

Pontifícia Universidade Católica
do Rio de Janeiro



Rewbenio Araújo Frota

**Aprendizado Regenerativo Autossupervisionado
com Aplicações em Petrofísica**

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia Elétrica, do Departamento de Engenharia Elétrica da PUC-Rio.

Orientador : Prof^a. Marley Maria Bernardes Rebuszi Vellasco

Coorientador: Prof. Guilherme de Alencar Barreto

Rio de Janeiro
Fevereiro de 2025



Rewbenio Araújo Frota

**Aprendizado Regenerativo Autossupervisionado
com Aplicações em Petrofísica**

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo:

Prof^a. Marley Maria Bernardes Rebuzzi Vellasco

Orientador

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Guilherme de Alencar Barreto

DETI - UFC

Prof. Alexandre Gonçalves Evsukoff

COPPE-UFRJ

Prof. Cleyton de Carvalho Carneiro

USP

Prof^a. Karla Tereza Figueiredo Leite

UERJ

Dr^a. Paula Yamada Burkle

PETROBRAS

Rio de Janeiro, 14 de Fevereiro de 2025

Todos os direitos reservados. A reprodução, total ou parcial do trabalho, é proibida sem a autorização da universidade, do autor e dos orientadores.

Rewbenio Araújo Frota

Rewbenio Araújo Frota graduou-se em Engenharia Elétrica pela Universidade Federal do Ceará (UFC) em 2003 e obteve o título de mestre em Engenharia de Teleinformática pela mesma universidade em 2005. Em 2006, concluiu uma especialização *lato sensu* em Engenharia de Petróleo pela Universidade Federal da Bahia (UFBA). Ingressou na Petrobras em 2005, onde atua de Avaliação de Reservatórios de Óleo e Gás.

Ficha Catalográfica

Frota, Rewbenio Araújo

Aprendizado Regenerativo Autossupervisionado com Aplicações em Petrofísica / Rewbenio Araújo Frota; orientador: Marley Maria Bernardes Rebuzzi Vellasco; coorientador: Guilherme de Alencar Barreto. – 2025.

149 f.: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2025.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Aprendizado Regenerativo. 3. Self-Organizing Maps. 4. Aprendizado Profundo. 5. Regressão Múltipla. 6. Análise de Perfis Petrofísicos em Poços. I. Vellasco, Marley M. B. R. II. Barreto, Guilherme A. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. IV. Título.

CDD: 621.3

Dedico esta tese à minha mãe,
Antônia de Castro,
a gigante em cujos ombros eu me apoiei;
e a meus filhos, Paulo e Lara,
que me encantam e encorajam desde o início.

Agradecimentos

Aos meus orientadores, Prof^a. Marley Vellasco e Prof. Guilherme Barreto, pela troca de experiências e estímulo durante a realização deste trabalho. Foi uma honra trabalhar com esses dois eminentes pesquisadores da área de inteligência computacional em nosso país. Agradeço especialmente a confiança dispensada pela Profa. Marley, que me aceitou como orientando em plena pandemia, quando fomos apresentados por uma tela de vídeo conferência, e também pela ajuda, paciência e compreensão durante as dificuldades pessoais que tive ao longo do curso. Ao Prof. Guilherme, agradeço igualmente a confiança em ter aceitado coorientar este trabalho quando ele já estava em andamento, retomando assim a parceria que tivemos durante meu curso de mestrado.

À PETROBRAS pela liberação e suporte para a realização de todo o curso de doutorado, em especial pela autorização de publicação dos resultados da pesquisa. Gostaria de um agradecimento especial aos gestores que apoiaram minha indicação ao doutorado, Klayton Jonne Vieira de Melo e Samia Freire Lima Furtado, assim como àqueles que me acompanharam durante os quatro anos do curso, Diego Vale da Nóbrega e Ana Luísa Maciel.

Aos meus colegas da PETROBRAS pela troca de conhecimento e pela manutenção de um estimulante ambiente de trabalho. Em especial, destaco minha amiga, Candida Menezes de Jesus, pela curadoria dos dados que utilizei na pesquisa e por me ensinar um pouco sobre os mistérios e encantos dos perfis de imagem de paredes de poços, e ao colega Fábio Moreira de Souza pela ajuda com o ambiente de computação de alto desempenho que acelerou uma etapa importante do trabalho. Aos colegas Frederico Bastos Schuab, Aline Christianne Brasil Caires, Guilherme de Almeida Fernandes e Vagner Rosmann Carvalho, que me ajudaram a entender melhor os custos financeiros e impacto do uso dos dados de perfis de imagem em decisões em tempo real durante a perfuração.

Aos meus colegas da PUC-Rio, cujo companheirismo foi fundamental para que superássemos momentos delicados do estudo remoto durante a pandemia de COVID-19.

A todos os professores e funcionários do Departamento de Engenharia Elétrica pelo apoio e esclarecimentos em diversos momentos.

Aos professores da PUC-Rio e convidados que participaram da Comissão Examinadora.

Por último, e por serem mais importantes, dedico a todos os amigos e familiares que muitas vezes confiam mais em mim do que eu mesmo. Especialmente, quero agradecer à minha mãe, Antônia de Castro Araújo, que continua sendo meu exemplo de superação, fé e boa vontade, e que me ajudou muito em momentos pessoais delicados que coincidiram com o curso de doutorado; aos meus dois filhos, Paulo e Lara, que são minha principal motivação para aprender e seguir evoluindo nesta jornada junto com eles; e à minha namorada Ana Kedyna Ribeiro de Souza, pelo amor e apoio.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Frota, Rewbenio Araújo; Vellasco, Marley M. B. R.; Barreto, Guilherme A. **Aprendizado Regenerativo Autossupervisionado com Aplicações em Petrofísica**. Rio de Janeiro, 2025. 149p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

A área de Reservatórios de Óleo e Gás coleta grandes volumes de dados para apoiar a tomada de decisões, a modelagem e a otimização da extração de hidrocarbonetos. No entanto, métodos tradicionais de análise, que se baseiam em modelos matemáticos e estatísticos clássicos, muitas vezes restringem a capacidade de extrair informações valiosas desses dados. Diante dos altos custos de aquisição e da complexidade dos dados, técnicas avançadas de aprendizado de máquina, como Aprendizado de Representações, Aprendizado Generativo e Aprendizado Regenerativo, surgem como alternativas para maximizar o aproveitamento da informação contida nos dados. Esta tese propõe uma nova metodologia autossupervisionada para aprendizado regenerativo bidirecional intermodalidades, utilizando Mapas Auto-organizáveis de Kohonen (ou rede SOM). O objetivo é permitir a associação, geração e regeneração de dados entre modalidades distintas, explorando inter-relações não lineares e lidando com dados potencialmente incompletos. A metodologia baseia-se em uma adaptação conceitual da rede SOM, transformando-a em uma memória heteroassociativa para mapeamentos entre espaços de dados de origem e de destino, com modalidades diferentes. Essa abordagem permite: (1) associar representações de modalidades distintas em uma mesma rede SOM, (2) gerar dados de menor complexidade a partir de dados mais complexos, (3) regenerar dados complexos a partir de dados simples e (4) lidar com valores faltantes. A validação foi realizada em duas aplicações até então inéditas de geração/regeração bidirecional de dados na Petrofísica de Reservatórios. A aplicação Tipo-I é a predição multivariada de perfis convencionais a partir de perfis especiais de imagem acústica, e a aplicação Tipo-II é a regeneração de perfis de imagem acústica a partir de um conjunto de perfis convencionais. Foram utilizados dados reais oriundos de poços perfurados em reservatórios carbonáticos altamente heterogêneos do pré-sal brasileiro. A metodologia demonstrou bom desempenho

na geração e regeneração bidirecionais de dados, permitindo a extração de informações valiosas de dados complexos e potencialmente incompletos. Os resultados evidenciaram a capacidade da rede SOM em explorar inter-relações entre modalidades e regerar dados coerentes, mesmo em cenários desafiadores. A metodologia mostrou-se capaz de lidar bem com dados com valores faltantes, inclusive com a capacidade de regerar imagens (aplicação Tipo-II) a partir de um número muito limitado de perfis convencionais. Outro produto da tese é um conjunto de métricas de avaliação das imagens obtidas para a aplicação Tipo-II, que são intuitivas para um intérprete petrofísico. A presente tese apresenta uma contribuição significativa para a área, ao permitir alavancar a extração de informações contidas nos diversos tipos de dados adquiridos na indústria de Óleo e Gás. A metodologia autossupervisionada baseada em SOM mostrou-se eficaz para aplicações de Aprendizado Regenerativo, que lhe confere potencial para reduzir custos, acelerar decisões em tempo real, aumentar a eficiência da utilização e integração dos dados. Ademais, a metodologia em questão pode ser estendida a outros contextos de mapeamento intermodalidades, ampliando seu impacto em diferentes áreas do conhecimento.

Palavras-chave

Aprendizado Regenerativo; Self-Organizing Maps; Aprendizado Profundo; Regressão Múltipla; Análise de Perfis Petrofísicos em Poços.

Abstract

Frota, Rewbenio Araújo; Vellasco, Marley M. B. R (Advisor); Barreto, Guilherme A (Co-Advisor). **Self-supervised Regeneration Learning with Applications to Petrophysics**. Rio de Janeiro, 2025. 149p. Tese de Doutorado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

The Oil and Gas Reservoirs segment acquires large volumes of data to support decision-making, modeling, and optimizing the extraction of hydrocarbons. However, traditional analysis methods, which rely on classical mathematical and statistical models, often restrict the ability to extract valuable information from these data. Given the high acquisition costs and the complexity of the data, advanced machine learning techniques, such as Representation Learning, Generative Learning, and Regeneration Learning, emerge as promising alternatives to maximize the utilization of the information contained in the data. This thesis proposes a novel self-supervised methodology for bidirectional intermodal regenerative learning using Kohonen Self-Organizing Maps (SOM). The goal is to enable the association, generation, and regeneration of data between distinct modalities, exploring non-linear interrelationships and handling potentially incomplete data. The methodology is based on a conceptual adaptation of the SOM, transforming it into a heteroassociative memory for mappings between source and target data spaces with different modalities. This approach allows: (1) associating representations of distinct modalities in the same SOM, (2) generating data of lower complexity from more complex data, (3) regenerating complex data from simple data, and (4) handling missing values. Validation was performed in two previously unexplored applications of bidirectional data generation/regeneration in Reservoir Petrophysics. Type-I application involves the multivariate prediction of conventional logs from special acoustic image logs, and Type-II application involves the regeneration of acoustic image logs from a set of conventional logs. The analysis drew upon real data from wells drilled in highly heterogeneous carbonate reservoirs in the Brazilian pre-salt. The methodology demonstrated efficacy in bidirectional data generation and regeneration, enabling valuable information to be extracted from complex and potentially incomplete data. The results obtained highlighted the ability of the SOM to explore interrelationships between modalities and regenerate coherent data, even in challenging scenarios. The methodology proved capable of

handling data with missing values, including the ability to regenerate images (Type-II application) from a very limited number of conventional logs. A further outcome of the thesis is a set of evaluation metrics for the images obtained in Type-II application, which are intuitive for a petrophysical interpreter. This thesis presents a significant contribution to the field by enabling the leveraging of information contained in the various types of data acquired in the Oil and Gas industry. The self-supervised SOM-based methodology proved effective for Regenerative Learning applications, offering potential to reduce costs, accelerate real-time decision-making, and increase the efficiency of data utilization and integration. Furthermore, the methodology can be extended to other intermodal mapping contexts, broadening its impact across different areas of knowledge.

Keywords

Regeneration Learning; Self-Organizing Maps; Deep Learning; Multi-Output Regression; Petrophysics Log Analysis.

Sumário

1	Introdução	21
1.1	Sobre Perfis Petrofísicos	25
1.2	Estado da Arte	28
1.3	Hipótese	28
1.4	Objetivos	29
1.5	Organização da Tese	32
2	Modelos para Aprendizado de Representações	34
2.1	Redes neurais autocodificadoras	34
2.2	Mapas de Kohonen	39
2.3	Associação da rede SOM com Aprendizado Profundo	45
2.4	Considerações finais	46
3	Memória Associativa Baseada em SOM para Dados de Diferentes Modalidades	47
3.1	Mapeamento Heteroassociativo Multimodalidade	47
3.2	Detalhamento do Arcabouço Proposto	49
3.3	Tipos de aplicação	51
3.4	Rede SOM Expandida para Aprendizado Regenerativo Bidirecional	51
3.5	Prova de Conceito: Dados MNIST	57
3.6	Considerações finais	62
4	Aplicação Tipo-I em Petrofísica	64
4.1	Dados e Notação	64
4.2	Modelo de Aprendizado Profundo para Comparação	67
4.3	Variante BERGSOM para Aplicação Tipo-I	68
4.4	Variante p -BERGSOM para Aplicação Tipo-I	83
4.5	Comparação com o estado da arte (SOTA) em Petrofísica	89
4.6	Considerações finais	93
5	Aplicação Tipo-II em Petrofísica	94
5.1	Dados e pré-processamento	94
5.2	Considerações gerais	95
5.3	Regeração de Imagens a partir de Perfis Básicos	99
5.4	Discussão sobre a regeração de imagens a partir de perfis básicos	118
5.5	Considerações finais	125
6	Conclusões e trabalhos futuros	130
7	Referências bibliográficas	136
A	Dataset e preparação dos dados	146
A.1	Dataset	146
A.2	Preparação de dados	146
A.3	Publicação dos Dados	148

Lista de figuras

Figura 1.1	Passos do paradigma do aprendizado regenerativo: 1) Destino-para-representação; 2) Origem-para-representação; 3) Representação-para-destino (regeração).	23
Figura 1.2	Três fontes de dados sobre um intervalo do reservatório num mesmo poço: (a) perfis básicos a poço aberto; (b) perfil de imagem acústica de parede de poço; (c) medições de vazão e pressão durante um teste de formação. Todos os dados foram adquiridos no mesmo intervalo do poço e estão anonimizados.	25
Figura 1.3	Aplicação Tipo-I: geração de dados mais simples a partir de dados mais complexos. Ilustração de solução usando técnicas de aprendizado profundo.	31
Figura 1.4	Aplicação Tipo-II: geração de dados mais complexos a partir de dados menos complexos. Ilustração de solução usando técnicas de aprendizado profundo.	31
Figura 1.5	Metodologia autossupervisionada proposta nesta tese. Ilustre-se seu uso na solução de dois problemas de geração de dados petrofísicos. Destaca-se a Aplicação Tipo-II, que explora o aprendizado regenerativo.	32
Figura 2.1	Arquitetura da rede neural autocodificadora.	35
Figura 2.2	Arquitetura da rede neural do tipo autocodificador variacional.	37
Figura 2.3	Exemplo ilustrativo de um autocodificador variacional convolucional (β -CoVAE).	39
Figura 2.4	Ilustração conceitual do mapeamento $\Phi : \mathcal{X} \rightarrow \mathcal{A}$ implementado pela rede SOM.	40
Figura 3.1	Ilustração conceitual do mapeamento heteroassociativo implementado pela rede SOM expandida.	49
Figura 3.2	Ilustração da sequência de etapas gerais sequenciais que compõem a metodologia proposta.	50
Figura 3.3	Uma ilustração dos 4 passos da metodologia proposta (exemplo da variante BERGSOM).	54
Figura 3.4	Ilustração das variantes BERGSOM (a) e p -BERGSOM (b) para aplicação Tipo-I com dados do conjunto MNIST.	58
Figura 3.5	Resultados da variante BERGSOM para aplicação Tipo-I com dados do conjunto MNIST.	59
Figura 3.6	Resultados da variante p -BERGSOM para aplicação Tipo-I com dados do conjunto MNIST.	60
Figura 3.7	Esquemático das variantes BERGSOM (a) e p -BERGSOM (b) para aplicação Tipo-II com dados do conjunto MNIST.	61
Figura 3.8	Demonstração da variante BERGSOM para aplicação Tipo-II com dados do conjunto MNIST. O termo “IMG->IMG” indica que as imagens são recuperadas pelo mapeamento BERGSOM alimentado com a representação da própria imagem real para a escolha do vencedor (similar ao que se faz na Aplicação Tipo-I).	62

Figura 3.9	Demonstração da variante p -BERGSOM para aplicação Tipo-II com dados do conjunto MNIST. Teste com ($N = 7$) imagens a regerar a partir da informação do rótulo desejado.	62
Figura 4.1	Disposição relativa dos poços na base de dados utilizada. Apresenta-se um dos conjuntos de validação, com 4 poços para treino/-validação (B até D) e o poço restante (A) para teste. Por exemplo, os resultados apresentados para os experimentos no poço A são obtidos com modelos treinados com dados dos poços B até D.	65
Figura 4.2	Esquemático do modelo β -CoVAE utilizado na Etapa 1 para os dados de perfis de imagem (dados origem).	66
Figura 4.3	Ilustração do modelo SimpleNet personalizado para regressão 7C.	68
Figura 4.4	Aplicação Tipo-I com a variante BERGSOM.	70
Figura 4.5	Reconstruções de imagens por diferentes modelos β -CoVAE - variando $dim(\mathcal{Z})$. Intervalo 2129,0 - 2134,9 m do poço A.	72
Figura 4.6	Predições para perfil DEN no poço A (SOM com 100 unidades) - sensibilidade a $dim(\mathcal{Z})$.	73
Figura 4.7	Análise conjunta de para predição BERGSOM para sete perfis no poço A (métricas RMSE e MAE) - sensibilidade a $dim(\mathcal{Z})$ em comparação com modelo SimpleNet (linha vermelha tracejada).	74
Figura 4.8	Avaliação da variante BERGSOM segundo a métrica de correlação de Pearson para previsão de 7C no poço A - sensibilidade a $dim(\mathcal{Z})$.	76
Figura 4.9	Ilustração de representações conjuntas origem-destino em uma rede SOM real, planar, com 36 unidades em vizinhança hexagonal.	77
Figura 4.10	Predição do perfil DEN no poço A ($dim(\mathcal{Z}) = 16$) - sensibilidade ao tamanho da rede SOM.	78
Figura 4.11	Predição do perfil NEU no poço A ($dim(\mathcal{Z}) = 16$) - sensibilidade ao tamanho da rede SOM.	79
Figura 4.12	Desempenho da predição BERGSOM ($dim(\mathcal{Z}) = 16$) para sete perfis no poço A (métricas RMSE e MAE) - sensibilidade ao tamanho da rede SOM. Para referência, as métricas da rede SimpleNet são mostradas nas linhas amarelas tracejadas).	80
Figura 4.13	Análise de análise de correlação de Pearson para previsão de 7C no poço A - sensibilidade ao tamanho do mapa.	81
Figura 4.14	Predições BERGSOM de todos os perfis para o poço A - SOM pequeno (10×10 unidades) - $dim(\mathcal{Z}) = 16$.	82
Figura 4.15	Anomalias observadas nas predições com o aumento de $dim(\mathcal{Z})$ - SOM grande (poço A) com vizinhança tipo bolha.	84
Figura 4.16	Análise das Matrizes U de cada SOM expandido com o aumento de $dim(\mathcal{Z})$.	85
Figura 4.17	Aplicação Tipo-I com a variante p -BERGSOM.	86
Figura 4.18	Predições p -BERGSOM de todos os perfis para o poço A - SOM pequeno (10×10 unidades) - $dim(\mathcal{Z}) = 16$.	89
Figura 4.19	Ampliação do perfil de densidade (DEN) e das predições (BERGSOM e p -BERGSOM) para o intervalo (2194,5 - 2244,5 m) do poço A (p -BERGSOM gera $N(= 100)$ predições) - SOM pequeno e $dim(\mathcal{Z}) = 16$.	90

Figura 5.1	Imagem real e imagens reconstruídas pelos diversos modelos β -CoVAE variando-se $dim(\mathcal{Z})$. Intervalo 2177,4 - 2184,9 m do poço A.	95
Figura 5.2	Métricas MSE, LPIPS e SSIM para as imagens da Figura 5.1. Métricas calculadas com referência à imagem real no lado esquerdo.	98
Figura 5.3	Métrica MSE para os modelos β -CoVAE em todos os poços de teste, variando-se $dim(\mathcal{Z})$.	99
Figura 5.4	Aplicação Tipo-II com a variante BERGSOM.	101
Figura 5.5	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.	102
Figura 5.6	Imagens regeradas BERGSOM \times imagens β -CoVAE correspondentes. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.	105
Figura 5.7	Aplicação Tipo-II com a variante p -BERGSOM.	108
Figura 5.8	Perfis de imagem acústica regeradas pela variante p -BERGSOM usando sete perfis básicos de entrada. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.	109
Figura 5.9	Detalhamento do efeito de variação de imagens regeradas pelos mesmos vencedores (BERGSOM \times p -BERGSOM). Rede SOM com 80×90 unidades.	110
Figura 5.10	Desempenho do modelo: imagens regeradas \times imagens β -CoVAE correspondente (p -BERGSOM). Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.	111
Figura 5.11	Comparação de métricas AQI - BERGSOM \times p -BERGSOM. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.	112
Figura 5.12	Resultado típico da variante BERGSOM para a aplicação Tipo-II. Intervalo 2177,4 - 2184,9 m do poço A. Imagem real e imagem β -CoVAE de referência à esquerda são comparadas com as imagens BERGSOM variando-se o tamanho da rede SOM (com $dim(\mathcal{Z}) = 16$).	113
Figura 5.13	Perfis de imagem acústica regeradas pela variante p -BERGSOM usando subconjuntos de perfis de entrada. Intervalo 2177,4 - 2184,9 m do poço A.	117
Figura 5.14	Desempenho do modelo: imagens geradas \times imagem real, conforme o Subconjunto de perfis.	118
Figura 5.15	Métricas z_{IQA} e b_{IQA} no intervalo do poço A que vem sendo usado como exemplo neste capítulo. 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.	122
Figura 5.16	Métricas AQI utilizadas nesta tese: z_{IQA} , b_{IQA} (propostas), MSE, SSIM e LPIPS mostradas como valores indexados em profundidade no intervalo do poço A que vem sendo usado como exemplo neste capítulo. As setas indicam a direção em que cada métrica indica melhora na imagem.	123
Figura 5.17	Perfis de imagem acústica geradas pela variante BERGSOM usando subconjuntos de perfis de entrada. Intervalo 2177,4 - 2184,9 m do poço A.	124
Figura 5.18	Comportamento da métrica MSE para experimento de regeração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).	125

Figura 5.19	Comportamento da métrica SSIM para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).	126
Figura 5.20	Comportamento da métrica LPIPS para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).	127
Figura 5.21	Comportamento da métrica z_{IQA} para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).	128
Figura 5.22	Comportamento da métrica b_{IQA} para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).	129
Figura A.1	Possíveis casos de incompatibilidade de profundidade entre os perfis básicos e o perfil de imagem acústica. O subíndice “img” refere-se ao perfil de imagem e o subíndice “bsc” refere-se aos perfis básicos.	147
Figura A.2	Intervalo comum de profundidade para ambas as fontes de dados.	148
Figura A.3	Ajuste para tratamento de diferenças de taxa de amostragem em profundidade das diferentes modalidades de perfis (básicos e imagem acústica).	148

Lista de tabelas

Tabela 1.1	Relação entre perfis básicos e perfis especiais de imagem acústica de poços.	27
Tabela 2.1	Seleção de trabalhos que exploram a rede SOM como memória associativa.	44
Tabela 2.2	Seleção de trabalhos que associam a rede SOM ao aprendizado profundo.	44
Tabela 4.1	Conjuntos de validação para avaliação do modelo.	65
Tabela 4.2	Taxas de compactação aplicadas aos dados de origem.	71
Tabela 4.3	Avaliação da variante BERGSOM e da rede SNet segundo a métrica RMSE, variando-se $\dim(\mathcal{Z})$ e SOM pequeno.	75
Tabela 4.4	Avaliação da variante BERGSOM e da rede SNet segundo a métrica MAE, variando-se $\dim(\mathcal{Z})$ e SOM pequeno.	75
Tabela 4.5	Tempos de execução em segundos (s) nas etapas de treinamento (soma dos tempos para as redes β -CoVAE e SOM no modelo BERGSOM) e inferência (Conjunto 1).	83
Tabela 4.6	Métrica RMSE das previsões p -BERGSOM para cada perfil em todos os poços de teste.	87
Tabela 4.7	Métrica MAE das previsões p -BERGSOM para cada perfil em todos os poços de teste.	87
Tabela 4.8	Referências SOTA para previsão de perfis/grandezas petrofísicos.	91
Tabela 5.1	Resultados para a aplicação Tipo-II com $\dim(\mathcal{Z}) = 16$. As setas indicam a direção de melhor desempenho para cada métrica de AQL. Os melhores valores gerais de métrica para cada conjunto estão destacados. Legenda do tamanho da rede SOM utilizada: 'P' para mapa pequeno, 'M' para mapa médio e 'G' para mapa grande.	113
Tabela 5.2	Composição de subconjuntos de perfis básicos.	116

Lista de algoritmos

Algoritmo 1	Treinamento sequencial da rede SOM	42
Algoritmo 2	Variante BERGSOM para aplicação Tipo-I	69
Algoritmo 3	Variante p -BERGSOM para aplicação Tipo-I	85
Algoritmo 4	Variante BERGSOM para aplicação Tipo-II	101
Algoritmo 5	Variante p -BERGSOM para aplicação Tipo-II	107
Algoritmo 6	Cálculo da métrica z -IQA (ambas as variantes)	121
Algoritmo 7	Cálculo da métrica b -IQA (ambas as variantes)	121

Lista de Abreviaturas

AE – *Autoencoder*

CoAE – *Convolutional Autoencoder*

CoVAE – *Convolutional Variational Autoencoder*

GAN – *Generative Adversarial Network*

MLP – *Multilayer Perceptron*

SOM – *Self-Organizing Maps*

VAE – *Variational Autoencoder*

CAL – Perfil de diâmetro do poço

DTC – Perfil de tempo de trânsito de ondas compressivas

PE – Perfil do fator fotoelétrico

RES – Perfil de indução medindo a resistividade da formação

NMR PhiT – Perfil de porosidade total medida por Ressonância Magnética Nuclear (RMN)

GR – Perfil *Gamma Ray*

DEN – Perfil *bulk density* (também referido como RHOB na literatura)

NEU – Perfil de porosidade de nêutrons (também referido como NPHI na literatura)

*Sou um explorador de silêncios, e
que tesouros encontrei nos silêncios que
eu possa contar com confiança?*

Khalil Gibran, O profeta.

1

Introdução

A área de reservatórios de petróleo e gás adquire dados de forma intensiva. Esses grandes volumes de dados são analisados e combinados por especialistas, tanto de Geologia como de Engenharia de Reservatórios, com o intuito de melhorar a compreensão e a modelagem dos reservatórios. Esses dados servem de insumo a diversas etapas da exploração de hidrocarbonetos: desde a tomada de decisões em tempo real na perfuração dos poços, até a modelagem computacional refinada para acompanhamento rotineiro dos campos produtores.

Um dos maiores desafios é a integração de dados, pois eles têm diferentes formatos (como séries temporais, imagens, texto etc), diferentes frequências de aquisição (como aquisição a cada cinco segundos ou a cada mês) e diferentes escalas espaciais (desde pequenas amostras de rocha em um dado poço a mapas sísmicos abrangendo campos inteiros). Em suma, os objetivos buscados ao se coletar todos esses dados são:

1. Melhorar a compreensão sobre os reservatórios;
2. Usar os dados em diversas aplicações práticas específicas, cuja meta final é retirar o máximo de hidrocarbonetos de forma segura dentro do tempo de concessão das jazidas. Como exemplos típicos de tais aplicações, podem-se citar: previsão de desempenho dos diversos poços produtores, decidir a posição de novos poços a perfurar, postergar o abandono de poços por baixa produtividade ou baixa economicidade.

No entanto, tradicionalmente, a análise desses dados depende da aplicação de métodos analíticos e estatísticos clássicos, por profissionais com variado grau de experiência, muitas vezes isoladamente ou com um nível de interdisciplinaridade abaixo do ideal. Isso faz com que boa parte da informação presente nos dados adquiridos não chegue a ser, de fato, utilizada ou assimilada nos modelos de reservatório.

Com o intuito de aumentar o grau de aproveitamento da informação contida nos dados, a área de reservatórios de hidrocarbonetos pode se beneficiar do uso de técnicas de aprendizado de máquinas, em especial técnicas como o aprendizado de representações (*representation learning*) (BENGIO; COURVILLE; VINCENT, 2013).

O aprendizado de representações ganhou relevância nos últimos anos no campo de estudos da inteligência artificial. Consiste em utilizar métodos de

aprendizado profundo (*deep learning*) (GOODFELLOW; BENGIO; COURVILLE, 2016), para extrair informações subjacentes aos dados (chamadas de *representações*). Os dados originais são transformados de modo a serem representados numa dimensão diferente, geralmente menor que a dimensão original dos dados. Uma imagem, por exemplo, pode ser representada por um vetor de características dela extraído por meio do processamento sucessivo através das camadas de uma rede neural convolucional. Neste sentido, podem-se citar dois objetivos principais do aprendizado de representações:

1. Ampliar o conhecimento sobre o problema em estudo;
2. Usar as representações obtidas para solucionar aplicações práticas (chamadas *downstream tasks*).

O conceito de *representação* permite conceber diferentes dados medidos de um mesmo ambiente ou contexto, um reservatório de hidrocarbonetos por exemplo, como complementares ou suplementares quando transformadas para algum espaço coerente. Assim, dados que em sua modalidade natural podem dificultar seu uso conjunto, passam a cooperar de forma direta por meio de suas representações.

O aprendizado de representações ganhou ainda mais relevância com a popularização das arquiteturas de aprendizado generativo (GOODFELLOW et al., 2014; BENGIO et al., 2013; KINGMA; WELING, 2013). Nesse tipo de aprendizado, busca-se modelar a distribuição subjacente aos dados de treinamento para, então, gerar novas amostras que se assemelhem aos dados usados para construir o modelo. Portanto, a qualidade dos dados gerados por esses modelos depende diretamente da qualidade das representações aprendidas.

No entanto, em muitos casos práticos, os modelos generativos não são capazes de modelar essas distribuições por vários motivos tais como, alta complexidade de mapeamento dos dados, elevado custo de computacional, esparsidade dos dados. Um recente paradigma chamado aprendizado regenerativo (*regeneration learning*, RegL) (TAN et al., 2023) se propõe a enfrentar alguns desses desafios enfrentados pelos métodos tradicionais de aprendizado generativo condicional.

No aprendizado regenerativo, dados de destino Y (como texto ou imagens) geralmente exibem maior dimensionalidade e complexidade que os dados de origem X , além de poder conter informações não presentes em X . Isso torna muito difícil de aprender o mapeamento direto $X \mapsto Y$ de maneira eficaz.

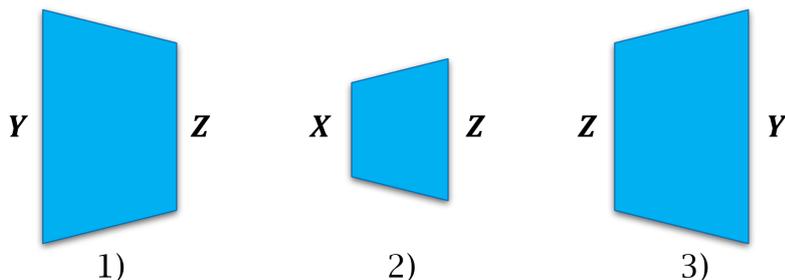


Figura 1.1: Passos do paradigma do aprendizado regenerativo: 1) Destino-para-representação; 2) Origem-para-representação; 3) Representação-para-destino (regeração).

A ideia central do aprendizado regenerativo¹, mostrado na Figura 1.1 é representar os dados de destino Y como uma camada intermediária denotada como Z e, em seguida, aprender os mapeamentos $X \mapsto Z$ e $Z \mapsto Y$, respectivamente. Assim, reduz-se a dificuldade da tarefa, pois tanto o mapeamento de X (origem) para Z (representação) como o de Z para Y (destino) são mais simples que o mapeamento direto de X para Y .

No contexto do aprendizado regenerativo, dois conceitos são importantes: modalidade de um dado e aprendizado cruzado intermodalidades.

Modalidade é o formato ou tipo do dado, por exemplo, imagem e texto são modalidades diferentes. Enquanto muitos trabalhos que envolvem aprendizado de máquinas tratam de dados do mesmo tipo, i.e., de mesma modalidade, recentemente têm emergido nessa área diversos estudos que operam com aprendizado conjunto entre diferentes modalidades (ŠKRLJ, 2024).

Vale ainda esclarecer o uso dos termos *dados simples* e *dados complexos*, que serão usados numerosas vezes no texto. Nesta tese, esses termos referem-se à relação de complexidade de conteúdo, ou densidade de informação² presente nos dados, entre as diferentes modalidades consideradas.

Deve-se ter em conta que uma dada modalidade de dados que registra aspectos da realidade tem intrinsecamente um limite quanto ao total da informação da realidade que ela pode codificar. Assim, pode-se coloquialmente definir uma densidade de informação como a razão entre a informação preservada no

¹Nesta tese utiliza-se o termo “aprendizado regenerativo” para traduzir *regeneration learning*, a exemplo da tradução “aprendizado generativo” para *generative learning* que popularizou-se em textos na língua portuguesa. Por coerência, enquanto dados são *gerados* por aprendizado generativo, esta tese trata de dados *regerados* por aprendizado regenerativo. Evita-se o uso do verbo regenerar, pois há outras aplicações diferentes daquelas tratadas nesta tese, por exemplo, quando dados corrompidos são recuperados, ou regenerados, por diferentes técnicas.

²No escopo desta tese, não se pretende explorar a definição formal de densidade de informação que se usa em teoria da informação. Pretende-se utilizar o sentido coloquial desse conceito, conforme detalhado no texto.

dado e a informação total. A partir dessa definição coloquial, percebe-se ser difícil precisar um valor absoluto para a densidade de informação, enquanto estabelecer uma escala relativa entre densidades para diferentes modalidades parece mais intuitivo. Por exemplo, ao registrar um ser humano, uma imagem capta mais informação que um conjunto de medidas como altura, peso corporal e cor da pele. Nos termos desta tese, o conjunto de medidas seria uma modalidade mais simples se comparada à imagem, considerada uma modalidade mais complexa.

Nesse sentido, um aspecto de uma modalidade que pode estar mais bem relacionada à sua complexidade relativa (ou densidade de informação) é sua dimensionalidade, i.e., o espaço numérico disponível para registrar as informações. Por exemplo, um vetor unidimensional será considerado simples em comparação a uma matriz bidimensional de valores, que será considerada mais complexa. Numa aplicação de regeneração de imagens a partir de uma sequência de vetores, os dados complexos são as imagens enquanto os dados mais simples são a série de vetores.

O aprendizado cruzado intermodalidades (*cross-modal learning*) (SKO-CAJ; LEONARDIS; KRUIJFF, 2012) relaciona informações aprendidas por diferentes modalidades de dados e permite aplicações tais como a geração cruzada de dados intermodalidades. O termo cruzada denota aqui o fato de, dado o aprendizado conjunto de duas modalidades, a fase de inferência habilitar a estimação ou geração de uma modalidade a partir da outra. Assim, por exemplo, o aprendizado regenerativo é um caso de geração cruzada intermodalidades no qual a modalidade mais complexa é regenerada por uma entrada de modalidade mais simples. Sob a ótica do aprendizado de representações, os diversos dados obtidos pela indústria trazem, de forma latente, sub-representações do reservatório, algumas delas exclusivas, outras com certo grau de sobreposição entre si (HUANG; PENG; YUAN, 2017). A Figura 1.2 traz um exemplo que ilustra esse conceito.

Na Figura 1.2 (a), mostra-se um conjunto de perfis geológicos básicos corridos num poço. Estes perfis fazem diferentes medições ao longo de algumas dezenas ou centenas de metros. Cada perfil geológico obtido é uma série de valores por profundidade, como se fosse uma série temporal em que o tempo é substituído pela profundidade³. No mesmo poço, pouco depois, pode ser adquirido um perfil de imagem da parede do poço (Figura 1.2 (b)). Este dado não é uma fotografia, mas uma abstração construída a partir de uma

³Cada valor registrado num perfil corresponde a uma profundidade específica e, portanto, a um tempo geológico relativo à época da deposição da rocha naquela profundidade. Assim, um perfil pode ser, de fato, interpretado como uma série temporal ao longo do tempo geológico (NICHOLS, 2009).

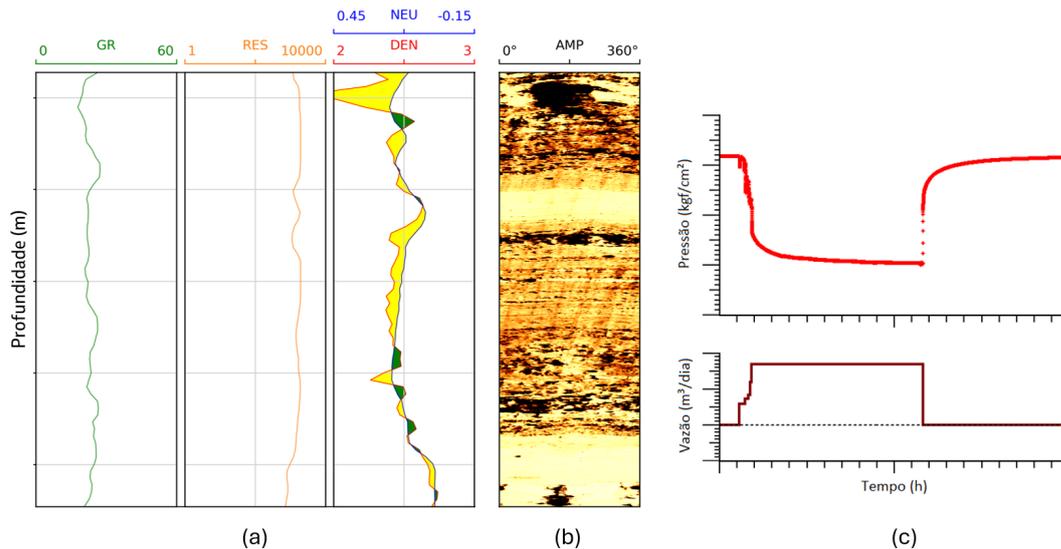


Figura 1.2: Três fontes de dados sobre um intervalo do reservatório num mesmo poço: (a) perfis básicos a poço aberto; (b) perfil de imagem acústica de parede de poço; (c) medições de vazão e pressão durante um teste de formação. Todos os dados foram adquiridos no mesmo intervalo do poço e estão anonimizados.

matriz (profundidade \times azimuth) com medidas de impedância acústica ao longo da parede do poço (MENEZES DE JESUS; COMPAN; SURMAS, 2016). Finalmente, algum tempo depois, pode ser realizado um teste de formação num intervalo comum ao trecho perfilado (Figura 1.2 (c)). Os dados obtidos num teste de formação são séries temporais de outras grandezas, como vazão e pressão do intervalo ao longo do período de teste. São, portanto, informações diferentes daquelas obtidas pelos perfis anteriores. Todas essas fontes de dados, porém, podem ser usadas para extrair, cada uma delas separadamente por meio de técnicas específicas para cada modalidade de dados, uma mesma característica do reservatório: a permeabilidade da rocha. Nesse sentido, a permeabilidade seria uma informação do reservatório que se pode inferir por três modalidades distintas de dados.

Diante do exposto, este trabalho trata do desenvolvimento, análise e aplicação prática de métodos de geração de dados de reservatório, explorando os conceitos acima de aprendizado de representações, em especial, o aprendizado regenerativo. A seguir, apresentam-se brevemente os perfis geológicos que serão usados nas aplicações desenvolvidas nesta tese.

1.1 Sobre Perfis Petrofísicos

Os dados que serão utilizados nas aplicações a seguir foram adquiridos em poços recém perfurados, e são chamados de perfis geológicos, ou petrofísicos. Esses perfis fornecem informações essenciais para a compreensão das

propriedades e do comportamento dos reservatórios de petróleo e gás (TIAB; DONALDSON, 2016). Os perfis mais comuns, doravante denominados perfis básicos, consistem em registros de propriedades da rocha, como a resistividade, a porosidade ou medições da saturação de fluido. Devido ao custo de aquisição relativamente baixo, os perfis básicos são adquiridos em quase todos os poços perfurados. Geralmente, são armazenados e interpretados como uma sequência de valores reais indexados à profundidade do poço. Segue uma sucinta descrição dos perfis básicos que serão utilizados nesta tese (LIU, 2017; ANDERSEN, 2011):

- *CAL*: medições do calibre do poço, i.e., o diâmetro do poço. Normalmente utilizado para controle de qualidade dos dados, evitando considerar, por exemplo, dados em trechos arrombados do poço.
- *GR*: perfil de raios gama, que mede a radioatividade natural da rocha.
- *DEN*: medição de densidade da rocha baseada na detecção de raios gama não absorvidos pela formação (também referido como RHOB na literatura).
- *NEU*: perfil de porosidade de nêutrons (também referido como NPFI na literatura).
- *DTC*: ou Delta-t Compressional, é uma medida de tempo de trânsito de ondas compressivas obtida por meio de perfis sônicos.
- *PE*: perfil do fator fotoelétrico, é uma medida sensível à litologia ou tipo de rocha (também referido como PeF na literatura).
- *RES*: perfil de indução que mede a resistividade profunda da formação.
- *NMR PhiT*: porosidade total medida por perfil de Ressonância Magnética Nuclear (RMN).

Por outro lado, os perfis especiais, como as imagens acústicas, fornecem maior conteúdo de informações do que os perfis básicos, o que justifica seu custo de aquisição mais alto. A ferramenta de imagem acústica emite uma onda ultrassônica de amplitude emitida A_e e mede a relação entre A_e e a amplitude A_r da onda recebida após a reflexão no poço. É usual expressar essa relação em decibéis (dB) (VALENTÍN, 2018)

$$\Lambda_\theta = 20 \log_{10} \left[\frac{A_e}{A_r(\theta)} \right], \quad (1-1)$$

na qual θ é o azimute de onde chega a onda refletida, e a grandeza Λ_θ mede o contraste de impedância acústica entre a rocha e o fluido de perfuração que preenche o poço (LIU, 2017).

Os dados de perfis de imagem acústica são normalmente armazenados como uma matriz de valores de amplitude da impedância acústica (profundidade \times azimute). Esse formato permite visualizar os dados como uma imagem usando uma escala de cores arbitrária apropriada⁴. Conforme ilustra a Figura 1.2 (b), esse modo de visualização para o perfil de imagem acústica permite identificar e interpretar diversas características geológicas, como fraturas, falhas, planos de acamamento e outros elementos estruturais como artefatos induzidos por perfuração. No entanto, devido aos seus custos de aquisição mais altos, esses registros especiais geralmente são limitados a poços selecionados durante as campanhas de perfuração.

As informações obtidas nos diferentes perfis são ora redundantes, ora complementares. Os volumes de rocha investigados variam de ferramenta a ferramenta, de modo que a análise integrada desses dados é essencial. Pode ocorrer ainda de uma ferramenta ser sensível a uma dada propriedade da rocha, mas não forneça uma medida dessa propriedade (ANDERSEN, 2011). A Tabela 1.1 traz um resumo do que se espera, *a priori*, de correlação entre os perfis básicos e os perfis de imagem acústica.

Tabela 1.1: Relação entre perfis básicos e perfis especiais de imagem acústica de poços.

Perfil Básico	Unidade	Correlação <i>a priori</i> com imagem acústica?
CAL	pol	NÃO
GR	GAPI	NÃO
DEN	g/cm ³	SIM
NEU	%	SIM
DTC	μ s/ft	SIM
PE	barns/electron	NÃO
RES	$\Omega \cdot$ m	NÃO
NMR PhiT	%	SIM

Conforme ilustra a Tabela 1.1, espera-se que apenas os perfis básicos relacionados à densidade ou à porosidade da rocha tenham boa correlação com o perfil de imagem acústica (LIU, 2017; ELLIS; SINGER, 2007; MENEZES DE JESUS; COMPAN; SURMAS, 2016).

⁴Por exemplo, nesta tese, as imagens são visualizadas usando o mapa de cores `afmhot`, disponível no pacote de código aberto Matplotlib para Python.

1.2

Estado da Arte

A grande maioria dos trabalhos na área de reservatórios que exploram o uso de técnicas de aprendizado de máquinas para geração de dados concentra-se nos seguintes casos:

- **Geração de dados de mesma modalidade:** por exemplo, previsão de vazão a partir de pressão (TIAN; HORNE, 2019); estimativa de valores faltantes de vazão em poços a partir da pressão de fluxo (TIAN; HORNE, 2015); gerar perfis básicos não adquiridos no poço a partir de outros (OLIVEIRA; CARNEIRO, 2021; LI; MISRA, 2019; AKKURT et al., 2018; AKINNIKAWA; LYNE; ROBERTS, 2018; SALEHI et al., 2017); melhorar qualidade de perfis de imagem em tempo real (TREVIZAN; MENEZES DE JESUS, 2023);
- **Geração de dados de modalidades diferentes (do mais complexo para o menos complexo):** por exemplo, estimativa de porosidade e permeabilidade a partir de perfis de imagem acústica (VALENTÍN et al., 2018) ou a partir de dois tipos de perfis de imagens, acústico e resistivo (BOM et al., 2021);

Ainda é pouco explorado o conceito de representações e seu potencial de geração cruzada de dados de reservatórios entre diferentes modalidades. Além disso, a geração cruzada de dados de reservatórios de modalidade ou natureza diferentes partindo do dado de menor complexidade para um dado de maior complexidade, ou seja o aprendizado regenerativo, é uma lacuna inexplorada em trabalhos na área de reservatórios que exploram técnicas de aprendizado de máquinas.

Este trabalho, portanto, pretende preencher essa lacuna, ao ampliar a gama de modelos que permitem a geração de dados de reservatórios, em especial, a geração cruzada de dados de diferentes modalidades e diferentes densidades de informação. A proposta, apresentada abaixo, tem dois eixos de contribuição: uma teórica com a inovação no campo de aprendizado de representações e aprendizado regenerativo e outra prática, com a solução de dois problemas ainda inéditos na área de petrofísica de reservatórios.

1.3

Hipótese

Esta tese tem como premissa que diferentes modalidades de dados adquiridos na indústria de óleo e gás, especialmente na área de reservatórios,

guardam relações que vão além do explorado nos métodos de análise tradicionais usados por geólogos e engenheiros de reservatórios. Tal premissa motivou as duas principais hipóteses de trabalho que são exploradas nesta tese, a saber:

- Métodos de aprendizado regenerativo podem ser usados para explorar o potencial gerativo de dados de modalidades complexas a partir de dados mais simples.
- O aprendizado auto-organizado de um mapeamento heteroassociativo entre dados de diferentes modalidades é uma solução para a geração bidirecional de dados de diferentes modalidades.

1.4

Objetivos

Neste trabalho, investiga-se a utilização de técnicas e conceitos de aprendizado de representações, aprendizado generativo e regenerativo em diferentes problemas de ciências de reservatórios. O objetivo é utilizar representações comuns oriundas de fontes de dados diferentes, potencialmente de modalidades diferentes, tais como imagens e sequências, para gerar novos dados, melhorar a qualidade de dados existentes, aumentar a densidade de informação em uma modalidade de dados a partir de outra e melhorar o desempenho em tarefas específicas, e.g., *downstream tasks*.

Pretende-se, em especial, explorar o paradigma de aprendizado regenerativo, para resolver problemas da área de reservatórios que demandam o mapeamento de dados menos complexos (ou com menor teor de informação) para dados mais complexos (i.e. com maior teor de informação). Os ganhos que o tema proposto podem proporcionar para a área de reservatórios e para a indústria de óleo e gás são listadas a seguir:

- Contribuir para o uso eficiente dos grandes volumes de dados de reservatório em suas diferentes modalidades, incluindo:
 - Desenvolver métodos de geração bidimensional cruzada entre diferentes modalidades;
 - Permitir que o método proposto de geração bidimensional seja capaz de funcionar a partir de dados incompletos;
 - Propor métricas de validação adequadas aos dados gerados.
- Contribuir para um melhor entendimento da relação entre dados de diferentes modalidades;
- Popularizar o conceito de representações na modelagem *data-driven* de reservatórios;

- Fomentar o desenvolvimento de modelos de reservatórios (tradicionais e *data-driven*) alternativos aos existentes;
- Permitir redução de custos em aquisição de dados futuros.

Para o propósito desta tese, as aplicações serão de duas naturezas, conforme a designação usada no jargão de aprendizado regenerativo:

Aplicação Tipo-I: geração de dados simples a partir de dados complexos (*data understanding*);

Aplicação Tipo-II: regeneração de dados complexos a partir de dados simples (*data regeneration*).

Para fins de validação, os vários modelos derivados a partir da metodologia desenvolvida nesta tese serão aplicados para solucionar dois problemas ainda inéditos na área de petrofísica (ilustradas nas Figuras 1.3 e 1.4). Esses problemas enquadram-se na classificação acima e são especificamente definidos como:

Aplicação Tipo-I: predição de perfis geológicos convencionais a partir de perfis de imagem acústica de parede dos poços;

Aplicação Tipo-II: regeneração de perfis de imagem acústica de parede de poços a partir de perfis geológicos convencionais, tais como raios gama, densidade, porosidade e resistividade.

A solução das duas aplicações acima são a principal contribuição científico-tecnológica desta tese.

A princípio, pode-se imaginar que levar adiante tais aplicações só pode ser viável com o uso de modelos complexos de aprendizado supervisionado, por exemplo baseado em aprendizado profundo, usando redes neurais convolucionais, como nas Figuras 1.3 e 1.4, nas quais ilustram-se as aplicações dadas como exemplo envolvendo perfis convencionais e perfis de imagem.

Esta tese propõe uma metodologia autossupervisionada para resolver os problemas acima. Assim, pode-se explorar as potencialidades do aprendizado de representações e, em particular para a aplicação Tipo-II, o aprendizado regenerativo. Os modelos autossupervisionadas que estão na base do método proposto⁵ são: autocodificador⁶ variacional (KINGMA; WELLING, 2013)

⁵A proposta prevê flexibilidade quanto ao método utilizado nas etapas de aprendizado de representações e aprendizado regenerativo. A rede SOM, porém, está no cerne da metodologia.

⁶Traduções para português já foram tentadas antes, como ‘rede auto-associativa’ (FROTA, 2005). Apesar de o termo em inglês ser usado com naturalidade em textos na língua portuguesa, optou-se pelo termo autocodificador no presente documento.

