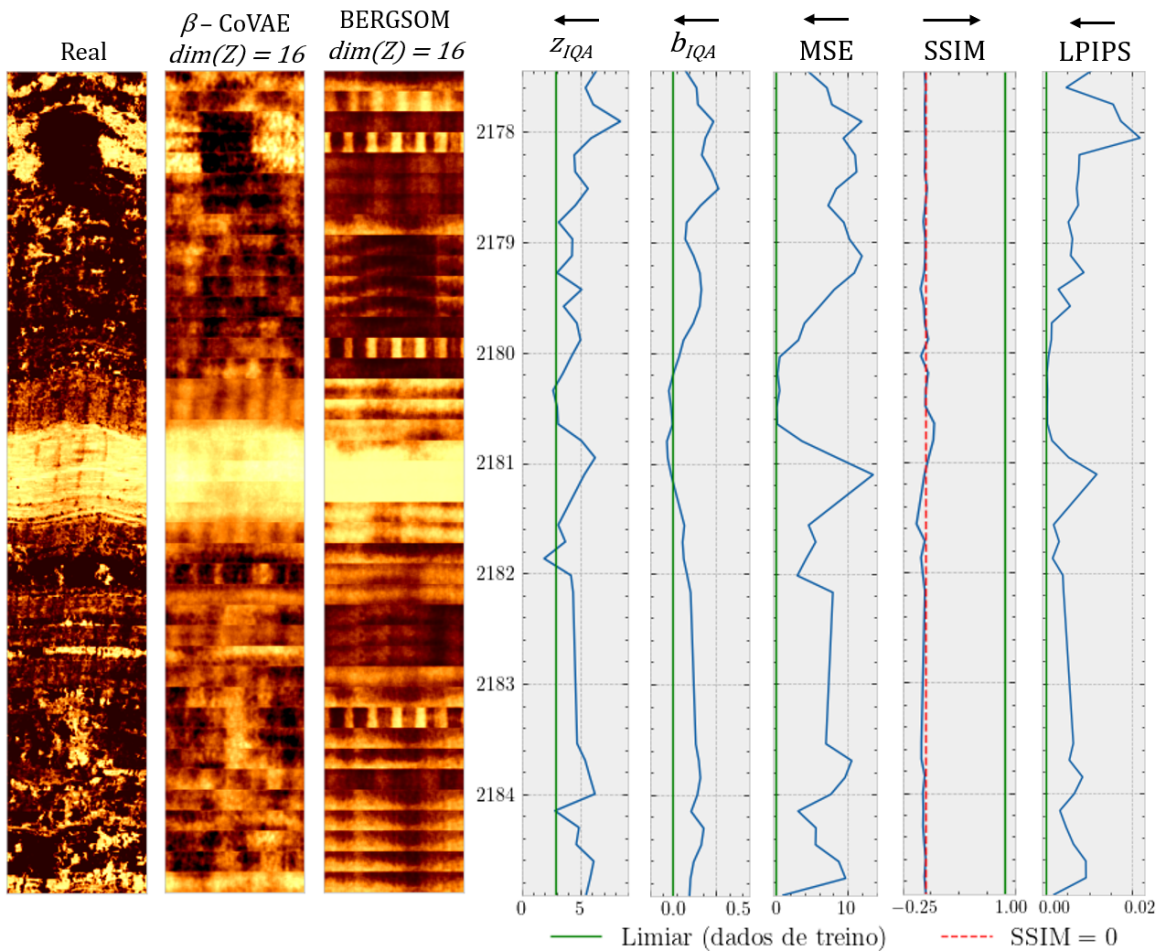


Fig. 8. Métricas IQA avaliadas no presente artigo: z_{IQA} , b_{IQA} (propostas), MSE, SSIM e LPIPS mostradas como valores indexados em profundidade no intervalo 2177,4 - 2184,9 m do poço A. As setas apontam a direção em que cada métrica indica melhora na imagem.

a 0,97, é tão distante dos valores obtidos com dados de teste, que optou-se por mostrar o valor de similaridade igual a zero (linha vermelha tracejada) para melhor referência. Ressalta-se a semelhança no comportamento das métricas MSE e z_{IQA} , o que é esperado dado que as imagens usadas para o cálculo da primeira são regeradas pelas representações usadas no cálculo da segunda.

V. CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho propôs duas novas métricas IQA, z_{IQA} (tipo FR) e b_{IQA} (tipo NR), calculadas a partir de uma rede SOM expandida usada na metodologia BERGSOM. As novas métricas IQA mostraram as seguintes vantagens em relação às métricas IQA tradicionais da visão computacional para a avaliação de imagens de perfis acústicos de parede de poços gerados por modelos generativos:

- 1) São mais rápidas de se calcular que a maioria das métricas tradicionais, em especial aquelas baseadas em aprendizado profundo como LPIPS;
- 2) São mais coerentes em relação a diferentes graus de degradação observado nas imagens;

- 3) Respondem de forma mais consistente a variações na imagem num mesmo poço. Especialmente a métrica b_{IQA} ;
- 4) A métrica b_{IQA} pode ser calculada em qualquer situação, mesmo quando não há imagens de referência para comparação.

Melhorias adicionais para métricas IQA específicas para imagens geológicas devem ser no sentido de dotar tais métricas de exprimir correlação com as estruturas geológicas presentes nas imagens.

AGRADECIMENTOS

Os autores agradecem à PETROBRAS pela permissão para publicar os resultados mostrados neste artigo. Este estudo foi financiado em parte pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, e pelas agências brasileiras CNPq (bolsas 304497/2017-7 e 313251/2023-1) e FAPERJ (E-26/201.110/2022).

REFERÊNCIAS

- [1] C. Menezes de Jesus, A. L. Martins Compan, and R. Surmas, “Permeability estimation using ultrasonic borehole image logs in dual-porosity carbonate reservoirs,” *Petrophysics*, vol. 57, no. 06, pp. 620–637, 2016.
- [2] A. Izadpanahi, N. Kumar, C. C. Gaeta Tassinari, M. Ali, T. Ahmad, and M. A. Sampaio, “A review of carbon storage in saline aquifers: Key obstacles and solutions,” *Geoenergy Science and Engineering*, vol. 250, p. 213806, 2025.
- [3] S. Salahshoor and S. Afzal, “Subsurface technologies for hydrogen production from fossil fuel resources: A review and techno-economic analysis,” *International Journal of Hydrogen Energy*, 2022.
- [4] F. Nath, M. N. Mahmood, E. Ofosu, and A. Khanal, “Enhanced geothermal systems: A critical review of recent advancements and future potential for clean energy production,” *Geoenergy Science and Engineering*, vol. 243, p. 213370, 2024.
- [5] R. A. Frota, G. A. Barreto, M. M. B. R. Vellasco, and C. Menezes de Jesus, “New cloth unto an old garment: Som for regeneration learning,” in *Advances in Self-Organizing Maps, Learning Vector Quantization, Interpretable Machine Learning, and Beyond*, T. Villmann, M. Kaden, T. Geweniger, and F.-M. Schleich, Eds. Springer Nature Switzerland, 2024, pp. 1–10.
- [6] W. Trevizan and C. MENEZES DE JESUS, “Application of GAN to Resolution Enhancement of LWD Real-Time Image Logs to Support Decision Making,” *Petrophysics*, vol. 64, no. 06, pp. 890–899, 12 2023.
- [7] Z. Wang and A. C. Bovik, “Mean squared error: Love it or leave it? a new look at signal fidelity measures,” *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, 2009.
- [8] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [9] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, “The unreasonable effectiveness of deep features as a perceptual metric,” in *IEEE/CVF*, 2018, pp. 586–595.
- [10] R. A. Frota, M. M. Vellasco, G. A. Barreto, and C. Menezes de Jesus, “Heteroassociative mapping with self-organizing maps for probabilistic multi-output prediction,” in *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024, pp. 1–6.
- [11] R. A. Frota, G. A. Barreto, M. M. B. R. Vellasco, and C. Menezes de Jesus, “Rede som para aprendizado de representações multimodais com aplicação em petrofísica,” in *Proceedings of the XXV Congresso Brasileiro de Automática*. Rio de Janeiro, RJ: SBA, 2024, p. 3494.
- [12] T. Kohonen, “The self-organizing map,” *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.
- [13] X. Tan, T. Qin, J. Bian, T.-Y. Liu, and Y. Bengio, “Regeneration learning: A learning paradigm for data generation,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, Mar. 2024, pp. 22 614–22 622.
- [14] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [15] M. H. Jácomo, G. A. Hartmann, T. B. Rebelo, N. H. Mattos, A. Batezelli, and E. P. Leite, “Mineralogical modeling and petrophysical properties of the barra velha formation, santos basin, brazil,” *Petrophysics*, vol. 64, no. 04, pp. 518–543, 08 2023.
- [16] K. Seshadrinathan, T. N. Pappas, R. J. Safranek, J. Chen, Z. Wang, H. R. Sheikh, and A. C. Bovik, “Chapter 21 - image quality assessment,” in *The Essential Guide to Image Processing*, A. Bovik, Ed. Boston: Academic Press, 2009, pp. 553–595.
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012.
- [18] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.