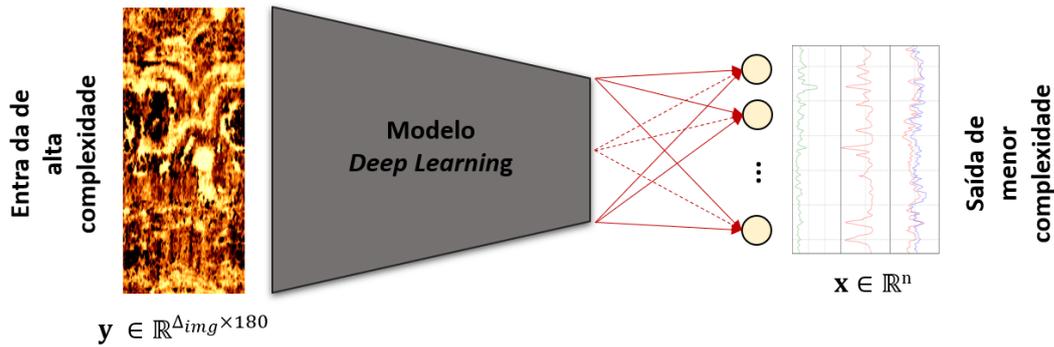


Figura 1.3: Aplicação Tipo-I: geração de dados mais simples a partir de dados mais complexos. Ilustração de solução usando técnicas de aprendizado profundo.

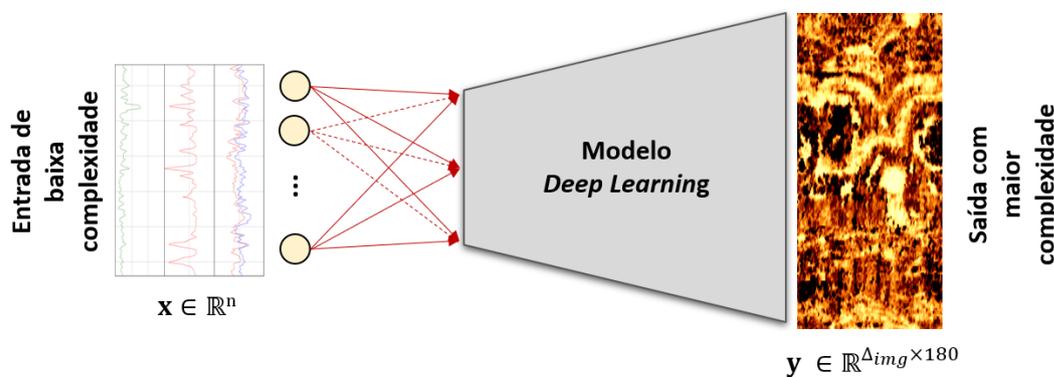


Figura 1.4: Aplicação Tipo-II: geração de dados mais complexos a partir de dados menos complexos. Ilustração de solução usando técnicas de aprendizado profundo.

(VAE) e mapas auto-organizáveis (SOM⁷) (KOHONEN, 1990). Cabe ressaltar que, apesar de ser naturalmente um algoritmo não supervisionado, a maneira como a rede SOM será utilizada permite que seja classificada aqui como um modelo autossupervisionado. A integração desses modelos, como ilustra a Figura 1.5, pode resolver ao mesmo tempo ambos os tipos de aplicações, e a maneira como isso pode ser feito é a principal contribuição teórica desta tese.

Modelos de aprendizado profundo (LI et al., 2022; KASEB; KHALED; GALAL, 2022), como nas Figuras 1.3 e 1.4, caso existentes na literatura, podem servir como base de comparação para os resultados obtidos pela nova metodologia a ser proposta nesta tese.

⁷Sigla que vem da nomenclatura em inglês *self-organizing maps*.

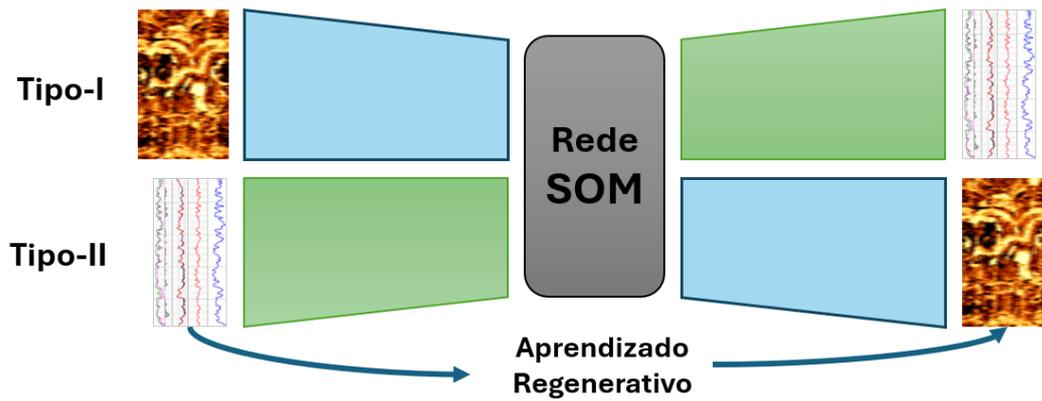


Figura 1.5: Metodologia autossupervisionada proposta nesta tese. Ilustra-se seu uso na solução de dois problemas de geração de dados petrofísicos. Destaca-se a Aplicação Tipo-II, que explora o aprendizado regenerativo.

1.5 Organização da Tese

Além deste capítulo introdutório, a tese está estruturada em cinco capítulos e dois apêndices.

O Capítulo 2 explora o tema do aprendizado de representações, com ênfase nos modelos de redes neurais que servirão de base para a metodologia proposta. A motivação para o uso desses modelos é fundamentada na revisão de trabalhos anteriores relevantes, que apresentam uma interface significativa com a presente proposta, seja pelo uso de modelos semelhantes ou por abordarem aplicações que envolvem dados de diferentes modalidades.

O Capítulo 3 descreve a proposta desta tese. Inicialmente, a proposta é apresentada de forma conceitual e geral. Em seguida, são definidos os dois tipos de aplicação que podem ser abordados pela metodologia proposta. A notação e os passos da proposta são formalizados de maneira abrangente, e as duas variantes da metodologia são definidas: BERGSOM e p -BERGSOM. Uma descrição detalhada dos modelos abordados no Capítulo 2 é elaborada, explicando como são utilizados em cada etapa da metodologia. Ao final, são utilizados dados da base MNIST (LECUN; CORTES; BURGESS, 2010) para ilustrar a metodologia e suas variantes em cada tipo de aplicação.

O Capítulo 4 é dedicado à aplicação Tipo-I em um problema real de petrofísica de reservatórios, focando na geração de um conjunto completo de perfis básicos a partir de um perfil de imagem acústica. Os resultados das variantes BERGSOM e p -BERGSOM para a aplicação Tipo-I são apresentados, e a metodologia proposta é comparada com um modelo convolucional profundo da literatura. Ao final, os resultados são contextualizados com base em trabalhos recentes correlatos na área de petrofísica, acompanhados de discussões

sobre os achados e considerações acerca do uso da metodologia proposta em aplicações do Tipo-I.

O Capítulo 5 é dedicado à aplicação Tipo-II no contexto da petrofísica de reservatórios, focando na geração de um perfil de imagem acústica a partir de um conjunto completo de perfis básicos. Os resultados das variantes BERGSOM e p -BERGSOM para essa aplicação mais complexa, caracterizada como aprendizado regenerativo, são apresentados. Demonstra-se que a proposta atual é uma generalização da ideia de imputação de valores faltantes utilizando a rede SOM, evidenciando sua flexibilidade ao permitir regerar um perfil de imagem a partir de subconjuntos que contêm apenas parte dos perfis básicos originais utilizados no treinamento dos modelos. Ao final, são discutidos os resultados e apontadas as especificidades da metodologia proposta para esse tipo de aplicação.

O Capítulo 6 apresenta as conclusões deste trabalho e indica possíveis direções futuras para novos desenvolvimentos a partir dele.

O Apêndice A detalha os dados utilizados, além de explicar as decisões de pré-processamento adotadas para as aplicações desenvolvidas.

O Apêndice B apresenta a produção científica do autor, juntamente com seus orientadores e demais coautores, ao longo do período do curso de doutoramento.

2

Modelos para Aprendizado de Representações

Pode-se dizer que o aprendizado de representações (AR) é uma evolução de duas técnicas clássicas: extração de características e engenharia de características (*feature engineering*), no sentido em que essas técnicas buscam encontrar o melhor conjunto de atributos numéricos a partir dos dados de entrada para obter o máximo de desempenho de um modelo de aprendizado de máquinas em dada tarefa. O AR, porém, torna-se não só uma técnica, mas um ponto focal do aprendizado de máquinas quando não apenas o desempenho, mas a própria solução do problema, depende da definição de uma representação apropriada para os dados (CHARTE et al., 2020).

Com o advento do aprendizado profundo (LECUN; BENGIO; HINTON, 2015), AR passou a ser integrada aos modelos de redes neurais, como, de fato, Rumelhart, Hinton e Williams (1986) já sugeriam em seu trabalho seminal sobre o algoritmo de aprendizado via retropropagação de erros. Desde então, AR tem servido como base conceitual e prática para um grande número de aplicações envolvendo o uso de algoritmos de aprendizado de máquinas.

Este capítulo tratará das duas redes neurais autossupervisionadas que serão utilizadas no desenvolvimento desta tese. A motivação para a escolha desses modelos será fundamentada oportunamente. A primeira delas é modelo de rede neural multicamadas do tipo autocodificadora; a segunda é o clássico mapa de Kohonen, que é originalmente um algoritmo de aprendizado não supervisionado. Nesta tese, ambas são treinadas num modo híbrido que passou a ser conhecido como *autossupervisionado*. Destaca-se o potencial dessas redes no aprendizado de representações e seu uso conceitual em aplicações diversas.

2.1

Redes neurais autocodificadoras

A arquitetura hoje conhecida como rede autocodificadora¹ (*autoencoder*, AE) foi originalmente proposta como um método para inicializar redes neurais do tipo MLP (BALLARD, 1987). Kramer (1991) trouxe à luz a interpretação da representação interna dos autocodificadores como uma extração não-linear de componentes principais, o que está no fulcro do aprendizado de representações. Desde então, o uso dos AEs vem crescendo, passando desde as hoje clássicas aplicações de detecção de falhas (PETSCHKE et al., 1995) até discri-

¹Também referida no texto como autocodificador.

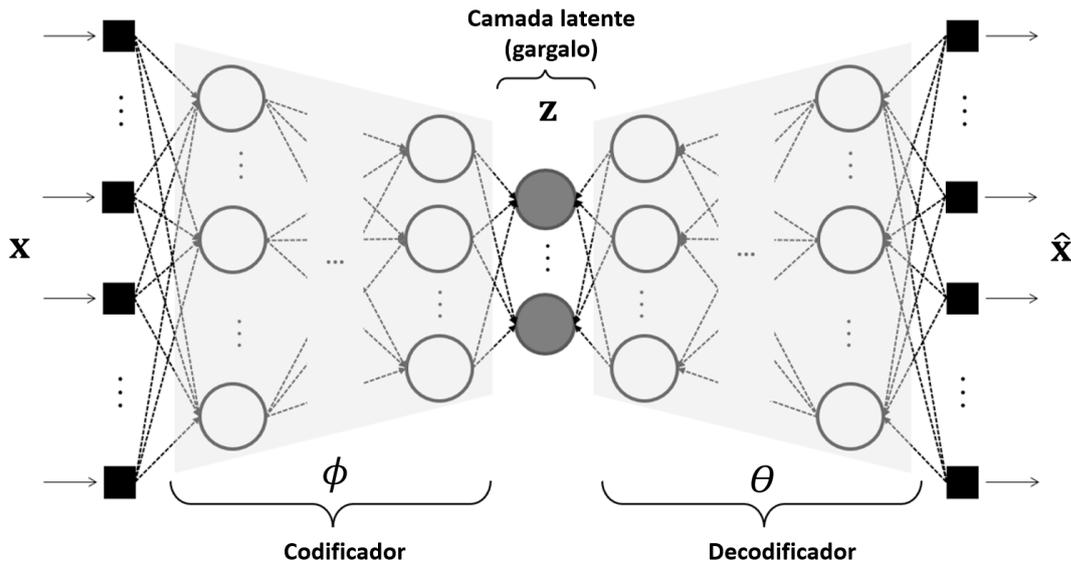


Figura 2.1: Arquitetura da rede neural autocodificadora.

minação de classes usando a representação latente aprendida pelo AE (LIMA et al., 2019).

O autocodificador é um tipo de rede neural multicamadas cujo objetivo é reproduzir o vetor de atributos de entrada na camada de saída. Esse mapeamento é feito impondo uma arquitetura na qual existe uma contração da dimensão de entrada, geralmente na camada intermediária da rede. Esta arquitetura é alternativamente chamada de rede gargalo (*bottleneck*), pois a camada central tem dimensão menor que as camadas anteriores e posteriores, conforme mostra a Figura 2.1.

Um AE pode ser dividido em duas partes, um codificador (parametrizado por uma função de mapeamento ϕ) que transforma o dado de entrada para uma dimensão latente normalmente com dimensão reduzida, e um decodificador (parametrizado por uma função de mapeamento θ), que tenta reconstruir o vetor de entrada original a partir da representação latente. Como não se faz uso de dados rotulados associados aos dados de entrada para seu treinamento, considera-se, no presente trabalho, que os autocodificadores pertencem ao conjunto de algoritmos de aprendizado autossupervisionado, apesar de ser treinado como uma rede MLP padrão.

A função objetivo utilizada no treinamento o AE requer a minimização da diferença entre a entrada fornecida $\mathbf{x} \in \mathbb{R}^p$ e a saída gerada (ou reconstruída) pela rede, $\hat{\mathbf{x}} \in \mathbb{R}^p$, que tenta reconstruir a entrada. Essa diferença é conhecida como *erro de reconstrução*, $d(\mathbf{x}, \hat{\mathbf{x}})$. Uma das métricas utilizadas para o cálculo do erro de reconstrução é o erro médio quadrático (2-1):

$$d(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{N} \sum_{i=1}^p (x_i - \hat{x}_i)^2, \quad (2-1)$$

na qual x_i e \hat{x}_i são as componentes do vetor \mathbf{x} e $\hat{\mathbf{x}}$, respectivamente e N é o total de instâncias no conjunto de dados de treinamento.

Por questão de uniformidade de notações, define-se novamente a função de custo como na equação a seguir:

$$\mathcal{L}(\theta, \phi; \mathbf{x}) = d(\mathbf{x}, \hat{\mathbf{x}}), \quad (2-2)$$

na qual se especifica que o erro a ser minimizado depende de ϕ , que é o mapeamento do codificador, e de θ , que é o mapeamento do decodificador. Os parâmetros que definem os mapeamentos ϕ e θ são aprendidos a partir dos dados, ao minimizar a perda definida pela Equação 2-2.

Desta forma, o AE é capaz de construir um espaço latente \mathcal{Z} cujos pontos estão biunivocamente associados às entradas usadas no treinamento. Ou seja, para cada entrada \mathbf{x} o codificador mapeia uma única representação \mathbf{z} no espaço latente, enquanto o decodificador mapeia essa representação biunivocamente a uma reconstrução $\hat{\mathbf{x}}$.

Vale destacar que uma rede do tipo autocodificadora com camada interna do tipo gargalo com mapeamentos ϕ e θ lineares, i.e., cujas unidades têm funções de ativação lineares, produz resultados equivalentes ao método PCA (COTTRELL; MUNRO; ZIPSER, 1987).

Apesar de sua utilidade como ferramenta para redução de dimensionalidade e para construção de representações compactas para diversos tipos de dados, o AE convencional não tem bom desempenho quando utilizado como modelo generativo de novos dados (REZENDE; MOHAMED; WIERSTRA, 2014), pois sua capacidade de extrapolação para informações fora do conjunto de treinamento é muito limitada. Para fins de aprendizado generativo, foi proposta uma alteração conceitual na arquitetura do AE de modo a permitir uma interpretação estocástica do espaço latente e, portanto, fazer inferência generativa a partir dele.

2.1.1

Autocodificador Variacional (VAE)

Uma forma de permitir que uma arquitetura do tipo AE possa funcionar para o aprendizado generativo é regularizar o espaço das representações por meio da imposição de que ele aprenda alguma distribuição definida *a priori* $p(\mathbf{z})$ (KINGMA; WELING, 2013). Essa modificação dá origem ao autocodificador variacional (*variational autoencoder*, VAE), cuja arquitetura é mostrada na Figura 2.2 com o objetivo de compará-la à arquitetura básica do AE mostrada na Figura 2.1.

Ao observar a Figura 2.2, nota-se a inclusão de uma camada para

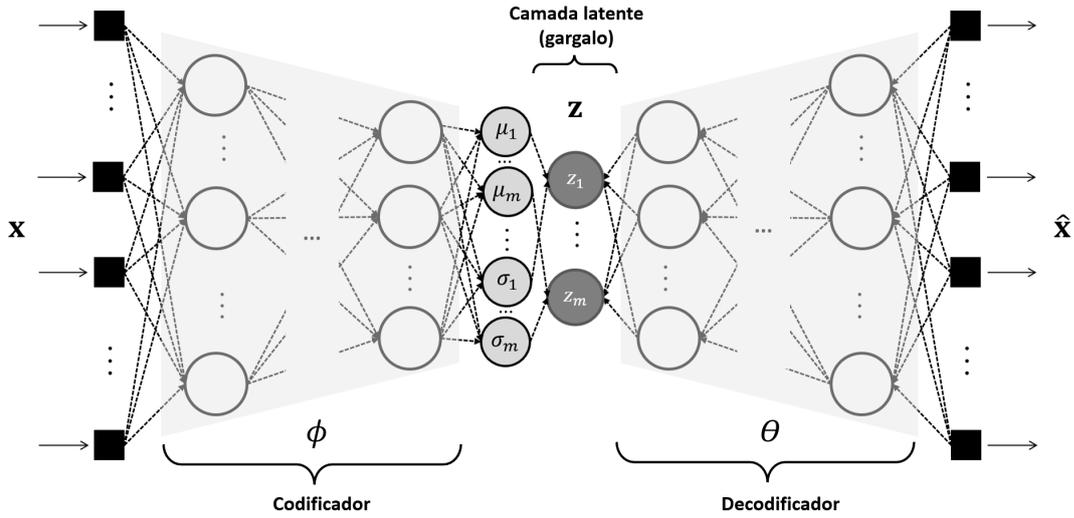


Figura 2.2: Arquitetura da rede neural do tipo autocodificador variacional.

aprender os parâmetros da distribuição imposta ao espaço latente. No caso ilustrado, $p(\mathbf{z})$ é definida por uma função gaussiana multivariada, i.e., $\mathbf{z} = p(\mathbf{z}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, na qual $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)$ e $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$. Assim, a representação latente na camada gargalo é calculada conforme a Equação 2-3:

$$z_i = \mu_i + \sigma_i \varepsilon, \quad (2-3)$$

para $i = 1, \dots, m$, e $\varepsilon \sim \mathcal{N}(0, 1)$. A Equação 2-3 torna o espaço latente estocástico e representa uma operação que vem sendo chamada de truque da reparametrização (*reparametrization trick*). Ao prover uma expressão determinística para uma variável estocástica, o truque da reparametrização permite realizar o treinamento do VAE da mesma maneira que se treina um AE convencional, usando, por exemplo, o método do gradiente descendente via algoritmo de retropropagação de erros.

Com essa alteração, cada vetor de entrada \mathbf{x} dos dados de treinamento a ser reconstruído fica associado não a um ponto único no espaço latente \mathcal{Z} , mas a uma distribuição. Isso permite, por exemplo, amostrar um ponto qualquer no espaço das representações que, ao passar pelo decodificador do VAE, gerará um vetor de saída que não esteja presente nos dados de treinamento, mas que guarde coerência em relação aos dados originais. Assim, pode-se dizer que os VAEs são modelos generativos estocásticos, pois não mapeiam cada entrada em um único ponto, mas a uma distribuição (CHARTE et al., 2020).

Ao assumir uma interpretação estocástica do VAE, suas duas partes podem ser interpretadas como:

1. **Codificador:** um modelo de mapeamento parametrizado por ϕ , $q_\phi(\mathbf{z} | \mathbf{x})$
2. **Decodificador:** um modelo generativo parametrizado por θ , $p_\theta(\mathbf{x} | \mathbf{z})$

A função objetivo do VAE é definida em termos do *evidence lower bound* (ELBO), também chamado de *variational lower bound* ou *variational free energy* (DAMM et al., 2020), como

$$\mathcal{L}(\theta, \phi; \mathbf{x}) = -D_{KL}[q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})] + \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x}|\mathbf{z})] \quad (2-4)$$

na qual se quer maximizar $\mathcal{L}(\theta, \phi; \mathbf{x})$ com respeito aos parâmetros de ϕ e θ . $D_{KL}[q_\phi(\mathbf{z}|\mathbf{x})||p(\mathbf{z})]$ refere-se à divergência de Kullback-Leibler entre o mapeamento aprendido pelos dados, $q_\phi(\mathbf{z}|\mathbf{x})$, e a distribuição definida *a priori*, $p(\mathbf{z})$.

Maximizar o segundo termo do lado direito da Equação 2-4 equivale a minimizar o erro de reconstrução definido anteriormente para o AE regular. Reduzir o termo D_{KL} (ou maximizar $-D_{KL}$) significa aproximar a distribuição do modelo de mapeamento $q_\phi(\mathbf{z}|\mathbf{x})$ da verdadeira distribuição posterior $p_\theta(\mathbf{x}|\mathbf{z})$, funcionando, então, como regularizador do mapeamento ϕ .

Foi mostrado posteriormente por Higgins et al. (2017) que um fator ponderador $\beta > 1$ sobre o termo de regularização D_{KL} na Equação 2-4 melhora o desempenho do VAE em tarefas de aprendizado generativo. Esse tipo de VAE tem sido chamado de β -VAE, cuja função objetivo fica, então, mais geral:

$$\mathcal{L}(\theta, \phi; \mathbf{x}) = -\beta D_{KL}[q_\phi(\mathbf{z} | \mathbf{x})||p(\mathbf{z})] + \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})}[\log p_\theta(\mathbf{x} | \mathbf{z})]. \quad (2-5)$$

Nas aplicações de interesse para o presente trabalho, as entradas dos modelos do tipo β -VAE utilizados são imagens, convém, portanto, apresentar um outro tipo de β -VAE adequado para esse tipo de dado, dotado de camadas convolucionais.

Neste trabalho, utilizam-se VAEs do tipo β -VAE com $\beta = 1,15^2$

2.1.2

Autocodificador Variacional Convolutacional

Neste trabalho, refere-se a autocodificadores variacionais convolucionais³ (CoVAEs) para descrever VAEs que usam camadas convolucionais para extração de atributos à medida em que a rede se aprofunda até a camada central (gargalo) de representações latentes. Uma estrutura semelhante é usada pelo decodificador para reconstruir a entrada. A Figura 2.3 ilustra esse tipo de autocodificador:

Exceto pelo uso das camadas convolucionais, o restante das características do β -VAE regular são preservadas no β -CoVAE.

²O valor para o hiperparâmetro β foi definido por testes de validação prévios, a partir do comportamento das curvas de erro referentes à componente D_{KL} .

³Usa-se a sigla CoVAE a partir da nomenclatura em inglês *convolutional variational autoencoders*, ou ainda β -CoVAE por coerência.

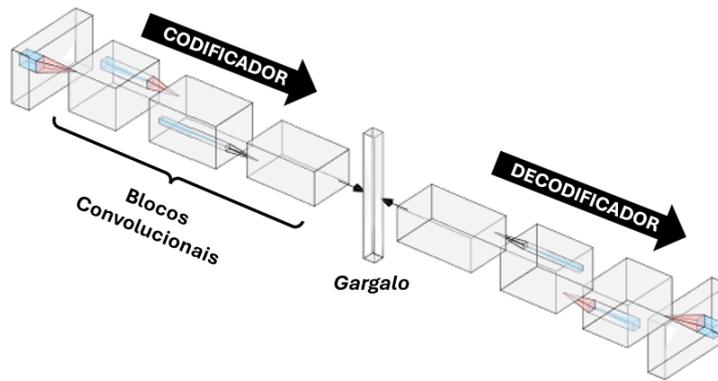


Figura 2.3: Exemplo ilustrativo de um autocodificador variacional convolucional (β -CoVAE).

2.1.3

O uso de modelos β -CoVAE no presente trabalho

Os objetivos apresentados no Capítulo 1 demandam o uso de um modelo de aprendizado de representações que seja, ao mesmo tempo, um modelo generativo, pois pretende-se efetuar um ciclo bidirecional de geração / regeneração de dados por meio da metodologia a ser proposta.

Apesar de ser um modelo clássico, o autocodificador variacional cumpre esses dois objetivos e provê ainda todo o arcabouço conceitual necessário para formalizar a metodologia sem perda de generalidade, caso se escolha outro modelo para as tarefas de aprendizado de representações e geração/regeneração de dados.

Além disso, o uso do modelo β -CoVAE permite uma interessante inovação que será explorada oportunamente; a saber, o uso de *atributos latentes probabilísticos* na etapa de aprendizado heteroassociativo.

2.2

Mapas de Kohonen

O mapa auto-organizável (KOHONEN, 2001; KOHONEN, 1990), ou rede SOM, é uma rede neural clássica de aprendizado não supervisionado que aprende um mapeamento não linear $\Phi: \chi \rightarrow \mathcal{A}^4$ do espaço de entrada contínuo χ para um espaço de saída discreto \mathcal{A} , definido por um arranjo geométrico em cujos nós estão posicionados os neurônios, ou unidades, da rede (HAYKIN, 1999).

É habitual ilustrar o mapeamento da rede SOM graficamente como na Figura 2.4, que também destaca a unidade vencedora i^* (em preto) e a sua vizinhança imediata (em coloração cinza).

⁴Não confundir o mapeamento Φ da rede SOM, com os parâmetros ϕ do codificador de um AE.

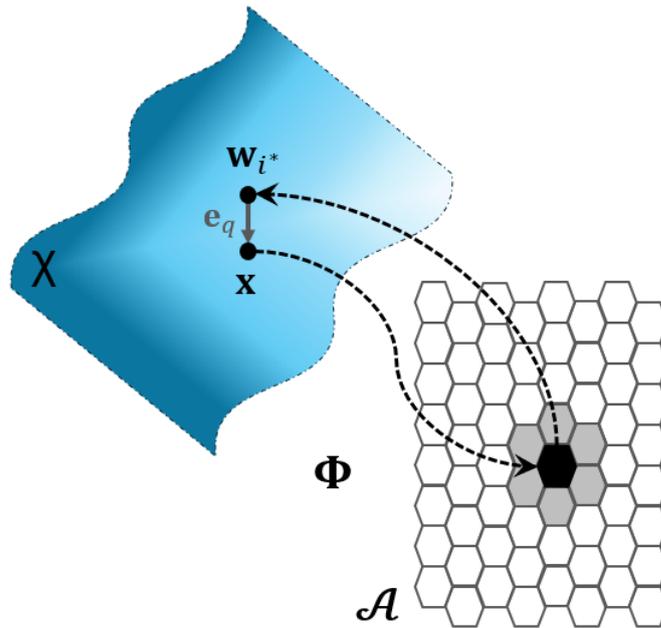


Figura 2.4: Ilustração conceitual do mapeamento $\Phi : \mathcal{X} \rightarrow \mathcal{A}$ implementado pela rede SOM.

Na rede SOM, cada neurônio $i \in \mathcal{A}$ carrega um vetor de pesos \mathbf{w}_i que tem a mesma dimensão do espaço de entrada χ . Assim, os pesos de todas os neurônios encontram-se no espaço de entrada original e podem servir de protótipos para diferentes sub-regiões em χ após a convergência do treinamento da rede. A Figura 2.4 ilustra um vetor de entrada \mathbf{x} que ativa o neurônio vencedor (i^*) no mapa que, por sua vez, está associado a um vetor de pesos \mathbf{w}_{i^*} no espaço original de entrada.

As propriedades acima comentadas justificam a utilização da rede SOM em tarefas de agrupamento de dados. Do mesmo modo, a rede SOM é amplamente utilizada para visualizar dados de elevada dimensão numa grade de dimensão reduzida (normalmente 2D). O fato de a rede SOM prover representações coerentes dos dados de entrada por meio dos protótipos de suas unidades tem motivado o emprego da rede SOM, inclusive, em modelos generativos como no trabalho de Kim e Jung (2019).

A rede SOM realiza uma quantização vetorial (QV) do espaço de entrada no mapa \mathcal{A} , ou seja, cada vetor de pesos da rede SOM funciona como o vetor representativo para um subconjunto de padrões de entrada. QV é uma técnica de compressão que envolve perda de informações. Quanto menor o número de unidades do SOM, maior será essa perda. A distância quadrática entre \mathbf{x} e \mathbf{w}_{i^*} calculada segundo alguma métrica é chamada de *erro quadrático de quantização* (e_q^2), cujo o somatório para todas as instâncias de treinamento atua como uma espécie de função objetivo que se quer minimizar no treinamento

da rede SOM (KOHONEN; NIEMINEN; HONKELA, 2009). A métrica usual é a distância euclidiana, conforme a Equação 2-6:

$$e_q^2 = \|\mathbf{e}_q\|^2 = \|\mathbf{x} - \mathbf{w}_{i^*}\|^2. \quad (2-6)$$

Abaixo são mostradas as equações que regem o treinamento da rede SOM. Inicialmente, a escolha do neurônio vencedor, i^* no instante t , se dá por uma busca pelo vetor protótipo mais próximo. Formalmente, tal busca é implementada pela seguinte expressão:

$$i^*(t) = \arg \min_{i \in \mathcal{A}} \|\mathbf{x}(t) - \mathbf{w}_i(t)\|, \quad (2-7)$$

na qual t indica a iteração de treinamento, $\mathbf{x}(t)$ e $\mathbf{w}_{i^*}(t)$ representam a entrada e os pesos na rede na iteração t .

Em seguida, os pesos do neurônio vencedor e daqueles em sua vizinhança são atualizados:

$$\Delta \mathbf{w}_i(t) = \alpha(t) h(i^*, i; t) [\mathbf{x}(t) - \mathbf{w}_i(t)], \quad (2-8)$$

na qual $0 < \alpha(t) < 1$ denota a taxa de aprendizagem, e $h(i^*, i; t)$ é a função de vizinhança, geralmente do tipo gaussiana:

$$h(i^*, i; t) = \exp\left(-\frac{\|\mathbf{r}_i - \mathbf{r}_{i^*}\|^2}{\vartheta^2(t)}\right), \quad (2-9)$$

na qual $\vartheta(t)$ controla o raio de abrangência da função de vizinhança; \mathbf{r}_i e \mathbf{r}_{i^*} são, respectivamente, as posições dos neurônios i e i^* no arranjo geométrico da rede SOM.

Os parâmetros α e ϑ devem decair com o decorrer das iterações de treinamento para garantir a convergência do algoritmo. Uma maneira comum de fazer esse tipo de redução é usar uma regra de decaimento exponencial, mostrada a seguir para o caso da taxa de aprendizagem:

$$\alpha(t) = \alpha_0 \left(\frac{\alpha_T}{\alpha_0}\right)^{(t/T)}, \quad (2-10)$$

na qual α_0 e α_T ($\alpha_T \ll \alpha_0$)⁵ são os valores inicial e final de α , e T é o número máximo de iterações. A velocidade de decaimento é controlada pelo parâmetro T , que simboliza o número máximo de iterações de treinamento.

O Algoritmo 1 sumariza o treinamento sequencial da rede SOM:

2.2.1

O uso da rede SOM no presente trabalho

A despeito do enorme número de trabalhos que usam a rede SOM ou variantes dela desde sua proposição, nos últimos anos ela vinha sendo aplicada

⁵Kohonen (KOHONEN, 1990) sugere α_0 de 0,01 ou 0,02 e que α_T pode ser zerado após um certo número de épocas.

Algoritmo 1: Treinamento sequencial da rede SOM

```

1 Definir o tamanho do mapa (número de neurônios);
2 Inicializar os pesos  $\mathbf{w}$ ;
3 enquanto não alcançar máximo de épocas faça
4   para  $t \leftarrow 1$  até  $p$  faça
5     Amostrar um vetor  $\mathbf{x}$  do conjunto de dados;
6     Encontrar o neurônio vencedor usando a Equação 2-7;
7     Atualizar os pesos da rede conforme a Equação 2-8

```

sobretudo como método de agrupamento e visualização de dados, que são apenas duas de suas muitas funcionalidades.

Neste trabalho, explora-se a rede SOM como um algoritmo autossupervisionado capaz de aprender um mapeamento entre dados de diferentes modalidades usando o princípio de memória associativa, porém ampliando esse conceito para mapeamento heteroassociativo, como será detalhado no Capítulo 3. Essa é uma das mais interessantes funcionalidades da rede SOM, tanto que Yin (2008) chega a afirmar que o termo *self-organizing maps* “vem da definição e do objetivo original de Kohonen, a saber, a memória associativa”.

Para usar a rede SOM como um estimador baseado em memória associativa, utiliza-se para seu treinamento um vetor de entrada $\mathbf{x} \in \mathbb{R}^{(m+n)}$ que incorpora valores de entrada \mathbf{x}_{in} e respectivos valores de saída desejada \mathbf{x}_{out} . Matematicamente, tem-se a seguinte definição:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_{in} \\ \mathbf{x}_{out} \end{bmatrix} = [x_1^{in}, \dots, x_m^{in}, x_1^{out}, \dots, x_n^{out}]^T. \quad (2-11)$$

Ao se treinar a rede SOM com uma entrada como definida pela Equação 2-11, faz-se que a rede implicitamente codifique a distribuição conjunta $P(\mathbf{x}^{in}, \mathbf{x}^{out})$ dos dados de treinamento em seus protótipos, ou seja, $P(\mathbf{w}_i^{in}, \mathbf{w}_i^{out})$. Portanto, em termos práticos, o vetor de pesos do i -ésimo neurônio apresenta a mesma estrutura do vetor de entrada aumentado descrito pela Equação 2-11:

$$\mathbf{w}_i = \begin{bmatrix} \mathbf{w}_i^{in} \\ \mathbf{w}_i^{out} \end{bmatrix} = [w_{i_1}^{in}, \dots, w_{i_m}^{in}, w_{i_1}^{out}, \dots, w_{i_n}^{out}]^T. \quad (2-12)$$

Barreto e Araujo (2004) demonstraram que o uso da rede SOM como estimador baseado em memória associativa equivale ao aprendizado supervisionado como o das redes MLP e RBF e o classificam como aprendizado autossupervisionado.

Cabe destacar que a ideia de incluir num mesmo vetor informações de entrada e informações de saída para treinamento da rede SOM remonta aos anos

1980 (RITTER; SCHULTEN, 1986; RITTER; KOHONEN, 1989; RITTER; MARTINETZ; SCHULTEN, 1989). Uma versão temporal de memória associativa usando a rede SOM, chamada de *vector-quantized temporal associative memory* (VQTAM), foi proposta para aplicações de previsão de séries temporais (LENDASSE et al., 2005; BARRETO; ARAUJO, 2002; BARRETO; ARAUJO, 2001), controle (BARRETO; ARAUJO; RITTER, 2003) e identificação de sistemas (BARRETO; ARAUJO, 2004).

Nos últimos anos, têm sido publicados trabalhos usando a rede SOM para memória associativa. A Tabela 2.1 traz uma seleção com alguns dos mais recentes, destacando-se suas aplicações, dimensões dos vetores de entrada e saída e geometria da rede SOM.

Tabela 2.1: Seleção de trabalhos que exploram a rede SOM como memória associativa.

Referência	Aplicação	$dim(\mathbf{x}_{in})$	$dim(\mathbf{x}_{out})$	Geometria SOM
Mukwachi, Arnonkijpanich e Sarakorn (2024)	Modelagem magnetotelúrica 2D	$\mathbb{R}^{N_d \times 3}$, $N_d = N_S \times N_T$ $N_T (= 31)$ número de períodos de amostragem	$\mathbb{R}^{N_d \times 2}$	40×40
Limtrakul e Arnonkijpanich (2017)	Modelagem cinemática	$N_S (= 21)$ número de estações de amostragem $\mathbb{R}^{(n_y + n_u) \times 6}$ $n_u (= 2)$ ordem de entrada do sistema $n_y (= 3)$ ordem de saída do sistema	\mathbb{R}^6	15×15
Wongswiworaphon, Arnonkijpanich e Pathummakul (2015)	Estimativa de peso de animais	\mathbb{R}^2 , métrica bidimensional extraída da imagem do animal	\mathbb{R}^1 , peso do animal	13×13
Barreto e Araujo (2004)	Identificação de sistemas SISO/MIMO	$\mathbb{R}^{n_y + n_u}$ $n_u (= 1)$ ordem de entrada do sistema $n_y (= 2)$ ordem de saída do sistema	\mathbb{R}^1 para SISO $\mathbb{R}^{N_2 - N_1 + 1}$ para MIMO	1×52
Barreto e Araujo (2002)	Identificação direta/inversa de atuador hidráulico.	$n_y (= 2)$ ordem de entrada do sistema $\mathbb{R}^{n_u + n_y}$ $n_u (= 2)$ ordem de entrada do sistema $n_y (= 3)$ ordem de saída do sistema	\mathbb{R}^1 para a saída ou \mathbb{R}^{n_u} para a entrada	1×500
Barreto e Araujo (2001)	Predição de séries temporais caóticas	\mathbb{R}^p p é o atraso da série	\mathbb{R}^1 , predição 1 passo à frente	1×400

Tabela 2.2: Seleção de trabalhos que associam a rede SOM ao aprendizado profundo.

Referência	Modelo	Aplicação	Rede Profunda	RegL?	Dataset
Li, He e Tang (2024)	SOMGAN	Aprendizado generativo com mitigação do colapso modal	GANs com restrição topológica (via rede SOM)	Não	MNIST, CIFAR-10, STL-10, CelebA, ImageNet
Muliukov, Rodriguez e Miramond (2023)	ReD-SOM	Reconstrução de sinal multimodal	CoVAE e conexões hebbianas	Sim	MNIST + Spoken MNIST, MNIST + Spoken MNIST + FMNIST
Sakkari et al. (2022)	G-UDSOM	Extração de características	AE Convolutacional (CoAE) e DCSOM	Não	MNIST, STL-10, AFHQ, USC-SIPI
Forest et al. (2021)	DESOM	Análise de agrupamentos	AE e CoAE	Não	MNIST, Fashion-MNIST, USPS, REUTERS-10k
Aly e Almotairi (2020)	DCSOM	Extração de características	SOM Convolutacional	Não	MNIST, variações do MNIST
		Reconhecimento de dígitos manuscritos			

Os trabalhos citados na Tabela 2.1, em sua maioria, exploram a memória associativa usando SOM em aplicações autoassociativas (KOHONEN, 2001), nas quais os dados de entrada e saída têm a mesma modalidade. O trabalho de Wongsriworaphon, Arnonkijpanich e Pathumnakul (2015) é um dos poucos a propor uma aplicação que envolve diferentes modalidades, na qual a imagem de entrada é transformada (sem o uso de aprendizado profundo) em uma medida bidimensional para previsão de um único valor de saída. Porém, além dessa proposta não envolver aprendizado regenerativo, tampouco se propõe a realizar a geração ou regeneração cruzada de dados de maneira bidirecional.

Diante do exposto, esta tese preenche essas lacunas ao utilizar o princípio de memória associativa em uma abordagem bidirecional envolvendo diferentes modalidades (i.e., tipos) de dados que se enquadra no paradigma de aprendizado regenerativo, pois se propõe a gerar (ou regerar) dados complexos a partir de entradas muito mais simples em termos de conteúdo de informação.

2.3

Associação da rede SOM com Aprendizado Profundo

Já foi mencionado que, nesta tese, a rede SOM será usada em conjunção com redes profundas, em particular autocodificadores variacionais, em problemas complexos de petrofísica.

Trabalhos recentes chegam mesmo a utilizar a rede SOM em conjunto com algum modelo de aprendizado profundo como AE ou redes generativas adversariais (*generative adversarial networks*, GANs) (GOODFELLOW et al., 2014). Na maioria dos casos, as redes são treinadas juntas, a partir da minimização ponta-a-ponta de uma função objetivo composta por uma parcela da rede SOM e uma parcela da rede com aprendizado profundo. Essas são abordagens mais complexas que a proposta neste trabalho.

Ao incorporar a rede SOM a uma arquitetura supervisionada de aprendizado profundo, busca-se explorar a capacidade da rede SOM em preservar relações de distância entre as instâncias no espaço de entrada nas relações de distância entre os protótipos da rede. Esses modelos usam a rede SOM para impor uma ordem ao espaço de representações dos dados, por exemplo, no caso de autocodificadores, o espaço latente \mathcal{Z} .

A Tabela 2.2 mostra alguns dos mais recentes, destacando-se suas aplicações, tipos de redes profundas usadas junto com a rede SOM e base de dados utilizadas nos testes.

Dentre os trabalhos mostrados na Tabela 2.2, apenas um trata de aprendizado entre diferentes modalidades de dados (MULIUKOV; RODRIGUEZ; MIRAMOND, 2023), mas é uma aplicação muito simplificada com a base

de dados MNIST e Spoken MNIST (KHACEF; RODRIGUEZ; MIRAMOND, 2019). O modelo ReD-SOM emprega dois módulos DESOM unidos por conexões hebbianas que comunicam o espaço latente de cada modalidade. Nesse sentido, assim como o presente trabalho, o modelo ReD-SOM é capaz de realizar aprendizado regenerativo.

Os modelos citados são baseados em arquiteturas complexas e apenas uma delas se propõe a realizar aprendizado regenerativo. A otimização de ponta-a-ponta impõe a essas arquiteturas um considerável peso computacional que nem sempre se justifica pela complexidade das aplicações que buscam solucionar.

Na presente tese a rede SOM é utilizada em conjunto com um autocodificador convolucional variacional, não em uma arquitetura com otimização de ponta-a-ponta, mas como módulos sequenciais. Além de exigir menor poder computacional, a metodologia a ser proposta torna-se mais flexível, permitindo o uso de outras redes na etapa de aprendizado de representações e aprendizado regenerativo.

Em suma, ao aplicar memória associativa baseada em rede SOM associada ao poder de aprendizado de representações e aprendizado regenerativo de modelos de aprendizado profundo, a metodologia proposta pretende fornecer uma abordagem eficaz e relativamente leve em termos computacionais para solucionar problemas complexos de aprendizado regenerativo.

2.4

Considerações finais

Este capítulo apresentou os modelos de redes neurais que servirão de base para a metodologia proposta, com ênfase em seu uso para aprendizado de representações. Destacou-se, especialmente, a utilização da rede SOM como estimador baseado em memória associativa, dada sua relevância para a proposta desta tese. Uma revisão bibliográfica foi realizada para contextualizar a escolha das redes utilizadas e identificar lacunas que serão abordadas pela presente proposta.

No Capítulo 3, a metodologia será descrita e formalizada, além de ilustrar seu funcionamento por meio de experimentos simplificados.

3

Memória Associativa Baseada em SOM para Dados de Diferentes Modalidades

Neste capítulo será fundamentada a ideia central da proposta desta tese. A metodologia ora desenvolvida é baseada em uma modificação conceitual do uso da rede SOM para memória associativa que permite sua expansão para o caso mais geral de memória de associação entre dados de diferentes modalidades. Assim, este novo mecanismo de memória torna-se capaz de aprender um mapeamento não linear entre diferentes modalidades de dados, aqui chamados de dados de origem e dados de destino.

Em seguida, a metodologia é apresentada em etapas gerais, sem vinculação imediata a nenhum modelo específico. A seguir, duas variantes são propostas para esta metodologia, ambas são detalhadas e contextualizadas em termos das aplicações de interesse para esta tese.

Ao final, um experimento simplificado com dados MNIST é apresentado para ilustração do funcionamento da metodologia, sem objetivo de aferição de desempenho.

3.1

Mapeamento Heteroassociativo Multimodalidade

Nesta tese, revisita-se uma modificação conceitual originalmente proposta por Ritter e Schulten (1986) que permite à rede SOM aprender mapeamentos entrada-saída. Para tanto, aumenta-se a entrada da rede $\mathbf{x} \in \mathbb{R}^{m+n}$ para incorporar tanto os valores originais de entrada quanto os valores de saída correspondentes:

$$\mathbf{x} = [x_1^{in}, \dots, x_m^{in}, x_1^{out}, \dots, x_n^{out}]^T, \quad (3-1)$$

m e n são as dimensões dos dados de entrada e de saída, respectivamente.

A rede SOM é, então, implicitamente forçada a codificar esta relação entrada-saída nos pesos do i -ésimo protótipo \mathbf{w}_i , $\forall i \in \mathbf{A}$, o que também refletirá o efeito de aumento observado em \mathbf{x} :

$$\mathbf{w}_i = [w_{i,1}^{in}, \dots, w_{i,m}^{in}, w_{i,1}^{out}, \dots, w_{i,n}^{out}]^T. \quad (3-2)$$

Propostos originalmente para controle (BARRETO; ARAUJO; RITTER, 2003), os mapeamentos de entrada-saída revelaram-se úteis noutros cenários, como previsão de séries temporais (BARRETO; ARAUJO, 2002; BARRETO; ARAUJO, 2001) e na identificação de sistemas dinâmicos (BARRETO;

ARAÚJO, 2004). Uma contribuição teórica desta tese é propor uma maneira de alavancar o alcance deste mapeamento para além da clássica relação entrada-saída de causa e efeito, extrapolando-o para mapear relações mais gerais, como as interdependências espaciais ou semânticas, em fontes de dados com modalidades diversas. Esse tipo de arquitetura passará a ser chamada de *rede SOM expandida*.

Considere-se um espaço de origem χ e um espaço de destino Υ contendo dados de duas modalidades diferentes, mas oriundos do mesmo contexto no mundo real. Portanto, $X \in \chi$ e $Y \in \Upsilon$ compreendem dois tipos de dados adquiridos sobre o mesmo objeto ou no mesmo local, partilhando informações semânticas apesar dos seus diferentes formatos de gravação e armazenamento. Por exemplo, uma mensagem de voz em áudio e sua transcrição; um exame médico de imagem e uma lista de medidas clínicas relacionadas com essa mesma área de investigação; ou, no contexto das geociências do petróleo, perfis de imagem de parede de poço e perfis regulares adquiridos na mesma porção desse poço.

Neste cenário, pode-se aplicar a rede SOM expandida, não no contexto causa-efeito da identificação e controle de sistemas dinâmicos, mas no contexto origem-destino da aprendizagem de representações multimodais cruzadas (YAO et al., 2023). Assim, passa-se a utilizar uma entrada aumentada para o treinamento da rede SOM expandida:

$$\mathbf{x}_{aug} = [x_1^s, \dots, x_m^s, x_1^t, \dots, x_n^t]^T, \quad (3-3)$$

m e n passam a indicar as dimensões dos dados origem (s , do inglês *source*) e destino (t , do inglês *target*), respectivamente.

Quando a rede SOM expandida é usada dessa maneira, ela se torna um estimador bidirecional entre as diferentes modalidades de dados. A Figura 3.1 mostra como a entrada aumentada \mathbf{x} força automaticamente a rede SOM a associar dois espaços de dados diferentes χ e Υ de conteúdo, modalidade e dimensões diferentes, simplesmente ao partilhar a localização do neurônio o vencedor e utilizar esta informação para recuperar $\mathbf{w}_{i^*}^t$ a partir de \mathbf{x}^s e, inversamente, $\mathbf{w}_{i^*}^s$ a partir de \mathbf{x}^t . Note-se que \mathbf{x}^t e \mathbf{w}^t não se encontram no espaço χ , mas em Υ .

A presente proposta pode ser resumida em sua forma mais geral pelas seguintes etapas, ilustradas na Figura 3.2:

- **Etapa 1 (aprendizado de representações):** aprender representações compactas significativas \mathbf{t} e \mathbf{z} dos espaços destino e origem (Υ e χ) a partir dos respectivos conjuntos de dados Y e X ;

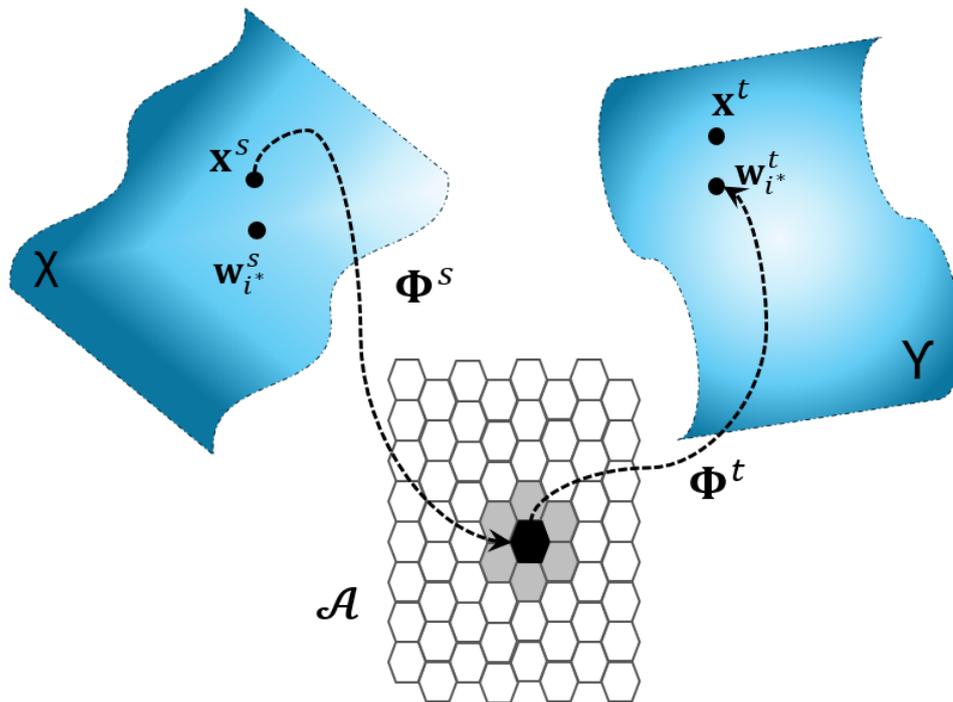


Figura 3.1: Ilustração conceitual do mapeamento heteroassociativo implementado pela rede SOM expandida.

- **Etapa 2 (mapeamento intermodalidades):** aprender um mapeamento intermodalidades Φ de \mathbf{z} e \mathbf{t} com a rede SOM expandida;
- **Etapa 3 (inferência cruzada):** gerar \mathbf{t} a partir de \mathbf{z} por meio do mapeamento Φ ou, sem nenhum passo adicional, gerar \mathbf{z} a partir de \mathbf{t} , pois o mapeamento Φ (Etapa 2) permite inferência bidirecional;
- **Etapa 4 (restauração ao espaço original):** restaurar \mathbf{t} (ou \mathbf{z}) para os espaços originais.

3.2

Detalhamento do Arcabouço Proposto

O cerne da metodologia proposta é a implementação do mapeamento heteroassociativo multimodalidade mencionado na Etapa 2 acima. Esse mapeamento é realizado pela rede SOM, com base na alteração conceitual descrita anteriormente segundo a qual é possível que a rede SOM aprenda um mapeamento não linear heteroassociativo entre dados de origem e dados de destino.

A metodologia, contudo, depende do aprendizado prévio de um conjunto de representações adequadas para os dados origem e destino (Etapa 1). Diversos modelos podem ser utilizados para essa tarefa, inclusive métodos lineares de redução de dimensionalidade, como PCA, ou não lineares, como os autocodificadores. Qual modelo utilizar, vai depender do problema a ser

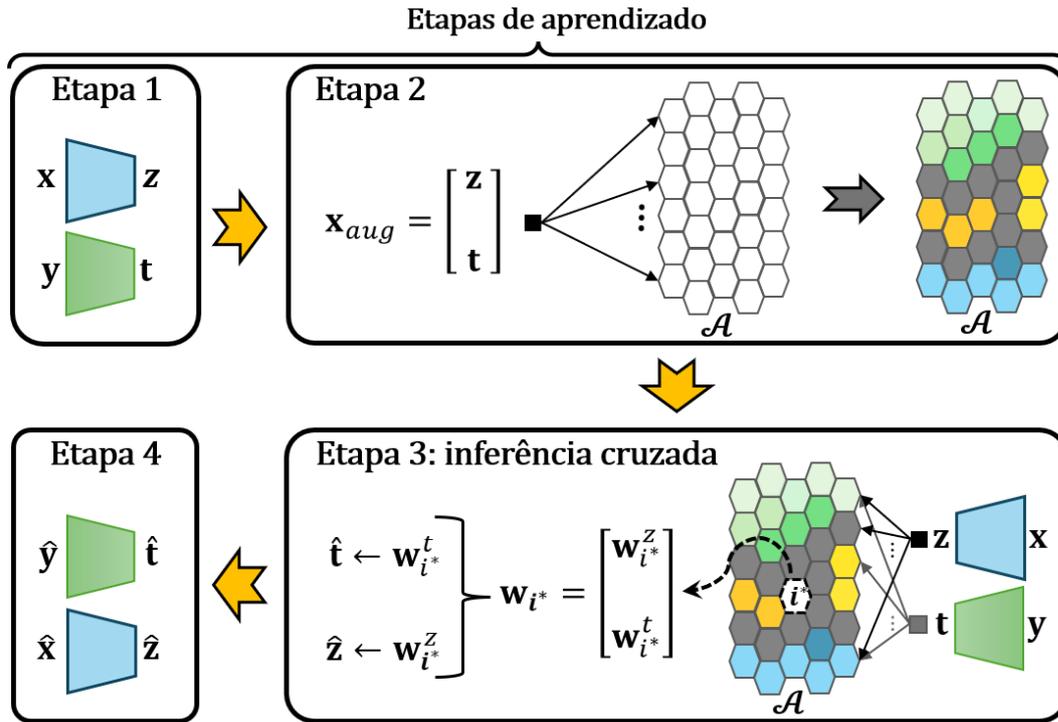


Figura 3.2: Ilustração da sequência de etapas gerais sequenciais que compõem a metodologia proposta.

resolvido e da sua necessidade em relação às etapas preditivas/regenerativas (Etapas 3 e 4). Na aplicação que será desenvolvida a seguir, por exemplo, escolheu-se um modelo β -CoVAE para aprender a representação de dados obtidos como imagens, enquanto não houve necessidade de transformação para o dado tipo série de valores (perfis básicos), que já tinham dimensão reduzida.

As Etapas 3 e 4 guardam, portanto, uma estreita relação com a Etapa 1, pois elas precisam ser capazes de retornar aos espaços originais dos dados origem e destino a partir das representações aprendidas.

A metodologia proposta, portanto, é flexível quanto às Etapas 1, 3 e 4. O mapeamento intermodalidades, contudo, depende da rede SOM. O uso da rede SOM nesse contexto é motivado pelos seguintes fatores:

- a memória heteroassociativa é construída pela simples concatenação de vetores, sem necessidade de alteração na complexidade do treinamento da rede;
- a rede SOM já se mostrou capaz de fazer mapeamentos similares mais simples, em contexto autoassociativo, com sucesso em diversas tarefas com variado grau de complexidade (e.g. Tabela 2.1);
- o mapeamento heteroassociativo é, por definição, bidirecional (e potencialmente multidirecional se há mais modalidades presentes em \mathbf{x}_{aug}) sem

- a necessidade de se retrainar a rede SOM expandida para cada sentido da geração cruzada de dados que se deseje realizar;
- a propriedade de preservação de relações de distância entre instâncias do espaço original no mapa de neurônios pode beneficiar o processo de mapeamento intermodalidades.

3.3

Tipos de aplicação

Ao tratar do mapeamento entre dois espaços de dados, origem e destino, de modalidades diferentes, e permitir o trânsito entre eles, por meio da geração ou regeneração de dados entre diferentes domínios, dois tipo de aplicação são possíveis:

- **Tipo-I:** Estimar dados de destino a partir de dados de origem (ou representações destes). Em especial, nesta tese, esse cenário envolve estimar dados mais simples a partir de dados mais complexos, que pode ser entendida como uma tarefa de *compreensão dos dados* (TAN et al., 2024);
- **Tipo-II:** Estimar dados de origem a partir de dados de destino (ou representações destes). Em especial, nesta tese, esse cenário envolve estimar dados complexos a partir de dados mais simples, que se define como uma tarefa de *aprendizado regenerativo* (TAN et al., 2024).

3.4

Rede SOM Expandida para Aprendizado Regenerativo Bidirecional

As quatro etapas gerais descritas acima serão aqui desdobradas em passos. Cada passo será descrito com maiores detalhes, levando em consideração os modelos específicos implementados em cada etapa, bem como as características das aplicações de interesse.

A partir de agora, utiliza-se a seguinte notação, levando em conta que se dispõe de p exemplos para treinamento dos modelos:

- $\mathbf{x} \in \mathbb{R}^l$, no espaço origem \mathcal{X} , de modo que $X = \{\mathbf{x}_1, \dots, \mathbf{x}_p\}$, $X \subset \mathcal{X}$;
- $\mathbf{z} \in \mathbb{R}^m$, a representação compacta do espaço origem, de modo que $Z = \{\mathbf{z}_1, \dots, \mathbf{z}_p\}$, $Z \subset \mathcal{Z}$;
- $\mathbf{y} \in \mathbb{R}^n$, no espaço destino \mathcal{Y} , de modo que $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_p\}$;
- $\mathbf{r} \in \mathbb{R}^v$, a representação compacta do espaço-destino, de modo que $T = \{\mathbf{r}_1, \dots, \mathbf{r}_p\}$, $T \subset \mathcal{T}$.
- p é o número de exemplos no conjunto de dados de treinamento.

- q é o número de exemplos no conjunto de dados de teste.
- D é o número total de exemplos no banco de dados ($D = p + q$).

Os passos podem ser resumidas da seguinte forma:

Passo 1: aprender uma representação compacta Z do espaço de origem χ a partir de X , usando um modelo β -CoVAE. Nas aplicações examinadas nesta tese, o espaço de destino já tem dimensão suficientemente baixa e dispensa essa etapa, ou seja, $\mathcal{T} = \Upsilon$.

Passo 2: aprender um mapeamento heteroassociativo entre Y e a representação do espaço de origem Z com a rede SOM expandida, montando o vetor de entrada da rede SOM como:

$$\mathbf{x}_{aug} = \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} = [z_1, \dots, z_m, y_1, \dots, y_n]^T. \quad (3-4)$$

Originalmente, a rede SOM é um modelo com aprendizado não supervisionado. Porém, o uso da entrada aumentada, associando representações de origem e dados de destino, faz com que a rede SOM expandida seja classificada como um modelo com aprendizado *autosupervisionado*. Após o treinamento, os pesos de qualquer unidade i da rede SOM expandida passam a refletir as mesmas componentes contidas em \mathbf{x}_{aug} :

$$\mathbf{w}_i = \begin{bmatrix} \mathbf{w}_i^z \\ \mathbf{w}_i^y \end{bmatrix} = [w_{i,1}^z, \dots, w_{i,m}^z, w_{i,1}^y, \dots, w_{i,n}^y]^T, \quad (3-5)$$

na qual as componentes \mathbf{w}_i^z e \mathbf{w}_i^y refletem as modalidades de dados presentes em \mathbf{x}_{aug} .

Passo 3: prever Y a partir de Z usando o mapeamento aprendido pela rede SOM expandida.

As previsões $\hat{\mathbf{y}}$ são dadas por uma função $G_y(\mathbf{w}_{i^*}^y)$ que, no presente caso, é a matriz identidade de ordem n , $\mathbf{I}_{n \times n}$: $\hat{\mathbf{y}} \leftarrow \mathbf{I}_{n \times n} \cdot \mathbf{w}_{i^*}^y$.

É necessário, portanto, selecionar as componentes apropriadas do vetor de pesos da unidade de melhor correspondência i^* , $\mathbf{w}_{i^*}^y$. Para tanto, define-se uma matriz de extração \mathbf{P}^y de dimensão $n \times (m + n)$:

$$\mathbf{P}^y = [\mathbf{0}_{n \times m} | \mathbf{I}_{n \times n}]. \quad (3-6)$$

na qual $\mathbf{0}_{n \times m}$ é a matriz nula de dimensão $n \times m$, e ‘|’ é um operador para concatenação de matrizes. Assim, $\mathbf{w}_{i^*}^y$ pode ser extraída de \mathbf{w}_{i^*} por meio de

$$\mathbf{w}_{i^*}^y_{[n \times 1]} = \mathbf{P}^y_{[n \times (m+n)]} \cdot \mathbf{w}_{i^*}_{[(m+n) \times 1]}. \quad (3-7)$$

Esse passo permite solucionar a aplicação do Tipo-I.

Passo 4: usando a rede SOM expandida, regerar X a partir de Y , via Z , restaurar o dado no espaço de origem usando o decodificador do modelo β -CoVAE treinado na Passo 1. Portanto, o mapeamento heteroassociativo (Passo 2) permite inferência bidirecional.

No presente caso, G_z é parametrizada como o decodificador do modelo β -CoVAE treinado no Passo 1, de modo que $\hat{\mathbf{x}} \leftarrow dec(\mathbf{w}_{i^*}^z)$.

Assim como no Passo 3, é necessário selecionar as componentes corretas do protótipo \mathbf{w}_{i^*} com melhor correspondência na rede SOM expandida. Define-se uma matriz de extração \mathbf{P}^z de dimensão $m \times (m + n)$:

$$\mathbf{P}^z = [\mathbf{I}_{m \times m} | \mathbf{0}_{m \times n}], \quad (3-8)$$

na qual $\mathbf{I}_{m \times m}$ é a matriz identidade de ordem m e $\mathbf{0}_{m \times n}$ é a matriz nula de dimensão $m \times n$. Assim, $\mathbf{w}_{i^*}^z$ pode ser selecionada por meio de

$$\mathbf{w}_{i^*}^z_{[m \times 1]} = \mathbf{P}^z_{[m \times (m+n)]} \cdot \mathbf{w}_{i^*}_{[(m+n) \times 1]}, \quad (3-9)$$

que permite obter as reconstruções $\hat{\mathbf{x}}$ regeradas por $G_z(\mathbf{w}_{i^*}^z)$.

Esse passo permite resolver a aplicação do Tipo-II.

Para uma melhor compreensão de todo o processo, uma representação esquemática dos quatro passos da metodologia proposta é mostrada na Figura 3.3.

Os Passos 1 e 2 são etapas de aprendizado, para treinar os modelos utilizados na metodologia proposta¹, enquanto os Passos 3 e 4 são etapas de inferência (predição ou regeneração). Isso define a metodologia proposta, que passará a ser chamada de *Estimador SOM Regenerativo Bidirecional* (BERGSOM, do inglês *Bidirectional Expanded ReGenerative SOM*).

Além do uso de mapeamento heteroassociativo de origem para destino, outra inovação desta tese é o uso de componentes da distribuição latente aprendidas pelo modelo β -CoVAE (usado no Passo 1). Essas componentes passam a compor a entrada \mathbf{x}_{aug} que alimenta o mapeamento heteroassociativo, em vez de usar diretamente as representações aprendidas \mathbf{z} como no caso

¹Ambos os modelos, β -CoVAE e SOM, são treinados separadamente. Na variante BERGSOM, as representações \mathbf{z} são amostradas no início e permanecem fixas ao longo do treinamento da rede SOM. O mesmo acontece na variante p -BERGSOM para a Aplicação Tipo-I. Amostrarmos o valor das representações \mathbf{z} a cada iteração de treinamento é uma alternativa que não foi explorada no presente trabalho, mas pode ser objeto de investigações futuras. Os experimentos foram executados em Python 3.10.9, com o auxílio de várias bibliotecas de código aberto, incluindo Scikit-Learn, Keras, Tensorflow, Pandas e Numpy. Para implementar o SOM, utilizou-se o pacote IntraSOM (de Gouvêa et al., 2023).

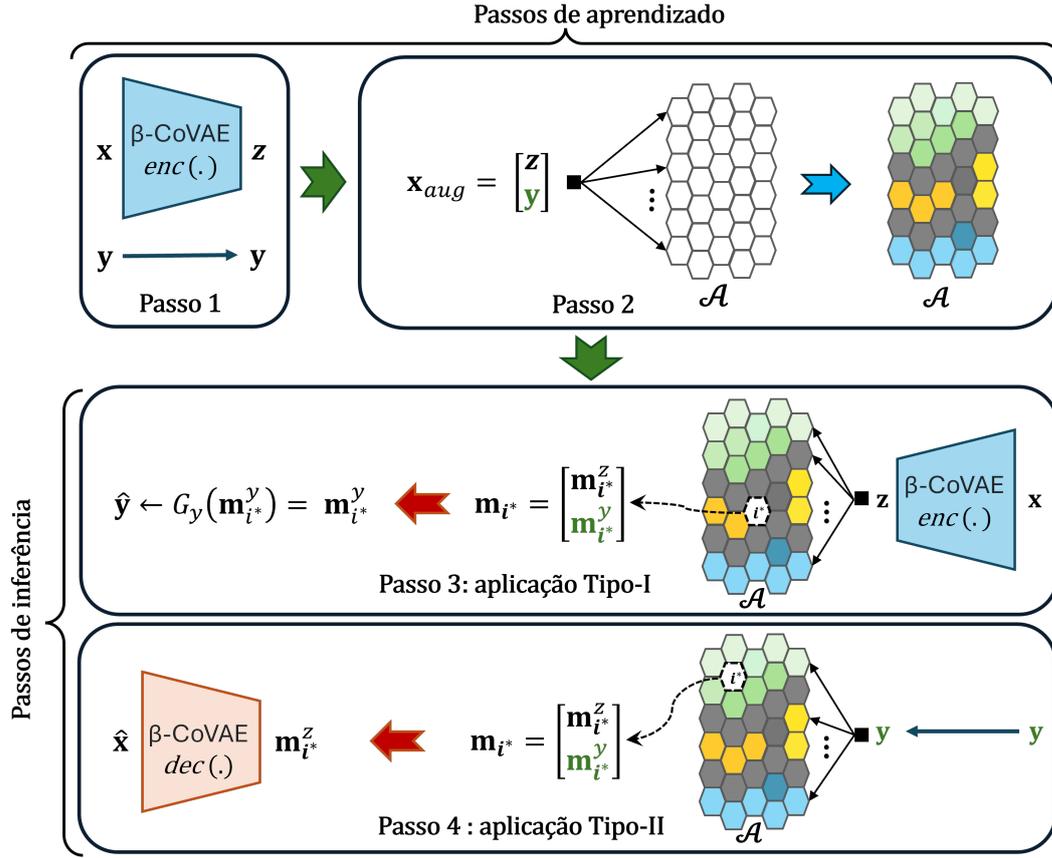


Figura 3.3: Uma ilustração dos 4 passos da metodologia proposta (exemplo da variante BERGSOM).

anterior. Assim, define-se uma outra composição de entrada possível para treino da rede SOM expandida:

$$\mathbf{x}_{aug} = \begin{bmatrix} \mathbf{z}(\mu, \sigma) \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mu^z \\ \sigma^z \\ \mathbf{y} \end{bmatrix} = [\mu_1^z, \dots, \mu_m^z, \sigma_1^z, \dots, \sigma_m^z, y_1, \dots, y_n]^T, \quad (3-10)$$

na qual os componentes de $\mathbf{z}(\mu, \sigma) \in \mathbb{R}^{2m}$, μ^z e σ^z , são fornecidos pelo codificador do modelo β -CoVAE usado no Passo 1. O vetor protótipo de uma unidade i na variante p -BERGSOM passa a mapear três tipos de informação: $\mathbf{w}_i = [\mathbf{w}_i^{\mu^z}, \mathbf{w}_i^{\sigma^z}, \mathbf{w}_i^y]^T$.

Portanto, essa variante da metodologia proposta será chamada de *Estimador SOM Regenerativo Bidirecional Probabilístico* (p -BERGSOM, do inglês *Probabilistic Bidirectional Expanded ReGenerative SOM*), pois a entrada aumentada incorpora características de distribuição latente aprendidas pelo modelo β -CoVAE.

O uso das componentes latentes probabilísticas permite que se façam múltiplas previsões. Isso é possível porque podem-se gerar N amostras no espaço latente para cada entrada \mathbf{x} usando as componentes latentes probabilísticas.

Ambas as variantes da metodologia proposta, seja BERGSOM ou p -BERGSOM, permitem inferência bidirecional cruzada entre as modalidades mapeadas.

Para a aplicação do Tipo-I, treina-se uma rede SOM expandida utilizando a entrada aumentada, definida pela Equação 3-4 para a variante BERGSOM ou pela Equação 3-10 para a p -BERGSOM. Nessa aplicação, o efeito probabilístico do p -BERGSOM se faz notar apenas na etapa de inferência. A etapa de inferência para BERGSOM e p -BERGSOM é detalhada a seguir:

- Obter a representação $\mathbf{z}(\mathbf{x})$ para um dado de origem \mathbf{x} por meio do modelo β -CoVAE já treinado, que é apresentada como entrada para a rede SOM expandida já treinada. Na etapa de inferência apenas a representação do dado de origem é apresentada à rede SOM expandida.
- **BERGSOM:** a representação do dado de origem é $\mathbf{z}(\mathbf{x}) \leftarrow \mathbf{z}$.
- **p -BERGSOM:** a representação do dado de origem $\hat{\mathbf{z}}$ é amostrada por meio das componentes probabilísticas de $\mathbf{z}(\mu, \sigma)$ usando o truque da reparametrização:

$$\mathbf{z}(\mathbf{x}) \leftarrow \hat{\mathbf{z}} = \boldsymbol{\mu}_z + \boldsymbol{\sigma}_z \odot \boldsymbol{\varepsilon}, \quad (3-11)$$

na qual $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}_{m \times 1}, \mathbf{I}_{m \times m})$ e \odot representa a multiplicação elemento a elemento (produto de Hadamard).

- Encontrar o neurônio vencedor i^* da rede SOM expandida usando para tal apenas a componente \mathbf{w}_i^z . Uma maneira de implementar essa seleção usando somente $\mathbf{z}(\mathbf{x})$ e \mathbf{w}_i^z é definir uma alteração da métrica de distância:

$$\text{dist}(\mathbf{z}(\mathbf{x}), \mathbf{w}_i) = (\mathbf{z}(\mathbf{x})_{aug} - \mathbf{w}_i)^T \mathbf{S}^z (\mathbf{z}(\mathbf{x})_{aug} - \mathbf{w}_i), \quad (3-12)$$

na qual \mathbf{S}^z é uma matriz de seleção de dimensão $(m + n) \times (m + n)$ definida por:

$$\mathbf{S}^z = \begin{bmatrix} \mathbf{I}_{m \times m} & \mathbf{0}_{m \times n} \\ \mathbf{0}_{n \times m} & \mathbf{0}_n \end{bmatrix}, \quad (3-13)$$

e $\mathbf{z}(\mathbf{x})_{aug}$ é definido de modo a ter coerência dimensional com \mathbf{w}_i :

$$\mathbf{z}(\mathbf{x})_{aug} = \begin{bmatrix} \mathbf{z}(\mathbf{x}) \\ \mathbf{0}_{n \times 1} \end{bmatrix}. \quad (3-14)$$

Portanto, o neurônio vencedor i^* é selecionado por meio de

$$i^* = \arg \min_{i \in \mathcal{A}} \{ \text{dist}(\mathbf{z}(\mathbf{x}), \mathbf{w}_i) \}. \quad (3-15)$$

- A partir do protótipo do vencedor \mathbf{w}_{i^*} , recuperar a componente destino correspondente \mathbf{w}_i^y utilizando a matriz de extração \mathbf{P}^y (Equações 3-6 e 3-7);
- Gerar a previsão: $\hat{\mathbf{y}} \leftarrow G_y(\mathbf{w}_{i^*}^y) = \mathbf{w}_{i^*}^y = \mathbf{P}^y \cdot \mathbf{w}_{i^*}$.

Para a aplicação do Tipo-II, treina-se a variante BERGSOM usando entradas expandidas definidas pela Equação 3-4 e a variante p -BERGSOM com entradas definidas pela Equação 3-10. Portanto, ao contrário da aplicação do Tipo-I, são treinadas duas redes SOM expandidas diferentes, uma para cada variante da metodologia. A inferência é então revertida, pois o objetivo é regerar dados complexos (de origem) a partir de dados mais simples (de destino), porém segue os mesmos passos para ambas as variantes:

- Apresentar apenas o vetor destino \mathbf{y} como entrada para a rede SOM expandida previamente treinada e encontrar o neurônio vencedor i^* usando somente a componente destino \mathbf{w}_i^y . Uma maneira de implementar essa seleção usando somente \mathbf{y} e \mathbf{w}_i^y é definir uma alteração da métrica de distância:

$$\text{dist}(\mathbf{y}, \mathbf{w}_i) = (\mathbf{y}_{aug} - \mathbf{w}_i)^T \mathbf{S}^y (\mathbf{y}_{aug} - \mathbf{w}_i), \quad (3-16)$$

na qual \mathbf{S}^y é uma matriz de seleção de dimensão $(m+n) \times (m+n)$ definida por:

$$\mathbf{S}^y = \begin{bmatrix} \mathbf{0}_{m \times m} & \mathbf{0}_{m \times n} \\ \mathbf{0}_{n \times m} & \mathbf{I}_{n \times n} \end{bmatrix}, \quad (3-17)$$

e \mathbf{y}_{aug} é definido de modo a ter coerência dimensional com \mathbf{w}_i :

$$\mathbf{y}_{aug} = \begin{bmatrix} \mathbf{0}_{m \times 1} \\ \mathbf{y} \end{bmatrix}. \quad (3-18)$$

Portanto, o neurônio vencedor i^* é selecionado por meio de

$$i^* = \arg \min_{i \in \mathcal{A}} \{\text{dist}(\mathbf{y}, \mathbf{w}_i)\}. \quad (3-19)$$

- A partir de \mathbf{w}_{i^*} , recuperar a componente correspondente da representação do dado de origem:
 - **BERGSOM:** extrair a componente relativa ao dado origem do protótipo vencedor usando a matriz de extração \mathbf{P}^z , definida pela Equação 3-8, usando a Equação 3-9: $\mathbf{w}_{i^*}^z = \mathbf{P}^z \cdot \mathbf{w}_{i^*}$;
 - p -**BERGSOM:** extrair as componentes latentes probabilísticas $\mathbf{w}_{i^*}^{\mu z}$ e $\mathbf{w}_{i^*}^{\sigma z}$ do protótipo vencedor, usando as matrizes de extração apropriadas \mathbf{P}_μ^z e \mathbf{P}_σ^z definidas a seguir

$$\mathbf{P}_\mu^z = [\mathbf{I}_{m \times m} \mid \mathbf{0}_{m \times m} \mid \mathbf{0}_{m \times n}], \quad (3-20)$$

e

$$\mathbf{P}_\sigma^z = [\mathbf{0}_{m \times m} \mid \mathbf{I}_{m \times m} \mid \mathbf{0}_{m \times n}]. \quad (3-21)$$

Portanto, as respectivas componentes extraídas são

$$\mathbf{w}_{i^*}^{\mu_z}_{[m \times 1]} = \mathbf{P}_\mu^Z_{[m \times (2m+n)]} \cdot \mathbf{w}_{i^*}^{[(2m+n) \times 1]} \quad (3-22)$$

e

$$\mathbf{w}_{i^*}^{\sigma_z}_{[m \times 1]} = \mathbf{P}_\sigma^Z_{[m \times (2m+n)]} \cdot \mathbf{w}_{i^*}^{[(2m+n) \times 1]} \quad (3-23)$$

– Obter uma representação do dado de origem a partir das componentes acima:

- **BERGSOM:** $\mathbf{z}(\mathbf{y}) \leftarrow \mathbf{w}_{i^*}^z$;
- **p -BERGSOM:** amostrar uma representação $\hat{\mathbf{z}}$ a partir das componentes probabilísticas usando o truque da reparametrização²:

$$\mathbf{z}(\mathbf{y}) \leftarrow \hat{\mathbf{z}} = \mathbf{w}_{i^*}^{\mu_z} + \mathbf{w}_{i^*}^{\sigma_z} \odot \boldsymbol{\varepsilon}, \quad (3-24)$$

na qual $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}_{m \times 1}, \mathbf{I}_{m \times m})$ e \odot representa o produto de Hadamard.

– Reegerar uma previsão correspondente dos dados de origem:

$$\hat{\mathbf{x}} \leftarrow G_z(\mathbf{z}(\mathbf{y})) = \text{dec}(\mathbf{z}(\mathbf{y})), \quad (3-25)$$

na qual $\text{dec}(\cdot)$ denota o decodificador do β -CoVAE (Passo 1).

Para ilustrar o funcionamento da metodologia proposta, ela é aplicada a seguir em um problema simples bastante conhecido usando a base de dados MNIST (LECUN; CORTES; BURGES, 2010).

3.5

Prova de Conceito: Dados MNIST

No presente experimento de prova de conceito, será usada a base de dados MNIST³ em tarefas de geração cruzada bidirecional de dados multimodalidades. Primeiro será previsto o dígito a partir de uma representação de sua imagem (aplicação Tipo-I), depois a imagem do dígito será regerada a partir da informação sobre ele (aplicação Tipo-II).

As imagens serão transformadas usando um modelo β -CoVAE como o que foi descrito anteriormente e que será usado depois na aplicação principal desta tese⁴. Os dígitos serão representados por um espaço latente com oito di-

²Ambas as operações de amostragem expressas nas Equações 3-11 e 3-24 são realizadas pela função de reparametrização embutida do modelo β -CoVAE utilizado no Passo 1 da metodologia proposta.

³Os dados foram obtidos via biblioteca Tensorflow para Python.

⁴A implementação desse modelo será detalhada no Capítulo 4.

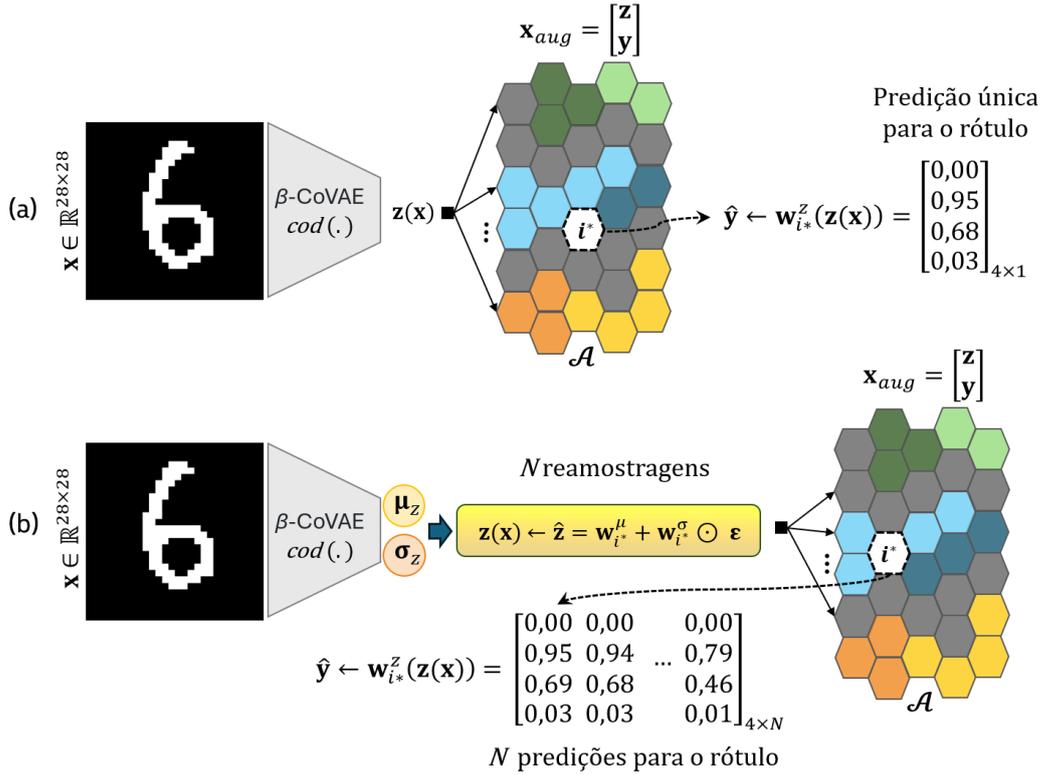


Figura 3.4: Ilustração das variantes BERGSOM (a) e p -BERGSOM (b) para aplicação Tipo-I com dados do conjunto MNIST.

mensões⁵, ou seja, o espaço latente do modelo $\beta\text{-CoVAE}$ terá oito componentes, $\dim(\mathcal{Z}) = 8$.

Nesse experimento, \mathbf{x} é uma das imagens de um dígito, \mathbf{z} é sua representação no espaço latente e \mathbf{y} é o rótulo codificado desse dígito. Os rótulos dos dígitos serão identificados por meio de uma codificação em base binária com quatro componentes. Ou seja: *Dígito 0* = $[0, 0, 0, 0]^T$, *Dígito 1* = $[0, 0, 0, 1]^T$, *Dígito 2* = $[0, 0, 1, 0]^T$, *Dígito 3* = $[0, 0, 1, 1]^T$ e assim por diante até *Dígito 9* = $[1, 0, 0, 1]^T$. Portanto, de acordo com a notação introduzida neste capítulo, temos $m = 8$, $n = 4$ e $p = 60000$ é o número de elementos no conjunto de treinamento.

3.5.1 MNIST: Aplicação Tipo-I

No contexto dos dados MNIST, a aplicação Tipo-I consiste em prever o rótulo do dígito (dado mais simples) a partir da imagem (dado mais complexo), ou seja, neste caso é uma aplicação de classificação. A Figura 3.4 ilustra essa aplicação para ambas as variantes da metodologia proposta.

Na variante BERGSOM, seguindo os passos descritos anteriormente, o

⁵A escolha do valor de $\dim(\mathcal{Z}) = 8$ foi arbitrária.

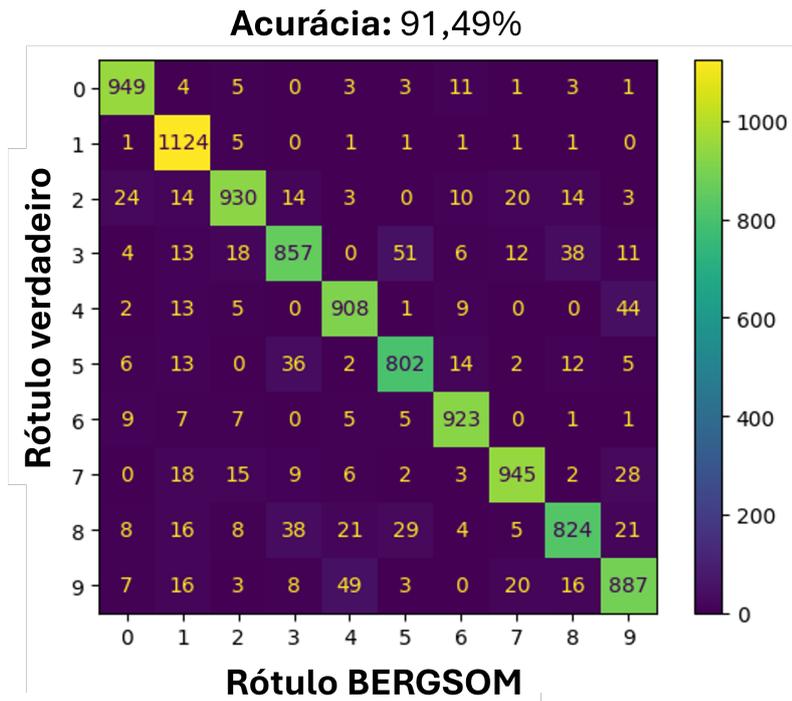


Figura 3.5: Resultados da variante BERGSOM para aplicação Tipo-I com dados do conjunto MNIST.

modelo foi capaz de prever o dígito a partir da imagem com elevado grau de acurácia, conforme mostrado na Figura 3.5.

A aplicação Tipo-I para os dados MNIST é do tipo “um para um”, i.e., cada imagem de entrada deve sempre ser associada a um único rótulo. Porém, devido à variabilidade na grafia das imagens na base MNIST, há chances de um dado dígito ser identificado de forma errônea. Nesse tipo de problema, a variante p -BERGSOM revela sua principal razão de ser, que é permitir herdar a variabilidade dos dados da imagem de entrada e incorporá-la às predições de saída. Os resultados são apresentados na Figura 3.6 para ilustrar o funcionamento da metodologia. O rótulo definido pela variante p -BERGSOM é aquele mais frequente dentre os N rótulos previstos por essa variante da metodologia.

Não se deve ressaltar o pequeno ganho de acurácia obtido pela variante p -BERGSOM, mas sim o fato de esse rótulo ser resultado da moda sobre um grande número de predições ($N = 100$), enquanto a variante BERGSOM só tem uma chance de acertar. Adicionalmente, esse resultado ilustra outra característica importante da variante p -BERGSOM, que é ser capaz de fornecer uma avaliação da qualidade das predições em termos de medidas de incerteza a elas associadas.

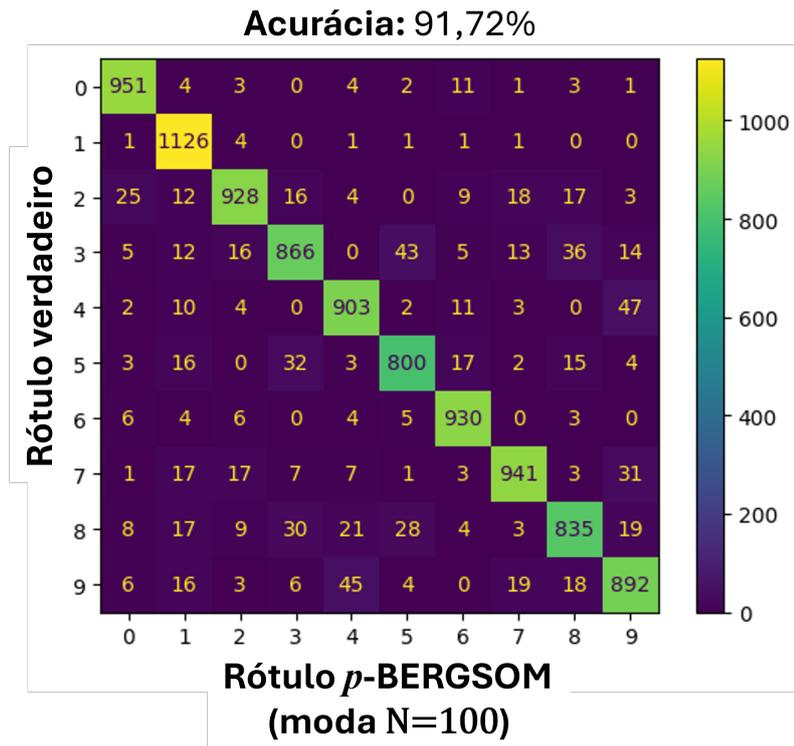


Figura 3.6: Resultados da variante p -BERGSOM para aplicação Tipo-I com dados do conjunto MNIST.

3.5.2

MNIST: Aplicação Tipo-II

No contexto dos dados MNIST, a aplicação Tipo-II consiste em gerar imagens dos dígitos (informação mais complexa) a partir de seu rótulo (dado mais simples), ou seja, neste caso é uma aplicação de aprendizado regenerativo. A Figura 3.7 ilustra essa aplicação para ambas as variantes da metodologia proposta.

Destaca-se na Figura 3.7 o fato de a rede SOM treinada para a aplicação Tipo-II ser diferente em cada uma das variantes. A variante BERGSOM (a) treina a rede SOM expandida usando a Equação 3-4 enquanto a variante p -BERGSOM (b) usa a Equação 3-10. Para maior clareza dessa importante diferença, as equações são incluídas no topo da figura, e as redes SOM treinadas são mostradas em cores diferentes.

Uma demonstração dos resultados para a variante BERGSOM é mostrada na Figura 3.8 enquanto uma demonstração da variante p -BERGSOM é mostrada na Figura 3.9.

A Figura 3.8 mostra um teste com a variante BERGSOM que foi solicitada a prever 16 imagens a partir de 16 rótulos escolhidos aleatoriamente dentre os exemplos disponíveis no conjunto de teste. Para comparação, na

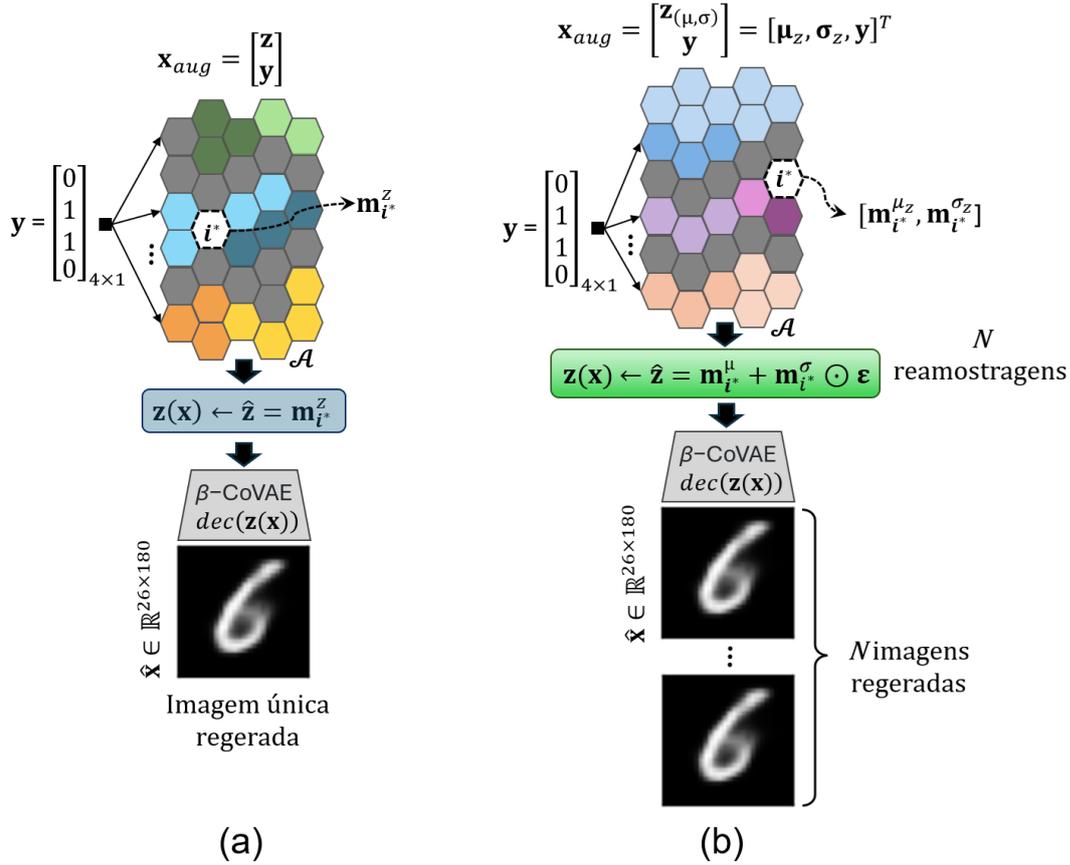


Figura 3.7: Esquemático das variantes BERGSOM (a) e p -BERGSOM (b) para aplicação Tipo-II com dados do conjunto MNIST.

linha central, mostram-se também as imagens recuperadas pelo mapeamento BERGSOM quando se utiliza a própria imagem real para a escolha do vencedor (como se faria na Aplicação Tipo-I), indicada como “IMG->IMG”. Na linha inferior, as respectivas imagens reais são mostradas. Vale ressaltar que, apesar do acerto em todos os dígitos da amostra acima, houve também casos de regeneração defeituosa. Na amostra da 3.8, chama a atenção o fato de que o uso do rótulo para prever a imagem (linha superior) errar menos que o uso da própria imagem como entrada (linha do meio, que errou um dos dígitos, destacado, quando a saída regenerada deveria ser um 9, mas foi um regenerado um 5).

A aplicação Tipo-II para os dados MNIST é do tipo “um para muitos”, i.e., dado um rótulo, por exemplo, o dígito 6, diversas imagens podem ser geradas. Nesse caso, a variante p -BERGSOM mostra seu potencial, por usar a variabilidade aprendida a partir dos dados para gerar diferentes imagens de saída.

Na Figura 3.9 são apresentadas as $N(= 7)$ imagens de dígitos regeneradas a partir da informação de seus rótulos. Vale destacar que, nos testes com a

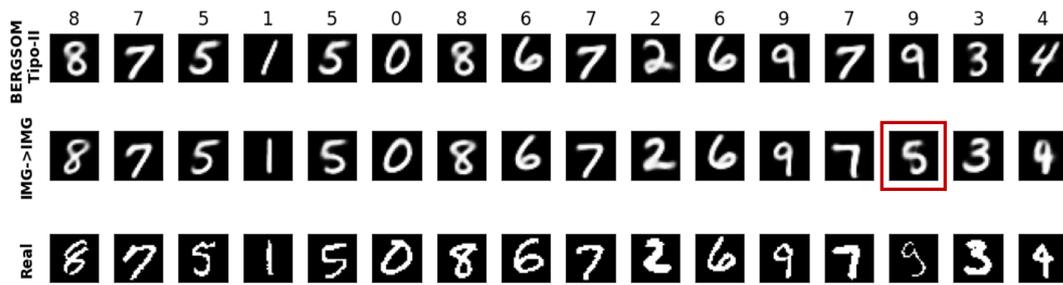


Figura 3.8: Demonstração da variante BERGSOM para aplicação Tipo-II com dados do conjunto MNIST. O termo “IMG->IMG” indica que as imagens são recuperadas pelo mapeamento BERGSOM alimentado com a representação da própria imagem real para a escolha do vencedor (similar ao que se faz na Aplicação Tipo-I).

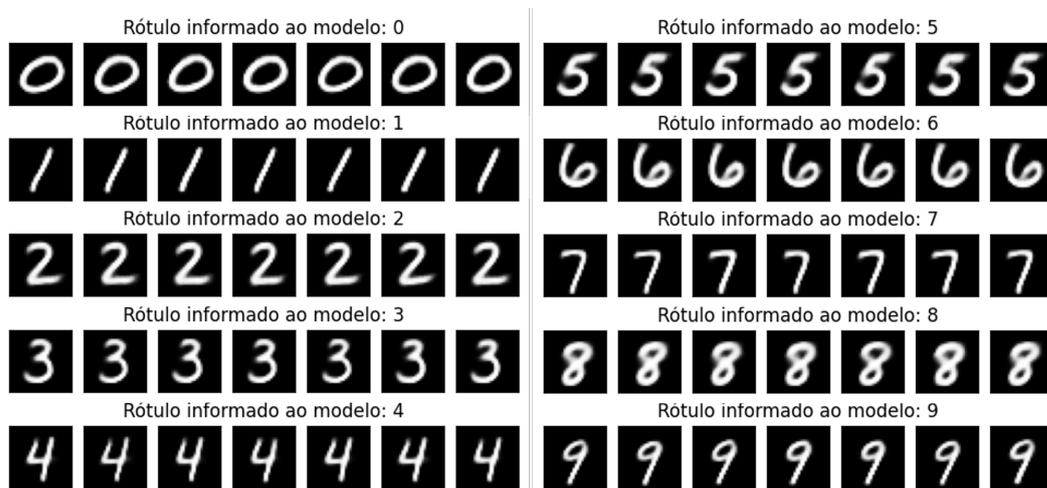


Figura 3.9: Demonstração da variante p -BERGSOM para aplicação Tipo-II com dados do conjunto MNIST. Teste com ($N = 7$) imagens a regerar a partir da informação do rótulo desejado.

variante p -BERGSOM, não houve erro na regeneração das imagens, ou seja, o dígito regerado sempre foi coerente com o rótulo informado ao modelo.

3.6

Considerações finais

Este capítulo introduziu e formalizou uma nova metodologia que utiliza a rede SOM como um estimador baseado em memória associativa. Essa abordagem expandiu a aplicação da técnica para a associação de dados numéricos de diferentes modalidades, permitindo o aprendizado de um mapeamento não linear entre dados de origem e destino, mesmo quando não há uma relação de causa e efeito entre eles. A metodologia foi apresentada em etapas gerais, sem vinculação a um modelo específico, para ilustrar sua flexibilidade.

Em seguida, as variantes da metodologia proposta foram detalhadas e contextualizadas em relação às aplicações desenvolvidas nesta tese. Na

sequência, o capítulo se encerra com um experimento que aplica a metodologia no conjunto de dados MNIST, sem a intenção de avaliar o desempenho, mas como prova de conceito, para proporcionar uma melhor compreensão da proposta.

Nos Capítulos 4 e 5, detalham-se experimentos realizados para avaliação da metodologia proposta na solução de problemas reais no âmbito da petrofísica de reservatórios.

4

Aplicação Tipo-I em Petrofísica

Neste capítulo¹, descreve-se o primeiro experimento para avaliar a metodologia proposta, para ambas as variantes, BERGSOM e *p*-BERGSOM, em um problema da área de petrofísica de reservatórios:

Tipo-I Geração de dados mais simples a partir de dados mais complexos: predição de um conjunto completo de perfis básicos a partir de perfis de imagens acústicas.

A predição de um conjunto de perfis básicos a partir de imagens acústicas era, até o momento, uma aplicação inédita para a área de petrofísica de reservatórios. Os motivos para isso são econômicos: perfis básicos são menos custosos de se adquirir e é incomum que seja necessária sua estimação a partir de perfis mais custosos como os perfis de imagem. Porém, há de se ter em conta o caráter da presente tese, cujo fim é a estimação bidirecional entre representações de dados de diferentes modalidades, que vai além de um ramo específico e justifica plenamente esta primeira aplicação.

Além do aspecto específico desse experimento no campo da petrofísica, as técnicas aqui propostas podem ser extrapoladas para uma vasta gama de aplicações que explorem relações entre diferentes modalidades de dados em outras áreas do conhecimento, em especial aplicações de estimação de valores numéricos a partir de imagens, ou regressão multivariada a partir de imagens, que seguem como problemas desafiadores na área de aprendizado de máquinas.

4.1

Dados e Notação

Os experimentos deste capítulo e do seguinte usam um conjunto de dados de perfis básicos e de perfis de imagem relativos a cinco poços, nomeados de A a E, que foram perfurados em carbonatos altamente heterogêneos de um reservatório da camada pré-sal da formação Barra Velha (JÁCOMO et al., 2023). Os poços estão nas proximidades de duas acumulações principais de hidrocarbonetos dessa formação e abrangem uma área ampla de aproximadamente 80 km^2 . A disposição relativa dos poços no campo é ilustrada na Figura 4.1.

As linhas de contorno apresentadas na Figura 4.1 indicam as localizações onde ocorrem os maiores volumes de porosidade. A Figura 4.1 também elucidada

¹Este capítulo serviu de base para a preparação dos artigos 1, 3, 5 (em parte) e 6 citados no Apêndice B.

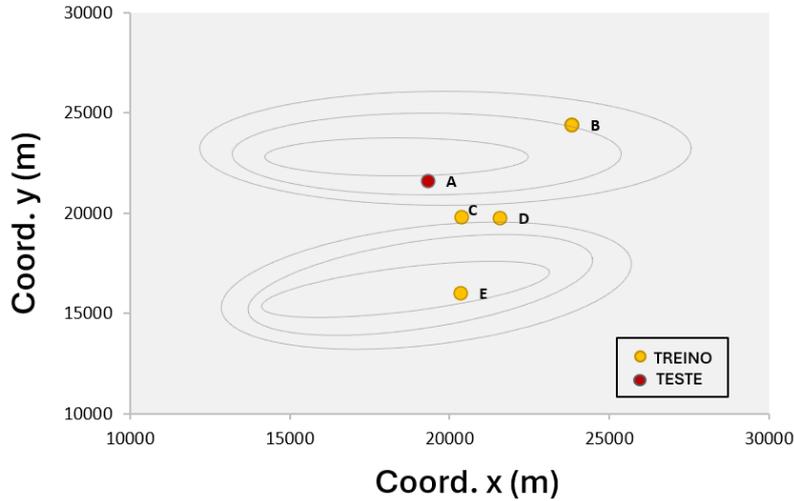


Figura 4.1: Disposição relativa dos poços na base de dados utilizada. Apresenta-se um dos conjuntos de validação, com 4 poços para treino/validação (B até D) e o poço restante (A) para teste. Por exemplo, os resultados apresentados para os experimentos no poço A são obtidos com modelos treinados com dados dos poços B até D.

Tabela 4.1: Conjuntos de validação para avaliação do modelo.

Conjunto	Poços treino/validação	Poço teste
1	B, C, D, E	A
2	A, C, D, E	B
3	A, B, D, E	C
4	A, B, C, E	D
5	A, B, C, D	E

a estratégia de validação cruzada empregada nos experimentos, que sempre reserva um poço para teste. Os cinco poços foram divididos em conjuntos de treinamento/validação e testes (Tabela 4.1). Cada conjunto de treino/validação é composto por quatro poços, e o quinto poço restante é usado para o teste cego. Este procedimento de validação cruzada é mais realista e preciso do que alguns encontrados em outros trabalhos, nos quais o conjunto de testes é composto por intervalos não vistos dos próprios poços de treinamento.

Os dados de perfis básicos, $Y|\mathbf{y} \in \mathbb{R}^n$, já apresentam dimensão suficientemente pequena $n = 7$, então optou-se por não implementar a Etapa 1 da metodologia proposta, que corresponde ao aprendizado de representações, para estes dados.

No entanto, para o conjunto de imagens acústicas², $X|\mathbf{x} \in \mathbb{R}^{26 \times 180}$, é

²Favor verificar no Apêndice A o motivo para esse tamanho de imagem, e como se fez o pareamento entre uma janela da imagem e uma leitura de perfis básicos neste trabalho.

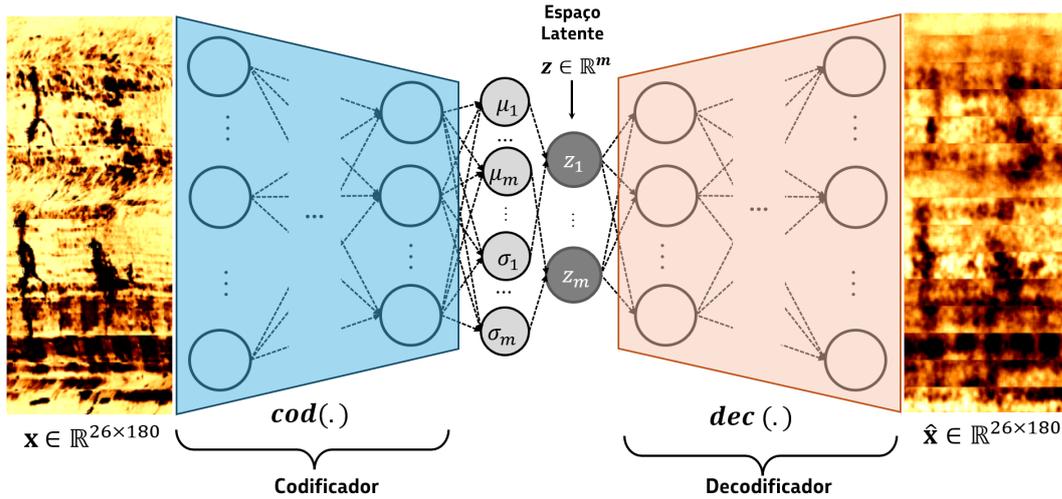


Figura 4.2: Esquemático do modelo β -CoVAE utilizado na Etapa 1 para os dados de perfis de imagem (dados origem).

necessário realizar o aprendizado de representações. Utilizou-se um modelo β -CoVAE personalizado para aprender uma representação compacta ($\mathbf{z} \in \mathcal{Z}$) do espaço origem χ a partir dos dados em X , conforme ilustra a Figura 4.2. A escolha de um autocodificador variacional foi também motivada pelo objetivo geral do trabalho, que é construir uma metodologia de geração bidirecional entre dados de diferentes modalidades. Convém recordar que o modelo β -CoVAE, além de construir um espaço de representações das imagens, provê um modelo generativo para perfis de imagem na aplicação Tipo-II.

A implementação utilizada para o modelo β -CoVAE consiste em três camadas convolucionais tanto no codificador (respectivamente 64, 32 e 16 filtros 3×3) quanto no decodificador (respectivamente 16, 32 e 64 filtros 3×3), conectadas por duas camadas densas. A camada mais interna contém as representações latentes. A camada final do modelo, também densa, é responsável pela reconstrução. As camadas convolucionais têm uma função de ativação `selu` e pesos inicializados com normalização He (HE et al., 2015). Não se utilizam camadas de *dropout* nem de *pooling*. Todas as camadas densas são lineares.

Após a conclusão do treinamento do β -CoVAE, cada imagem \mathbf{x} com 26×180 pixels é apresentada ao codificador β -CoVAE, $cod(\cdot)$, para gerar sua respectiva representação. Portanto, à cada imagem corresponderá não apenas um vetor de perfis básicos \mathbf{y} , mas também um vetor \mathbf{z} com a representação obtida via codificador, além do vetor de componentes probabilísticas $\mathbf{z}_{(\mu, \sigma)}$, conforme arquitetura de camadas mostrada na Figura 4.2.

Assim, em termos da notação que vem sendo utilizada, mantêm-se as seguintes convenções para os experimentos deste capítulo e do seguinte:

- $\mathbf{x} \in \mathbb{R}^{26 \times 180}$ refere-se a uma imagem acústica de parede de poço (dado origem);
- $\mathbf{z} \in \mathbb{R}^m$ é uma representação de \mathbf{x} no espaço latente (o valor de $m = \dim(\mathcal{Z})$ será objeto de avaliação, pois tem impacto no desempenho da metodologia proposta);
- $\mathbf{z}_{(\mu, \sigma)} \in \mathbb{R}^{2m}$ é uma representação de \mathbf{x} em componentes probabilísticas;
- o espaço destino (que contém os perfis básicos) e o espaço de representações do dado de destino serão iguais, i.e., segundo a notação definida no Capítulo 3, $\Upsilon = \mathcal{T}$;
- $\mathbf{y} \in \mathbb{R}^n$ refere-se a um vetor contendo dados de perfis básicos (dado destino), com $n = 7$;
- p é a quantidade de elementos no conjunto de treinamento e q é a quantidade de elementos no conjunto de teste.

O que se tratou na presente seção será válido também para o Capítulo 5. Por questão de clareza na exibição dos experimentos realizados neste capítulo e no seguinte, maiores detalhes sobre os dados de poço que foram utilizados estão no Apêndice A.

4.2

Modelo de Aprendizado Profundo para Comparação

Como a aplicação de previsão de um conjunto completo de perfis básicos a partir de imagens acústicas de poços (doravante referida também como *predição 7C*) é inteiramente nova, é necessário desenvolver um modelo alternativo que permita extrair conclusões válidas sobre o desempenho da metodologia proposta. Para este propósito, desenvolveu-se uma arquitetura de rede neural convolucional profunda baseado na rede SimpleNet (HASANPOUR et al., 2016), que é um modelo bem estabelecido de rede neural convolucional profunda, capaz de fazer boas previsões com uma arquitetura computacionalmente leve. Esta rede foi originalmente proposta para classificação de objetos em imagens convencionais, portanto deve-se fazer adaptações para seu uso na tarefa de regressão multivariada.

SimpleNet é uma rede neural profunda com 13 camadas convolucionais cujos filtros têm tamanhos variados (principalmente 3×3 , exceto nas 11^a e 12^a camadas, cujos filtros são 1×1). Camadas de *max pooling* usam filtros de dimensão 2×2 . A rede SimpleNet também usa normalização em lotes antes de cada função de ativação ReLU e aplica *dropout* para evitar *overfitting*. Do modelo original utilizado para classificação de imagens, a única modificação foi a substituição da última camada de classificação (de ativação **softmax** por

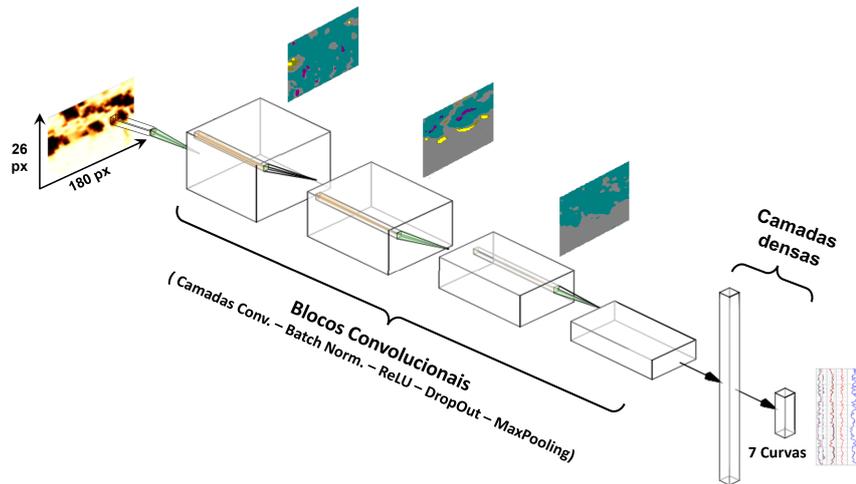


Figura 4.3: Ilustração do modelo SimpleNet personalizado para regressão 7C.

uma camada linear densa para realizar a tarefa de regressão multivariada. O modelo é ilustrado na Figura 4.3.

Apesar de ser um modelo aqui construído para fins de comparação e validação de resultados, um esforço considerável foi despendido para otimizar os hiperparâmetros da rede SimpleNet, com o intuito de prever o conjunto de sete perfis básicos básicos a partir perfis de imagem acústica. O próprio desenvolvimento da SimpleNet para este fim é, em si, também uma inovação deste trabalho no campo da petrofísica e no campo da regressão multivariada a partir de imagens. Porém, explorar mais profundamente o potencial de modelos CNN supervisionados regulares para a presente aplicação Tipo-I está fora do escopo deste trabalho, cujo foco é em desenvolver modelos de geração cruzada bidirecional de dados. Sugerem-se métodos de aprendizado de máquinas automático (AutoML), como Q-NAS (SZWARCMAN; CIVITARESE; VELLASCO, 2022), para investigar arquiteturas convolucionais profundas que alcancem um desempenho ainda melhor na aplicação Tipo-I usando aprendizado supervisionado.

4.3

Variante BERGSOM para Aplicação Tipo-I

A variante BERGSOM da metodologia proposta foi detalhadamente definida no Capítulo 3, porém, está sumarizada a seguir no contexto da aplicação Tipo-I como Algoritmo 2.

A Figura 4.4 ilustra os passos de inferência que solucionam a aplicação do Tipo-I usando a variante BERGSOM da metodologia proposta. Destaca-se o procedimento de predição única para os sete perfis gerada a partir da componente $\mathbf{w}_{i^*}^y$ do protótipo \mathbf{w}_{i^*} da unidade vencedora na rede SOM

Algoritmo 2: Variante BERGSOM para aplicação Tipo-I

-
- 1 Tomar conjunto de treinamento com dados de imagens
 $X = \{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ e de perfis básicos $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_p\}$;
 - 2 **Passo 1:** treinar um modelo β -CoVAE com dados de X e gerar conjunto de representações $Z = \{\mathbf{z}_1, \dots, \mathbf{z}_p\}$;
 - 3 **Passo 2:** usar conjuntos Z (gerado no Passo 1) e Y para treinar uma rede SOM expandida usando Equação 3-4;
 - 4 **Inferência:** tomar conjunto de teste com dados de imagens
 $X^{teste} = \{\mathbf{x}_1^{teste}, \dots, \mathbf{x}_q^{teste}\}$;
 - 5 **para** $t \leftarrow 1$ **até** q **faça**
 - 6 Selecionar \mathbf{x}_t^{teste} ;
 - 7 Encontrar uma representação $\mathbf{z}_t^{teste} = \text{cod}(\mathbf{x}_t^{teste})$ via modelo β -CoVAE (treinado no Passo 1);
 - 8 Apresentar \mathbf{z}_t^{teste} à rede SOM expandida (treinada no Passo 2) para definir unidade vencedora i^* , usando Equação 3-15;
 - 9 Extrair $\mathbf{w}_{i^*}^y$ do protótipo \mathbf{w}_{i^*} , usando a Equação 3-7;
 - 10 Gerar a previsão: $\hat{\mathbf{y}}_t \leftarrow G_y(\mathbf{w}_{i^*}^y) = \mathbf{I}_{n \times n} \cdot \mathbf{w}_{i^*}^y$;
 - 11 Avaliar variante BERGSOM, por meio de métricas de regressão multivariada usando como referência os dados reais de perfis básicos
 $Y^{teste} = \{\mathbf{y}_1^{teste}, \dots, \mathbf{y}_q^{teste}\}$.
-

expandida.

Dois hiperparâmetros principais que influenciam os resultados da metodologia proposta de forma mais significativa: 1) a dimensão $\dim(\mathcal{Z})$ das representações das imagens, ou seja, o valor de m na Figura 4.2) das representações da imagem; 2) o tamanho da rede SOM utilizada.

Quanto ao tamanho adequado para a dimensão da representação da imagem, $\dim(\mathcal{Z})$, deve ser tal que seja capaz de regerar imagens de boa qualidade na aplicação Tipo-II, enquanto mantém seu potencial preditivo de perfis básicos na aplicação Tipo-I.

Quanto à rede SOM, deve ser encontrada uma arquitetura adequada em termos de geometria e número de unidades de modo a obter um equilíbrio entre a qualidade dos dados previstos/regenerados e a complexidade da rede.

Portanto, a seguir examina-se o desempenho da metodologia proposta na aplicação Tipo-I variando dois aspectos principais: (a) a influência da taxa de compressão aplicada aos dados de origem, ou seja, a dimensão $\dim(\mathcal{Z})$ do espaço de representação da imagem (valores da Tabela 4.2); (b) a influência do tamanho da rede SOM.

As análises³ sobre os hiperparâmetros que influenciam a metodologia pro-

³A implementação foi codificada em linguagem Python e utilizou Tensorflow (ABADI et al., 2015) e Keras (CHOLLET et al., 2015) para modelos de aprendizagem profunda (β -CoVAE e SimpleNet) e IntraSOM (de Gouvêa et al., 2023) para a rede SOM.

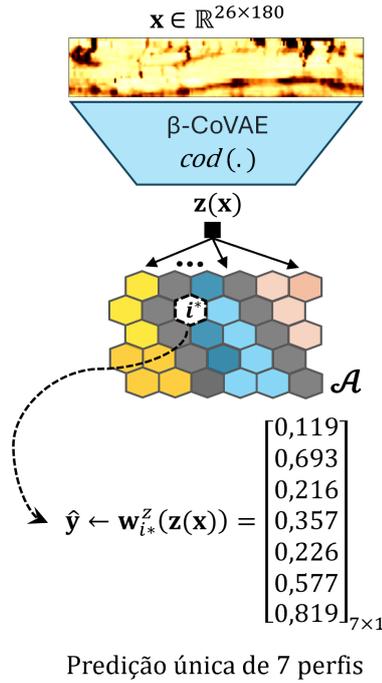


Figura 4.4: Aplicação Tipo-I com a variante BERGSOM.

posta são levadas a cabo na Seção 4.3.1 e aplicam somente a variante BERGSOM usando o Algoritmo 2, sem perda de generalidade sobre as conclusões obtidas.

4.3.1

Análise de $\dim(\mathcal{Z})$

Inicialmente, avalia-se o impacto da redução de dimensionalidade realizada pelo β -CoVAE ao extrair características de imagens de entrada. Representações latentes significativas no espaço \mathcal{Z} são fundamentais não apenas para permitir predições⁴ coerentes de perfis básicos com a rede SOM expandida, mas também para preservar a capacidade de recuperar boas imagens a partir dessas representações compactas para a aplicação Tipo-II. A Figura 4.5 mostra a imagem real comparada às imagens reconstruídas a partir de diferentes representações obtidas para um trecho do poço A usando sete modelos β -CoVAE (variando-se $\dim(\mathcal{Z})$) treinados com os poços do Conjunto 1.

O trecho da imagem real do poço A na Figura 4.5 possui uma ampla gama de características geológicas e não geológicas, portanto é uma excelente amostra para avaliar o processo de reconstrução dos modelos β -CoVAE. Todas as reconstruções preservam o aspecto geral da imagem real, mas é perceptível

⁴Neste trabalho, usa-se preferencialmente o termo ‘predição’ para indicar as estimativas multivariadas obtidas por meio do mapeamento heteroassociativo com a rede SOM para diferenciar dos métodos tradicionais de regressão supervisionada, quando optou-se por utilizar o termo ‘previsão’.

Tabela 4.2: Taxas de compactação aplicadas aos dados de origem.

Dimensão original	$dim(\mathcal{Z})$	Taxa de compressão
4680 (= 26 × 180)	16	292,50
4680	32	146,25
4680	64	73,13
4680	128	36,56
4680	256	18,28
4680	512	9,14
4680	1024	4,57

uma melhora no nível de detalhe nas reconstruções a partir de $dim(\mathcal{Z}) = 64$. No entanto, torna-se difícil perceber maiores diferenças nas reconstruções obtidas a partir de representações de dimensões superiores, por exemplo $dim(\mathcal{Z}) > 256$. Em termos práticos, considerando custo computacional e dificuldades inerentes ao tratamento de dados de elevada dimensão, pode-se prosseguir com o experimento mesmo com $dim(\mathcal{Z}) = 16$, pois todas as representações obtidas com os diferentes modelos β -CoVAE cumprem sua função primária de reconstruir boas imagens⁵.

A seguir, faz-se uma análise semelhante, variando $dim(\mathcal{Z})$ e observando o impacto nas predições dos perfis. Para essa análise, o tamanho do SOM é de 10×10 unidades, que é considerado um mapa de tamanho pequeno, com formato toroidal e vizinhança do tipo gaussiana. Mantém-se o tamanho do mapa constante nessa fase do experimento.

A Figura 4.6 mostra o comportamento das predições do perfil densidade (DEN) para o poço A geradas com a variante BERGSOM com diferentes valores de $dim(\mathcal{Z})$ variados de 16 a 1024. A primeira trilha à esquerda compara o perfil real e a previsão do modelo SimpleNet (SNet), enquanto as trilhas restantes a comparam com as predições BERGSOM.

Visualmente, é difícil identificar diferenças no comportamento das diversas predições BERGSOM obtidas com a variação de $dim(\mathcal{Z})$. Vale destacar, porém, que todas as predições obtidas pela metodologia proposta são muito próximas das obtidas pela rede SNet.

Quantitativamente, pode-se afirmar que não há diferenças significativas entre as predições de perfis obtidas pela variante BERGSOM à luz das métricas padrão adotadas neste trabalho (RMSE, MAE e correlação de Pearson). A Figura 4.7 pode ilustrar essa observação. Ele mostra uma ligeira redução de

⁵No Capítulo 5, essa impressão visual será confirmada por meio de métricas apropriadas de avaliação de qualidade de imagens.

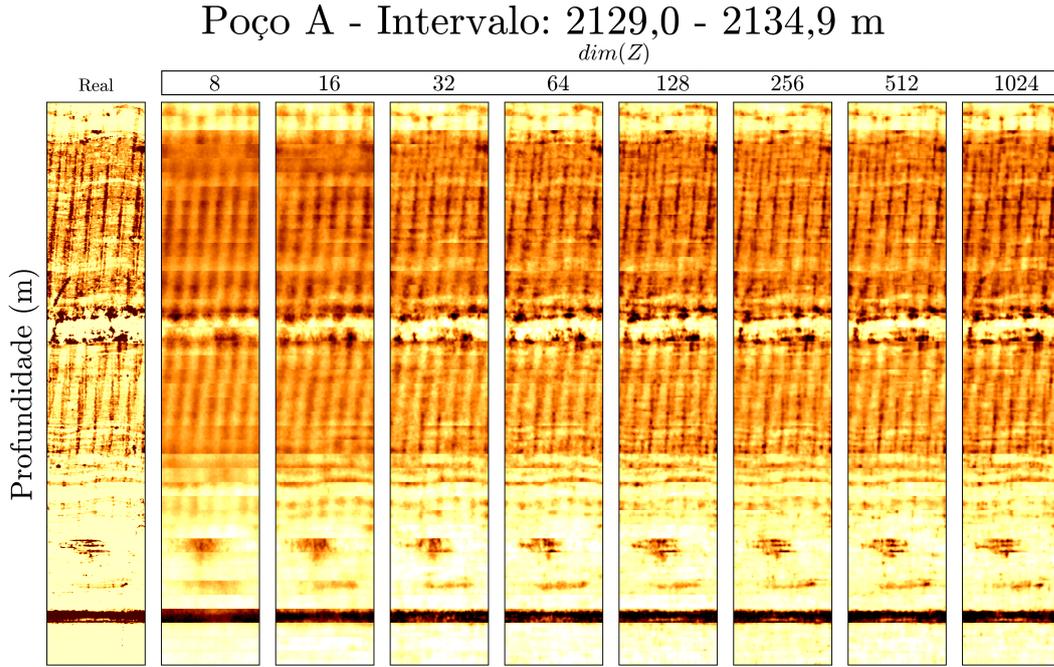


Figura 4.5: Reconstruções de imagens por diferentes modelos β -CoVAE - variando $dim(\mathcal{Z})$. Intervalo 2129,0 - 2134,9 m do poço A.

ambas as métricas RMSE e MAE enquanto aumenta os valores de $dim(\mathcal{Z})$ até um certo ponto, a partir do qual não há mais melhora. Esse comportamento se repete em todos os demais conjuntos de poços, apenas com pequenas alterações na localização da inflexão da curva, conforme mostrado nas Tabelas 4.3 e 4.4. Deve-se observar que RMSE e MAE da previsão SNet são usados como referências porque o modelo SimpleNet não aprende a partir de \mathbf{z} .

Os valores destacados nas Tabelas 4.3 e 4.4⁶ levam a duas observações sobre o desempenho da variante BERGSOM para a aplicação do Tipo-I: (i) a variante BERGSOM tem desempenho superior ao modelo SNet para todos os poços independente da métrica, exceto para a métrica MAE no poço A; (ii) maiores valores de $dim(\mathcal{Z})$ não favorecem a melhora do desempenho da variante BERGSOM. Destaca-se que mesmo a associação da representação mais compacta do experimento, $dim(\mathcal{Z}) = 8$ com uma rede SOM expandida de tamanho considerado pequeno, oferece um desempenho competitivo, tanto

⁶As métricas calculadas nesta seção são baseadas em uma única rede SOM expandida treinada para cada conjunto de treinamento, o que gera uma instância da metodologia BERGSOM para estimar todos os perfis em um dado poço de teste. As métricas mostradas nas Tabelas 4.3 e 4.4 para cada poço de teste são valores médios agregados para as sete curvas. Deve-se notar que não há como gerar uma previsão de variação ou estimativa de incerteza para as predições da variante BERGSOM usando a mesma rede SOM, seria necessário treinar diversas redes e aplicar a metodologia para cada uma delas. Porém, na variante p -BERGSOM, será possível fornecer essa estimativa de incerteza, pois essa variante pode gerar um número arbitrário de N predições para quaisquer dos perfis básicos a partir de uma mesma imagem de entrada.

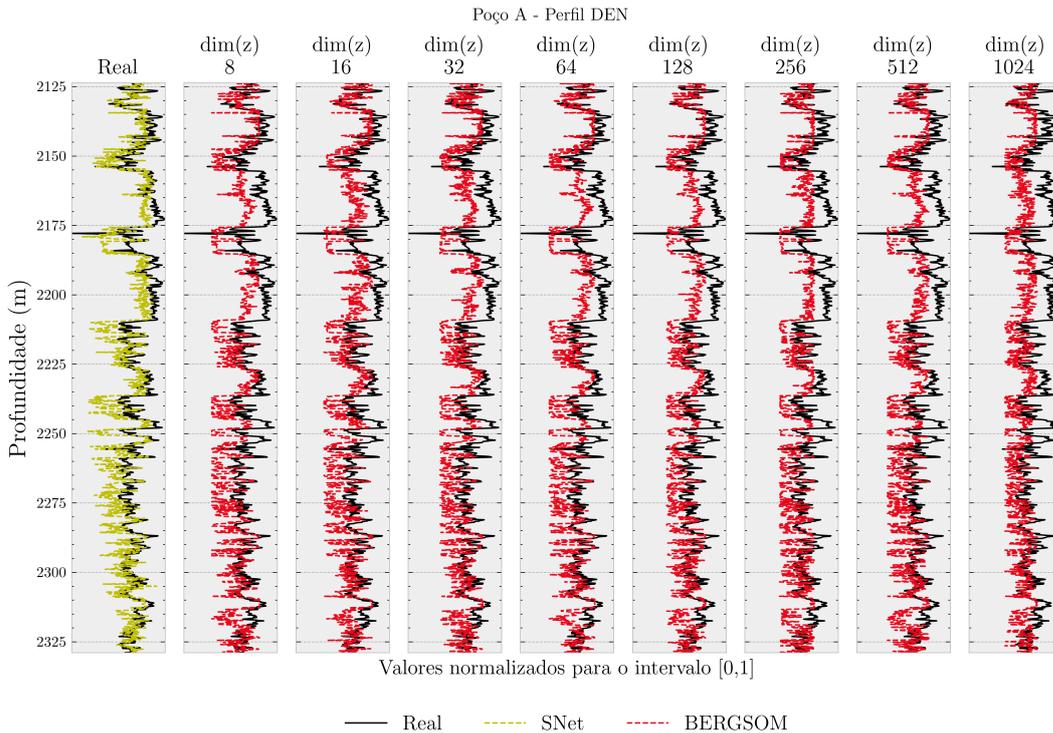


Figura 4.6: Predições para perfil DEN no poço A (SOM com 100 unidades) - sensibilidade a $\dim(\mathcal{Z})$.

visualmente quanto pelas métricas quantitativas, de forma alguma inferior ao apresentado pelo modelo SimpleNet. Outrossim, esses resultados indicam que valores moderados $\dim(\mathcal{Z})$ (16 a 64) oferecem melhores resultados quantitativos. Essas duas observações reforçam a decisão de usar valores menores para $\dim(\mathcal{Z})$ no decorrer dos experimentos deste capítulo.

No geral, a Figura 4.8 considera o desempenho desse primeiro experimento para todos os perfis no poço A, em termos da correlação de Pearson. Basicamente, a variante BERGSOM tem desempenho um pouco melhor que o SimpleNet para a maioria dos perfis preditos para o poço A, exceto para os perfis DEN e NMR PhiT. Embora o aumento da dimensão não mostre ganho significativo em relação às melhores predições, obtidas nos perfis diretamente relacionadas à porosidade, como DEN, NEU e NMR PhiT, parece melhorar o desempenho em alguns dos piores (como GR e PE) no poço A. Vale destacar, como se viu na Tabela 4.4, que o poço A foi o único no qual a rede SNet superou a variante BERGSOM em termos de métrica MAE.

4.3.2

Análise do tamanho do mapa do SOM

A rede SOM realiza quantização vetorial (QV) em seu processo de mapeamento. A perda de informação induzida pela quantização, que é o

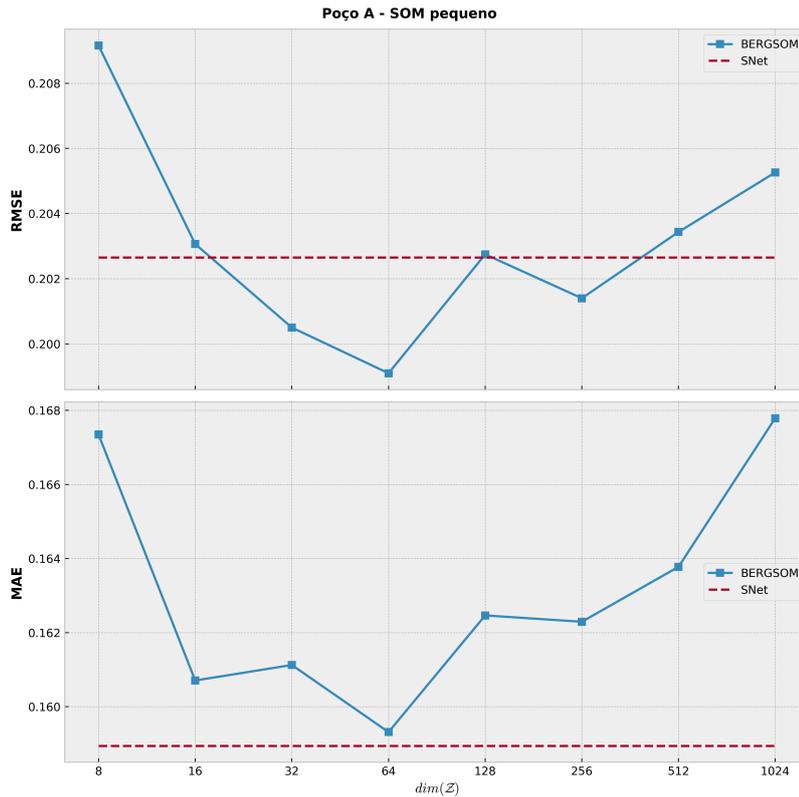


Figura 4.7: Análise conjunta de para predição BERGSOM para sete perfis no poço A (métricas RMSE e MAE) - sensibilidade a $dim(\mathcal{Z})$ em comparação com modelo SimpleNet (linha vermelha tracejada).

processo de discretização de uma região contínua do espaço por um vetor protótipo, depende do tamanho do mapa. Quanto mais unidades, menor a perda, e vice-versa. Esse compromisso entre nível de quantização vetorial obtido e as perdas associadas tem impacto nas predições dos perfis. Além disso, existe o problema de uma mesma unidade ser eleita como representativa de uma grande parte do espaço de entrada, pois uma rede SOM mais compacta tenderá a selecionar o mesmo neurônio para vários padrões de entrada que estiverem localizados na mesma região do espaço coberta pelo vetor protótipo. Para todas as entradas que ativam a mesma unidade, a rede SOM fornece o mesmo peso \mathbf{w}_{i*}^y e, por conseguinte, gera a mesma predição para cada uma dessas entradas. Portanto, o número de unidades no mapa pode afetar fortemente o processo de previsão de saída. A seguir, examina-se a sensibilidade dos resultados da previsão ao tamanho da rede SOM.

A cada vetor de componentes básicos de perfil corresponde uma representação \mathbf{z} relativa a uma imagem 26×180 pixels. Portanto, uma rede SOM composta por K unidades aprende e armazena um total de K representações conjuntas dos espaços origem e destino. Após o estágio generativo, $dec(\mathbf{z})$, cada \mathbf{w}_i^z corresponderá a um *template* de imagem acústica e a um vetor *template*

Tabela 4.3: Avaliação da variante BERGSOM e da rede SNet segundo a métrica RMSE, variando-se $\dim(\mathcal{Z})$ e SOM pequeno.

Poço	8	16	32	64	128	256	512	1024	SNet
A	0,205	0,208	0,205	0,200	0,202	0,202	0,202	0,205	0,203
B	0,162	0,156	0,153	0,154	0,154	0,157	0,156	0,162	0,184
C	0,180	0,175	0,178	0,175	0,179	0,179	0,178	0,181	0,220
D	0,186	0,187	0,186	0,183	0,186	0,187	0,191	0,189	0,207
E	0,203	0,192	0,190	0,190	0,188	0,186	0,183	0,188	0,203

Tabela 4.4: Avaliação da variante BERGSOM e da rede SNet segundo a métrica MAE, variando-se $\dim(\mathcal{Z})$ e SOM pequeno.

Poço	8	16	32	64	128	256	512	1024	SNet
A	0,164	0,167	0,165	0,160	0,163	0,162	0,163	0,167	0,159
B	0,124	0,120	0,119	0,118	0,120	0,125	0,123	0,129	0,142
C	0,141	0,137	0,139	0,138	0,141	0,141	0,140	0,143	0,170
D	0,142	0,137	0,138	0,136	0,137	0,139	0,143	0,141	0,153
E	0,159	0,152	0,151	0,150	0,148	0,148	0,144	0,149	0,160

de perfis básicos. Para ilustrar esse processo, apresenta-se na Figura 4.9 um exemplo com uma rede SOM de tamanho muito reduzido, apenas trinta e seis unidades ($K = 36$), com vizinhanças hexagonais em mapa planar.

A Figura 4.9 destaca quatro unidades na rede SOM com seus respectivos pesos divididos para melhor compreensão. A porção $\mathbf{w}_{i^*}^z$ do peso é transformada novamente em formato de imagem usando o decodificador do β -CoVAE, $dec(\mathbf{w}_{i^*}^z)$. Cada uma das sete componentes da porção $\mathbf{w}_{i^*}^y$ dos pesos é mostrada em cor diferente nos gráficos de barras. Essa visualização ajuda a entender como o procedimento de mapeamento heteroassociativo pode ser usado para efetuar inferência cruzada entre dados de diferentes modalidades. O pequeno tamanho do mapa afeta o nível de detalhe preservado nos pesos, assim, as imagens regeneradas carecem de maiores detalhes, especialmente de textura; uma das razões é a perda significativa no processo de quantização já que, em um mapa tão pequeno, um único neurônio tem que representar uma grande região do espaço de representações (e, portanto, do espaço original das imagens). No entanto, ainda é possível observar como a preservação da topologia se faz presente, pois ao observar as componentes dos pesos, verifica-se que diferentes porções do mapa representam coerentemente diferentes regiões do espaço de entrada aumentado.

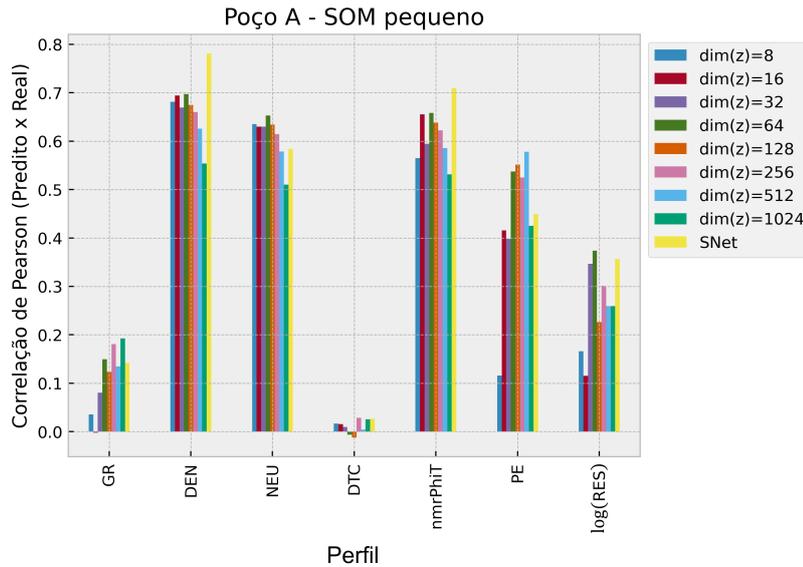


Figura 4.8: Avaliação da variante BERGSOM segundo a métrica de correlação de Pearson para previsão de 7C no poço A - sensibilidade a $dim(\mathcal{Z})$.

Para avaliar a sensibilidade do método proposto ao tamanho da rede SOM, realiza-se um experimento semelhante ao anterior, agora mantendo fixa a dimensão de representação em $dim(\mathcal{Z}) = 16$ e variando o tamanho da rede SOM em três níveis arbitrários: Pequeno ($10 \times 10 = 100$ unidades), Médio ($30 \times 50 = 1500$ unidades) e Grande ($80 \times 90 = 7200$ unidades). O objetivo não é selecionar o tamanho de mapa mais adequado, mas sim entender o efeito desse hiperparâmetro no experimento. Todos os mapas possuem uma geometria toroidal com uma vizinhança gaussiana hexagonal entre as unidades da rede SOM.

A Figura 4.10 mostra a previsão BERGSOM para os três tamanhos considerados para a rede SOM em comparação com a previsão SNet e o perfil real de densidade (DEN) no poço A. É pequena a variação visual para as três previsões BERGSOM. O que se percebe é um comportamento mais errante das previsões BERGSOM, que apresentam amplitudes de variação crescentes à medida que o mapa aumenta. Isso fica mais claro quando se observa o perfil NEU no mesmo poço (Figura 4.11) no qual se percebe um comportamento mais adequado para o mapa pequeno. Essa intuição visual de perda de desempenho com o aumento do tamanho do mapa é confirmada por métricas quantitativas, como mostra a Figura 4.12.

A Figura 4.12 mostra uma ligeira degradação nas métricas RMSE e MAE à medida que o mapa SOM aumenta. A métrica de correlação de Pearson também indica esse comportamento das previsões BERGSOM para o poço A, conforme mostrado na Figura 4.13.

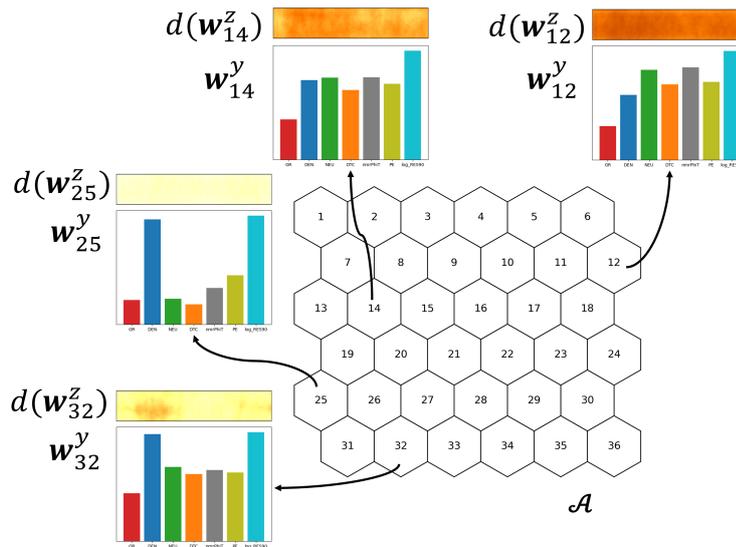


Figura 4.9: Ilustração de representações conjuntas origem-destino em uma rede SOM real, planar, com 36 unidades em vizinhança hexagonal.

A Figura 4.13 considera o desempenho das predições BERGSOM nesse segundo experimento para todos os perfis do poço A em termos da métrica de correlação de Pearson (ρ). Há apenas uma ligeira diferença na correlação para os sete perfis ao aumentar o tamanho do mapa. Pela métrica ρ , no poço A, mapas menores em geral parecem ter melhor desempenho, como se vê nos perfis NMR PhiT e PE, além do perfil NEU, no qual as predições BERGSOM obtidas com mapa pequeno mostram-se superiores ao modelo SNet.

A Figura 4.14 ilustra todas as predições BERGSOM para o poço A. Verifica-se que tanto SNet quando a variante BERGSOM não são capazes de aprender os perfis que não são diretamente correlacionados com a densidade ou a porosidade da rocha, como o perfil RES. Isso já era esperado, pois essas curvas têm baixa ou nenhuma correlação *a priori* com as imagens acústicas, que são a entrada para esse experimento. No entanto, os perfis relacionados à densidade/porosidade apresentam bons resultados, também conforme esperado. Destaca-se novamente o desempenho competitivo da variante BERGSOM com apenas 100 unidades na rede SOM e uma dimensão pequena para a representação latente das imagens.

4.3.3 Análise de Desempenho Computacional

Uma questão importante a ser discutida é o desempenho computacional dos modelos BERGSOM em comparação ao modelo de referência SimpleNet. A Tabela 4.5 apresenta uma comparação desses modelos em termos de tempo de processamento em segundos (s). Os tempos de inferência relatados são a média

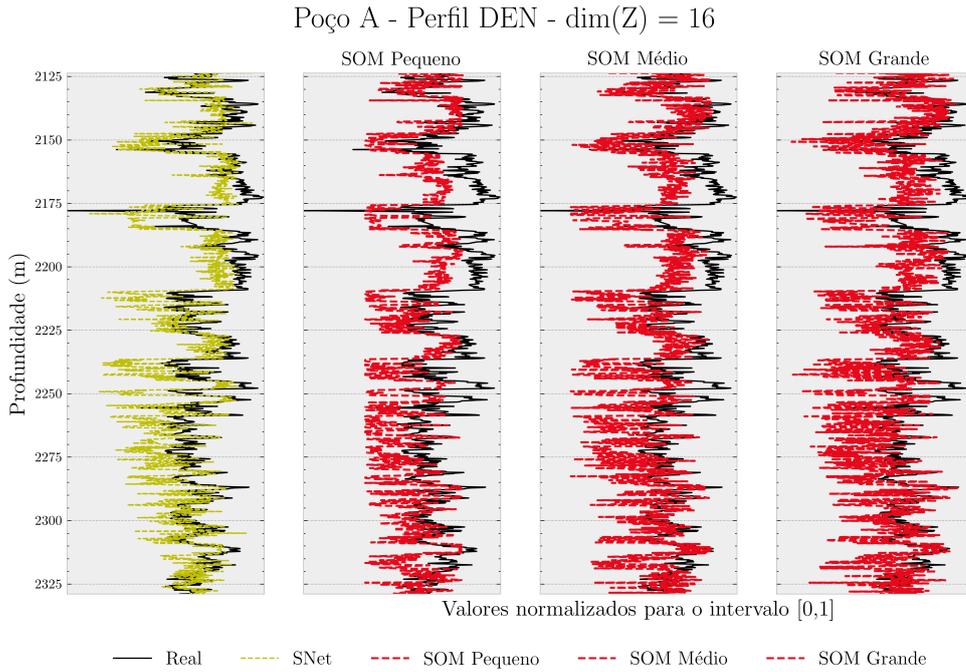


Figura 4.10: Predição do perfil DEN no poço A ($\dim(\mathcal{Z}) = 16$) - sensibilidade ao tamanho da rede SOM.

de 100 execuções de inferência para cada modelo. Uma estratégia de parada antecipada foi usada para o treinamento das redes SimpleNet e β -CoVAE. Todos os modelos usaram treinamento em lotes com tamanho 16. Em média, os modelos β -CoVAE foram treinados por menos de 100 épocas, enquanto a rede SimpleNet foi treinada em menos de 200 épocas. O treinamento do SOM foi controlado por um hiperparâmetro chamado `train_len_factor`, definido com valor unitário nos experimentos. Os tempos reportados foram obtidos usando um PC regular, com um processador Intel(R) Core(TM) i5-1245U 1,60 GHz de 12^a geração e com memória RAM total do sistema de 16 GB.

A Tabela 4.5, com foco em seus tempos de execução para procedimentos de treinamento e inferência. A metodologia BERGSOM requer o treinamento de uma rede β -CoVAE e uma rede SOM, de modo que somamos seus respectivos tempos de treinamento e inferência. Dois valores de $\dim(\mathcal{Z})$ são usados para ilustrar como esse hiperparâmetro afeta o desempenho: $\dim(\mathcal{Z}) = 16$ e $\dim(\mathcal{Z}) = 1024$ (o valor máximo investigado neste trabalho para esse hiperparâmetro). Considera-se mais importante avaliar o tempo de inferência, porque isso permite uma comparação imediata do desempenho de cada modelo em condições operacionais reais, i.e., quando os modelos já estão treinados e prontos para uso. Comparando-se o tempo total (treino + inferência), o modelo BERGSOM se destaca como o mais eficiente, para ambos os valores de $\dim(\mathcal{Z})$. O tempo demandado pela metodologia BERGSOM é influenciado

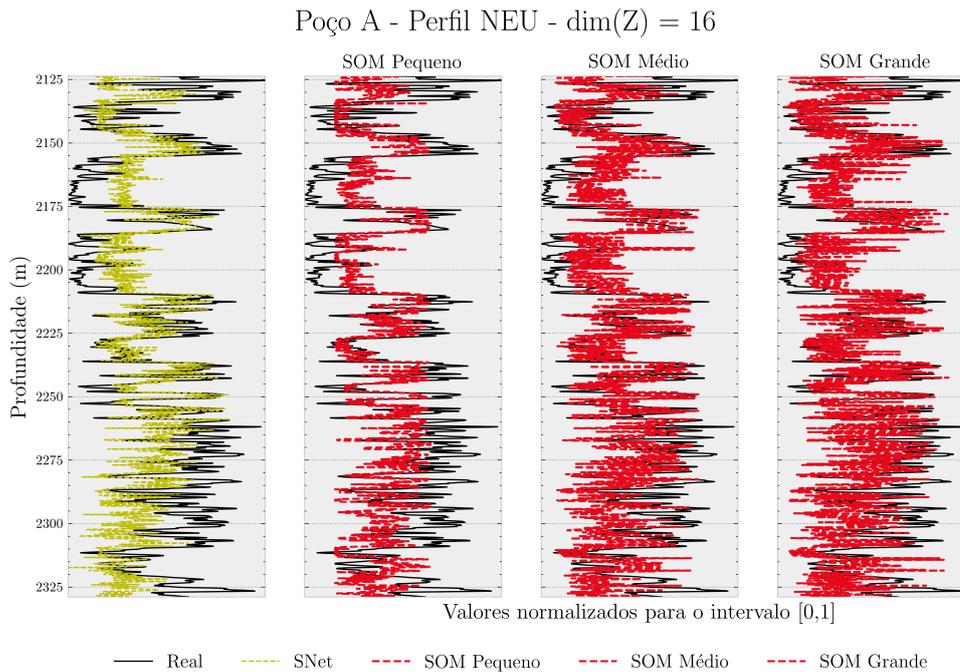


Figura 4.11: Predição do perfil NEU no poço A ($\dim(\mathcal{Z}) = 16$) - sensibilidade ao tamanho da rede SOM.

principalmente pelo tempo necessário para treinar o modelo β -CoVAE, que foi de 26,63 s para $\dim(\mathcal{Z}) = 16$, enquanto a rede SOM necessita de apenas 0,72 s. Vale destacar que, dentre os modelos apresentados nessa comparação de desempenho, somente a metodologia BERGSOM é capaz de inferência bidirecional, uma vantagem que poderia compensar casos em que ela fosse mais custosa em termos de tempo de treinamento. Além disso, uma vez que a estrutura é flexível quanto aos modelos utilizados para aprender as representações dos dados, algumas aplicações podem utilizar modelos mais leves nesta fase, como PCA. Em uma aplicação real na qual O modelo treinado já estará disponível, a metodologia proposta apresenta o tempo de inferência mais rápido. A Tabela 4.5 mostra que a metodologia BERGSOM ganha eficiência com valores mais baixos de $\dim(\mathcal{Z})$, que reforça o argumento de se utilizar a menor representação de imagem possível.

4.3.4

Anomalias encontradas

Tendo em vista os resultados obtidos nos experimentos anteriores, vale destacar o comportamento de uma rede SOM grande (80×90 unidades) quando se aumenta a dimensão da representação da imagem. Neste caso, foi usada vizinhança do tipo bolha para a rede SOM, em que menos unidades são atualizadas em cada passo do treinamento em comparação à vizinhança

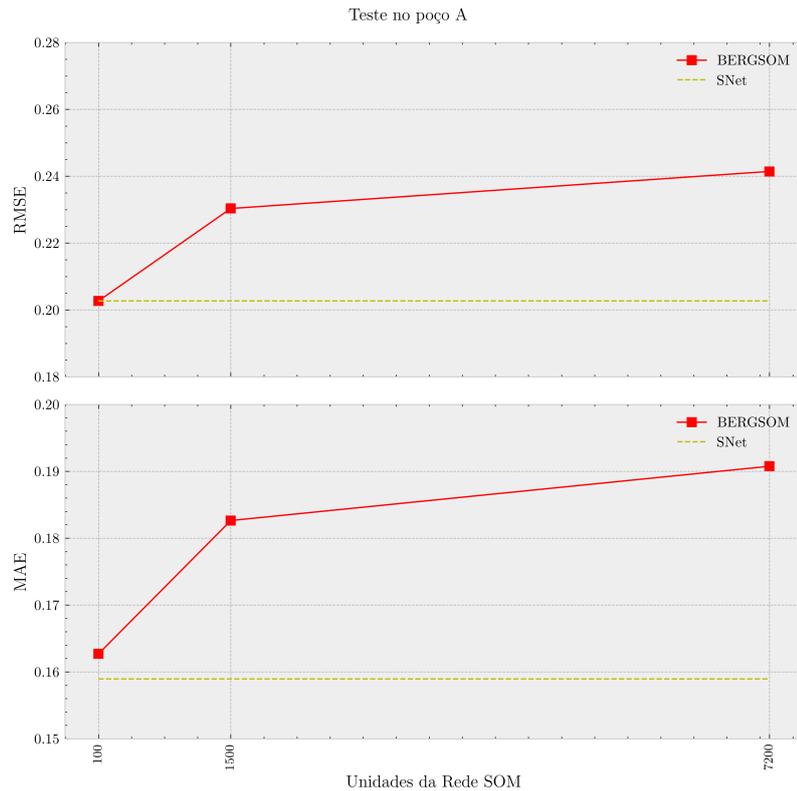


Figura 4.12: Desempenho da predição BERGSOM ($\dim(\mathcal{Z}) = 16$) para sete perfis no poço A (métricas RMSE e MAE) - sensibilidade ao tamanho da rede SOM. Para referência, as métricas da rede SimpleNet são mostradas nas linhas amarelas tracejadas).

gaussiana. A Figura 4.15 mostra que, para esse tamanho de mapa, o aumento em $\dim(\mathcal{Z})$ causa uma degradação acelerada nos resultados. O que não ocorreu no primeiro experimento descrito na Subseção 4.3.1 que usou um mapa pequeno ao variar $\dim(\mathcal{Z})$. Na Figura 4.15 as setas mostram degradação nas predições com o aumento de $\dim(\mathcal{Z})$.

Observou-se também, que o aumento de $\dim(\mathcal{Z})$ no mapa grande faz com que a rede tenha muitas unidades mortas, i.e., neurônios que jamais foram escolhidos como vencedores nem estiveram na vizinhança dos vencedores. Portanto, os pesos dessas unidades permanecem com os valores aleatórios alocados na inicialização da rede SOM (ver Figura 4.16).

Na Figura 4.16, mostra-se a Matriz-U para as diferentes redes SOM expandidas de tamanho grande construídas com valores crescentes para $\dim(\mathcal{Z})$. Na visualização da Matriz-U, valores baixos são codificados em cores frias (mais para o azul escuro) e indicam que unidades próximas no mapa também têm pesos próximos no espaço de entrada aumentado, enquanto valores altos são codificados em cores quentes (mais para o vermelho) indicam que essas unidades vizinhas no mapa têm seus respectivos pesos distantes no espaço de

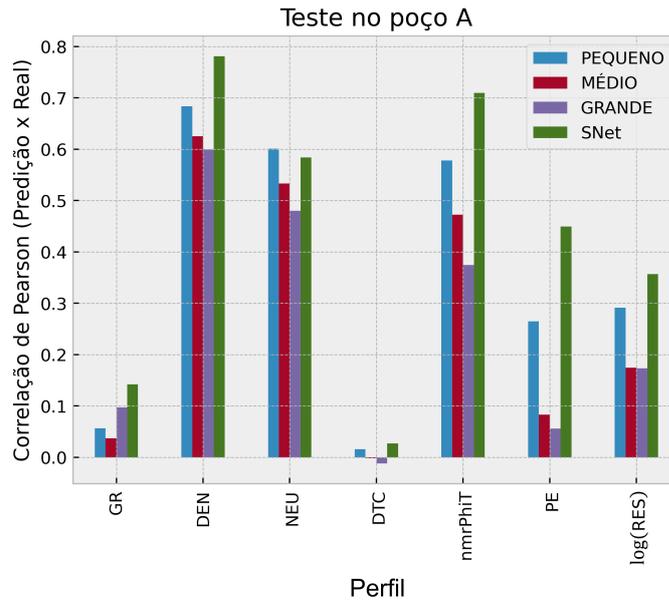


Figura 4.13: Análise de análise de correlação de Pearson para previsão de 7C no poço A - sensibilidade ao tamanho do mapa.

entrada. A partir de $\dim(\mathcal{Z}) = 512$ percebe-se o crescimento de unidades mortas, conforme visto em azul marinho e indicado pelas setas vermelhas. Outra forma de verificar a mesma informação, é observar o histograma dos valores que compõem as diversas Matrizes-U. Novamente, as setas vermelhas mostram a tendência de unidades mortas com distâncias muito baixas entre si e distâncias maiores em relação aos demais neurônios do mapa.

Curiosamente, ao observar as previsões na Figura 4.15, percebe-se pelo valor congelado em algumas previsões marcadas com setas a partir de $\dim(\mathcal{Z}) = 256$, que são devidas exatamente às unidades mortas selecionadas como vencedoras na etapa de inferência cruzada.

Especula-se aqui que essa degradação do desempenho do modelo com o aumento da dimensionalidade do espaço das representações \mathcal{Z} deve-se a dois fenômenos:

1. **Maldição da dimensionalidade:** modelos baseados em discriminação por meio de métricas de distância, como a rede SOM, sofrem com o efeito da concentração das normas (FRANCOIS; WERTZ; VERLEYSSEN, 2007) em espaços de dimensão elevada. Em tais casos, as normas, ou distâncias entre os pontos, concentram-se em uma faixa muito estreita, de modo que o conceito de vizinho mais próximo degrada-se rapidamente com o aumento da dimensionalidade, i.e., o ponto mais próximo indicado pela métrica pode não ser, de fato, um vizinho. Isso é crucial para a abordagem proposta, pois a rede SOM é fundamentada no conceito de

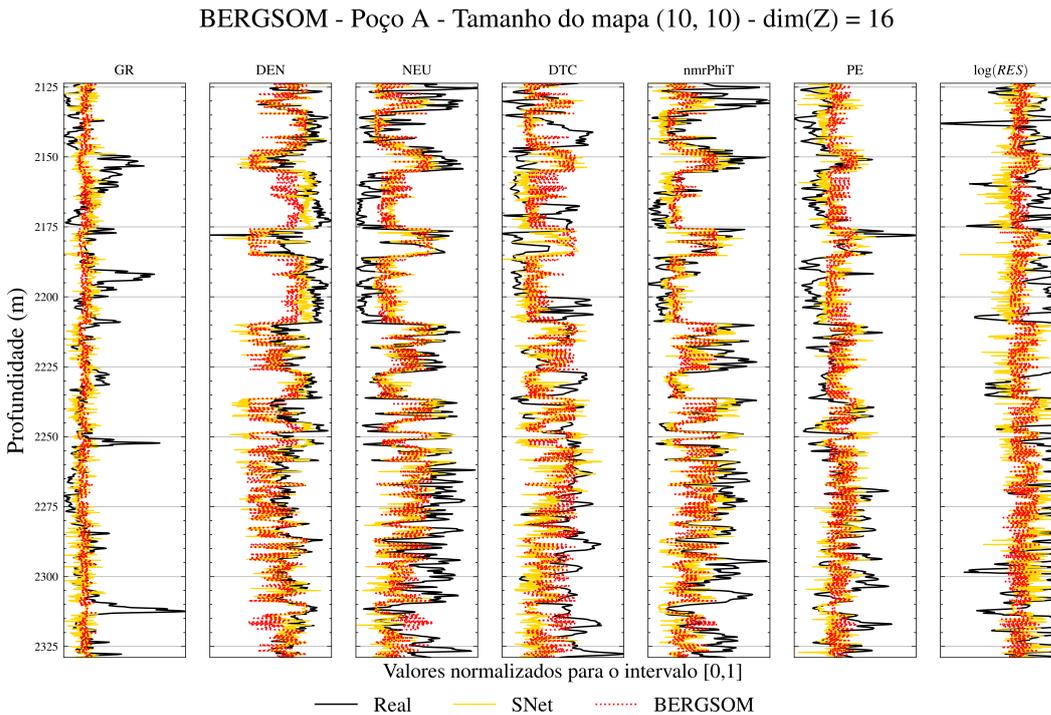


Figura 4.14: Predições BERGSOM de todos os perfis para o poço A - SOM pequeno (10×10 unidades) - $\dim(\mathcal{Z}) = 16$.

unidade vencedora escolhida por uma métrica de distância. Por outro lado, isso explicaria o bom desempenho observado nos experimentos com valores reduzidos de $\dim(\mathcal{Z})$ e se torna um incentivo teórico para buscar o uso de representações compactas.

2. Contraste entre as dimensionalidades das componentes de \mathbf{x}_{aug} :

as componentes da entrada aumentada, \mathbf{x}_{aug} (\mathbf{z} para as representações dos dados de origem e \mathbf{y} para dados de destino) têm diferentes dimensionalidades. Quando $\dim(\mathbf{z}) \gg \dim(\mathbf{y})$, \mathbf{z} terá muito mais relevância na escolha das unidades vencedoras e na ordenação do mapa, que tenderá a priorizar a informação das imagens em detrimento da informação dos perfis básicos. Na etapa de inferência, portanto, a informação contida na imagem pode induzir à escolha de um vencedor que não seja o mesmo que seria obtido se fosse usada diretamente a componente \mathbf{y} . Nos experimentos que variam $\dim(\mathcal{Z})$, percebe-se um grande contraste entre as dimensões das parcelas que compõem \mathbf{x}_{aug} : 8 contra 7 no melhor caso, e 1024 contra 7 no pior caso. Essa discrepância deve estar relacionada sobretudo ao efeito da seleção de unidades mortas na fase de inferência cruzada intermodalidades, que leva o modelo a fazer predições completamente erradas.

Os resultados dos experimentos desta seção, acerca dos hiperparâmetros

Tabela 4.5: Tempos de execução em segundos (s) nas etapas de treinamento (soma dos tempos para as redes β -CoVAE e SOM no modelo BERGSOM) e inferência (Conjunto 1).

Modelo	Treino (s/época)	Inferência (s/poço)	Etapas inferência	Bidirecional?
BERGSOM $dim(\mathcal{Z}) = 16$	27,35	0,59	$\mathbf{x} \rightarrow \mathbf{z} \rightarrow \mathbf{y}$	SIM
BERGSOM $dim(\mathcal{Z}) = 1024$	99,57	0,85	$\mathbf{x} \rightarrow \mathbf{z} \rightarrow \mathbf{y}$	SIM
SimpleNet	522,47	30,75	$\mathbf{x} \rightarrow \mathbf{y}$	NÃO

$dim(\mathcal{Z})$ e tamanho da rede SOM, permitem escolher um valor de $dim(\mathcal{Z}) = 16$ e uma rede SOM pequena (10×10 unidades) para continuar com os experimentos finais deste capítulo, com a aplicação da variante p -BERGSOM para a aplicação Tipo-I.

4.4

Variante p -BERGSOM para Aplicação Tipo-I

A variante p -BERGSOM da metodologia proposta foi detalhadamente definida no Capítulo 3, porém, está sumarizada a seguir no contexto da aplicação Tipo-I como Algoritmo 3.

A Figura 4.17 ilustra os passos de inferência que solucionam a aplicação do Tipo-I usando a variante p -BERGSOM da metodologia proposta. As cores utilizadas para a rede SOM são as mesmas usadas na Figura 4.4, para ressaltar que ambas as variantes usam uma mesma rede SOM expandida na etapa de inferência. Destaca-se o procedimento de amostragem da variante p -BERGSOM, realizado usando a Equação 3-11, que permite gerar N predições reamostrando $\hat{\mathbf{z}}$ a partir das componentes latentes probabilísticas aprendidas pelo modelo β -CoVAE. Um protótipo da rede SOM expandida \mathbf{w}_i mostrado na Figura 4.17 incorpora apenas dois tipos de informação, $(\mathbf{w}_{i^*}^z, \mathbf{w}_{i^*}^y)$. Assim, na variante p -BERGSOM, é necessário reamostrar os valores $\hat{\mathbf{z}}$ para poder apresentá-los à rede SOM e obter, assim, N predições para cada um dos sete perfis usando a etapa de inferência intermodal cruzada.

Nos experimentos a seguir, a variante p -BERGSOM será aplicada usando uma rede SOM pequena, com 100 unidades, e uma representação das imagens com dimensão 16.

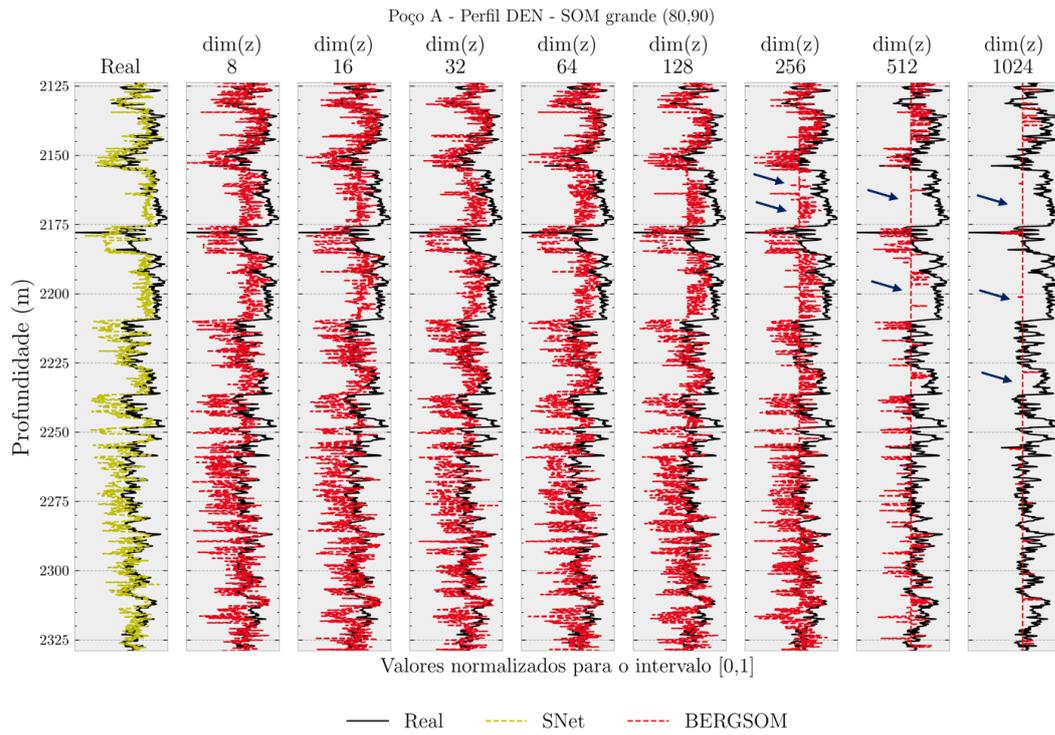


Figura 4.15: Anomalias observadas nas predições com o aumento de $dim(\mathcal{Z})$ - SOM grande (poço A) com vizinhança tipo bolha.

A Tabela 4.6 mostra a avaliação das predições para a variante p -BERGSOM para cada perfil em todos os poços de teste usando a métrica RMSE. A informação de desvio padrão associada a cada erro médio é indicada por meio do símbolo ‘ \pm ’. O mesmo procedimento é repetido para a métrica MAE mostrada na Tabela 4.7.

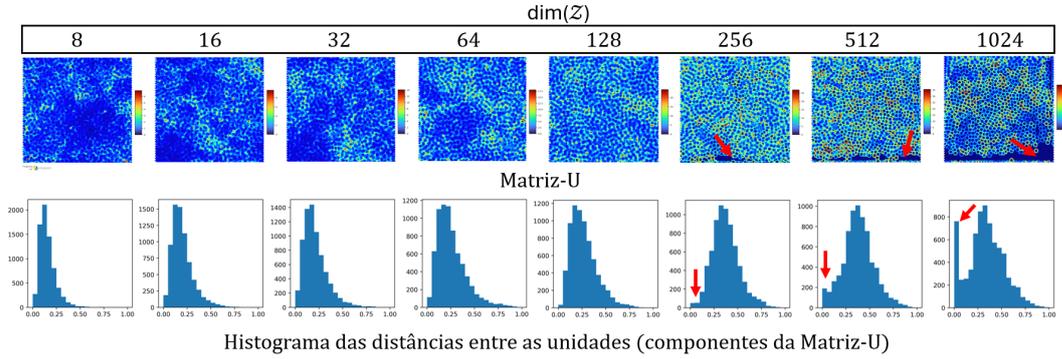


Figura 4.16: Análise das Matrizes U de cada SOM expandido com o aumento de $\dim(Z)$.

Algoritmo 3: Variante p -BERGSOM para aplicação Tipo-I

- 1 Tomar conjunto de treinamento com dados de imagens $X = \{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ e de perfis básicos $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_p\}$;
 - 2 **Passo 1:** treinar um modelo β -CoVAE com dados de X e gerar conjunto de representações $Z = \{\mathbf{z}_1, \dots, \mathbf{z}_p\}$;
 - 3 **Passo 2:** usar conjuntos Z (gerado no Passo 1) e Y para treinar uma rede SOM expandida usando Equação 3-4;
 - 4 **Inferência:** tomar conjunto de teste com dados de imagens $X^{teste} = \{\mathbf{x}_1^{teste}, \dots, \mathbf{x}_q^{teste}\}$;
 - 5 **para** $t \leftarrow 1$ **até** q **faça**
 - 6 Selecionar \mathbf{x}_t^{teste} ;
 - 7 Obter a representação probabilística $\mathbf{z}_{(\mu, \sigma)}(\mathbf{x}_t^{teste}) = [\mu^{z_t}, \sigma^{z_t}]^T$ via modelo β -CoVAE (treinado no Passo 1);
 - 8 Definir número N de previsões a gerar;
 - 9 **para** $j \leftarrow 1$ **até** N **faça**
 - 10 Amostrar $\mathbf{z}(\mathbf{x}_t^{teste})$ usando o truque da reparametrização (Equação 3-11);
 - 11 Apresentar $\mathbf{z}(\mathbf{x}_t^{teste})$ à rede SOM expandida (treinada no Passo 2) para definir unidade vencedora i^* , usando Equação 3-15;
 - 12 Extrair $\mathbf{w}_{i^*}^y$ do protótipo \mathbf{w}_{i^*} , usando a Equação 3-7;
 - 13 Gerar a previsão: $\hat{\mathbf{y}}_t(j) \leftarrow G_y(\mathbf{w}_{i^*}^y) = \mathbf{I}_{n \times n} \cdot \mathbf{w}_{i^*}^y$;
 - 14 Avaliar variante p -BERGSOM, por meio de métricas de regressão multivariada usando como referência os dados de perfis básicos $Y^{teste} = \{\mathbf{y}_1^{teste}, \dots, \mathbf{y}_q^{teste}\}$.
-

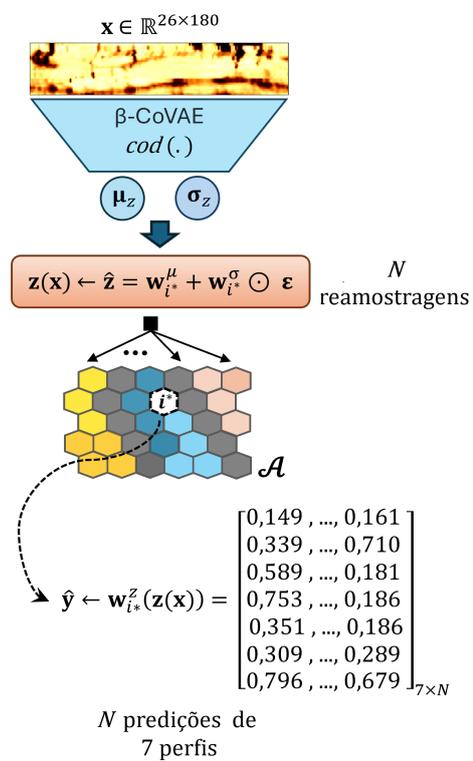


Figura 4.17: Aplicação Tipo-I com a variante p -BERGSOM.

Tabela 4.6: Métrica RMSE das predições *p*-BERGSOM para cada perfil em todos os poços de teste.

Poço	GR	DEN	NEU	DTC	NMR PhiT	PE	log(RES)
A	0,143±5,0E-04	0,190±5,5E-04	0,204±6,0E-04	0,252±7,2E-04	0,210±4,3E-04	0,160±6,0E-04	0,232±9,3E-04
B	0,124±5,6E-04	0,116±5,5E-04	0,136±5,8E-04	0,156±6,3E-04	0,134±7,1E-04	0,217±6,5E-04	0,224±7,5E-04
C	0,142±6,0E-04	0,131±6,5E-04	0,195±8,9E-04	0,249±1,0E-03	0,141±8,3E-04	0,170±8,9E-04	0,208±6,0E-04
D	0,130±7,2E-04	0,154±6,3E-04	0,149±7,8E-04	0,218±9,7E-04	0,182±5,3E-04	0,185±4,5E-04	0,315±1,2E-03
E	0,139±4,8E-04	0,183±5,5E-04	0,145±6,9E-04	0,251±4,3E-04	0,193±6,1E-04	0,186±4,2E-04	0,218±6,2E-04

Tabela 4.7: Métrica MAE das predições *p*-BERGSOM para cada perfil em todos os poços de teste.

Poço	GR	DEN	NEU	DTC	NMR PhiT	PE	log(RES)
A	0,093±4,5E-04	0,153±6,0E-04	0,166±5,9E-04	0,208±6,1E-04	0,166±5,8E-04	0,126±5,3E-04	0,211±6,3E-04
B	0,090±4,1E-04	0,089±5,7E-04	0,109±8,4E-04	0,128±1,1E-03	0,100±5,3E-04	0,196±5,5E-04	0,165±8,5E-04
C	0,105±7,4E-04	0,101±6,6E-04	0,153±7,2E-04	0,226±7,3E-04	0,112±6,1E-04	0,121±5,7E-04	0,147±8,5E-04
D	0,092±4,1E-04	0,138±8,0E-04	0,139±1,0E-03	0,203±1,0E-03	0,153±8,8E-04	0,135±4,6E-04	0,203±4,9E-04
E	0,093±4,7E-04	0,146±4,6E-04	0,119±5,8E-04	0,218±4,8E-04	0,161±6,1E-04	0,168±5,1E-04	0,169±6,2E-04

A variante p -BERGSOM foi desenvolvida para superar algumas limitações percebidas na variante BERGSOM com respeito à avaliação de incertezas em suas predições:

- As predições da variante BERGSOM são únicas e, para obter uma avaliação de incerteza sobre essas predições, seria necessário simular N modelos BERGSOM para cada poço de teste. Nesse caso, a variabilidade das predições seria em parte extraída da variação estocástica dos conjuntos de treinamentos usados para cada um desses N modelos e das condições iniciais do treinamento (pesos iniciais). Não haveria, de fato, uma transferência da variabilidade de uma modalidade dos dados, no caso dados de origem, para as predições.
- Ao permitir que uma mesma realização de um modelo baseado na rede SOM heteroassociativa seja capaz de gerar um número arbitrário N de predições, a variante p -BERGSOM apresenta-se como um método capaz de transferir a incerteza das imagens de entrada para as predições de múltiplos perfis. Isso é uma característica desejável para uma aplicação do Tipo-I, pois se trata de um problema inverso mal posto, do tipo “um para muitos”.
- Além das métricas de erro médio e desvio padrão, podem-se usar estatísticas intervalares para cada predição de perfil. A Figura 4.19 ilustra essa possibilidade em um intervalo do perfil DEN no poço A.

A Figura 4.18 ilustra todas as $N(= 100)$ predições p -BERGSOM sobrepostas (em verde) para cada um dos sete perfis do poço A. As predições únicas obtidas pela variante BERGSOM (em vermelho) também são mostradas e verifica-se que elas estão dentro do intervalo de variação das predições p -BERGSOM. Mais uma vez vale ressaltar que o desempenho da metodologia proposta (em ambas as variantes) usando uma rede SOM de tamanho pequeno e com $\dim(\mathcal{Z}) = 16$ é comparável ao do modelo de referência SNet (curva tracejada em amarelo). Outrossim, a variante p -BERGSOM incorpora a vantagem de permitir a estimativa de estatísticas intervalares para avaliar a incerteza das predições. Um gráfico detalhado mostrando apenas o perfil DEN é mostrado na Figura 4.19 para destacar a avaliação da incerteza na predição de múltiplos perfis a partir de imagens.

A Figura 4.19 é uma visualização ampliada do perfil DEN para o intervalo (2194,5 - 2244,5 m) do poço A. Na primeira faixa, à esquerda, comparam-se o perfil DEN real com as predições únicas de BERGSOM e SNet e as $N(= 100)$ predições p -BERGSOM (em verde). Para manter a referência, o perfil real também está presente nas demais faixas 2 (ao centro) e 3 (à direita), sempre em

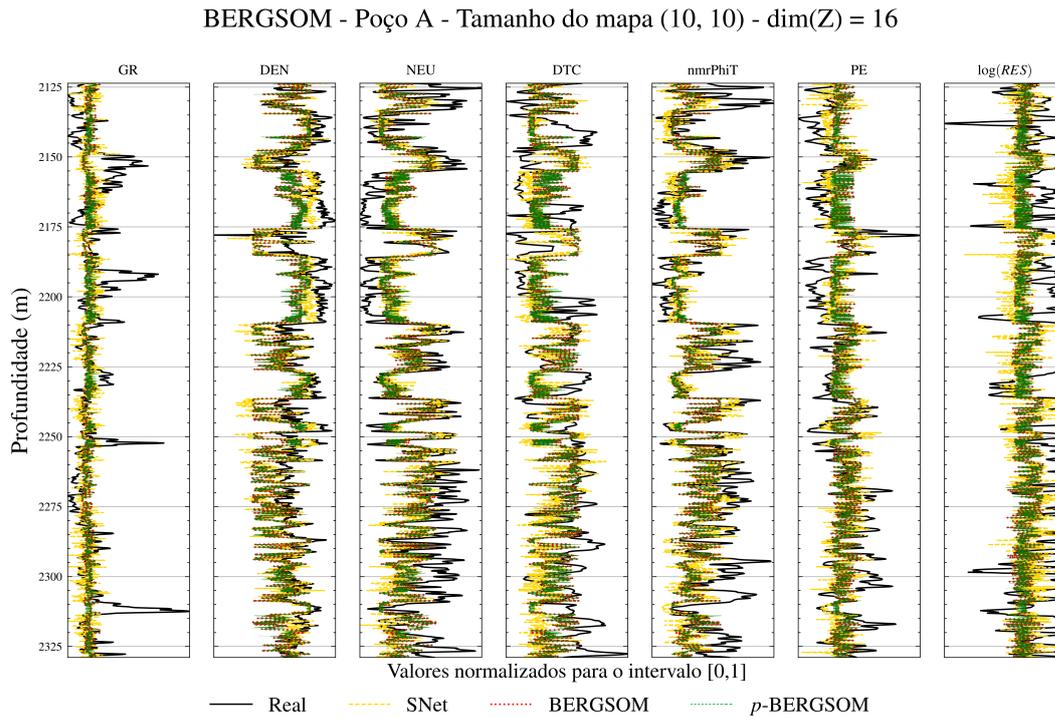


Figura 4.18: Predições p -BERGSOM de todos os perfis para o poço A - SOM pequeno (10×10 unidades) - $\dim(Z) = 16$.

traço contínuo na cor preta. A primeira faixa destaca a grande correspondência das predições únicas BERGSOM e SNet. Na segunda faixa, comparam-se a previsão simples de BERGSOM e a previsão p -BERGSOM média (sobre $N (= 100)$ predições) juntamente com a faixa interquartil (IQR) para as N predições probabilísticas de p -BERGSOM. Finalmente, na terceira faixa, mostra-se a previsão p -BERGSOM média, destacando a faixa mínimo-máximo dos valores das predições p -BERGSOM. Vale ressaltar a proximidade das predições BERGSOM e p -BERGSOM média. A vantagem conceitual da variante p -BERGSOM é a incorporação de informações de incerteza herdadas dos dados de imagem. As estatísticas intervalares fornecidas pela variante p -BERGSOM tornam-no um método de estimativa de múltiplas saídas mais significativo quando comparado à variante BERGSOM ou ao modelo supervisionado profundo SNet, pois ambos fornecem apenas uma previsão de perfis para uma determinada imagem de entrada.

4.5

Comparação com o estado da arte (SOTA) em Petrofísica

Ao comparar os resultados quantitativos dos modelos aqui desenvolvidos (tanto aqueles baseados na rede SOM expandida, quanto o modelo baseado na SimpleNet) para a tarefa de predição de curvas com trabalhos recentes relevantes na área de petrofísica de reservatórios, não se encontram termos

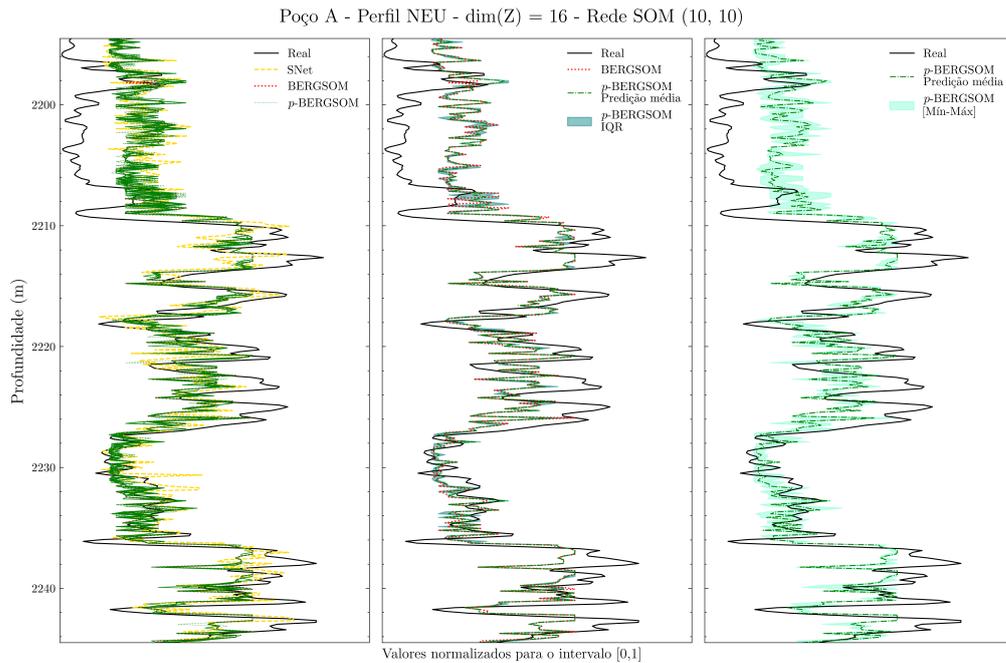


Figura 4.19: Ampliação do perfil de densidade (DEN) e das previsões (BERGSOM e p -BERGSOM) para o intervalo (2194,5 - 2244,5 m) do poço A (p -BERGSOM gera $N(= 100)$ previsões) - SOM pequeno e $\dim(\mathcal{Z}) = 16$.

de comparação diretos, pois ainda não havia tentativa anterior de resolver o mesmo problema. No entanto, encontram-se trabalhos correlacionados na literatura que podem pelo menos fornecer uma base comparativa aos resultados aqui reportados.

Tamoto, Gioria e Carneiro (2023) avaliam quatro modelos de aprendizado supervisionado para a tarefa de predição individual de três perfis de RMN: porosidade total (PhiT), porosidade efetiva (PhiE) e nível de fluido livre (FF). Chen et al. (2023) propõem um método de programação genética associado a princípios físicos para predição individual de dois perfis convencionais a partir de seis outras medições petrofísicas (principalmente perfis de RMN) em reservatórios carbonáticos heterogêneos. Li (2023) combina métodos baseados em dados e conhecimento do domínio do problema para prever a permeabilidade de reservatórios carbonáticos complexos. Bom et al. (2021) aplicam dois tipos de perfis de imagens, acústica e resistiva, como entradas para um modelo que combina aprendizado profundo e métodos bayesianos para fornecer funções de densidade de probabilidade separadamente para previsões de permeabilidade e porosidade. Simões et al. (2023) comparam o desempenho de três métodos de aprendizado de máquina (ML) ao prever uma das seguintes curvas básicas: densidade (DEN), porosidade de nêutrons (NEU) e curvas de lentidão compressiva (DTC) em reservatórios siliciclásticos. Kim (2022) propõe uma abordagem híbrida de aprendizado de máquina para gerar o perfil sônico de

Tabela 4.8: Referências SOTA para previsão de perfis/grandezas petrofísicas.

Referência	Reservatório	Validação	Entrada	Saída	Desempenho
Chen et al. (2023)	C	NB	BSC	k PE	$R_k^2 \in [0, 75 - 0, 85]$ $R_{PE}^2 \in [0, 93 - 0, 97]$
Li (2023)	C	B	BSC	k	$R_k^2 = 0, 87$
Simões et al. (2023)	R	B	BSC	DEN NEU DTC	$\rho_{DEN} = 0, 95$ $\rho_{NEU} = 0, 83$ $\rho_{DTC} = 0, 80$
Tamoto, Gioria e Carneiro (2023)	C	B	BSC	NMR PhiT NMR PhiE NMR FF	$R_{PhiT}^2 \in [0, 67 - 0, 85]$ $R_{PhiE}^2 \in [0, 67 - 0, 87]$ $R_{FF}^2 \in [0, 62 - 0, 75]$
Kim (2022)	R	B	BSC	DTS	$R_{DTS}^2 = 0, 931$
Bom et al. (2021)	C	NB	IMG	k PHI	$R_k^2 > 0, 9$ $R_{PHI}^2 > 0, 9$

cisalhamento (DTS) sintético usando como entrada outros cinco perfis básicos (GR, RHOB, NPHI, RT e DTC). A Tabela 4.8⁷ resume essas referências destacando as métricas R^2 e ρ usadas para validação da maioria deles.

Na Tabela 4.8, percebe-se que, embora os problemas tenham algum grau de relação com os experimentos ora realizados, nenhum deles permite termos de comparação perfeitos. Todos utilizaram métodos de aprendizagem supervisionada, o que já é um ponto a favor do caráter inovador desta tese, que usa uma abordagem autossupervisionada. Poucos lidam com a mesma informação de entrada (imagens de poço), e nenhum realizou aprendizagem para predição multivariada; em vez disso, eles previram uma curva/propriedade de cada vez.

Um aspecto fundamental que mostra a complexidade da previsão de 7C a partir de imagens acústicas diz respeito aos diversos artefatos nos dados de imagem. Apesar dos esforços para utilizar as melhores imagens possíveis em termos de qualidade, os poços ainda apresentam artefatos induzidos pela construção do poço, como *breakouts*, perfurações para retirada de amostras laterais e marcas criadas pelos patins das próprias ferramentas de aquisição dos perfis de imagem. Esses artefatos perturbam localmente a relação entre a imagem e os perfis básicos, cuja frequência de aquisição e nível de detalhamento são menores que os dos perfis de imagem (ver Apêndice A). Então, isso pode distorcer o processo de aprendizagem heteroassociativa. Tal problema não é verificado nas previsões que usam perfis regulares para predição univariada de

⁷Na Tabela 4.8, usa-se ‘C’ para reservatórios com predominância de rochas carbonáticas e ‘R’ para aqueles compostos por outros tipos de rocha (arenito, calcário etc); na coluna de validação, ‘B’ significa teste cego em poço separado não utilizado no treinamento, ‘NB’ indica que partes dos poços de treino, embora não utilizadas no treinamento, foram usadas para avaliação de desempenho; na coluna *Entrada*, ‘BSC’ significa uso de perfis básicos e ‘IMG’ para perfis de imagens. As métricas reportadas em (SIMÕES et al., 2023) correspondem a 90% das amostras com melhores resultados.

outros perfis básicos e é menos prejudicial nas previsões unidimensionais de perfis convencionais a partir de perfis de imagem.

Vale ressaltar que o experimento de regressão multivariada para prever as sete curvas básicas equivale a uma tentativa de extração de informações interpretáveis a partir de perfis de imagens de paredes de poços, o que é um problema desafiador. Tal objetivo enseja uma dificuldade adicional que é a tentativa de prever alguns perfis básicos que geralmente não são, *a priori*, correlacionadas com perfis de imagem acústicas, como os perfis GR ou RES. No entanto, a interrelação entre um perfil de imagem e os perfis básicos pode se comportar de maneira diferente quando se utilizam outros tipos de perfis de imagens, como aqueles de imagens resistivas.

Em suma, os resultados acima mostram que, numa tarefa desafiadora de predição cruzada intermodalidades, a abordagem autossupervisionada proposta, mesmo em configurações reduzidas (mapas menores e representações de dimensão reduzida), pode competir e até superar o desempenho de abordagens baseadas em redes neurais convolucionais profundas com arquitetura complexa. A seguir, discutem-se brevemente algumas justificativas teóricas para isso.

Ao concatenar dados de origem e de destino para mapeamento heteroassociativo com a rede SOM expandida, assume-se uma forte relação entre essas fontes de dados, o que é verdade no presente caso, como mostra a prática regular de especialistas em petrofísica de reservatórios. No entanto, ao contrário de abordagens supervisionadas, que consideram uma dependência funcional explícita entre entradas e saídas. A metodologia autossupervisionada proposta assume que é possível associá-las computacionalmente por meio de processos de aprendizado heteroassociativo.

Os métodos supervisionados usam um sinal formal, o erro de predição $\|\mathbf{y} - \hat{\mathbf{y}}\|$, como um sinal “professor”. A presente abordagem, por sua vez, pode estimar $\hat{\mathbf{y}}$ ($\leftarrow \mathbf{w}_{i^*}^y$) sem depender de tal sinal. Em vez disso, o mapeamento conjunto fornece a estimativa, usando o neurônio vencedor i^* , definido nas Equações 3-15, como a ponte entre as duas modalidades de dados.

A regra de aprendizagem da rede SOM foi projetada para reduzir o erro de quantização, ou seja, a distância no espaço original entre as entradas do mapa e os pesos de suas respectivas unidades. Adaptando-se a Equação 2-6 para o presente caso:

$$e_q(\mathbf{y}, \mathbf{w}_{i^*}^y) = \|\mathbf{y} - \mathbf{w}_{i^*}^y\|. \quad (4-1)$$

Usando a definição da Equação 4-1 e considerando a fase de convergência do treinamento do SOM, Barreto e Araujo (2004) mostram que um método de memória associativa baseado em rede SOM, como o proposto nesta tese, também é um algoritmo de aprendizagem baseado em erros. Assim, mé-

todos supervisionados e autossupervisionados são comparáveis em tarefas de predição/regressão uni- ou multivariadas.

4.6

Considerações finais

Neste capítulo, foi apresentada a primeira aplicação da metodologia proposta, abordando ambas as variantes, BERGSOM e p -BERGSOM, no contexto da petrofísica de reservatórios: a predição de um conjunto completo de perfis básicos a partir de perfis de imagens acústicas. Essa aplicação era inédita na área.

O capítulo enfatizou o caráter bidirecional da metodologia, que visa permitir a estimação cruzada entre dados de diferentes modalidades. Assim, a metodologia proposta, juntamente com o modelo de aprendizado profundo desenvolvido para comparação de resultados, pode ser estendida a diversas outras aplicações que demandem técnicas de regressão numérica a partir de imagens, como a avaliação de imagens médicas.

Além disso, foram discutidas características específicas e limitações encontradas para a metodologia apresentada no contexto da aplicação Tipo-I em dados petrofísicos.

A seguir, o Capítulo 5 dedica-se à aplicação do Tipo-II no contexto da petrofísica de reservatórios.

5 Aplicação Tipo-II em Petrofísica

Neste capítulo¹, descreve-se o segundo tipo de experimentos práticos para avaliar a metodologia proposta em um problema da área de petrofísica de reservatórios:

Tipo-II Geração de dados complexos a partir de dados mais simples: regeneração de perfis de imagens acústicas a partir de conjuntos de perfis básicos.

É uma aplicação mais desafiadora que aquela explorada no Capítulo 4, tanto que a criação de uma arquitetura concorrente para comparação de resultados, por exemplo usando aprendizado profundo, poderia por si só, ser objeto de outra tese. A metodologia porém, em suas duas variantes, mostrou-se capaz de regerar perfis de imagem coerentes, com um nível adequado de discriminação de estruturas que permite, por exemplo, solucionar demandas típicas para tomada de decisão em tempo real durante a perfuração dos poços.

Outrossim, demonstra-se conceitualmente que a metodologia proposta configura-se como um arcabouço geral no qual se inclui o já clássico uso da rede SOM em aplicações de imputação de dados faltantes.

O capítulo apresenta ainda uma funcionalidade difícil de implementar em modelos supervisionados puros, que é regerar perfis de imagem coerentes a partir de um número menor de perfis básicos do que aquele utilizado para treinamento da rede SOM expandida para o mapeamento heteroassociativo.

Os resultados obtidos indicam que a proposta deste trabalho com mapeamento heteroassociativo baseado na rede SOM expandida pode contribuir com outras aplicações em diversos campos do conhecimento que explorem o aprendizado regenerativo, quando há demanda de regerar dados complexos ou mais custosos a partir de dados mais simples, ou com baixo custo de obtenção.

5.1 Dados e pré-processamento

Por questão de clareza na exibição dos experimentos a seguir e seus resultados, sugere-se conhecer maiores detalhes sobre os dados de poço que foram utilizados no Apêndice A.

¹Este capítulo serviu de base para a preparação dos artigos 2 e 6 (em parte) citados no Apêndice B.

5.2

Considerações gerais

A regeneração de imagens acústicas com a metodologia de aprendizado heteroassociativo pressupõe associar a rede SOM a um modelo generativo. Neste trabalho, esse modelo é o decodificador do β -CoVAE usado para aprender as representações compactas das imagens que compõem o espaço das representações \mathcal{Z} . No entanto, a metodologia proposta permite variações nessa etapa, e outros modelos generativos podem ser utilizados.

Assim, as imagens reconstruídas pelo decodificador do β -CoVAE alimentado com imagens reais do poço serão a referência de desempenho para as imagens regeneradas a partir da metodologia proposta. A Figura 5.1 mostra diferentes reconstruções da imagem e um trecho do poço A usando diferentes modelos β -CoVAE segundo a dimensão da representação latente $dim(\mathcal{Z})$.

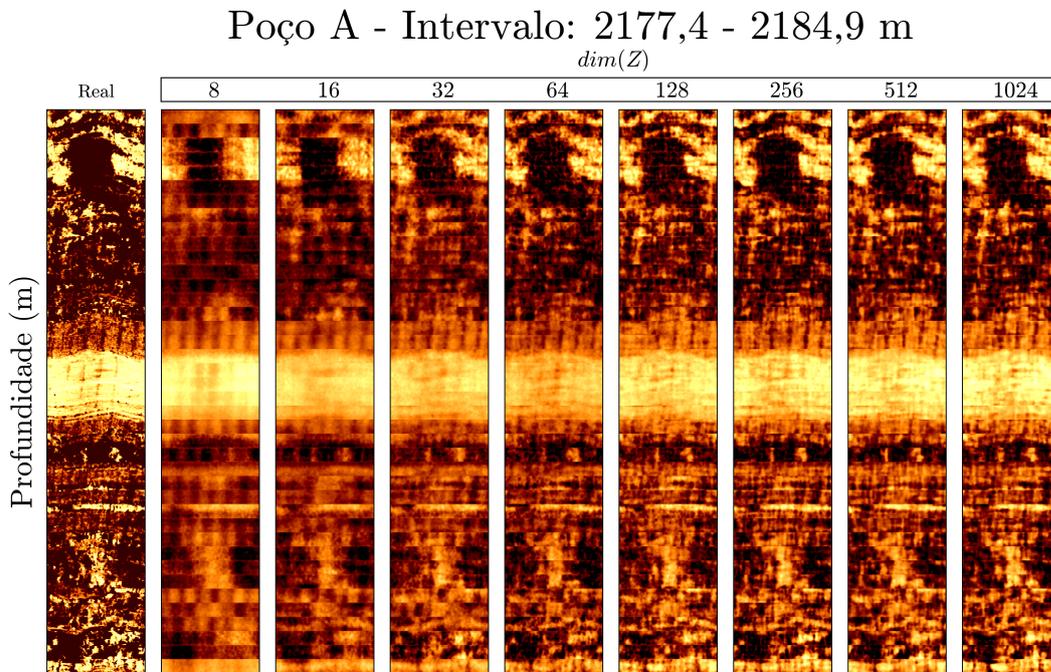


Figura 5.1: Imagem real e imagens reconstruídas pelos diversos modelos β -CoVAE variando-se $dim(\mathcal{Z})$. Intervalo 2177,4 - 2184,9 m do poço A.

As reconstruções mostradas na Figura 5.1 são obtidas diretamente pelos diferentes modelos β -CoVAE alimentados com as próprias imagens reais. Deve-se, portanto, ter em conta que qualquer imagem regenerada a partir dos perfis básicos usando a rede SOM expandida tem como limite máximo de qualidade a imagem reconstruída pelo decodificador correspondente do modelo β -CoVAE alimentado com a imagem real.

Conforme esperado intuitivamente a Figura 5.1 dá a impressão visual de tendência de melhora das imagens reconstruídas com o aumento de $dim(\mathcal{Z})$. No

entanto, comparar as imagens acima apenas visualmente não permite precisar a relação entre a qualidade das imagens e a dimensão do espaço latente dos 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 nos remete à necessidade do uso de métricas para avaliação de qualidade da imagem (AQI), em especial métricas de comparação pareada entre imagens, do tipo FR (do inglês *full reference*) quando existe uma imagem e referência.

Como as imagens regeneradas têm a mesma dimensão da imagem real, há algumas métricas clássicas da área de visão computacional que podem ser empregadas para comparação direta entre uma imagem de referência \mathbf{x}^{ref} e uma imagem cuja qualidade deve ser avaliada, em nosso caso $\hat{\mathbf{x}}$. O erro médio quadrático, MSE (Equação 5-1) é uma métrica inicial natural.

$$\text{MSE}(\mathbf{x}^{\text{ref}}, \hat{\mathbf{x}}) = \left\| \mathbf{x}^{\text{ref}} - \hat{\mathbf{x}} \right\|_2^2 \quad (5-1)$$

Porém, 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 (SESHADRINATHAN et al., 2009; WANG; BOVIK, 2009). Assim, outras métricas foram propostas com objetivo de ser perceptualmente mais realistas. Uma das mais usuais é o índice de similaridade estrutural (SSIM, do inglês *structural similarity*) (WANG et al., 2004), projetado para comparar padrões locais pixel a pixel, mas levando em conta informação estrutural da imagem. Matematicamente, este índice é definido como

$$\begin{aligned} \text{SSIM}(\mathbf{x}, \hat{\mathbf{x}}) &= l(\mathbf{x}, \hat{\mathbf{x}}) \cdot c(\mathbf{x}, \hat{\mathbf{x}}) \cdot s(\mathbf{x}, \hat{\mathbf{x}}) = \\ &= \left(\frac{2\mu_x\mu_{\hat{x}} + C_1}{\mu_x^2 + \mu_{\hat{x}}^2 + C_1} \right) \cdot \left(\frac{2\sigma_x\sigma_{\hat{x}} + C_2}{\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2} \right) \cdot \left(\frac{\sigma_{x\hat{x}} + C_3}{\sigma_x\sigma_{\hat{x}} + C_3} \right). \end{aligned} \quad (5-2)$$

O SSIM é definido pela Equação 5-2 como o produto de três características extraídas das imagens: luminescência $l(\mathbf{x}, \hat{\mathbf{x}})$, contraste $c(\mathbf{x}, \hat{\mathbf{x}})$ e estrutura $s(\mathbf{x}, \hat{\mathbf{x}})$. As estatísticas locais μ_x , $\mu_{\hat{x}}$, σ_x , $\sigma_{\hat{x}}$ e $\sigma_{x\hat{x}}$ são calculados com base numa janela de dimensão 11×11 que desliza pixel a pixel sobre as imagens \mathbf{x}^{ref} e $\hat{\mathbf{x}}$. As constantes C_1 , C_2 e C_3 são incluídas de modo a evitar problemas numéricos caso $\mu_x^2 + \mu_{\hat{x}}^2$, $\sigma_x^2 + \sigma_{\hat{x}}^2$ e $\sigma_x\sigma_{\hat{x}}$, respectivamente, assumam valores muito próximos de zero. Os autores (WANG et al., 2004) apresentam ainda uma forma mais simples da Equação 5-2, ao considerar $C_3 = C_2/2$:

$$\text{SSIM}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{(2\mu_x\mu_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2)}, \quad (5-3)$$

na qual sugerem-se os valores $C_1 = 0,01^2$ e $C_2 = 0,03^2$. A Equação 5-3 será utilizada para cálculo da métrica SSIM nas avaliações de desempenho a seguir.

Mais recentemente, com a ubiquidade de modelos baseados em aprendizado profundo na área de visão computacional, surgiram métricas de AQI baseadas nas representações (ou *embeddings*) aprendidos por modelos profundos em tarefas de classificação de imagens. Uma das mais usadas é a métrica LPIPS, do inglês *Learned Peceptual Image Patch Similarity* (ZHANG et al., 2018), que explora o fato de que as representações aprendidas por modelos profundos parecem capturar informações estruturais e semânticas das imagens que são importantes para a percepção humana. Assim, essa métrica foi projetada para superar as limitações de outras que não necessariamente concordam com a percepção humana sobre qualidade de imagens.

A métrica LPIPS pode ser calculada usando os *embeddings* aprendidos por diferentes modelos de aprendizado profundo, como Alex-Net (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) e VGG (SIMONYAN; ZISSERMAN, 2014), por uniformidade com o que se tornou praxe na literatura, optou-se aqui pela implementação fornecida pelos proponentes da métrica². Essa versão foi treinada em um grande conjunto de dados de julgamentos de similaridade perceptual humana, incluindo distorções tradicionais e outras inseridas por algoritmos como autocodificadores (ZHANG et al., 2018).

As métricas MSE e LPIPS são medidas de dissimilaridade entre a imagem avaliada e a imagem de referência, portanto as melhores imagens obtêm valores mais baixos. Por sua vez, a métrica SSIM é uma medida de similaridade, e as melhores imagens obtêm valores mais altos nessa métrica.

Neste trabalho, as imagens são reescaladas para o intervalo $[0, 1]$ tanto para o treinamento das diferentes redes neurais utilizadas, como para calcular quaisquer das métricas AQI para avaliação da metodologia proposta.

Aplicando-se as métricas acima descritas a imagens da Figura 5.1, em relação ao perfil de imagem acústica real de um trecho do poço A, obtém-se o comportamento mostrado na Figura 5.2, segundo as métricas MSE, LPIPS e SSIM.

A Figura 5.2 indica que as métricas acima descritas parecem estar adequadas ao tipo de imagem tratadas neste trabalho, pois a percepção visual de melhoria com o aumento de $\dim(\mathcal{Z})$ é em geral verificada nas métricas LPIPS e SSIM para o exemplo dado, enquanto que, apesar de oscilar, o MSE em geral decresce com o aumento de $\dim(\mathcal{Z})$. Vale ressaltar, porém, que esperavam-se valores mais elevados de similaridade estrutural, pois a partir de $\dim(\mathcal{Z}) = 128$ elas são todas muito parecidas com a original, mas o valor máximo de SSIM é de apenas 0,14, i.e., 14% de similaridade.

²Disponível em Perceptual Similarity Repository, acessado em 26/11/2024.

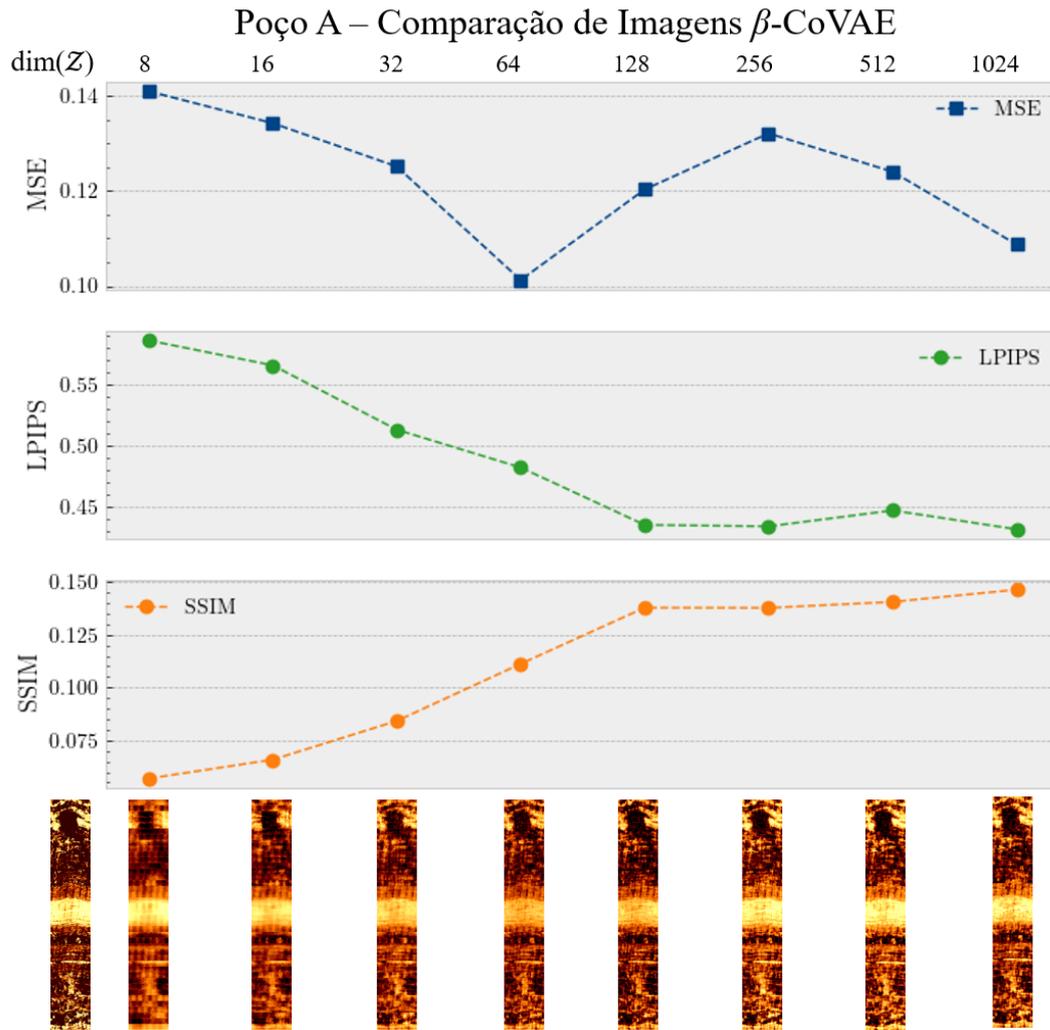


Figura 5.2: Métricas MSE, LPIPS e SSIM para as imagens da Figura 5.1. Métricas calculadas com referência à imagem real no lado esquerdo.

Para confirmar a percepção acima, foram calculadas as três métricas para todos os conjuntos de poços (sempre levando em conta o comprimento total com perfil de imagem disponível no poço utilizado para teste), comparando as imagens reais com aquelas reconstruídas pelos diferentes modelos β -CoVAE variando-se a dimensão do espaço de representações \mathcal{Z} . Os resultados são mostrados na Figuras 5.3. Para todos os poços, observam-se as métricas AQI em geral melhorando com o aumento de $\dim(\mathcal{Z})$, porém sem grandes diferenças a partir de $\dim(\mathcal{Z})$.

Qualitativamente, percebe-se na Figura 5.3, que pode-se reduzir a investigação para uma faixa de valores para $\dim(\mathcal{Z})$ entre 16 e 128, na qual o desempenho dos modelos de geração de imagem foram adequados, sem a necessidade de manusear representações de imagem com dimensão por demais elevada.

A seguir, a metodologia proposta será aplicada em suas duas variantes,

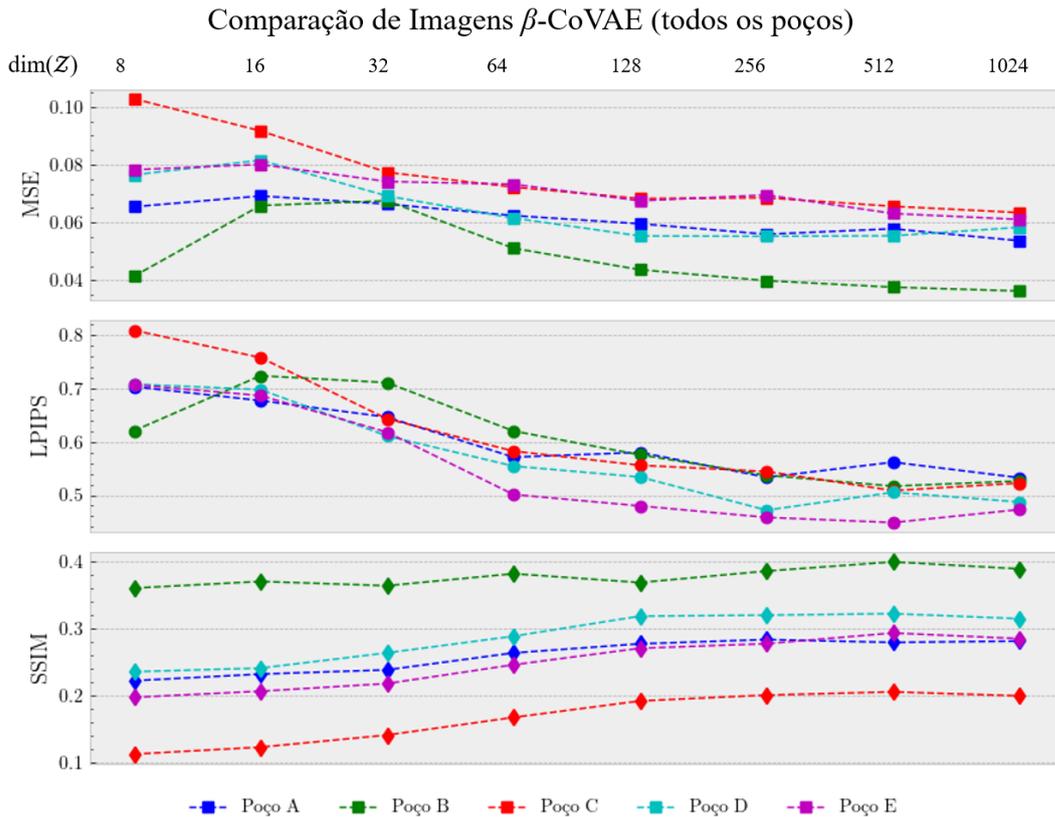


Figura 5.3: Métrica MSE para os modelos β -CoVAE em todos os poços de teste, variando-se $\dim(\mathcal{Z})$.

BERGSOM e p -BERGSOM para a aplicação Tipo-II. Os experimentos estão centrados na avaliação dos mesmos hiperparâmetros avaliados no Capítulo 4: dimensão do espaço das representações $\dim(\mathcal{Z})$ e tamanho da rede SOM usada no mapeamento heteroassociativo³. Serão usados sete perfis de entrada⁴: GR, DEN, NEU, DTC, NMR PhiT, PE, $\log(RES)$. Para todos os experimentos, a rede SOM tem as seguintes configurações: vizinhança hexagonal gaussiana, geometria toroidal e inicialização randômica dos pesos.

5.3

Regeração de Imagens a partir de Perfis Básicos

O termo “regeração” é empregado para indicar a geração de dados de alta densidade a partir de uma fonte de dados de baixa densidade, como definido no paradigma de aprendizado regenerativo (TAN et al., 2023). Na aplicação Tipo-II, a aprendizagem regenerativa ocorre quando as variantes BERGSOM ou p -BERGSOM alimentam um modelo generativo (no presente

³O tamanho da rede SOM em três níveis arbitrários: Pequeno ($10 \times 10 = 100$ unidades), Médio ($30 \times 50 = 1500$ unidades) e Grande ($80 \times 90 = 7200$ unidades).

⁴Na aplicação Tipo-II, as entradas são os perfis básicos usados como dados de destino da aplicação Tipo-I explorada no Capítulo 4, enquanto as saídas são os perfis de imagem acústica.

caso, o decodificador β -CoVAE) capaz de gerar (regerar) dados complexos a partir de entradas mais simples.

Imagens geológicas, como aquelas fornecidas por perfis acústicos de paredes de poços, são tipicamente avaliadas por especialistas humanos que julgam sua qualidade com base em seu potencial para identificar estruturas geológicas ou artefatos de perfuração. No entanto, em nossos experimentos, compara-se a semelhança entre uma imagem regerada e uma imagem de referência. Para a aplicação Tipo-II, deve-se ter em mente que a referência para avaliação de qualidade das imagens regeradas pela metodologia proposta é sempre a imagem produzida pelo modelo β -CoVAE correspondente alimentado diretamente com a imagem real. Portanto, utiliza-se uma série de métricas de avaliação de qualidade da imagem para avaliar essa semelhança: MSE, SSIM e LPIPS.

Nos experimentos a seguir, a exemplo do que se fez no Capítulo 4, será utilizado um valor fixo para a dimensão de representação em $\dim(\mathcal{Z}) = 16$ e será variado o tamanho da rede SOM nos mesmos três níveis arbitrários: SOM pequena ($10 \times 10 = 100$ unidades), SOM média ($30 \times 50 = 1500$ unidades) e SOM grande ($80 \times 90 = 7200$ unidades). Não se trata de uma busca exaustiva pelo tamanho de mapa ótimo, e sim um experimento para entender o efeito desse hiperparâmetro na aplicação Tipo-II nos moldes do que foi feito para a aplicação Tipo-I. A rede SOM tem os mesmos demais ajustes usados anteriormente: geometria toroidal com vizinhança gaussiana hexagonal entre suas unidades.

5.3.1

BERGSOM para Regeração de Perfis de Imagens

A variante BERGSOM da metodologia proposta foi detalhadamente definida no Capítulo 3, porém, será sumarizada a seguir como Algoritmo 4 no contexto da aplicação Tipo-II. Ilustra-se na Figura 5.4 a etapa de inferência da variante BERGSOM da metodologia proposta para resolver a aplicação do Tipo-II.

Destaca-se na Figura 5.4 a considerável capacidade de regeração da metodologia proposta, pois mesmo a variante BERGSOM (mais simples) é capaz de regerar imagens coerentes e complexas, de dimensão 26×180 , a partir de uma entrada reduzida com apenas 7 dimensões. Deve-se observar que, para uma determinada entrada, essa variante é capaz de regerar uma única imagem. Em comparação com a variante BERGSOM para a aplicação Tipo-I, podem-se usar os mesmos modelos treinados nos experimentos do Capítulo 4 (Figura 4.4) e simplesmente inverter o sentido do passo de inferência cruzada

Algoritmo 4: Variante BERGSOM para aplicação Tipo-II

- 1 Tomar conjunto de treinamento com dados de imagens $X = \{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ e de perfis básicos $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_p\}$;
 - 2 **Passo 1:** treinar um modelo β -CoVAE com dados de X e gerar conjunto de representações $Z = \{\mathbf{z}_1, \dots, \mathbf{z}_p\}$;
 - 3 **Passo 2:** usar conjuntos Z (gerado no Passo 1) e Y para treinar uma rede SOM expandida usando Equação 3-4;
 - 4 **Inferência:** tomar conjunto de teste com dados de perfis básicos $Y^{teste} = \{\mathbf{y}_1^{teste}, \dots, \mathbf{y}_q^{teste}\}$;
 - 5 **para** $t \leftarrow 1$ **até** q **faça**
 - 6 Selecionar \mathbf{y}_t^{teste} ;
 - 7 Apresentar \mathbf{y}_t^{teste} à rede SOM expandida (treinada no Passo 2) para definir unidade vencedora i^* , usando Equação 3-19;
 - 8 Extrair $\mathbf{w}_{i^*}^z$ do protótipo \mathbf{w}_{i^*} , usando a Equação 3-9;
 - 9 Definir a representação $\mathbf{z}_t^{teste} = \mathbf{w}_{i^*}^z$;
 - 10 Re gerar a previsão: $\hat{\mathbf{x}}_t \leftarrow G_z(\mathbf{z}_t^{teste}) = dec(\mathbf{z}_t^{teste})$, na qual $dec(\cdot)$ representa o decodificador do modelo β -CoVAE treinado no Passo 1;
 - 11 Avaliar variante BERGSOM, por meio de métricas IQA usando como referência as imagens correspondentes geradas pelos modelos β -CoVAE alimentados diretamente com imagens reais $X^{teste} = \{\mathbf{x}_1^{teste}, \dots, \mathbf{x}_q^{teste}\}$.
-

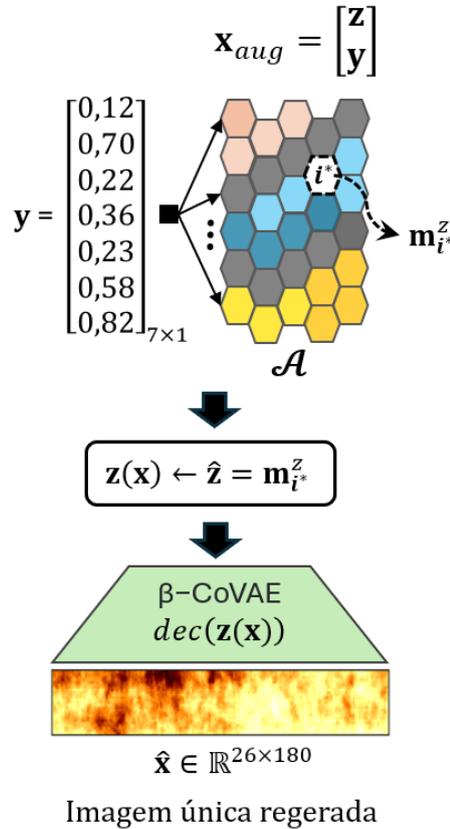


Figura 5.4: Aplicação Tipo-II com a variante BERGSOM.

intermodalidades.

A Figura 5.5 mostra imagens regeradas para um mesmo trecho do poço A usando a variante BERGSOM variando-se a dimensão do espaço das representações $dim(\mathcal{Z})$ e com uma rede SOM de tamanho grande (80×90 unidades).

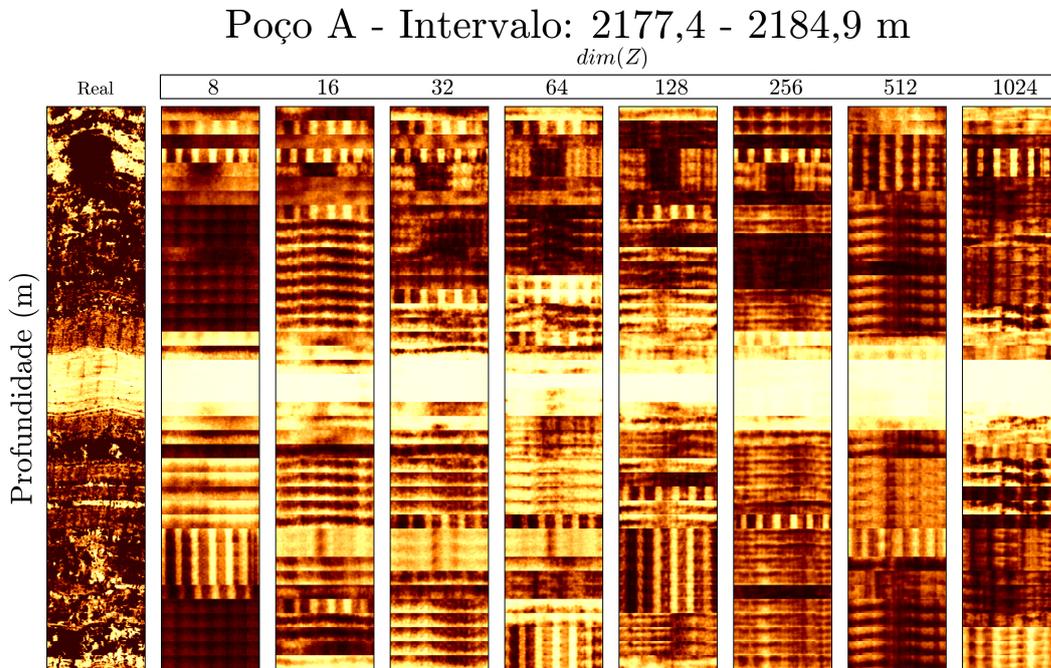


Figura 5.5: Perfis de imagem acústica regeradas pela variante BERGSOM usando sete perfis básicos de entrada. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.

Ao analisar visualmente os resultados mostrados na Figura 5.5, observa-se uma qualidade geral inferior à observada anteriormente, quando se reconstroem as imagens usando os modelos β -CoVAE da Seção 5.2. 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 qualidade apresentado, com as grandes estruturas e contraste entre camadas representados, pode ser suficiente para algumas aplicações do cotidiano da indústria de óleo e gás citadas no Capítulo 1, tendo em vista o enorme potencial desse tipo de regeração de dados complexos a partir de fontes de dados já disponíveis ou rápidas de se adquirir a um baixo custo. Na etapa de perfuração, por exemplo, destacam-se algumas das principais decisões em tempo real que se devem tomar com o apoio fundamental dos perfis de imagem são (TREVIZAN; MENEZES DE JESUS, 2023):

- **Isolamento de zonas produtoras ou injetoras:** zonas de produção ou injeção precisam ser isoladas adequadamente para evitar a comunicação indesejada entre elas;

- **Seleção dos pontos de cimentação:** a presença de fraturas e outras estruturas como *vugs* pode comprometer a cimentação posterior do poço, que é crucial para manter a integridade durante sua vida útil. A identificação prévia dessas estruturas, viável somente com perfis de imagem, permite a escolha de pontos de cimentação mais seguros e eficazes;
- **Necessidade de nova aquisição de dados:** em caso de falhas de aquisição em tempo real, pode ser necessário suspender a perfuração para aguardar por nova aquisição de dados com ferramentas diferentes (perfis de imagens a cabo) que podem ainda nem estar disponíveis na sonda, acarretando muitas horas ou mesmo dias de espera com elevado custo e atraso no cronograma de atividades.

As decisões acima precisam ser tomadas em tempo real, numa janela de tempo de algumas horas, mas impactam toda a vida útil do poço (que é da ordem de décadas), além de ter impacto em questões de segurança e preservação ambiental. Algumas dessas aplicações já foram alvo de estudos recentes em petrofísica:

- Orban et al. (2021) tratam das dificuldades de tomada de decisão em tempo real durante a perfuração de poços em reservatórios complexos, especificamente os carbonatos do pré-sal brasileiro, pois as ferramentas de perfilagem de imagem em tempo real (LWD, do inglês *Logging While Drilling*) não fornecem imagens de alta definição durante a perfuração. Além da baixa resolução em tempo real, outro problema é a dificuldade em compensar os erros causados pelo balanço da sonda de perfuração causado pelas ondas (efeito conhecido como *rig heave*). É proposta uma nova tecnologia de imagem de alta definição durante a perfuração, além de um algoritmo de compensação de efeitos do *heave*;
- Trevizan e MENEZES DE JESUS (2023) investigam o problema da baixa resolução das imagens de perfis de poços em tempo real obtidas durante a perfuração com foco em apoio à tomada de decisão. Essa limitação se deve à capacidade restrita de transmissão de dados em tempo real, que restringe a transmissão de imagens com apenas 50% da informação azimutal completa, especialmente para imagens resistivas. O trabalho propõe o uso de redes adversárias generativas (GANs), especificamente a arquitetura CycleGAN (ZHU et al., 2017), para reconstruir imagens de alta qualidade com dados da memória da ferramenta LWD (usualmente disponíveis alguns dias após o fim da perfuração) a partir dos dados de imagem LWD em tempo real, de menor resolução.

Nota-se que, a despeito das inovações propostas, nem todas as questões foram adequadamente endereçadas, especialmente a questão crítica de chegar a ser necessário adquirir novos dados com ferramentas diferentes.

A técnica proposta neste trabalho, pode fornecer meios de solucionar ou evitar os problemas acima mencionados. Sua principal vantagem, além da simplicidade dos algoritmos envolvidos, é ser baseada no uso de dados mais simples, os perfis básicos, que são baratos e disponíveis em tempo real. Conforme mostrado acima, os perfis básicos podem gerar imagens coerentes para a tomada de decisão, principalmente evitando a necessidade de suspensão de atividades de perfuração, que resulta em enormes custos e atrasos.

Sobre o aspecto visual das imagens regeradas pela variante BERGSOM, há três fatores que devem ser levados em consideração:

- **Nível de representação da imagem:** conforme discutido na Seção 5.2, a qualidade da imagem depende do nível de compressão imposto ao modelo β -CoVAE por meio do hiperparâmetro $\dim(\mathcal{Z})$. Observou-se que $16 \leq \dim(\mathcal{Z}) \leq 128$ regeneram imagens com qualidade aceitável;
- **Efeito de quantização vetorial obtido pela rede SOM:** similarmente ao discutido no Capítulo 4, mapas maiores tendem a fornecer imagens com menor perda de informação devida ao processo de quantização vetorial implícito à rede SOM. Isso se faz perceber quando ocorre repetição do neurônio vencedor para diferentes entradas com perfis básicos, o que leva à regeneração de trechos repetidos da imagem;
- **Procedimento de ajuste de profundidade entre perfis básicos e de imagem⁵:** o ajuste acentua o aspecto “quantizado” das imagens regeradas, uma vez que cada vetor de entrada de sete dimensões produz uma fatia de imagem de 26×180 pixels, que é empilhada verticalmente abaixo da anterior para completar toda a seção da imagem do poço.

O mesmo trecho do poço A mostrado anteriormente foi avaliado quanto às métricas AQI, mostradas na Figura 5.6.

Ao analisar visualmente os resultados mostrados na Figura 5.6, observa-se uma queda da similaridade (SSIM) com o aumento de $\dim(\mathcal{Z})$, o que não se verificou nos testes anteriores quando a métrica SSIM foi validada (Figura 5.2). Isso indica redução na qualidade da imagem gerada com o aumento da dimensão da representação utilizada para as imagens. Já se sabe que esse aumento de $\dim(\mathcal{Z})$ por si não piora a imagem, então o que deve ocorrer é uma degradação no mapeamento cruzado, similar ao relatado na Seção 4.3.4, no que diz respeito a falhas no treinamento da rede SOM devido

⁵Esse ajuste é descrito no Apêndice A.

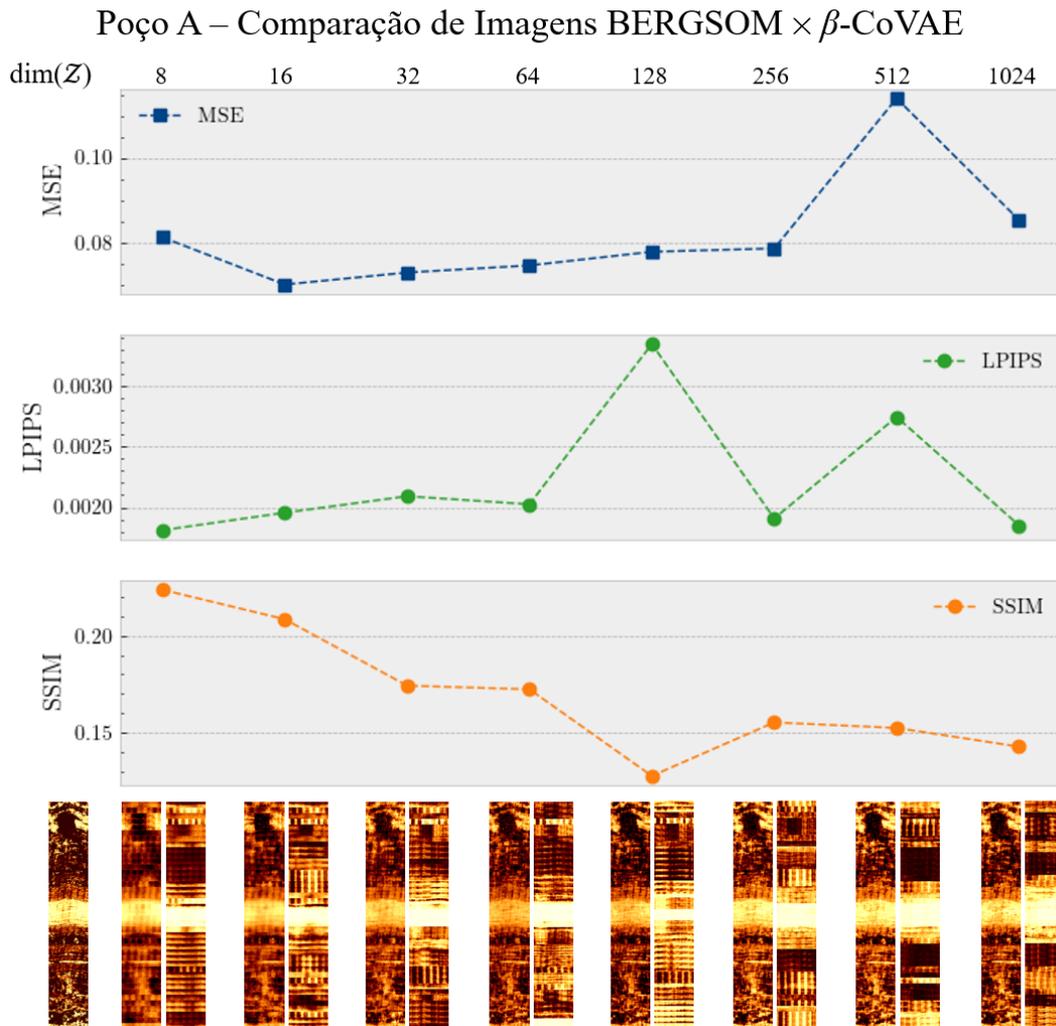


Figura 5.6: Imagens regeradas BERGSOM \times imagens β -CoVAE correspondentes. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.

à elevada dimensão dos vetores. Esta anomalia tem se mostrado pior em mapas com maior número de unidades, o que é ainda mais deletério, pois mapas grandes têm maior possibilidade de gerar menores erros de quantização e se aproveitar melhor da propriedade da rede SOM de preservação de topologia (ULTSCH; MÖRCHEN, 2005); portanto, teriam maior potencial de regeneração de imagens de melhor qualidade. Quanto às métricas MSE e LPIPS, vale destacar o comportamento estável para $16 \leq \dim(\mathcal{Z}) \leq 64$. Manter um valor relativamente baixo para $\dim(\mathcal{Z})$ favorece a metodologia proposta, por reduzir os impactos citados acima. Os resultados obtidos indicam que uso de representações mais compactas para as imagens favorecem a metodologia proposta, pois o mapeamento efetuado pela rede SOM é fortemente baseado em cálculo de distâncias, o que pode ser problemático em espaços de elevada dimensão (FRANCOIS; WERTZ; VERLEYSSEN, 2007).

Quanto à percepção visual, porém, o que mais parece afetar a qualidade

das imagens regeneradas é o problema de repetição de trechos da imagem devido à repetição de unidades vencedoras da rede SOM para trechos próximos no poço. Isso é, em geral, uma consequência da quantização vetorial inerente ao treinamento da rede SOM e tem relação com a quantidade de unidades da rede.

Para a aplicação Tipo-II, a variante BERGSOM traz certas limitações quanto à variabilidade das imagens regeneradas. Além disso, conceitualmente, é uma variante menos coerente com a natureza do problema, que é do tipo inverso e mal posto e permite relações do tipo “um para muitos”, i.e., um mesmo conjunto de perfis básicos pode estar associado a imagens diferentes. Propõe-se, então, o uso das componentes latentes probabilísticas, já usadas anteriormente na aplicação Tipo-I, para tornar a metodologia proposta mais flexível quanto à regeneração de imagens, ainda que um mesmo vencedor seja escolhido para trechos vizinhos de um poço. Para esta aplicação Tipo-II, a variante p -BERGSOM proposta adiciona ao mapeamento heteroassociativo da rede SOM expandida a capacidade de inferência variacional dos modelos β -CoVAE, ao permitir que a rede SOM incorpore as componentes latentes probabilísticas em seus pesos. Essa alteração é outra contribuição original deste trabalho e será explorada a seguir.

5.3.2

p -BERGSOM para Regeração de Perfis de Imagens

A variante p -BERGSOM da metodologia proposta foi detalhadamente definida no Capítulo 3, porém, está sumarizada a seguir no contexto da aplicação Tipo-II como Algoritmo 4.

A Figura 5.7 ilustra o uso da variante p -BERGSOM da metodologia proposta para resolver a aplicação do Tipo-II. Observa-se, em comparação com a Figura 5.4, que são usadas cores diferentes para ilustrar a rede SOM expandida já treinada. Isso indica que cada variante tem uma rede SOM diferente na aplicação do Tipo-II. A variante p -BERGSOM é treinada com uma entrada aumentada definida pela Equação 3-10. Nesse caso, um vetor protótipo \mathbf{w}_i codifica três informações ($\mathbf{w}_i^{\mu_z}$, $\mathbf{w}_i^{\sigma_z}$, \mathbf{w}_i^y). Portanto, ao contrário da variante BERGSOM, não se podem usar as mesmas redes SOM treinadas para os experimentos do Capítulo 4 e somente inverter o sentido do passo de inferência cruzada intermodalidades, pois os protótipos BERGSOM codificam somente informações não probabilísticas (\mathbf{w}_i^z , \mathbf{w}_i^y).

Além da grande capacidade de regeneração, o uso de componentes latentes probabilísticas permite à variante p -BERGSOM representar as imagens na rede SOM expandida por meio de $\mathbf{w}_i^{\mu_i}$ e $\mathbf{w}_i^{\sigma_z}$ e, portanto, regerar um número

Algoritmo 5: Variante p -BERGSOM para aplicação Tipo-II

-
- 1 Tomar conjunto de treinamento com dados de imagens
 $X = \{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ e de perfis básicos $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_p\}$;
 - 2 **Passo 1:** treinar um modelo β -CoVAE com dados de X e gerar conjunto de representações probabilísticas
 $Z^{(\mu, \sigma)} = \{\mathbf{z}_1^{(\mu, \sigma)}, \dots, \mathbf{z}_p^{(\mu, \sigma)}\}$;
 - 3 **Passo 2:** usar conjuntos $Z^{(\mu, \sigma)}$ (gerado no Passo 1) e Y para treinar uma rede SOM expandida usando Equação 3-10;
 - 4 **Inferência:** tomar conjunto de teste com dados de perfis básicos
 $Y^{teste} = \{\mathbf{y}_1^{teste}, \dots, \mathbf{y}_q^{teste}\}$;
 - 5 **para** $t \leftarrow 1$ **até** q **faça**
 - 6 Selecionar \mathbf{y}_t^{teste} ;
 - 7 Apresentar \mathbf{y}_t^{teste} à rede SOM expandida (treinada no Passo 2) para definir unidade vencedora i^* , usando Equação 3-19;
 - 8 Extrair as componentes probabilísticas do protótipo \mathbf{w}_{i^*} , $\mathbf{w}_{i^*}^{\mu_z}$ e $\mathbf{w}_{i^*}^{\sigma_z}$, usando as Equações 3-22 e 3-23, respectivamente;
 - 9 Definir a representação \mathbf{z}_t^{teste} usando o truque da reparametrização (Equação 3-24);
 - 10 Reegerar a previsão: $\hat{\mathbf{x}}_t \leftarrow G_z(\mathbf{z}_t^{teste}) = dec(\mathbf{z}_t^{teste})$, na qual $dec(\cdot)$ representa o decodificador do modelo β -CoVAE treinado no Passo 1;
 - 11 Avaliar variante p -BERGSOM, por meio de métricas IQA usando como referência as imagens correspondentes geradas pelos modelos β -CoVAE alimentados diretamente com imagens reais
 $X^{teste} = \{\mathbf{x}_1^{teste}, \dots, \mathbf{x}_q^{teste}\}$.
-

arbitrário (N) de imagens a partir de um mesmo vetor de entrada de 7 dimensões. Isso possibilita gerar diferentes imagens para poços com o mesmo conjunto de perfis básicos, o que é desejável nesse tipo de problema que permite relações “um para muitos”. Essa característica da variante p -BERGSOM enriquece as possibilidades de análise para um profissional petrofísico.

Para fins de comparação, repete-se o processo de regeneração de imagens para o mesmo intervalo do poço A que vem sendo objeto dos exemplos neste capítulo. A Figura 5.8 traz as imagens regeradas pela variante p -BERGSOM, que usa a representação probabilística $\mathbf{z}(\mu_z, \sigma_z)$ no treinamento da rede SOM. Percebe-se que ainda ocorre repetição de vencedores, porém, nesse caso as imagens geradas a partir das representações reamostradas usando a Equação 3-24 para um mesmo vencedor são sempre diferentes.

Ainda sobre a Figura 5.8, convém dar destaque ao desempenho da imagem regerada com $dim(\mathcal{Z}) = 16$, que foi capaz de aproximar a estrutura central no topo do trecho mostrado, o que só foi observada com o uso de dimensões maiores nas demais imagens.

Para mostrar mais claramente a diferença entre as imagens regeradas por

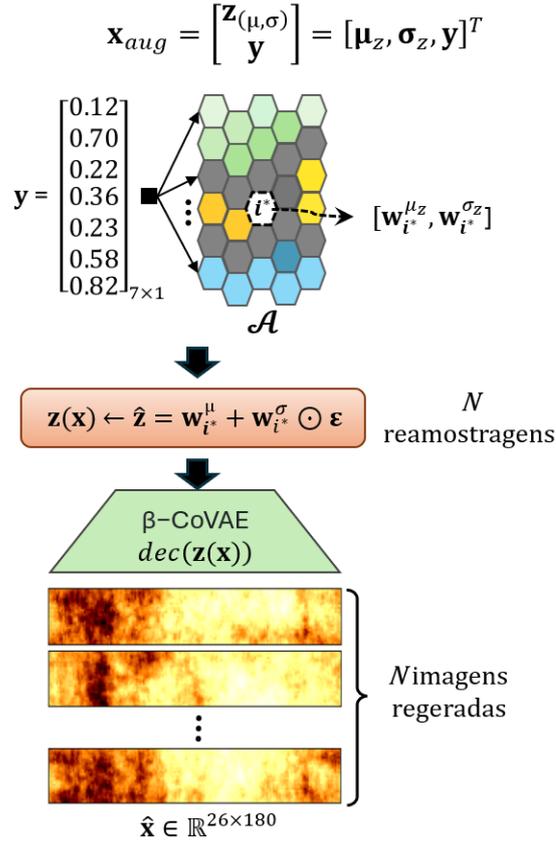


Figura 5.7: Aplicação Tipo-II com a variante p -BERGSOM.

cada uma das variantes, tomam-se duas imagens regeradas para $dim(\mathcal{Z}) = 32$ em ambas as variantes. As setas vermelhas na Figura 5.9 mostram partes em que uma mesma unidade vencedora da rede SOM em trechos consecutivos do poço levou à regeneração de segmentos repetidos que compõem o trecho completo do perfil de imagem acústica.

Observa-se nas porções em destaque da Figura 5.9 que a imagem regerada pela variante BERGSOM contém casos de repetição exata de segmentos quando a mesma unidade vencedora na rede SOM é escolhida para entradas diferentes de perfis. Já na imagem regerada pela variante p -BERGSOM, os segmentos, ainda que parecidos, jamais são exatamente iguais para uma mesma unidade vencedora selecionada.

A Figura 5.10 mostra as métricas AQI de desempenho para a variante p -BERGSOM. As métricas são calculadas comparando cada imagem p -BERGSOM com a respectiva imagem obtida diretamente do modelo β -CoVAE segundo o valor de $dim(\mathcal{Z})$. A Figura 5.11 mostra a diferentes métricas AQI de ambas as variantes em um mesmo gráfico, para se ter uma melhor ideia comparativa entre as variantes da metodologia proposta.

A Figura 5.11 traz as métricas AQI calculadas para ambas as metodologias no trecho do poço A que vem sendo usado nos exemplos deste capítulo.

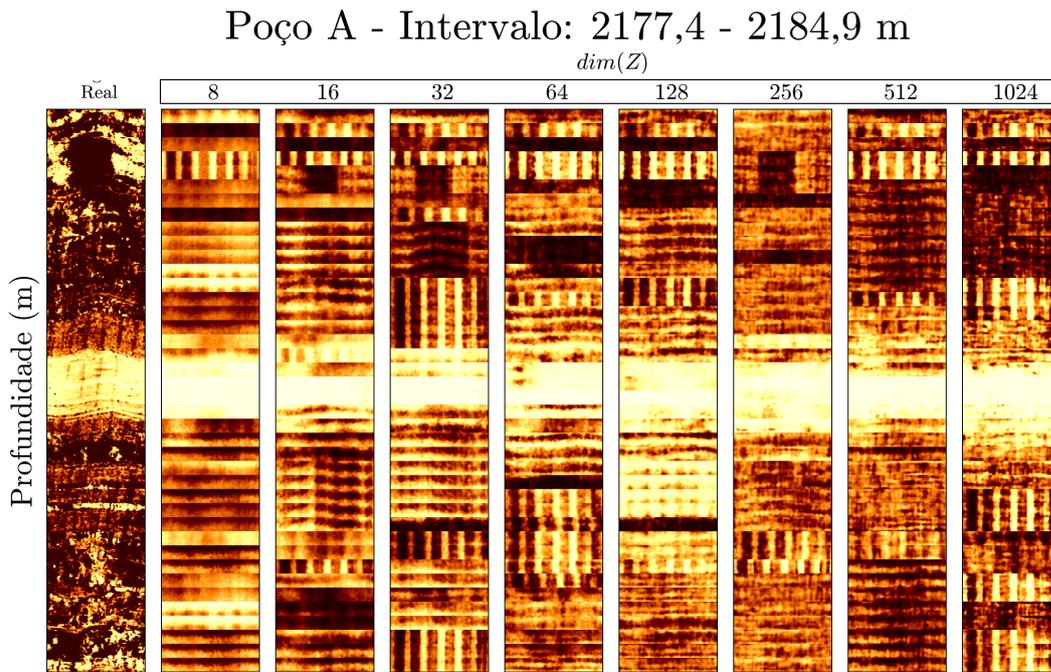


Figura 5.8: Perfis de imagem acústica regeneradas pela variante p -BERGSOM usando sete perfis básicos de entrada. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.

Observa-se ligeira melhora na métrica SSIM para a variante BERGSOM em comparação à variante p -BERGSOM.

Apesar das vantagens oferecidas pela variante p -BERGSOM para a aplicação Tipo-II, ela tem a desvantagem de duplicar a dimensão da representação da imagem $\mathbf{z}_{(\mu,\sigma)}$ usada para compor a entrada aumentada \mathbf{x}_{aug} , o que, como já vem sendo alertado ocasionalmente, pode trazer problemas no cálculo de normas em elevadas dimensões.

5.3.3

Comparativo entre as variantes: tamanho da rede SOM

Além do hiperparâmetro relacionado ao nível de compressão da imagem, é necessário avaliar o impacto do tamanho da rede SOM nas imagens regeneradas pela metodologia em suas duas variantes.

Três tamanhos de rede SOM são testados, arbitrariamente denominados PEQUENO ($10 \times 10 = 100$ unidades), MÉDIO ($30 \times 50 = 1500$ unidades) e GRANDE ($80 \times 90 = 7200$ unidades), para ilustrar o efeito desse hiperparâmetro na aplicação Tipo-II. O valor do hiperparâmetro $dim(\mathcal{Z})$ é fixado em 16 para os experimentos a seguir.

O experimento de comparação das variantes BERGSOM e p -BERGSOM para a aplicação Tipo-II variando-se o tamanho da rede SOM é mostrado na Figura 5.12:

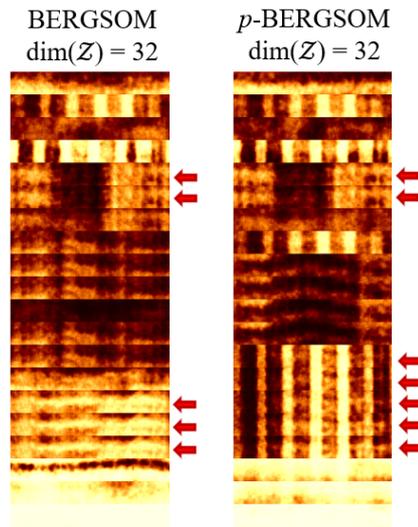


Figura 5.9: Detalhamento do efeito de variação de imagens regeradas pelos mesmos vencedores (BERGSOM \times p -BERGSOM). Rede SOM com 80×90 unidades.

A Figura 5.12 mostra a imagem real (faixa mais à esquerda, a partir de agora referida como faixa 1) e a imagem de referência correspondente (faixa 2), seguida das imagens regeradas obtidas pela variante BERGSOM (faixas 3 a 5). Os valores de todas as imagens são reescalados para o intervalo $[0,1]$. A imagem β -CoVAE de referência usada para calcular as métricas de AQI é mostrada na faixa 2. Pode-se perceber que, para a aplicação Tipo-II, aumentar o tamanho da rede SOM na variante BERGSOM beneficia a qualidade das imagens regeradas. Mapas maiores permitem um mapeamento com menor perda devido à quantização vetorial e, portanto, uma representação mais detalhada do espaço de entrada é codificada nos protótipos da rede SOM. Ao contrário do que se verificou nas aplicações do Tipo-I, a predominância da informação da imagem na entrada aumentada \mathbf{x}_{aug} é benéfica para a regeneração de imagens. Outrossim, pode-se inferir a partir da Figura 5.12 o quão desafiador é definir a melhor métrica de AQI para perfis de imagem, pois essas métricas podem não concordar entre si. Para este intervalo do poço C, o mapa de tamanho médio proporciona os melhores valores de métricas de AQI. Percebe-se que, em geral, as imagens geradas pela variante BERGSOM sofrem mais com a degradação visual causada pela repetição de unidades vencedoras em trechos próximos, em comparação com as imagens regeradas pela variante p -BERGSOM. As imagens p -BERGSOM têm um aspecto mais suave e parecem ser mais sensíveis a mudanças de amplitude em regiões maiores, em detrimento dessas variações em regiões menores. Considerando esses comportamentos diferentes das imagens regeradas por BERGSOM e p -BERGSOM, e dado que em cenários de aplicação real não é possível calcular métricas de AQI, recomenda-se ao petrofísico usar

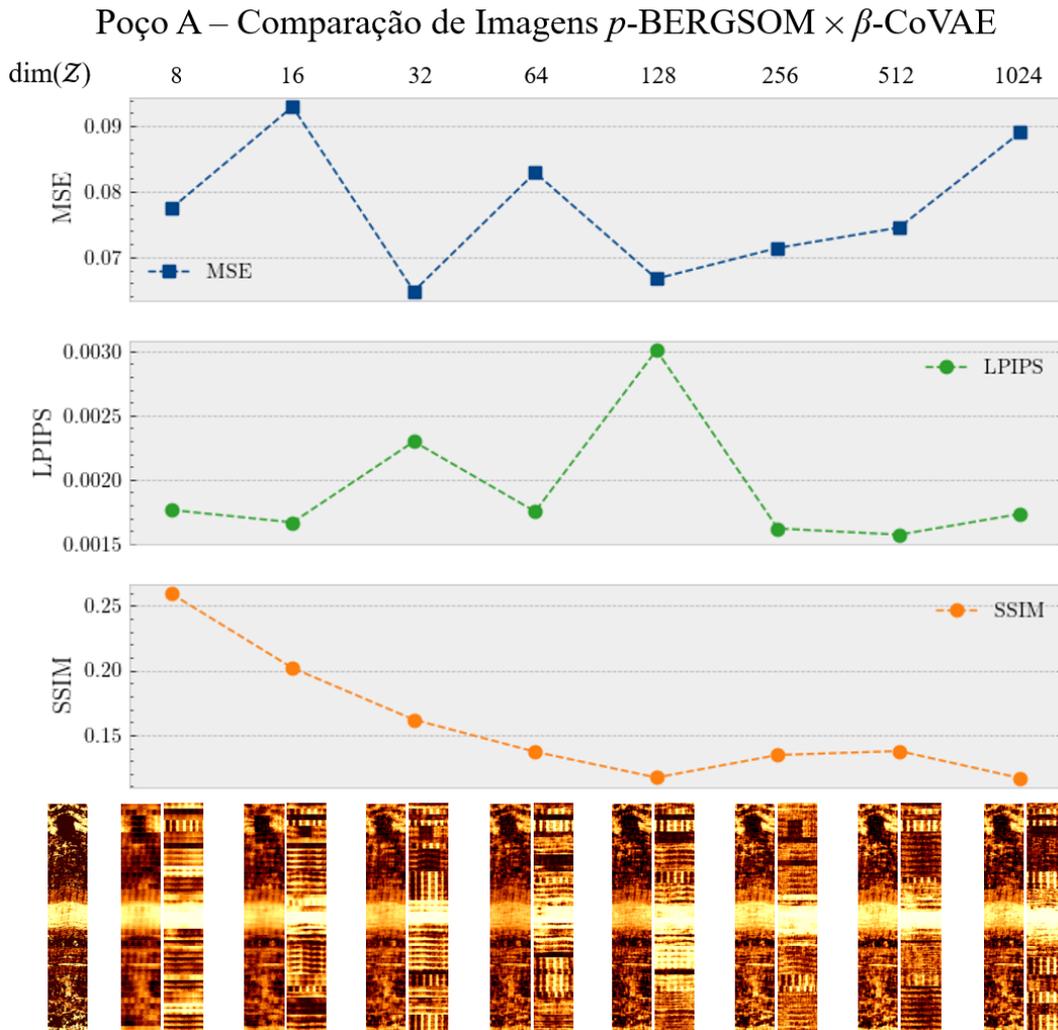


Figura 5.10: Desempenho do modelo: imagens regeradas \times imagens β -CoVAE correspondente (p -BERGSOM). Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.

ambos as variantes após definir um tamanho de mapa apropriado. Um resumo das métricas de AQI calculadas para o comprimento total todos os poços de teste de acordo com os cinco conjuntos de validação é mostrado na Tabela 5.1.

Os valores mostrados na Tabela 5.1 corroboram as impressões visuais discutidas ao analisar o intervalo do poço A apresentada na Figuras 5.12. O melhor desempenho, de acordo com as métricas MSE e LPIPS, é alcançado para mapas de tamanho pequeno e médio, enquanto a métrica SSIM favorece mapas médios e grandes. A comparação entre as variantes BERGSOM e p -BERGSOM continua sendo difícil, pois o melhor desempenho varia de acordo com o poço de teste e a métrica de AQI considerada. Isso reforça a recomendação de usar ambas as variantes para aplicações do mundo real. Cabe ressaltar que a variante p -BERGSOM pode gerar várias imagens para a mesma entrada, enriquecendo ainda mais o repertório de imagens possíveis

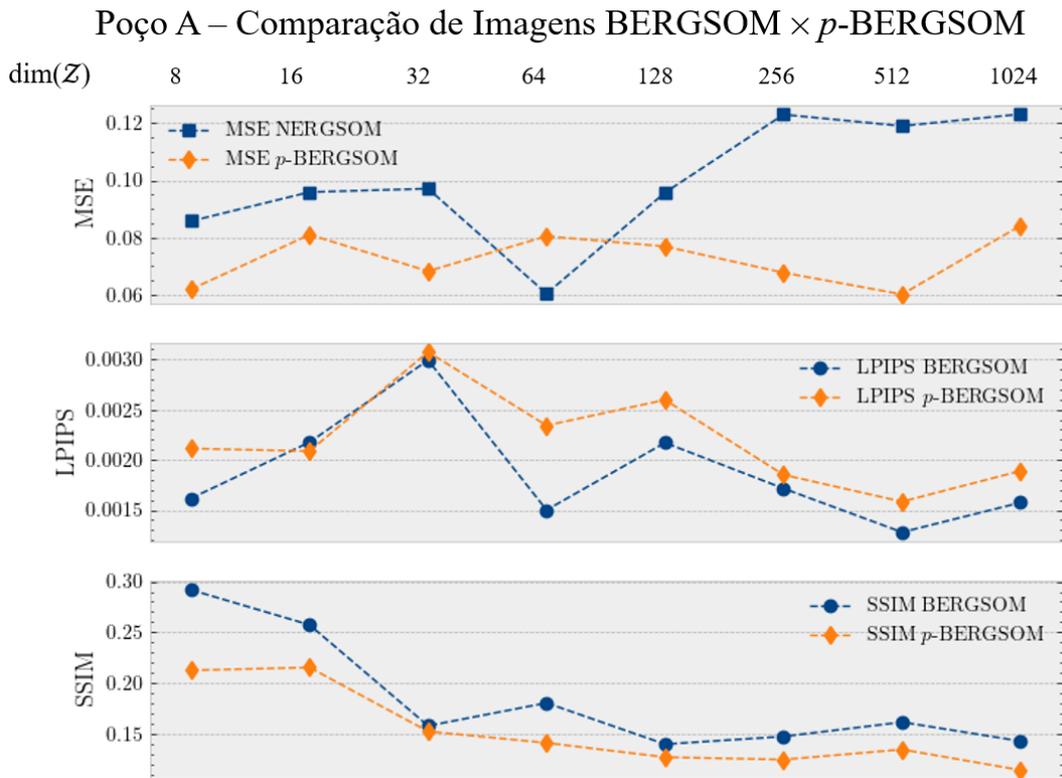


Figura 5.11: Comparação de métricas AQI - BERGSOM \times p -BERGSOM. Intervalo 2177,4 - 2184,9 m do poço A. Rede SOM com 80×90 unidades.

para análise do petrofísico.

A seguir, será apresentada a última contribuição deste capítulo, que agrega enorme valor ao método proposto, pois oferece grande potencial de regeneração de dados, com máximo aproveitamento de informação de dados complexos na indústria em geral.

5.3.4

Regeração de dados complexos a partir de entradas parciais

A presente metodologia para regeneração de dados usando a rede SOM expandida permite um tipo de inferência que não é possível em outros métodos, especialmente aqueles baseados somente em aprendizado supervisionado. Em tais métodos, a mesma composição de entrada empregada no treinamento (dimensão e tipo de entrada) deve ser usada no momento de empregar a rede em uma aplicação real. Por exemplo, se uma rede neural é treinada com exemplos de entrada compostos por sete componentes (\mathbb{R}^7) para gerar uma determinada saída, a rede não funcionará se for apresentada a ela uma entrada de teste com um número de componentes diferente de sete.

Por sua vez, a presente metodologia, uma vez treinada uma rede SOM com \mathbf{x}_{aug} , seja definido pela Equação 3-4 (variante BERGSOM) ou pela

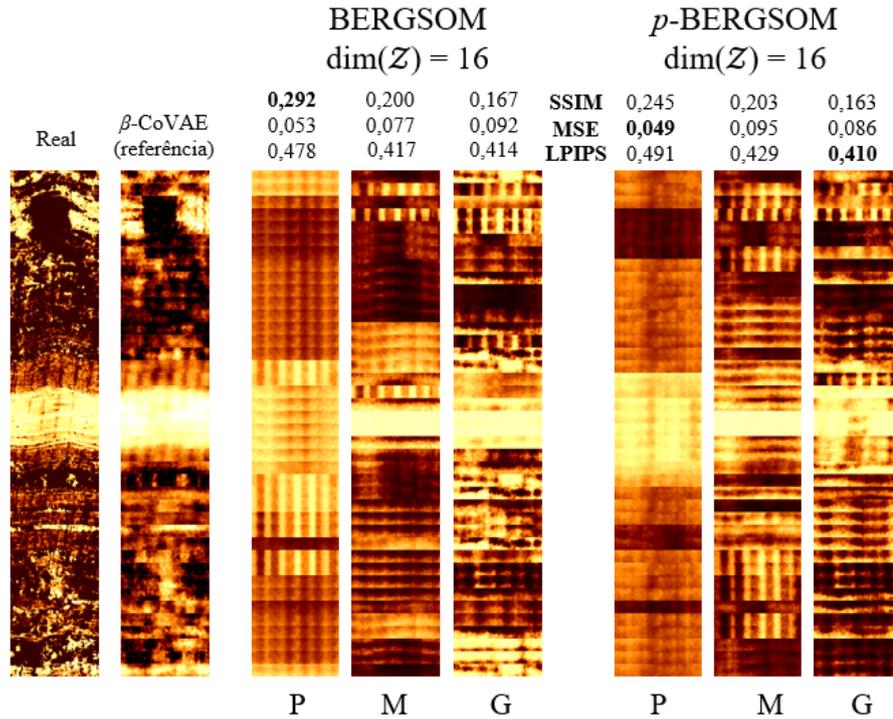


Figura 5.12: Resultado típico da variante BERGSOM para a aplicação Tipo-II. Intervalo 2177,4 - 2184,9 m do poço A. Imagem real e imagem β -CoVAE de referência à esquerda são comparadas com as imagens BERGSOM variando-se o tamanho da rede SOM (com $\dim(\mathcal{Z}) = 16$).

Tabela 5.1: Resultados para a aplicação Tipo-II com $\dim(\mathcal{Z}) = 16$. As setas indicam a direção de melhor desempenho para cada métrica de AQL. Os melhores valores gerais de métrica para cada conjunto estão destacados. Legenda do tamanho da rede SOM utilizada: ‘P’ para mapa pequeno, ‘M’ para mapa médio e ‘G’ para mapa grande.

Poço	Tamanho Rede SOM	BERGSOM			p -BERGSOM		
		MSE↓	LPIPS↓	ssim↑	MSE↓	LPIPS↓	ssim↑
A	P	0,056	0,082	0,093	0,076	0,083	0,090
	M	0,455	0,492	0,515	0,439	0,498	0,506
	G	0,471	0,398	0,335	0,427	0,413	0,327
B	P	0,041	0,054	0,068	0,050	0,057	0,065
	M	0,327	0,410	0,432	0,353	0,403	0,440
	G	0,428	0,413	0,356	0,484	0,423	0,348
C	P	0,070	0,084	0,094	0,071	0,088	0,101
	M	0,441	0,468	0,479	0,399	0,456	0,458
	G	0,438	0,361	0,320	0,440	0,373	0,314
D	P	0,072	0,097	0,097	0,084	0,096	0,093
	M	0,453	0,494	0,483	0,413	0,499	0,496
	G	0,397	0,368	0,325	0,417	0,370	0,321
E	P	0,050	0,091	0,074	0,084	0,094	0,084
	M	0,458	0,497	0,465	0,370	0,491	0,428
	G	0,337	0,343	0,279	0,393	0,303	0,262

Equação 3-10 (variante p -BERGSOM), permite recuperar a informação do vencedor usando qualquer subconjunto de perfis básicos na composição do vetor \mathbf{y} de entrada para a etapa de inferência cruzada intermodalidades.

Vale observar que trabalhos anteriores, que exploraram a ideia da rede SOM expandida para aplicações de controle e identificação de sistemas dinâmicos (BARRETO; ARAUJO, 2004; BARRETO; ARAUJO; RITTER, 2003) não mencionaram essa possibilidade.

Uma ideia similar, porém de escopo mais limitado, tem sido empregada há muitos anos para a imputação de valores faltantes usando SOM (SAMAD; HARP, 1992; FESSANT; MIDENET, 2002; WANG, 2003). Em tais aplicações, o método mais comum consiste em treinar a rede SOM com um conjunto de dados sem valores faltantes. Depois, no processo de imputação (inferência), a entrada apresentada à rede SOM para a escolha do vencedor contém apenas uma parte das componentes de treinamento (como na etapa de inferência da metodologia ora proposta). Portanto, a distância entre as unidades da rede SOM e a entrada com valores faltantes é calculada apenas usando as componentes correspondentes aos dados presentes nos pesos da rede. Assim, as componentes dos pesos do vencedor relacionadas aos dados faltantes servem de estimativa para aqueles valores. Técnicas mais recentes incorporam a imputação dos valores faltantes ainda durante o treinamento da rede SOM, o que pode melhorar a precisão para a imputação de valores faltantes na base de teste (REJEB; DUVEAU; REBAFKA, 2022; VATANEN et al., 2015). Todas as ideias acima baseiam-se em implicitamente imaginar o dado de entrada da rede SOM como uma espécie de entrada aumentada \mathbf{x}_{aug} composta pelos valores potencialmente faltantes $\mathbf{x}_{faltante}$ e valores presentes $\mathbf{x}_{presente}$:

$$\mathbf{x}_{aug} = \begin{bmatrix} \mathbf{x}_{presente} \\ \mathbf{x}_{faltante} \end{bmatrix}. \quad (5-4)$$

Ao verificar a Equação 5-4 e sua semelhança conceitual, por exemplo, com as Equações 3-4 ou 3-10, verifica-se que o uso da rede SOM para a imputação de valores faltantes pode ser visto como um caso particular da metodologia de inferência cruzada proposta neste trabalho. Nesse caso, um protótipo da rede SOM teria componentes correspondentes:

$$\mathbf{w}_i = \begin{bmatrix} \mathbf{w}_i^{presente} \\ \mathbf{w}_i^{faltante} \end{bmatrix}. \quad (5-5)$$

A metodologia aqui desenvolvida, porém, vai além: não só permite a imputação de valores faltantes (inclusive perfis básicos inteiros), mas também gerar imagens acústicas a partir de qualquer subconjunto de perfis básicos, no limite até mesmo usando uma única curva. Nesse sentido, no caso específico

da aplicação do Tipo-II, a metodologia permite:

- A **regeração** perfis de imagem a partir de um subconjunto qualquer de perfis básicos de entrada;
- Ao mesmo tempo, **predizer** valores para os perfis básicos faltantes para aquele poço.

A utilidade de tal propriedade é evidente: por exemplo, pode-se treinar uma rede SOM expandida usando dados de poços de um campo nos quais a maior quantidade possível de perfis (básicos ou especiais, como os de imagem) tenham sido adquiridos. A seguir, a rede SOM expandida pode ser usada para estimar os perfis existentes na base de poços de treinamento nos demais poços daquele campo que não tenham esses dados. A regeração usando como entrada somente uma parcela dos perfis usados no treinamento (que são normalmente aqueles com menor custo de aquisição) apresenta-se, portanto, como uma forma nova e eficaz de aproveitar ao máximo os custosos dados adquiridos em poços selecionados de um campo e fornecer boas estimativas desses dados aos demais poços.

A seguir, detalham-se os passos específicos para o uso da metodologia proposta para regerar um perfil de imagem a partir de um subconjunto de perfis básicos (seja na variante BERGSOM ou p -BERGSOM):

- Treinar a rede SOM expandida com o maior número possível de perfis básicos em \mathbf{y} para compor \mathbf{x}_{aug} , seja usando a Equação 3-4 (BERGSOM) ou 3-10 (p -BERGSOM). O vetor \mathbf{y} será, portanto, definido como:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_p \\ \mathbf{y}_f \end{bmatrix}, \quad (5-6)$$

na qual o subíndice p significa presente e o subíndice f significa faltante;

- Apresentar à rede SOM expandida apenas o vetor de perfis básicos presentes \mathbf{y}_p (um subconjunto dos perfis básicos usadas no treinamento da rede SOM) como entrada de teste;
- Encontrar o vencedor i^* usando somente a componente $\mathbf{w}_i^{y_p}$ correspondente à entrada de teste \mathbf{y}_p :

$$i^* = \arg \min_{i \in \mathcal{A}} \|\mathbf{y}_p - \mathbf{w}_i^{y_p}\|; \quad (5-7)$$

- A partir do protótipo do vencedor \mathbf{w}_{i^*} , recuperar a componente correspondente da representação da imagem, $\mathbf{w}_{i^*}^z$:
 - **BERGSOM**: extrair do protótipo vencedor a componente relativa à imagem usando usando a Equação 3-9;

- **p -BERGSOM:** extrair as componentes latentes probabilísticas $\mathbf{w}_{i^*}^{\mu_z}$ e $\mathbf{w}_{i^*}^{\sigma_z}$ do protótipo vencedor, usando as Equações 3-22 e 3-23.
- Obter uma representação do dado de origem a partir das componentes acima:
 - **BERGSOM:** $\mathbf{z}(\mathbf{y}_p) \leftarrow \mathbf{w}_{i^*}^z$;
 - **p -BERGSOM:** amostrar uma representação da imagem $\hat{\mathbf{z}}$ a partir das componentes probabilísticas $\mathbf{w}_{i^*}^{\mu_z}$ e $\mathbf{w}_{i^*}^{\sigma_z}$ usando o truque da reparametrização (Equação 3-24): $\mathbf{z}(\mathbf{y}_p) \leftarrow \hat{\mathbf{z}}$.
- A partir de $\mathbf{z}(\mathbf{y}_p)$, regerar a imagem correspondente usando uma função $G_z(\mathbf{y}_p)$ que é parametrizada como o decodificador $dec(\cdot)$ do modelo β -CoVAE adequado: $\hat{\mathbf{x}} \leftarrow G_z(\mathbf{z}(\mathbf{y}_p)) = dec(\mathbf{z}(\mathbf{y}_p))$.

Para ilustrar essa aplicação, será usado o modelo p -BERGSOM com $dim(\mathcal{Z}) = 16$, por ser um dos mais rápidos e porque a imagem correspondente a essa dimensão mostrada na Figura 5.8 foi considerada satisfatória. Foram compostos sete subconjuntos de perfis básicos para o experimento, conforme a Tabela 5.2. Os subconjuntos de perfis básicos (exceto o *Triple Combo*) foram selecionados com base no desempenho obtido nos diferentes poços na aplicação do Tipo-I, vai-se excluindo gradualmente o perfil que obteve em média a menor correlação na predição de perfis obtidos pela metodologia proposta naquela aplicação.

Tabela 5.2: Composição de subconjuntos de perfis básicos.

Subconjunto	Perfis presentes
1C	DEN
2C	DEN, NEU
3C	DEN, NEU, NMR PhitT
4C	DEN, NEU, DTC, NMR PhitT
<i>Triple Combo</i>	GR, DEN, NEU, $\log(RES)$
5C	DEN, NEU, DTC, NMR PhitT, PE
6C	GR, DEN, NEU, DTC, NMR PhitT, PE
7C	Todos

A Figura 5.13 mostra a regeneração de imagem usando subconjuntos de perfis básicos conforme a Tabela 5.2. Destaca-se o termo *Triple Combo* que é usado para ilustrar o conjunto dos quatro perfis básicos mais comuns, GR, DEN, NEU e $\log(RES)$ (SIMÕES et al., 2023), obtidos por ferramentas de perfilagem sensíveis a três princípios físicos: densidade, porosidade e resistividade (daí o termo *triple*).

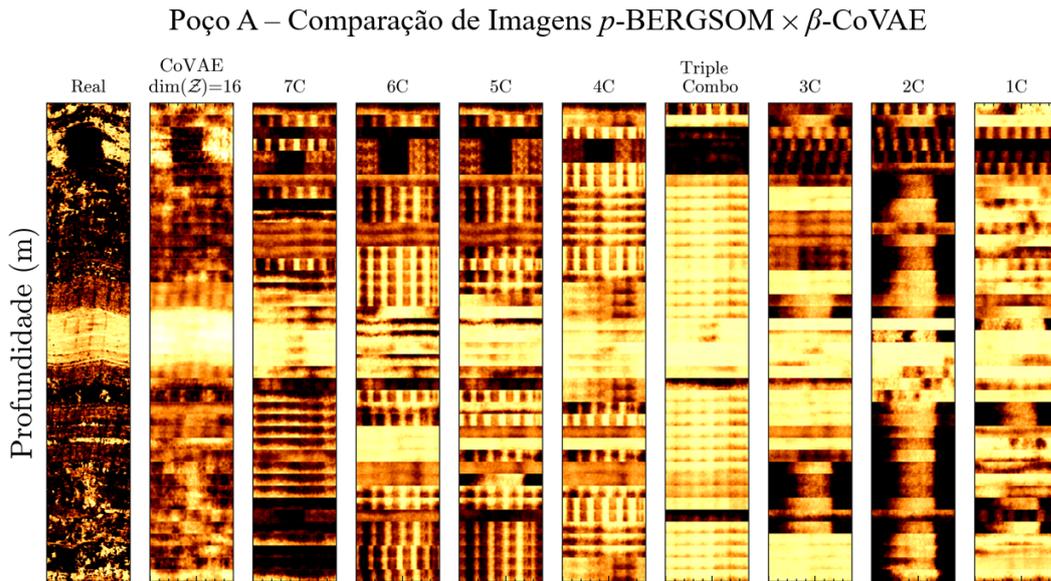


Figura 5.13: Perfis de imagem acústica regeneradas pela variante p -BERGSOM usando subconjuntos de perfis de entrada. Intervalo 2177,4 - 2184,9 m do poço A.

Com base na Figura 5.13, é possível atestar o potencial de recuperação da informação mais importante da imagem mesmo com um número reduzido de perfis. Visualmente, pelo menos até o subconjunto 4C, estruturas maiores da imagem continuam a ser regeneradas, como a estrutura escura no topo. Contudo, a partir somente três perfis, o perfil de imagem regenerado começa a mostrar estruturas não existentes. O uso de uma só curva de entrada é apenas ilustrativo dos limites do método, e não se acredita que tal extremo vá trazer alguma informação utilizável, a despeito de haver, pelo menos nos níveis de alta e baixa amplitude, uma recuperação qualitativa razoável mesmo nessa previsão extrema com uma só curva. A comparação das imagens regeneradas pelos subconjuntos 7C e 6C, e pelos subconjuntos 3C, ‘Triple Combo’ e 4C, parece indicar que a inclusão do perfil $\log(RES)$ nos subconjuntos leva à regeneração de imagens piores. A análise quantitativa com as métricas que vêm sendo aplicadas, no entanto, não permite uma coerência completa com a percepção visual. Essas métricas são mostradas na Figura 5.14.

Na Figura 5.14, enquanto o MSE parece de acordo com a percepção visual de piora com a redução do número de perfis de entrada, o mesmo não ocorre com as métricas LPIPS e SSIM que chegam a ter valor máximo para a imagem regenerada no subconjunto *Triple Combo*, que tem visualmente baixa qualidade. Este é mais um experimento que reforça a necessidade de encontrar uma métrica mais confiável para a validação dos experimentos.

No cômputo geral, os resultados apresentados mostram que a metodologia proposta, dentro do escopo e objetivos da aplicação Tipo-II, é capaz de

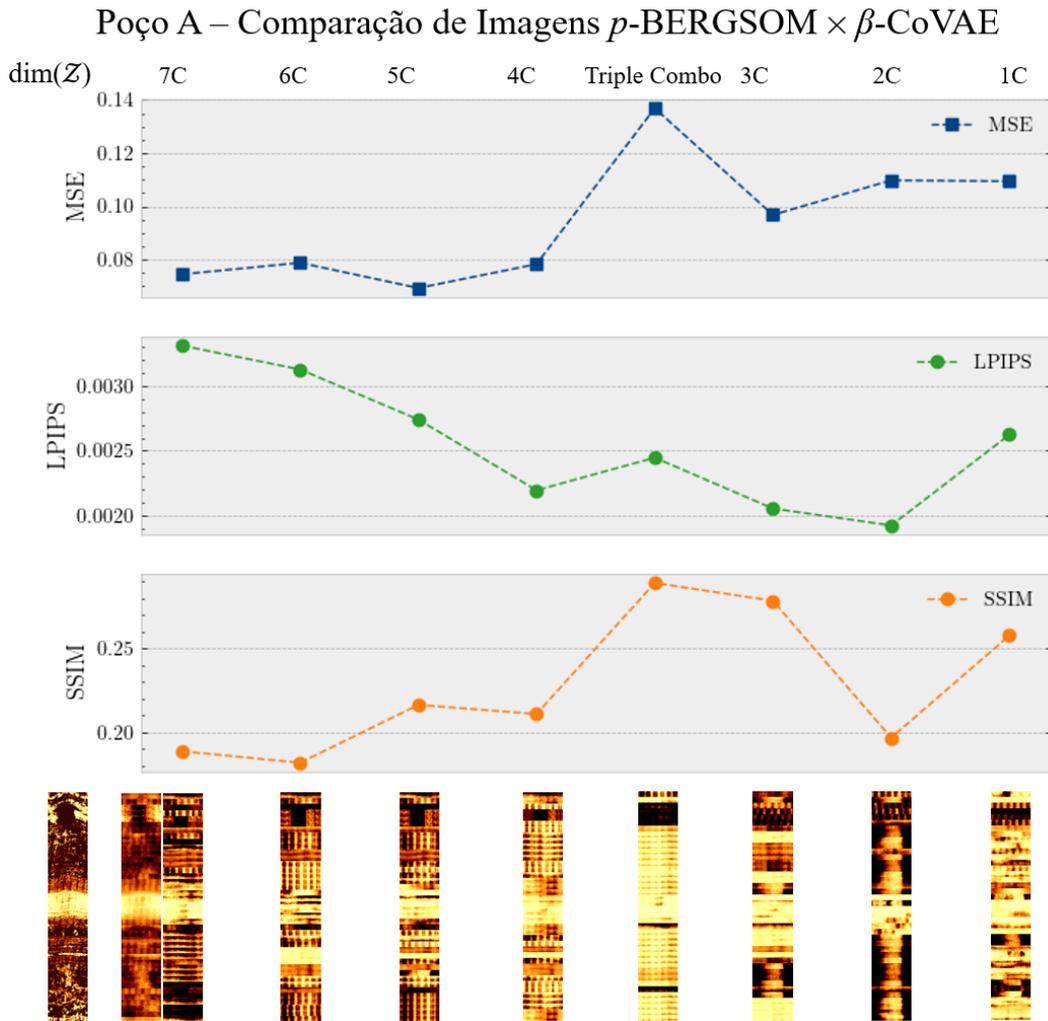


Figura 5.14: Desempenho do modelo: imagens geradas \times imagem real, conforme o Subconjunto de perfis.

solucionar satisfatoriamente o problema de regeneração de perfis de imagens a partir de perfis básicos e posiciona-se como uma técnica potencial em diversos outros campos de aplicação.

A seguir, discutem-se de forma mais ampla detalhes e limitações da metodologia proposta para a aplicação Tipo-II.

5.4

Discussão sobre a regeneração de imagens a partir de perfis básicos

Levando em conta todo o escopo da metodologia proposta, em especial a aplicação Tipo-II, algumas questões merecem uma ligeira discussão:

- **Seleção de perfis a incluir no conjunto de dados básicos que serão usados na composição de \mathbf{x}_{aug} :** esse é um tópico que sequer seria relevante caso a metodologia fosse baseada exclusivamente em em aprendizado supervisionado, pois nesse caso apenas as perfis com maior

correlação com a saída a ser prevista seriam relevantes. Porém, com a presente proposta, ocorre um mapeamento heteroassociativo entre diferentes modalidades de dados, e não é construído um modelo de regressão. A aplicação de regeneração de imagens usando subconjuntos de perfis ilustra isso, pois é possível usar muito menos perfis para uma previsão aceitável. No entanto, há de se destacar que a qualidade das imagens recuperadas no experimento com subconjuntos de perfis depende do tamanho do conjunto original de perfis usados no treinamento da rede SOM para o mapeamento heteroassociativo (em qualquer uma das variantes da metodologia). Ao adicionar mais perfis, incrementa-se a informação de contexto que ficará incorporada à rede SOM expandida. Os resultados indicam, porém, que perfis muito influenciados pelo fluido no poço (como o RES) podem confundir o mapeamento heteroassociativo, como se viu no experimento de regeneração a partir de subconjuntos de curvas bem como na pobre predição da curva $\log(RES)$ no Capítulo 4.

- **Qualidade geral das imagens regeradas pelo modelos BERGSOM e p -BERGSOM:** a depender do grau de exigência, pode-se dizer que alguns dos experimentos de regeneração de imagens pela metodologia proposta apresentaram baixa qualidade. Além das três razões indicadas no corpo do texto para uma menor percepção visual de qualidade das imagens, a questão da discrepância na cardinalidade dos componentes de \mathbf{x}_{aug} faz com que uma das informações (a imagem) tenha maior peso no cômputo das normas e, por conseguinte, afeta o cerne do treinamento dos modelos SOM que é a escolha dos vencedores. Essa questão, juntamente com a necessidade de manter baixa a dimensionalidade total de \mathbf{x}_{aug} , é a principal limitação da metodologia baseada em mapeamento heteroassociativo. Para a aplicação Tipo-II, observou-se melhoria das imagens com o aumento da rede SOM, em oposição à aplicação Tipo-I. Os comportamentos opostos têm a mesma explicação: mapas maiores favorecem a representação da imagem em detrimento da informação de perfis básicos nos protótipos da rede SOM. Como a aplicação Tipo-II tem maior grau de interesse, a abordagem com mapas maiores deve ser preferida.
- **Discrepâncias entre a percepção visual e métricas AQI:** os casos citados de discrepâncias entre as métricas AQI utilizadas, todas elas bem fundamentadas na área de visão computacional, e também sua discrepância com a percepção visual para este tipo de imagem geológica, indicam a necessidade de uma métrica que, além de verificar a similaridade geral da imagem (como o SSIM), seja capaz de avaliar uma imagem acústica regerada quanto à presença de estruturas geológicas

(como fraturas) e artefatos (como *breakouts*). Portanto, este trabalho parece indicar que as métricas usuais da literatura não são as mais adequadas para atestar a qualidade de imagens acústicas de paredes de poços geradas por modelos generativos de aprendizado de máquinas. A seguir, propõem-se possíveis direcionamentos sobre métricas para avaliação das imagens regeradas pela metodologia proposta.

5.4.1

Proposta de Métricas IQA baseadas na rede SOM

As métricas utilizadas até aqui para avaliar a qualidade das imagens regeradas pela metodologia proposta são do tipo FR (“full reference”), ou seja, há sempre uma imagem de referência livre de distorções. Quando não há imagens de referência disponíveis, podem-se usar métricas do tipo NR (“no reference”).

Em que pese o fato de que na etapa de desenvolvimento e validação da metodologia essas imagens de referência (as imagens fornecidas pelos modelos β -CoVAE usando imagens reais como entrada) estejam disponíveis, isso não é possível em casos reais de aplicação, quando não será possível calcular qualquer métrica do tipo FR. Além disso, pelo menos mais duas críticas podem ser feitas para as métricas AQI médias usadas até agora:

1. Elas podem divergir da percepção visual sobre as diversas imagens regeradas;
2. Um possível intérprete petrofísico em uma aplicação real está habituado a uma medida de qualidade associada à profundidade do perfil analisado, em vez de uma medida única para todo o trecho.

Nesta seção, são propostas métricas diretamente calculadas a partir da rede SOM treinada em quaisquer das variantes, BERGSOM ou p -BERGSOM que se propõem a solucionar essas duas questões. A primeira métrica⁶, chamada de z_{IQA} é do tipo FR e a outra, chamada b_{IQA} (ou “blind” IQA) do tipo NR. O procedimento para o cálculo de cada uma delas é descrito a seguir. O Algoritmo 6 mostra os passos para cálculo da métrica z_{IQA} .

Vale ressaltar no Algoritmo 6 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 7 mostra os passos para cálculo da métrica b_{IQA} .

⁶Optou-se por usar o acrônimo IQA, no inglês *image quality assessment* para as siglas que compõem as métricas propostas.

Algoritmo 6: Cálculo da métrica z -IQA (ambas as variantes)

-
- 1 **Modelos** tomar modelos β -CoVAE e SOM treinados para regeneração de perfis imagens;
 - 2 **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, $Z^{teste} = \{\mathbf{z}_1^{teste}, \dots, \mathbf{z}_q^{teste}\}$;
 - 3 **para** $t \leftarrow 1$ **até** q **faça**
 - 4 Selecionar \mathbf{y}_t^{teste} e \mathbf{z}_t^{teste} ;
 - 5 Apresentar \mathbf{y}_t^{teste} à rede SOM expandida para definir unidade vencedora i^* , usando Equação 3-19;
 - 6 Definir a representação $\hat{\mathbf{z}}_t$ obtida pelos pesos da unidade vencedora;
 - 7 Calcular a distância quadrática entre $\hat{\mathbf{z}}_t$ e \mathbf{z}_t^{teste} :

$$z_{IQA}(\mathbf{z}_t^{teste}, \hat{\mathbf{z}}_t) = \|\mathbf{z}_t^{teste} - \hat{\mathbf{z}}_t\|_2^2.$$
-

Algoritmo 7: Cálculo da métrica b -IQA (ambas as variantes)

-
- 1 **Modelos** tomar modelos β -CoVAE e SOM treinados para regeneração de perfis imagens;
 - 2 **Dados:** tomar conjunto de teste com dados de perfis básicos $Y^{teste} = \{\mathbf{y}_1^{teste}, \dots, \mathbf{y}_q^{teste}\}$;
 - 3 **para** $t \leftarrow 1$ **até** q **faça**
 - 4 Selecionar \mathbf{y}_t^{teste} ;
 - 5 Apresentar \mathbf{y}_t^{teste} à rede SOM expandida para definir unidade vencedora i^* , usando Equação 3-19;
 - 6 Selecionar a componente $\mathbf{w}_{i^*}^y$ correspondente à informação de perfis básicos;
 - 7 Calcular a distância quadrática entre $\mathbf{w}_{i^*}^y$ e \mathbf{y}_t^{teste} :

$$b_{IQA}(\mathbf{y}_t^{teste}, \mathbf{w}_{i^*}^y) = \|\mathbf{y}_t^{teste} - \mathbf{w}_{i^*}^y\|_2^2.$$
-

Além das métricas, a Figura 5.15 traz linhas verticais verdes que são limiares de qualidade para referência. Esses limiares são calculados da seguinte forma:

- **Limiar z IQA:** aplicar o Algoritmo 6 para os dados de treinamento. O limiar será um percentil arbitrário (P_α) dos valores de z_{IQA} calculados em dados de treino;
- **Limiar y IQA:** aplicar o Algoritmo 7 para os dados de treinamento. O limiar será um percentil arbitrário (P_α) dos valores de b_{IQA} calculados em dados de treino.

Para a Figura 5.15 foi usado o P_{95} , ou seja $\alpha = 0,95$. Para o trecho mostrado, verifica-se coerência entre as métricas, com destaque para a métrica 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

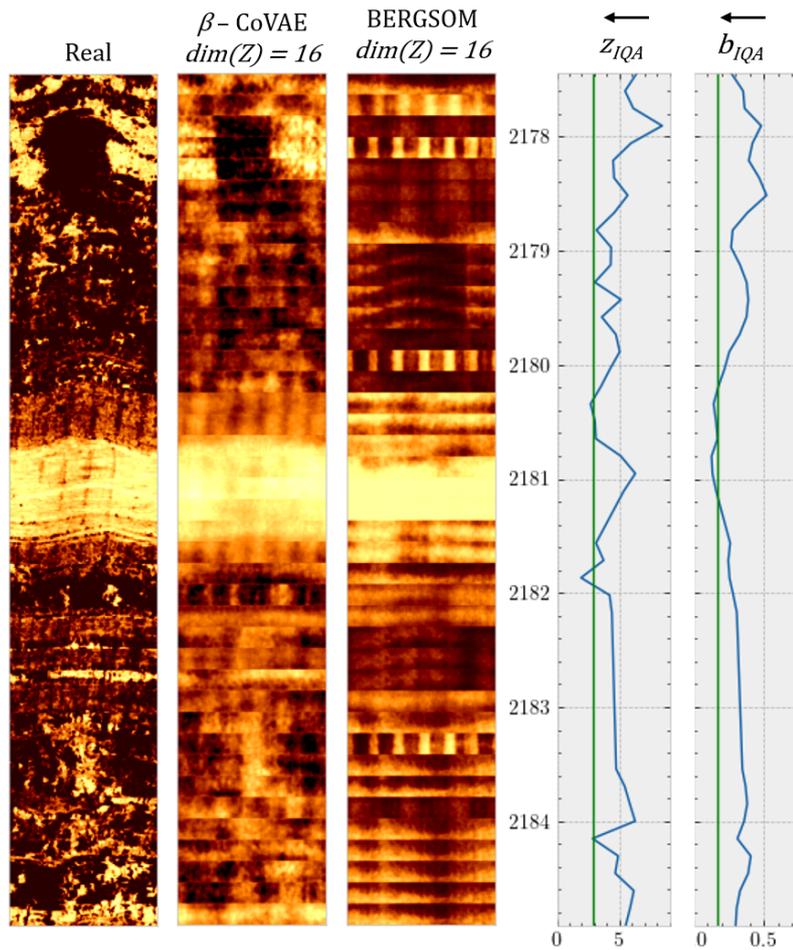


Figura 5.15: Métricas z_{IQA} e b_{IQA} no intervalo do poço A que vem sendo usado como exemplo neste capítulo. A métrica z_{IQA} é do tipo FR, enquanto a métrica b_{IQA} é do tipo NR, i.e., não depende de imagens de referência para seu cálculo.

gerada pelo modelo β -CoVAE. O mesmo não ocorreu com a métrica z_{IQA} , que aumentou em torno da profundidade 2181 m.

As métricas IQA do tipo FR usadas anteriormente também podem ser calculadas em forma de séries de valores indexados em profundidade do poço.

A Figura 5.16 mostra todas as métricas AQI utilizadas neste capítulo, 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 regeneração das imagens de 26×180 pixels, a imagem BERGSOM regenerada com a imagem correspondente gerada pelo modelo β -COVAE (no exemplo mostrado, foram regeneradas 40 imagens para a composição 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 propostas. Por exemplo, na região ao redor de 2181 m na qual as imagens

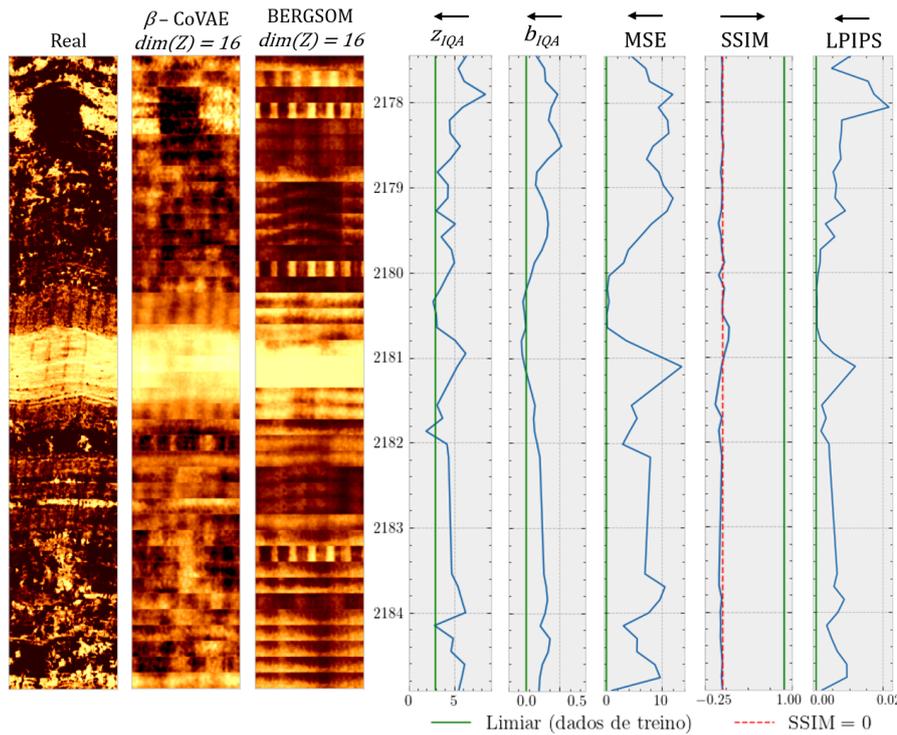


Figura 5.16: Métricas AQI utilizadas nesta tese: z_{IQA} , b_{IQA} (propostas), MSE, SSIM e LPIPS mostradas como valores indexados em profundidade no intervalo do poço A que vem sendo usado como exemplo neste capítulo. As setas indicam a direção em que cada métrica indica melhora na imagem.

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 chegou 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 geradas pelas representações usadas no cálculo da segunda.

Outro aspecto interessante a se avaliar nas diferentes métricas AQI utilizadas é seu comportamento na avaliação das imagens da Seção 5.3.4. Esse comportamento será ilustrado com base em uma série de figuras a seguir.

A Figura 5.17 mostra um resultado para o experimento de regeneração de perfis de imagens a partir de perfis básicos com aplicação da variante BERGSOM para o intervalo 2177,4 - 2184,9 m do poço A. As imagens mostradas serão usadas como base para o cálculo das métricas AQI a seguir.

Ao observar cuidadosamente as Figuras 5.18 a 5.22, percebe-se que apenas as métricas SSIM, z_{IQA} e b_{IQA} parecem ser sensíveis às variações nas imagens regeneradas pela metodologia proposta. A insensibilidade a imagens tão díspares é uma desvantagem grave para uma métrica desse tipo, principalmente

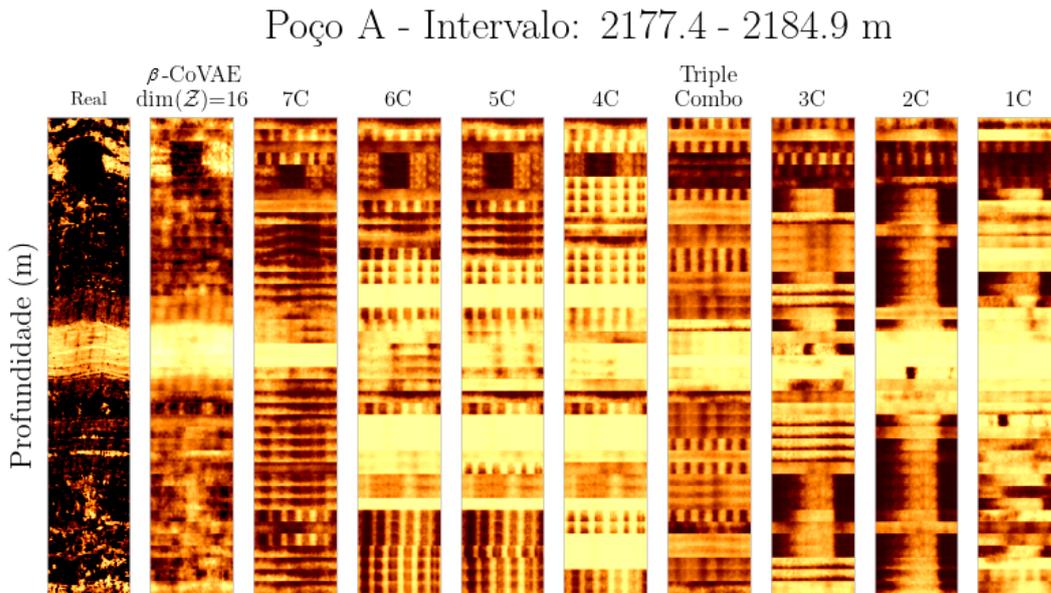


Figura 5.17: Perfis de imagem acústica geradas pela variante BERGSOM usando subconjuntos de perfis de entrada. Intervalo 2177,4 - 2184,9 m do poço A.

aquelas que exigem maiores custos computacionais. Destaca-se também que a semelhança apontada anteriormente entre as métricas MSE e z_{IQA} só parece ser válida para o subconjunto 7C, para os demais, a métrica MSE permanece semelhante enquanto a métrica z_{IQA} apresenta grandes variações, em geral coerentes com a visual degradação das imagens regeradas. Apesar de apresentar sensibilidade às variações nas imagens regeradas a partir dos diferentes subconjuntos, a métrica SSIM não variou de modo a atestar a degradação crescente nas imagens (que é visualmente perceptível pelo menos a partir do subconjunto 3C).

Diante do exposto, as análises levadas a cabo nesta seção apontam para as seguintes conclusões:

- A métrica proposta b_{IQA} traz as seguintes vantagens:
 1. é do tipo NR, o que permite seu cálculo em qualquer situação prática;
 2. é sensível às variações de imagem de maneira coerente com a degradação visual percebida (o mesmo pode ser dito da métrica z_{IQA} , com a desvantagem desta ser do tipo FR).
- Apesar de ser sensível às variações em imagens que se degradam, a métrica SSIM não se mostrou coerente no experimento de regeneração de imagens a partir de subconjuntos de perfis básicos;

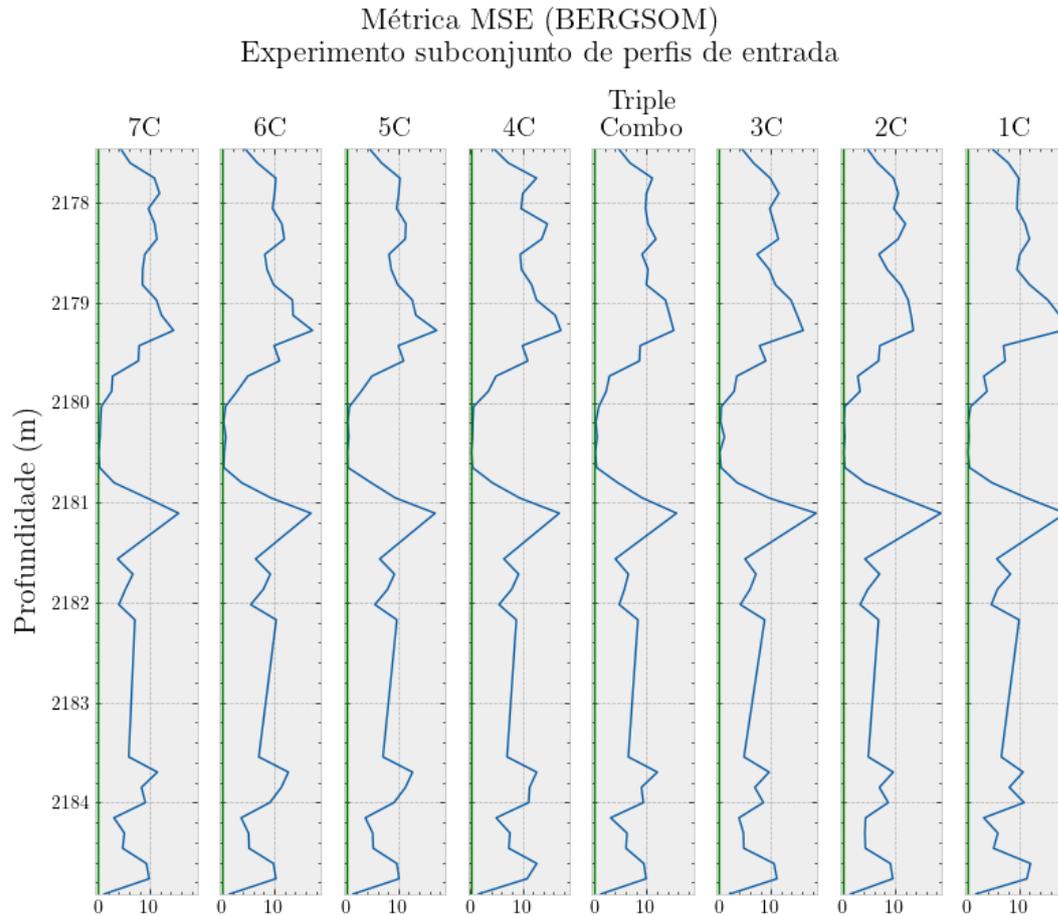


Figura 5.18: Comportamento da métrica MSE para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).

- Métricas MSE e LPIPS, a despeito da grande diferença em termos de custo computacional, apresentaram comportamentos semelhantes.

A partir do exposto, parece haver elementos para propor a métrica b_{IQA} como a mais adequada dentre as avaliadas para esse tipo de imagem.

5.5

Considerações finais

Este capítulo abordou a aplicação Tipo-II no campo da petrofísica de reservatórios, i.e., a regeneração de perfis de imagem acústica a partir de um conjunto abrangente de perfis básicos. Foram apresentados e discutidos os resultados do emprego das variantes BERGSOM e p -BERGSOM para essa aplicação mais complexa e que pode ser classificada como aprendizado regenerativo.

Destacou-se também o fato de a metodologia proposta ser uma generalização da técnica de imputação de valores faltantes utilizando a rede SOM. A metodologia, portanto, permite a regeneração de perfis de imagem acústica a par-

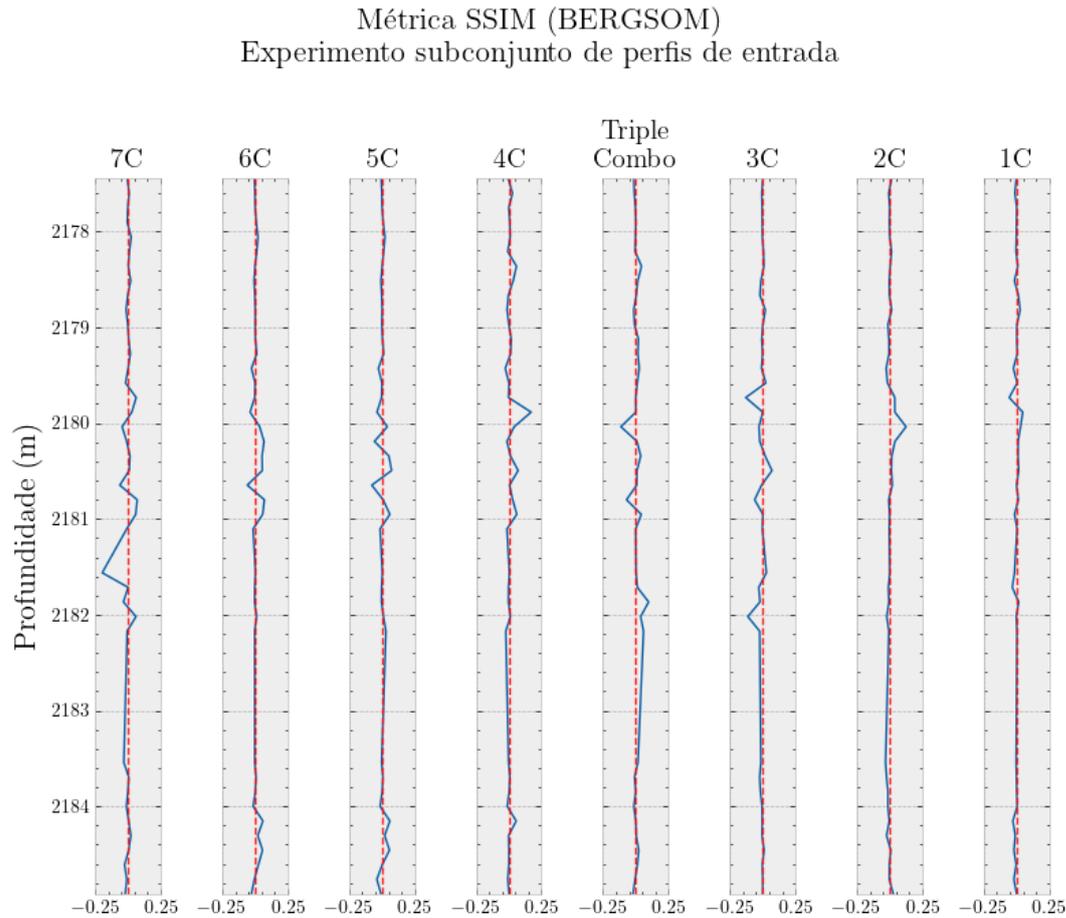


Figura 5.19: Comportamento da métrica SSIM para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).

tir de subconjuntos que contêm apenas uma parte dos perfis básicos originais empregados no treinamento da rede SOM. Tal capacidade de trabalhar com dados incompletos é uma característica valiosa para estudos de reservatórios, na qual a quantidade de dados disponíveis pode ser limitada ou incompleta.

Outro subproduto da metodologia proposta para aplicações que envolvam regeneração de imagens são duas métricas para avaliação de qualidade de imagens, z_{IQA} e b_{IQA} . Destacou-se a métrica b_{IQA} que, por ser do tipo NR, usa para seu cálculo dados que comumente estarão disponíveis em aplicações reais. Quanto à sua capacidade como indicador de qualidade das imagens, a métrica b_{IQA} mostrou desempenho melhor ou similar que outras métricas do tipo FR também avaliadas.

A seguir, o Capítulo 6 apresenta as conclusões desta tese e apresenta sugestões de trabalhos futuros baseados nos desenvolvimentos aqui mostrados.

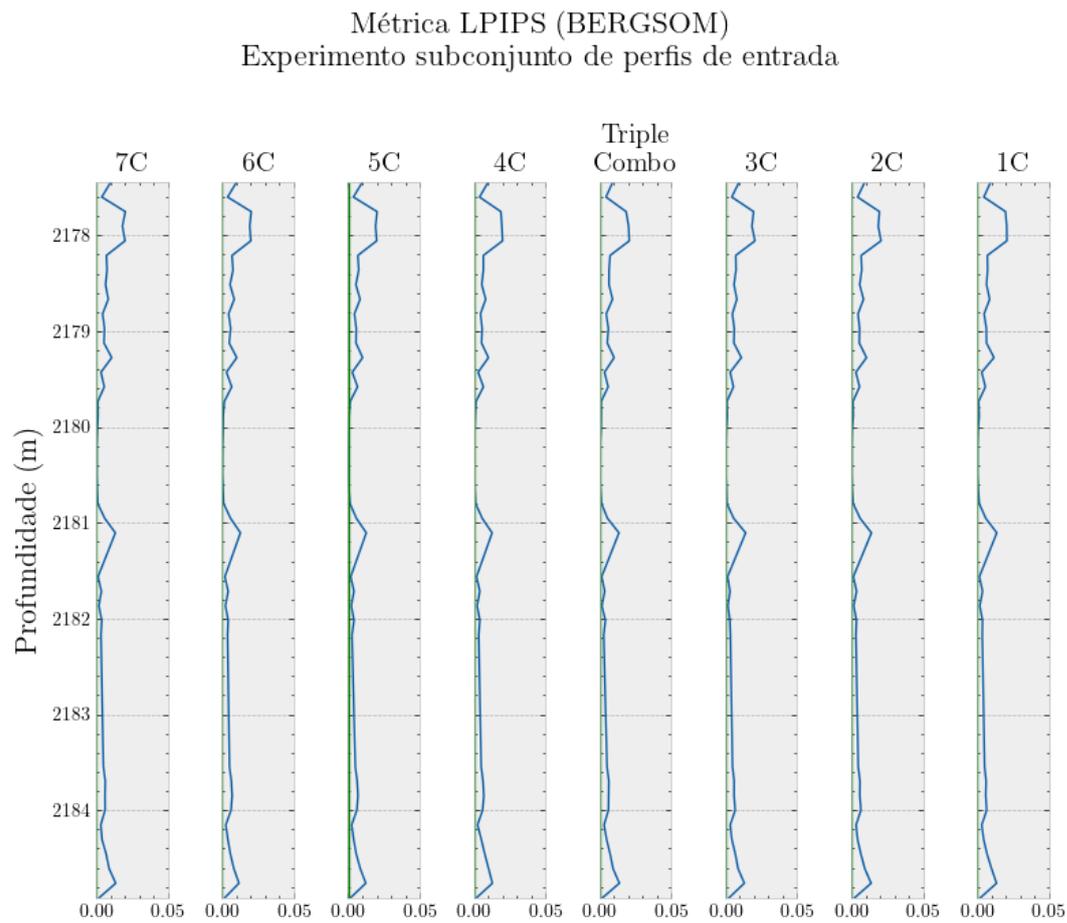


Figura 5.20: Comportamento da métrica LPIPS para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).

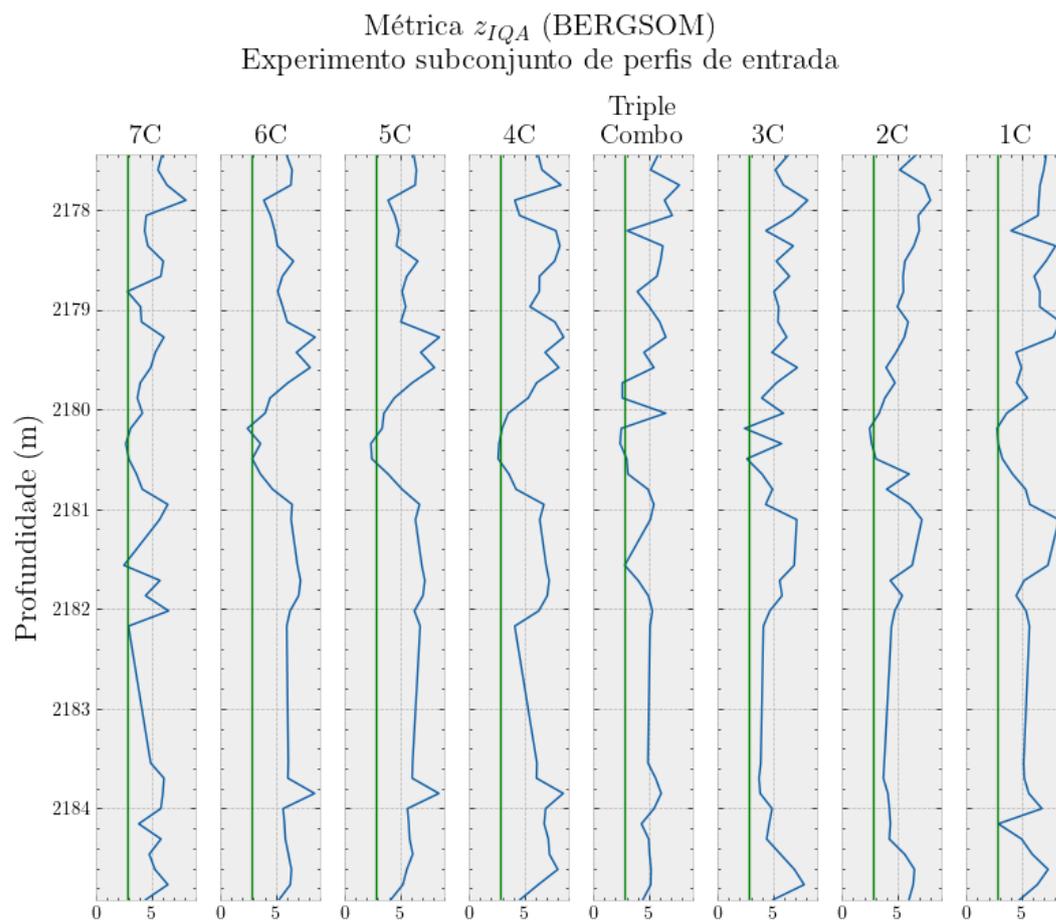


Figura 5.21: Comportamento da métrica z_{IQA} para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).

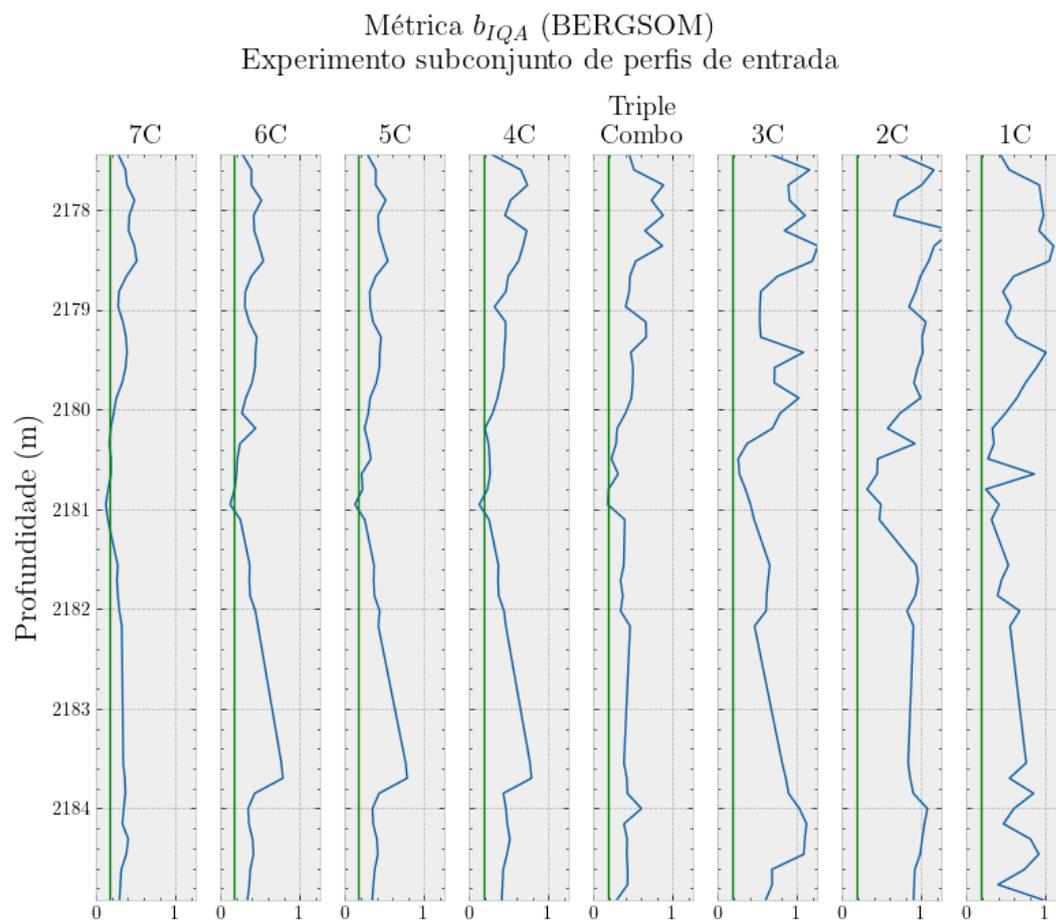


Figura 5.22: Comportamento da métrica b_{IQA} para experimento de regeneração de curvas com subconjuntos de perfis básicos de entrada (variante BERGSOM).

6

Conclusões e trabalhos futuros

Esta tese propôs uma nova metodologia autossupervisionada para aprendizado de representações entre dados de diferentes modalidades baseada no mapa auto-organizável de Kohonen (*Kohonen's self-organizing maps*, SOM). Para tal, foi proposta uma alteração conceitual em uma técnica clássica que usa a rede SOM para controle e identificação de sistemas dinâmicos na qual uma mesma rede SOM aprende conjuntamente uma relação não-linear entre dados de entrada e dados de saída. Propõe-se e desenvolve-se a ampliação deste conceito para o contexto mais geral de associação entre dados de um espaço de origem e dados de um espaço de destino.

Além disso, no cerne da proposta, está a inclusão da rede SOM no contexto do aprendizado de representações, notadamente no recente paradigma de aprendizado regenerativo. Isso acontece pelo fato de adaptar a rede SOM para aprender a associação bidirecional entre diferentes modalidades de dados, por meio das representações aprendidas para esses dados por algum modelo de aprendizado de representações, como um modelo baseado em aprendizado profundo, por exemplo. O aprendizado regenerativo ocorre quando o mapeamento cruzado entre dados de diferentes modalidades feito com a rede SOM alimenta um modelo generativo capaz de regerar um dado mais complexo (em termos de densidade de informações) a partir de um dado de menor complexidade que é fornecido como entrada à rede SOM.

Ao contrário da maioria das aplicações recentes que envolvem aprendizado generativo baseadas na rede SOM, na presente proposta esta rede não é apenas um coadjuvante na execução de tarefas de agrupamento ou visualização de dados. Na verdade, é o núcleo da metodologia para aprendizagem e inferência bidirecional entre dados de diferentes modalidades. A rede SOM permite que a metodologia proposta realize inferência entre diferentes modalidades de dados, incluindo a capacidade de realizar previsão de curvas e regeneração de imagens.

Duas aplicações no campo da geologia, especificamente em petrofísica de reservatórios de óleo e gás, são desenvolvidas para ilustrar as potencialidades da proposta:

1. **Aplicação Tipo-I:** predição de perfis geológicos convencionais a partir de perfis de imagem acústica de parede dos poços;

2. **Aplicação Tipo-II:** regeneração de perfis de imagem acústica de parede de poços a partir de perfis geológicos convencionais.

Ambas as aplicações práticas acima se propõem a resolver problemas até então intocados na área de petrofísica de reservatórios. A viabilidade de tais tarefas interessa à indústria de óleo e gás, mesmo a mais simples delas de prever um conjunto completo de perfis básicos a partir de imagens acústicas de poços, uma vez que existem muitas aplicações potenciais para esse tipo de predição numérica multivariada a partir de imagens, não apenas na área reservatórios.

Na primeira aplicação, a nova proposta totalmente não supervisionada apresenta um desempenho muito competitivo em comparação com um sólido modelo de aprendizado profundo usado como referencial, a rede SimpleNet, a um custo computacional muito menor.

Na segunda aplicação, a nova proposta mostra resultados promissores. No desenvolvimento do modelo de regeneração de imagens a partir de curvas básicas usando a proposta desta tese, surgiram duas importantes contribuições adicionais, uma na própria metodologia e outra no âmbito da aplicação em si:

- **Variante p -BERGSOM:** uma desvantagem da variante básica da metodologia proposta, BERGSOM, é que pode acontecer de um mesmo neurônio vencedor gerar exatamente a mesma imagem de saída (aplicação Tipo-II), o que não é fisicamente factível e nem desejável para a aplicação. Ao incorporar no mapeamento cruzado as componentes latentes probabilísticas, em termos de médias e desvios-padrão aprendidos para representar o espaço latente pelos modelos β -CoVAE, dota-se a rede SOM de uma capacidade de gerar imagens diferentes a partir de um mesmo vencedor. A aplicação prática parece se beneficiar dessa inovação, pois as imagens regeneradas demonstram maior similaridade com as imagens de referência;
- **Regeneração de imagens a partir de um subconjunto de perfis básicos:** mostra-se que a presente proposta engloba e estende a clássica aplicação de SOM para imputação de valores faltantes. Neste contexto, demonstra-se como imagens acústicas podem ser regeneradas por qualquer subconjunto de perfis básicos usados para treinar a rede SOM, em ambas as variantes da metodologia. Essa contribuição tem inúmeras aplicações potenciais em diversas áreas de atuação.

Vale salientar que a ideia básica contida na presente tese é portátil para várias outras aplicações para previsão de dados inter-domínios ou mesmo para geração/regeneração de dados, dependendo das especificidades do campo de aplicação.

O desenvolvimento da proposta permite citar algumas conclusões, nomeadamente:

- Na aplicação Tipo-I, a proposta de memória heteroassociativa entre diferentes modalidades baseada em aprendizado autossupervisionado pode realizar com sucesso a predição multivariada a partir de imagens sem o fardo computacionalmente intensivo de modelos supervisionados. A metodologia proposta tem desempenho semelhante a um modelo CNN profundo mais complexo e produz predições coerentes mesmo a partir de uma configuração muito compacta: um espaço de representação da imagem muito pequena (taxa de compressão de 292,5) e uma rede SOM de tamanho pequeno (100 unidades).
- Na aplicação Tipo-II, a proposta aqui desenvolvida mostrou-se capaz de regerar um perfil de imagem completo a partir de sete perfis convencionais usados como entrada. Especificamente, um vetor de sete componentes gera uma imagem coerente com 26×180 pixels. Adicionalmente, demonstrou-se que a metodologia permite regerar tais imagens a partir de vetores de entrada de dimensão ainda menor, no limite com apenas uma componente de entrada.
- Ainda no âmbito da aplicação Tipo-II, observou-se que métricas AQI consagradas para determinar a qualidade de imagens geradas ou regeneradas por modelos de aprendizado de máquinas podem não ser diretamente aplicáveis ao tipo de imagem tratada na aplicação Tipo-II. As métricas propostas baseadas nos erros de quantização fornecidos pela rede SOM expandida mostraram-se promissoras, principalmente a métrica b_{IQA} que, por ser do tipo NR, poderá ser sempre calculada em aplicações reais. Também cabe destacar que a apresentação de tais métricas como uma série de valores indexados em profundidade, como se fosse mais um perfil convencional, pode ser mais adequada para aplicações reais de petrofísica.
- Além das duas aplicações acima citadas, a rede SOM expandida também pode ser usada para melhorar a compreensão geral das relações entre os dados origem e dados de destino, contando com recursos consagrados dessa arquitetura de rede neural, como análise de *clusters* e visualização de dados. Esta funcionalidade não foi explorada no presente trabalho, mas acredita-se que é um importante subproduto do método que pode alavancar o desenvolvimento de outras soluções. Tal subproduto “gratuito” pode contribuir para melhorar a compreensão das relações entre os dados, o que seria difícil de conseguir, por exemplo, ao lidar com

modelos supervisionados do tipo caixa-preta baseados em aprendizado profundo.

O caráter inovador tanto da metodologia proposta quanto das aplicações em petrofísica abrem espaço para algumas sugestões de trabalhos futuros. A seguir essas sugestões são separadas de acordo com sua relação com a metodologia ou com as aplicações possíveis:

– **Relacionadas a variações na metodologia proposta:**

- Sugere-se investigar a memória associativa entre dados de diferentes modalidades usando outros algoritmos de quantização vetorial, desde os mais simples, como o K -médias, até redes crescentes, como as redes *Neural Gas* (MARTINETZ; BERKOVICH; SCHULTEN, 1993) ou *Growing SOM* (ALAHAKOON; HALGAMUGE; SRINIVASAN, 2000; RAUBER; MERKL; DITTENBACH, 2002). A premissa desta tese, fundamentada em trabalhos anteriores que exploraram a memória associativa com a rede SOM padrão, é que a quantização vetorial realizada por essa rede, que preserva as relações topológicas do espaço de entrada no mapa de neurônios, é a mais adequada para a metodologia proposta. Testes preliminares com o algoritmo K -médias, por exemplo, não apresentaram resultados tão promissores quanto os obtidos pela rede SOM, o que sugere preservação da topologia dessa rede é uma propriedade relevante para o tratamento de dados multimodalidades. No entanto, uma investigação mais abrangente ainda é necessária;
- Alguns métodos para adaptação de domínio, e.g., (LIU et al., 2019), são baseados em *clustering* e usam concatenação de dados dos diversos domínios, a exemplo do que se faz na memória associativa usando a rede SOM. Sugere-se investigar tais modelos como uma opção alternativa à rede SOM nessa fase de associação entre dados de modalidades diversas.

– **Relacionadas às aplicações em petrofísica:**

- A exemplo do que foi desenvolvido nesta tese, modelos existentes baseados em aprendizado profundo podem ser adaptados para a regressão multivariada a partir de imagens para a aplicação do Tipo-I. Avaliar e comparar extensivamente os diversos modelos dessa natureza não pertencem ao escopo desta tese, mas pode ser objeto de estudos futuros;

- Sugere-se investigar a aplicação dos modelos que associam a rede SOM ao aprendizado profundo citados no Capítulo 2 para as aplicações Tipo-I e Tipo-II;
 - Para a aplicação do Tipo-II, especialmente na área de petrofísica, não há modelos propostos anteriormente que permitam comparação com a metodologia proposta. O desenvolvimento de um segundo modelo para esta aplicação inédita não pertencia ao escopo desta tese. Sugere-se adaptar modelos generativos existentes, como as redes GAN condicionais (MIRZA; OSINDERO, 2014), que permitem impor ao módulo gerador de uma rede GAN uma informação adicional que condiciona a geração das imagens a partir do espaço latente. No caso da aplicação em petrofísica, essa informação pode ser o conjunto de perfis básicos. Modelos de difusão (DHARIWAL; NICHOL, 2021) têm sido usados com sucesso em diversas aplicações de visão computacional, inclusive na área de óleo e gás, e.g., (DURALL et al., 2023). Uma linha de pesquisa em modelos de difusão pode fornecer métodos adicionais de solução para a aplicação do Tipo-II, além de poder ser usada juntamente com a metodologia proposta para melhorar a definição das imagens regeneradas;
 - Ainda para a aplicação do Tipo-II, pode-se investigar melhor o uso de modelos de aprendizado profundo para aprender a relação entre os perfis básicos e as representações das imagens \mathcal{Z} (etapa origem-para-representação do aprendizado regenerativo). Tal modelo poderia, então, alimentar um modelo generativo, como o β -CoVAE e regerar as imagens;
 - Finalmente, trabalhos futuros podem aplicar o mapeamento heteroassociativo entre dados de diferentes modalidades proposto para problemas que envolvem mais de duas modalidades ou domínios. Por exemplo, uma extensão natural desta tese na área de reservatórios seria incluir, juntamente com os perfis básicos e de imagem, a informação sísmica na entrada aumentada \mathbf{x}_{aug} para o treinamento da rede SOM.
- **Outras sugestões de trabalhos futuros:**
- O uso de métricas AQI consagradas para imagens de paredes de poços se mostrou insuficiente em muitos casos. Apesar de esta tese ter proposto métricas baseadas na rede SOM, estas só podem ser calculadas com o uso da metodologia proposta. Sugere-se a investigação adicional de métricas de avaliação de qualidade específicas para esse tipo de imagem, que levem em conta a preservação de

estruturas geológicas e artefatos importantes, como os *breakouts*, e seu potencial para a estimativa de propriedades petrofísicas.

7

Referências bibliográficas

BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation learning: A review and new perspectives. **IEEE Trans. Pattern Anal. Mach. Intell.**, IEEE Computer Society, USA, v. 35, n. 8, p. 1798–1828, aug 2013. ISSN 0162-8828. Disponível em: <<https://doi.org/10.1109/TPAMI.2013.50>>. Citado na página 21.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. <<http://www.deeplearningbook.org>>. Citado na página 22.

GOODFELLOW, I. J. et al. Generative adversarial nets. In: **NIPS**. [S.l.: s.n.], 2014. Citado 2 vezes nas páginas 22 e 45.

BENGIO, Y. et al. Deep generative stochastic networks trainable by backprop. In: **International Conference on Machine Learning**. [S.l.: s.n.], 2013. Citado na página 22.

KINGMA, D. P.; WELING, M. Auto-encoding variational bayes. **arXiv preprint arXiv:1312.6114**, 2013. Citado 3 vezes nas páginas 22, 30 e 36.

TAN, X. et al. Regeneration learning: A learning paradigm for data generation. **arXiv preprint arXiv:2301.08846**, 2023. Citado 2 vezes nas páginas 22 e 99.

ŠKRLJ, B. Data modalities and representation learning. In: _____. **From Unimodal to Multimodal Machine Learning: An Overview**. Cham: Springer Nature Switzerland, 2024. p. 7–16. ISBN 978-3-031-57016-2. Disponível em: <https://doi.org/10.1007/978-3-031-57016-2_2>. Citado na página 23.

SKOCAJ, D.; LEONARDIS, A.; KRUIJFF, G.-J. M. Cross-modal learning. In: _____. **Encyclopedia of the Sciences of Learning**. Boston, MA: Springer US, 2012. p. 861–864. ISBN 978-1-4419-1428-6. Disponível em: <https://doi.org/10.1007/978-1-4419-1428-6_239>. Citado na página 24.

HUANG, X.; PENG, Y.; YUAN, M. Cross-modal common representation learning by hybrid transfer network. In: **Proceedings of the 26th International Joint Conference on Artificial Intelligence**. [S.l.]: AAAI Press, 2017. (IJCAI'17), p. 1893–1900. ISBN 9780999241103. Citado na página 24.

NICHOLS, G. **Sedimentology and Stratigraphy**. 2. ed. [S.l.]: John Wiley & Sons, 2009. ISBN 978-1-4051-3592-4. Citado na página 24.

MENEZES DE JESUS, C.; COMPAN, A. L. M.; SURMAS, R. Permeability estimation using ultrasonic borehole image logs in dual-porosity carbonate reservoirs. **Petrophysics**, SPWLA, v. 57, n. 06, p. 620–637, 2016. Citado 2 vezes nas páginas 25 e 27.

TIAB, D.; DONALDSON, E. C. **Petrophysics**. [S.l.]: Elsevier, 2016. ISBN 9780128031889. Citado na página 26.

LIU, H. **Principles and Applications of Well Logging**. Springer Berlin Heidelberg, 2017. ISBN 9783662549773. Disponível em: <<http://dx.doi.org/10.1007/978-3-662-54977-3>>. Citado 2 vezes nas páginas 26 e 27.

ANDERSEN, M. Introduction to wireline logging. **Oilfield Review**, v. 23, p. 59–60, 2011. ISSN 0923-1730. Disponível em: <<https://www.slb.com/resource-library/oilfield-review/defining-series/defining-logging>>. Citado 2 vezes nas páginas 26 e 27.

VALENTÍN, M. B. **Deep learning methods on geological reservoir borehole log images and applications**. Dissertação (Mestrado) — CBPF - Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, RJ, 03 2018. Disponível em: <<http://cbpfindex.cbpf.br/index.php?module=main&moduleFile=pubDetails&publd=13535&typeld=10>>. Citado na página 26.

ELLIS, D. V.; SINGER, J. M. **Well logging for earth scientists**. 2. ed. Dordrecht, Netherlands: Springer, 2007. Citado na página 27.

TIAN, C.; HORNE, R. N. Applying machine-learning techniques to interpret flow-rate, pressure, and temperature data from permanent downhole gauges. **SPE Reservoir Evaluation & Engineering**, SPE, v. 22, n. 02, p. 386–401, 2019. Citado na página 28.

TIAN, C.; HORNE, R. N. Machine learning applied to multiwell test analysis and flow rate reconstruction. In: ONEPETRO. **SPE Annual Technical Conference and Exhibition**. [S.l.], 2015. Citado na página 28.

OLIVEIRA, L. de; CARNEIRO, C. Synthetic geochemical well logs generation using ensemble machine learning techniques for the brazilian pre-salt reservoirs. **Journal of Petroleum Science and Engineering**, v. 196, p. 108080, 2021. ISSN 0920-4105. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0920410520311359>>. Citado na página 28.

LI, H.; MISRA, S. Long short-term memory and variational autoencoder with convolutional neural networks for generating nmr t2 distributions. **IEEE Geoscience and Remote Sensing Letters**, v. 16, n. 2, p. 192–195, 2019. Citado na página 28.

AKKURT, R. et al. Accelerating and enhancing petrophysical analysis with machine learning: a case study of an automated system for well log outlier detection and reconstruction. In: ONEPETRO. **SPWLA 59th Annual Logging Symposium**. [S.l.], 2018. Citado na página 28.

AKINNIKAWA, O.; LYNE, S.; ROBERTS, J. Synthetic well log generation using machine learning techniques. In: ONEPETRO. **SPE/AAPG/SEG Unconventional Resources Technology Conference**. [S.l.], 2018. Citado na página 28.

SALEHI, M. M. et al. Estimation of the non records logs from existing logs using artificial neural networks. **Egyptian Journal of Petroleum**, Elsevier, v. 26, n. 4, p. 957–968, 2017. Citado na página 28.

TREVIZAN, W.; MENEZES DE JESUS, C. Application of GAN to Resolution Enhancement of LWD Real-Time Image Logs to Support Decision Making. **Petrophysics**, v. 64, n. 06, p. 890–899, 12 2023. ISSN 1529-9074. Citado 3 vezes nas páginas 28, 102 e 103.

VALENTÍN, M. B. et al. Estimation of permeability and effective porosity logs using deep autoencoders in borehole image logs from the brazilian pre-salt carbonate. **Journal of Petroleum Science and Engineering**, Elsevier BV, v. 170, p. 315–330, nov. 2018. Disponível em: <<https://doi.org/10.1016/j.petrol.2018.06.038>>. Citado na página 28.

BOM, C. et al. Bayesian deep networks for absolute permeability and porosity uncertainty prediction from image borehole logs from brazilian carbonate reservoirs. **Journal of Petroleum Science and Engineering**, v. 201, p. 108361, 2021. ISSN 0920-4105. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0920410521000206>>. Citado 3 vezes nas páginas 28, 90 e 91.

FROTA, R. A. **Avaliação de algoritmos de redes neurais artificiais em tarefas de detecção de novidades: uma abordagem unificadora**. Dissertação (Mestrado) — Universidade Federal do Ceará, Fortaleza, CE, 07 2005. Disponível em: <<http://repositorio.ufc.br/handle/riufc/53309>>. Citado na página 30.

KOHONEN, T. The self-organizing map. **Proceedings of the IEEE**, v. 78, n. 9, p. 1464–1480, 1990. Citado 3 vezes nas páginas 31, 39 e 41.

LI, Z. et al. A survey of convolutional neural networks: Analysis, applications, and prospects. **IEEE Transactions on Neural Networks and Learning Systems**, v. 33, n. 12, p. 6999–7019, 2022. Citado na página 31.

KASEB, A.; KHALED, M.; GALAL, O. Convolutional neural networks for semantic segmentation: A recent survey. In: **2022 International Arab Conference on Information Technology (ACIT)**. [S.l.: s.n.], 2022. p. 1–7. Citado na página 31.

LECUN, Y.; CORTES, C.; BURGES, C. Mnist handwritten digit database. **ATT Labs [Online]**. Available: <http://yann.lecun.com/exdb/mnist>, v. 2, 2010. Disponível em: <<http://yann.lecun.com/exdb/mnist>>. Citado 2 vezes nas páginas 32 e 57.

CHARTE, D. et al. An analysis on the use of autoencoders for representation learning: Fundamentals, learning task case studies, explainability and challenges. **Neurocomputing**, v. 404, p. 93–107, 2020. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S092523122030624X>>. Citado 2 vezes nas páginas 34 e 37.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, Springer Science and Business Media LLC, v. 521, n. 7553, p. 436–444, maio 2015. Disponível em: <<https://doi.org/10.1038/nature14539>>. Citado na página 34.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, p. 533–536, 1986. ISSN 1476-4687. Disponível em: <<https://www.nature.com/articles/323533a0>>. Citado na página 34.

BALLARD, D. H. Modular learning in neural networks. In: **Proceedings of the sixth National Conference on Artificial Intelligence - Volume 1**. [S.l.: s.n.], 1987. p. 279–284. Citado na página 34.

KRAMER, M. A. Nonlinear principal component analysis using autoassociative neural networks. **AIChE Journal**, Wiley Online Library, v. 37, n. 2, p. 233–243, 1991. Citado na página 34.

PETSCHKE, T. et al. A neural network autoassociator for induction motor failure prediction. In: TOURETZKY, D.; MOZER, M.; HASSELMO, M. (Ed.). **Advances in Neural Information Processing Systems**. MIT Press, 1995. v. 8. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/1995/file/062ddb6c727310e76b6200b7c71f63b5-Paper.pdf>. Citado na página 34.

LIMA, L. et al. Unsupervised deep learning for facies pattern recognition on borehole images. In: OTC. **Offshore Technology Conference Brasil**. [S.l.], 2019. p. D011S013R001. Citado na página 35.

COTTRELL, G. W.; MUNRO, P.; ZIPSER, D. Learning internal representations from gray-scale images: An example of extensional programming. In: **Ninth Annual Conference of the Cognitive Science Society**. [S.l.: s.n.], 1987. p. 462–473. Citado na página 36.

REZENDE, D. J.; MOHAMED, S.; WIERSTRA, D. Stochastic backpropagation and approximate inference in deep generative models. In: XING, E. P.; JEBARA, T. (Ed.). **Proceedings of the 31st International Conference on Machine Learning**. Beijing, China: PMLR, 2014. (Proceedings of Machine Learning Research, 2), p. 1278–1286. Disponível em: <<https://proceedings.mlr.press/v32/rezende14.html>>. Citado na página 36.

DAMM, S. et al. The elbo of variational autoencoders converges to a sum of entropies. In: **International Conference on Artificial Intelligence and Statistics**. [s.n.], 2020. Disponível em: <<https://proceedings.mlr.press/v206/damm23a/damm23a.pdf>>. Citado na página 38.

HIGGINS, I. et al. beta-VAE: Learning basic visual concepts with a constrained variational framework. In: **International Conference on Learning Representations**. [s.n.], 2017. Disponível em: <<https://openreview.net/forum?id=Sy2fzU9gl>>. Citado na página 38.

KOHONEN, T. **Self-Organizing Maps**. 3rd. ed. [S.l.]: Springer-Verlag, 2001. Citado 2 vezes nas páginas 39 e 45.

HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. 2nd. ed. [S.l.]: Prentice-Hall, 1999. Citado na página 39.

KIM, H.; JUNG, S. H. SOGN: novel generative model using SOM. **Electronics Letters**, Institution of Engineering and Technology (IET), v. 55, n. 10, p. 597–599, maio 2019. Disponível em: <<https://doi.org/10.1049/el.2019.0202>>. Citado na página 40.

KOHONEN, T.; NIEMINEN, I. T.; HONKELA, T. On the quantization error in som vs. vq: A critical and systematic study. In: PRÍNCIPE, J. C.; MIIKKULAINEN, R. (Ed.). **Advances in Self-Organizing Maps**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. p. 133–144. ISBN 978-3-642-02397-2. Citado na página 41.

YIN, H. The self-organizing maps: Background, theories, extensions and applications. In: _____. **Computational Intelligence: A Compendium**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 715–762. ISBN 978-3-540-78293-3. Citado na página 42.

BARRETO, G.; ARAUJO, A. Identification and control of dynamical systems using the self-organizing map. **IEEE Transactions on Neural Networks**, v. 15, n. 5, p. 1244–1259, 2004. Citado 6 vezes nas páginas 42, 43, 44, 48, 92 e 114.

RITTER, H.; SCHULTEN, K. Topology conserving mappings for learning motor tasks. In: AMERICAN INSTITUTE OF PHYSICS. **AIP Conference Proceedings**. [S.l.], 1986. v. 151, p. 376–380. Citado 2 vezes nas páginas 43 e 47.

RITTER, H.; KOHONEN, T. Self-organizing semantic maps. **Biological Cybernetics**, Springer Science and Business Media LLC, v. 61, n. 4, p. 241–254, ago. 1989. Disponível em: <<https://doi.org/10.1007/bf00203171>>. Citado na página 43.

RITTER, H. J.; MARTINETZ, T. M.; SCHULTEN, K. J. Topology-conserving maps for learning visuo-motor-coordination. **Neural Networks**, v. 2, n. 3, p. 159–168, 1989. ISSN 0893-6080. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0893608089900014>>. Citado na página 43.

LENDASSE, A. et al. Vector quantization: a weighted version for time-series forecasting. **Future Generation Computer Systems**, v. 21, n. 7, p. 1056–1067, 2005. ISSN 0167-739X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167739X04000238>>. Citado na página 43.

BARRETO, G.; ARAUJO, A. Temporal associative memory and function approximation with the self-organizing map. In: **Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing**. [S.l.: s.n.], 2002. p. 109–118. Citado 3 vezes nas páginas 43, 44 e 47.

BARRETO, G.; ARAUJO, A. A self-organizing narx network and its application to prediction of chaotic time series. In: **IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)**. [S.l.: s.n.], 2001. v. 3, p. 2144–2149 vol.3. Citado 3 vezes nas páginas 43, 44 e 47.

BARRETO, G.; ARAUJO, A.; RITTER, H. Self-organizing feature maps for modeling and control of robotic manipulators. **Journal of Intelligent and Robotic Systems**, Springer Science and Business Media LLC, v. 36, n. 4, p. 407–450, 2003. Disponível em: <<https://doi.org/10.1023/a:1023641801514>>. Citado 3 vezes nas páginas 43, 47 e 114.

MUKWACHI, P.; ARNONKIJPANICH, B.; SARAORN, W. Simulating 2-d magnetotelluric responses using vector-quantized temporal associative memory artificial neural network-based approaches. **Geoscience Letters**, Springer Science and

Business Media LLC, v. 11, n. 1, mar. 2024. ISSN 2196-4092. Disponível em: <<http://dx.doi.org/10.1186/s40562-024-00328-8>>. Citado na página 44.

LIMTRAKUL, S.; ARNONKIJPANICH, B. Supervised learning based on the self-organizing maps for forward kinematic modeling of stewart platform. **Neural Computing and Applications**, Springer Science and Business Media LLC, v. 31, n. 2, p. 619–635, jun. 2017. ISSN 1433-3058. Disponível em: <<http://dx.doi.org/10.1007/s00521-017-3095-4>>. Citado na página 44.

WONGSRIWORAPHON, A.; ARNONKIJPANICH, B.; PATHUMNAKUL, S. An approach based on digital image analysis to estimate the live weights of pigs in farm environments. **Computers and Electronics in Agriculture**, v. 115, p. 26–33, 2015. ISSN 0168-1699. Citado 2 vezes nas páginas 44 e 45.

LI, W.; HE, Y.; TANG, Y. Improving mode exploring capability of generative adversarial nets by self-organizing map. **Neurocomputing**, Elsevier BV, v. 576, p. 127244, 2024. Citado na página 44.

MULIUKOV, A. R.; RODRIGUEZ, L.; MIRAMOND, B. Cortex inspired learning to recover damaged signal modality with red-som model. In: **2023 International Joint Conference on Neural Networks (IJCNN)**. [S.l.: s.n.], 2023. p. 01–09. Citado 2 vezes nas páginas 44 e 45.

SAKKARI, M. et al. Feature extraction-based deep self-organizing map. **Circuits, Systems, and Signal Processing**, Springer Science and Business Media LLC, v. 41, n. 5, p. 2643–2666, 2022. Citado na página 44.

FOREST, F. et al. Deep embedded self-organizing maps for joint representation learning and topology-preserving clustering. **Neural Computing and Applications**, Springer Science and Business Media LLC, v. 33, n. 24, p. 17439–17469, 2021. Citado na página 44.

ALY, S.; ALMOTAIRI, S. Deep convolutional self-organizing map network for robust handwritten digit recognition. **IEEE Access**, v. 8, p. 107035–107045, 2020. Citado na página 44.

KHACEF, L.; RODRIGUEZ, L.; MIRAMOND, B. **Written and spoken digits database for multimodal learning**. Zenodo, 2019. Disponível em: <<https://doi.org/10.5281/zenodo.3515935>>. Citado na página 46.

YAO, Y. et al. Cross-modal representation learning. In: LIU, Z.; LIN, Y.; SUN, M. (Ed.). **Representation Learning for Natural Language Processing**. Singapore: Springer Nature Singapore, 2023. p. 211–240. ISBN 978-981-99-1600-9. Citado na página 48.

TAN, X. et al. Regeneration learning: A learning paradigm for data generation. In: **Proceedings of the AAAI Conference on Artificial Intelligence**. [S.l.: s.n.], 2024. v. 38, p. 22614–22622. Citado na página 51.

de Gouvêa, R. C. T. et al. Intrasom: A comprehensive python library for self-organizing maps with hexagonal toroidal maps training and missing data handling. **Software Impacts**, v. 17, p. 100570, 2023. ISSN 2665-9638. Citado 2 vezes nas páginas 53 e 69.

JÁCOMO, M. H. et al. Mineralogical modeling and petrophysical properties of the barra velha formation, santos basin, brazil. **Petrophysics**, v. 64, n. 04, p. 518–543, 08 2023. ISSN 1529-9074. Disponível em: <<https://doi.org/10.30632/PJV64N4-2023a3>>. Citado na página 64.

HE, K. et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: **Proceedings of the IEEE International Conference on Computer Vision (ICCV)**. [S.l.: s.n.], 2015. Citado na página 66.

HASANPOUR, S. H. et al. **Lets keep it simple, Using simple architectures to outperform deeper and more complex architectures**. 2016. Citado na página 67.

SZWARCMAN, D.; CIVITARESE, D.; VELLASCO, M. Quantum-inspired evolutionary algorithm applied to neural architecture search. **Applied Soft Computing**, v. 120, p. 108674, 2022. ISSN 1568-4946. Citado na página 68.

ABADI, M. et al. **TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems**. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>. Citado na página 69.

CHOLLET, F. et al. **Keras**. 2015. <<https://keras.io>>. Citado na página 69.

FRANCOIS, D.; WERTZ, V.; VERLEYSEN, M. The concentration of fractional distances. **IEEE Transactions on Knowledge and Data Engineering**, v. 19, n. 7, p. 873–886, 2007. Citado 2 vezes nas páginas 81 e 105.

TAMOTO, H.; GIORIA, R.; CARNEIRO, C. Prediction of nuclear magnetic resonance porosity well-logs in a carbonate reservoir using supervised machine learning models. **Journal of Petroleum Science and Engineering**, v. 220, p. 111169, 2023. ISSN 0920-4105. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S092041052201021X>>. Citado 2 vezes nas páginas 90 e 91.

CHEN, S. et al. Use of symbolic regression for developing petrophysical interpretation models. **Petrophysics**, v. 64, n. 02, p. 174–190, 04 2023. ISSN 1529-9074. Disponível em: <<https://doi.org/10.30632/PJV64N2-2023a3>>. Citado 2 vezes nas páginas 90 e 91.

LI, X. Permeability calculation of complex carbonate reservoirs based on data mining techniques. **Petrophysics**, v. 64, n. 01, p. 92–106, 02 2023. ISSN 1529-9074. Disponível em: <<https://doi.org/10.30632/PJV64N1-2023a7>>. Citado 2 vezes nas páginas 90 e 91.

SIMÕES, V. et al. Comparative study of machine-learning-based methods for log prediction. **Petrophysics**, v. 64, n. 02, p. 192–212, 04 2023. ISSN 1529-9074. Disponível em: <<https://doi.org/10.30632/PJV64N2-2023a4>>. Citado 3 vezes nas páginas 90, 91 e 116.

KIM, J. Synthetic shear sonic log generation utilizing hybrid machine learning techniques. **Artificial Intelligence in Geosciences**, v. 3, p. 53–70, 2022. ISSN 2666-5441. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2666544122000259>>. Citado 2 vezes nas páginas 90 e 91.

SESHADRINATHAN, K. et al. Chapter 21 - image quality assessment. In: BOVIK, A. (Ed.). **The Essential Guide to Image Processing**. Boston: Academic Press, 2009. p. 553–595. ISBN 978-0-12-374457-9. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780123744579000214>>. Citado na página 96.

WANG, Z.; BOVIK, A. C. Mean squared error: Love it or leave it? a new look at signal fidelity measures. **IEEE Signal Processing Magazine**, v. 26, n. 1, p. 98–117, 2009. Citado na página 96.

WANG, Z. et al. Image quality assessment: from error visibility to structural similarity. **IEEE Transactions on Image Processing**, v. 13, n. 4, p. 600–612, 2004. Citado na página 96.

ZHANG, R. et al. The unreasonable effectiveness of deep features as a perceptual metric. In: **2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2018. p. 586–595. Citado na página 97.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F. et al. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2012. v. 25. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>. Citado na página 97.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014. Citado na página 97.

ORBAN, N. et al. Assessment of new high-definition borehole imaging-while-drilling technology: Learnings from pre-salt carbonates of Brazil. In: **SPWLA Annual Logging Symposium 2021**. [s.n.], 2021. p. D021S016R003. Disponível em: <<https://doi.org/10.30632/SPWLA-2021-0050>>. Citado na página 103.

ZHU, J.-Y. et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: **2017 IEEE International Conference on Computer Vision (ICCV)**. IEEE, 2017. p. 2242–2251. Disponível em: <<http://dx.doi.org/10.1109/ICCV.2017.244>>. Citado na página 103.

ULTSCH, A.; MÖRCHEN, F. **ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM**. [S.l.]: University of Marburg, 2005. v. 46. Citado na página 105.

SAMAD, T.; HARP, S. A. Self-organization with partial data. **Network: Computation in Neural Systems**, Taylor & Francis, v. 3, n. 2, p. 205–212, 1992. Citado na página 114.

FESSANT, F.; MIDENET, S. Self-organising map for data imputation and correction in surveys. **Neural Computing & Applications**, Springer, v. 10, p. 300–310, 2002. Citado na página 114.

WANG, S. Application of self-organising maps for data mining with incomplete data sets. **Neural Computing & Applications**, Springer, v. 12, p. 42–48, 2003. Citado na página 114.

REJEB, S.; DUVEAU, C.; REBAFKA, T. Self-organizing maps for exploration of partially observed data and imputation of missing values. **Chemometrics and Intelligent Laboratory Systems**, Elsevier BV, v. 231, p. 104653, dez. 2022. Disponível em: <<https://doi.org/10.1016/j.chemolab.2022.104653>>. Citado na página 114.

VATANEN, T. et al. Self-organization and missing values in SOM and GTM. **Neurocomputing**, v. 147, p. 60–70, 2015. ISSN 0925-2312. Advances in Self-Organizing Maps Subtitle of the special issue: Selected Papers from the Workshop on Self-Organizing Maps 2012 (WSOM 2012). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231214007127>>. Citado na página 114.

MARTINETZ, T.; BERKOVICH, S.; SCHULTEN, K. 'neural-gas' network for vector quantization and its application to time-series prediction. **IEEE Transactions on Neural Networks**, v. 4, n. 4, p. 558–569, 1993. Citado na página 133.

ALAHAKOON, D.; HALGAMUGE, S.; SRINIVASAN, B. Dynamic self-organizing maps with controlled growth for knowledge discovery. **IEEE Transactions on Neural Networks**, v. 11, n. 3, p. 601–614, 2000. Citado na página 133.

RAUBER, A.; MERKL, D.; DITTENBACH, M. The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data. **IEEE Transactions on Neural Networks**, v. 13, n. 6, p. 1331–1341, 2002. Citado na página 133.

LIU, H. et al. Structure-preserved unsupervised domain adaptation. **IEEE Transactions on Knowledge and Data Engineering**, v. 31, n. 4, p. 799–812, 2019. Citado na página 133.

MIRZA, M.; OSINDERO, S. **Conditional Generative Adversarial Nets**. 2014. Disponível em: <<https://arxiv.org/abs/1411.1784>>. Citado na página 134.

DHARIWAL, P.; NICHOL, A. Diffusion models beat GANs on image synthesis. In: RANZATO, M. et al. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2021. v. 34, p. 8780–8794. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2021/file/49ad23d1ec9fa4bd8d77d02681df5cfa-Paper.pdf>. Citado na página 134.

DURALL, R. et al. Deep diffusion models for seismic processing. **Computers & Geosciences**, v. 177, p. 105377, 2023. ISSN 0098-3004. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S009830042300081X>>. Citado na página 134.

WANG, H.; TOKSÖZ, M. N.; FEHLER, M. C. **Borehole Acoustic Logging – Theory and Methods**. 1. ed. [S.l.]: Springer, 2020. Citado na página 146.

MENEZES DE JESUS, C. et al. **U.S. Patent No. 20170082767A1**. 2017. U.S. Patent and Trademark Office, Washington, DC. <https://patents.google.com/patent/US20170082767A1/>. Citado na página 146.

MCDONALD, A. Data quality considerations for petrophysical machine-learning models. **Petrophysics**, SPWLA, v. 62, n. 06, p. 585–613, 2021. Citado na página 146.

A

Dataset e preparação dos dados

A.1

Dataset

Neste trabalho, utilizam-se dados de cinco poços, denominados A a E. Cada poço dispõe de um conjunto de dados de imagens acústicas de alta resolução e um conjunto de dados complementar contendo sete perfis petrofísicos básicos. Os perfis básicos foram descritos no Capítulo 1.

A.2

Preparação de dados

Os dados de perfis de imagem da parede do poço devem ser processados para diversas correções devido tanto às condições de poço como a peculiaridades de aquisição, como correções de velocidade, orientação e excentricidade da ferramenta (WANG; TOKSÖZ; FEHLER, 2020; MENEZES DE JESUS et al., 2017). No entanto, esta etapa de preparação de dados de imagem está fora do escopo desta tese, e já se parte dos dados corrigidos previamente. Demais pré-processamentos na imagem são os relacionados ao aprendizado de máquina, como dimensionar os valores da imagem para uma escala $[0, 1]$ no momento oportuno.

Para os perfis petrofísicos básicos, no entanto, é necessária uma preparação prévia dos dados, segundo algumas práticas recomendadas pela literatura em petrofísica (MCDONALD, 2021):

- valores negativos (se houver) nas curvas de resistividade são descarregados. Os valores restantes são convertidos para uma escala logarítmica;
- todas as curvas são submetidas à remoção de *outliers* por inspeção visual e cortes de valores baseados no perfil CAL;
- todos os perfis básicos são normalizados dimensionados para o intervalo $[0, 1]$ usando a normalização min-max;
- aplicou-se uma transformação logarítmica decimal aos dados do perfil de resistividade.

Como os experimentos exigem relacionar um conjunto de valores de perfis (um vetor) a um determinado trecho da imagem da parede do poço (uma matriz de valores reais), duas tarefas de pré-processamento dos dados são necessárias para garantir essa correspondência:

1. O perfil de imagem e os perfis básicos de poços podem ser adquiridos em intervalos diferentes (em termos de profundidade), como mostra a Figura A.1).
2. Os dois tipos de perfis (imagem e convencionais) possuem diferentes taxas de amostragem ou aquisição (em profundidade), como mostra a Figura A.3.

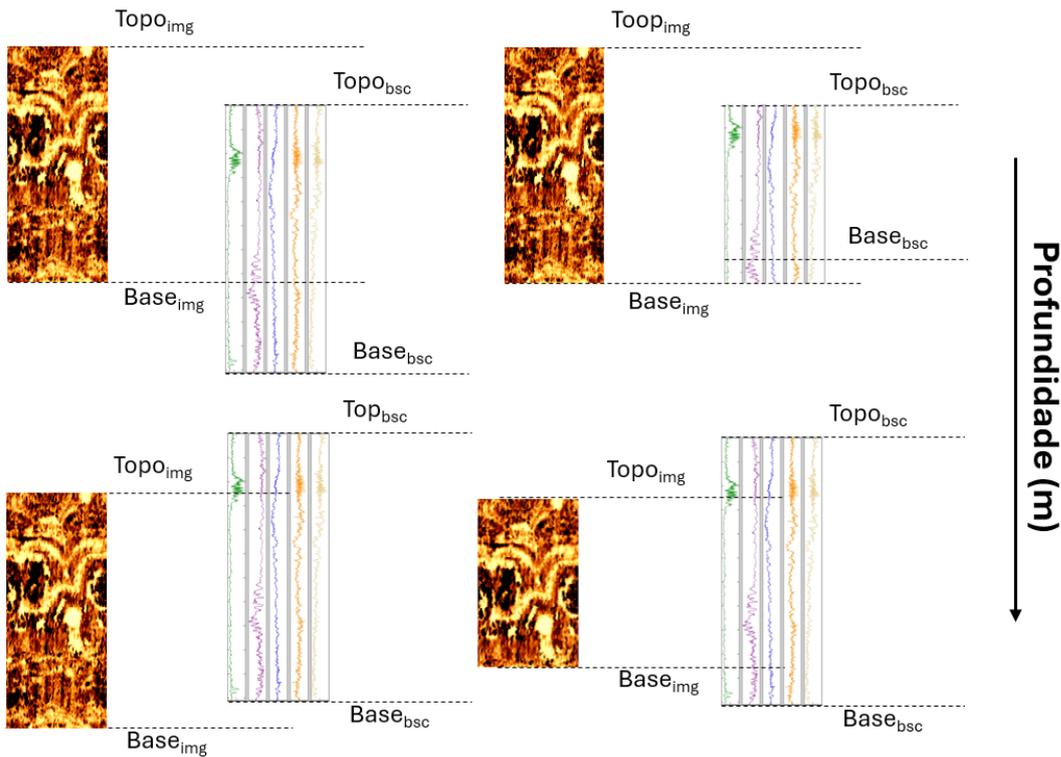


Figura A.1: Possíveis casos de incompatibilidade de profundidade entre os perfis básicos e o perfil de imagem acústica. O subíndice “img” refere-se ao perfil de imagem e o subíndice “bsc” refere-se aos perfis básicos.

Para resolver o primeiro problema, determina-se um intervalo mútuo de profundidade no qual ambos os tipos de perfis são adquiridos simultaneamente, e apenas este intervalo válido é considerado para a construção do conjunto de dados de modelagem, como mostra a Figura A.2.

A segunda questão da taxa de aquisição (em profundidade) dos diferentes tipos de perfil. No presente caso, uma imagem acústica tem uma nova aquisição (uma nova linha da imagem) a cada 0,0058 m, enquanto o perfil regular tem uma nova leitura a cada 0,1524 m. Assim, uma imagem cuja altura em pixels seja de Δ_{img} (26 pixels) corresponde a um único vetor \mathbf{y} de valores de perfis básicos (pontos coloridos na Figura A.3 ilustrando um caso de quatro curvas). Para lidar com essa diferença na resolução dos dados em termos de

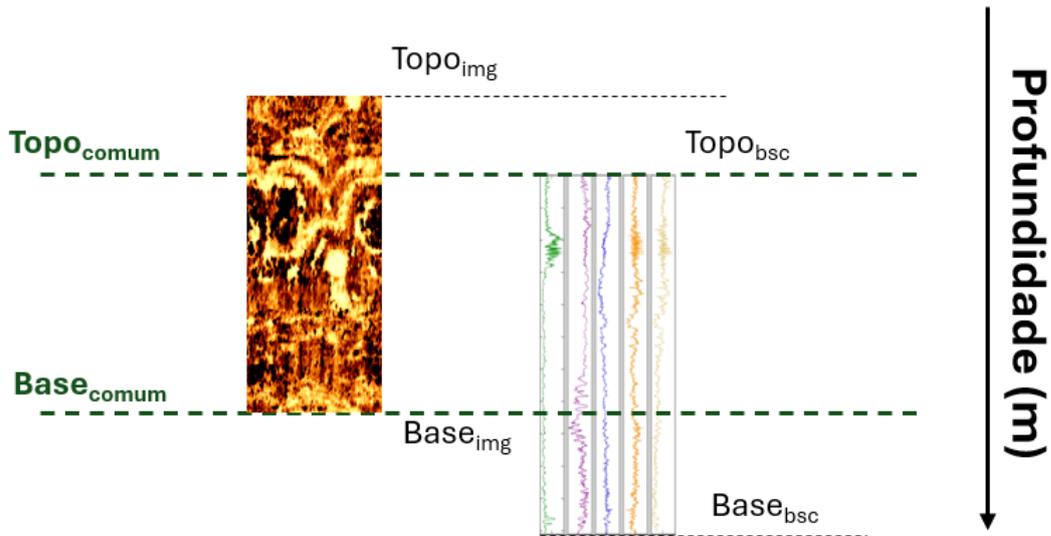


Figura A.2: Intervalo comum de profundidade para ambas as fontes de dados.

profundidade, optou-se por centralizar a imagem na profundidade exata de uma aquisição do perfil convencional, de modo que se a imagem correspondente seja formada por 13 pixels acima e 13 pixels abaixo desta profundidade. Portanto, o banco de dados possui p imagens (\mathbf{x} na Figura A.3) de 26×180 pixels correspondentes a p vetores m -dimensionais (\mathbf{y} na Figura A.3, e m é o número de perfis básicos no conjunto de dados).

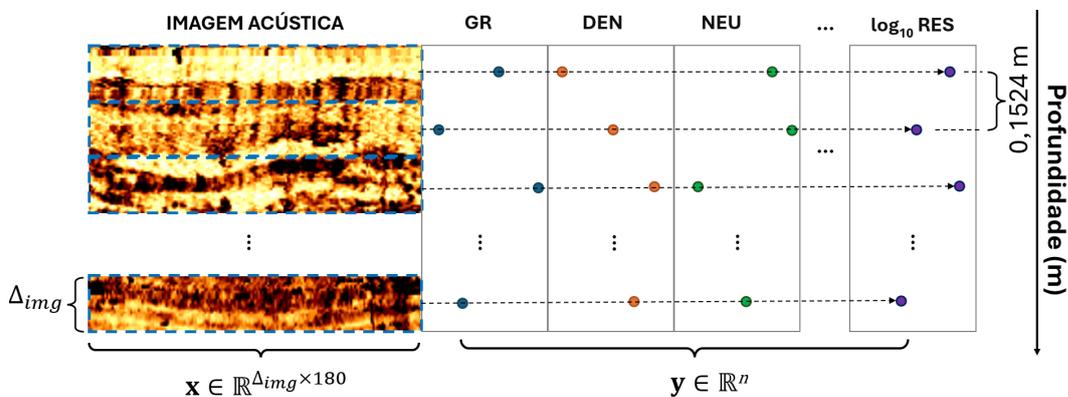


Figura A.3: Ajuste para tratamento de diferenças de taxa de amostragem em profundidade das diferentes modalidades de perfis (básicos e imagem acústica).

A.3 Publicação dos Dados

Os dados utilizados nesta tese foram publicados num *online* repositório chamado WAID (do inglês *Wellbore Acoustic Image Database*)¹.

¹Disponível em Wellbore Acoustic Image Database.

B Publicações

Durante o desenvolvimento do presente curso de doutorado, foi desenvolvida a seguinte produção bibliográfica:

1. Frota, Rewbenio A., Vellasco, M.M.B.R., Barreto, G.A. & Menezes de Jesus, C., **Heteroassociative Mapping with Self-Organizing Maps for Probabilistic Multi-output Prediction**. 2024 International Joint Conference on Neural Networks (IJCNN). [S.l.]: IEEE. June 2024. p. 1–6.
2. Frota, Rewbenio A., Barreto, G.A., Vellasco, M.M.B.R. & Menezes de Jesus, C., **New Cloth Unto an Old Garment: SOM for Regeneration Learning**. In: **Advances in Self-Organizing Maps, Learning Vector Quantization, Interpretable Machine Learning, and Beyond**. Springer Nature Switzerland, 2024. p. 1–10. ISBN: 9783031671593 ISSN: 2367-3389.
3. Frota, Rewbenio A., Barreto, G.A., Vellasco, M.M.B.R. & Menezes de Jesus, C., **Rede SOM para Aprendizado de Representações Multimodais com Aplicação em Petrofísica**. Anais do XXV Congresso Brasileiro de Automática (CBA). Rio de Janeiro: SBA. 2024.
4. Frota, Rewbenio A., Tanscheit, R., & Vellasco, M. (2023). **Fuzzy logic for control of injector wells flow rates under produced water reinjection**, *Journal of Petroleum Science and Engineering*, Volume 215, Part A, pp. 110574.¹

Os trabalhos abaixo foram submetidos:

1. Frota, Rewbenio A., Barreto, G.A., Vellasco, M.M.B.R. & Menezes de Jesus, C., **Bidirectional Multimodal Petrophysical Logs Generation with Self-supervised Regeneration Learning**, *Engineering Applications of Artificial Intelligence*. Submetido em 26/07/2024.²
2. Frota, Rewbenio A., Barreto, G.A., Vellasco, M.M.B.R. & Menezes de Jesus, C., **Self-Organizing Maps for Cross-modal Representation Learning in Petrophysics Applications**, *Geoenergy Science and Engineering*, Special Issue on Advances in Machine Learning Applications for Subsurface Energy Resources. Submetido em 30/09/2023.³

¹Qualis A1.

²Qualis A1.

³Previamente intitulado *Journal of Petroleum Science and Engineering* (Qualis A1).