

TexVar: Um Autoencoder Variacional para representação e interpretação de texturas mamográficas

Lucca Lemos Lago¹, Alexei Manso Correa Machado^{1,2}

¹Departamento de Ciência da Computação - PUC Minas

²Departamento de Anatomia e Imagem - Universidade Federal de Minas Gerais

l1llago@sga.pucminas.br, alexeimcmachado@gmail.com

Abstract. *The classification of breast density through X-ray exams is still considered the main screening mechanism for the early detection of breast cancer, since the fibroglandular tissue can hide initial tumors. The objective of this article is to propose an interpretable model for the classification of breast densities in mammographic images. The proposed architecture consists of a variational autoencoder (VAE) composed by convolutional and dense layers with a latent space of 32 variables. The model allows for explaining the meaning of the latent variables through the texture transition between the 4 classes of the Breast Image Reporting and Data System (BI-RADS) scale. Experiments using the IRMA public database composed of 5024 images demonstrated the ability of the VAE to reduce the dimensionality of the problem to a space where the most discriminating variables can be visually interpreted, supporting computer-aided diagnosis.*

Resumo. *A classificação da densidade mamária através de exames de raio X ainda é considerado o principal mecanismo para detecção precoce do câncer de mama, uma vez que o tecido fibroglandular pode esconder tumores iniciais. O objetivo deste artigo é propor um modelo interpretável para a classificação de densidades mamárias em imagens mamográficas. A arquitetura proposta é constituída por um autoencoder variacional (VAE) de camadas convolucionais e densas com espaço latente de 32 variáveis. O modelo permite examinar o significado das variáveis através da transição de textura entre as 4 classes da escala Breast Image Reporting and Data System (BI-RADS). Experimentos utilizando a base de dados pública IRMA composta de 5024 imagens demonstraram a capacidade do VAE em reduzir a dimensionalidade do problema para um espaço onde as variáveis mais discriminantes podem ser interpretadas de forma visual, fundamentando o processo de diagnóstico.*

1. Introdução

Embora exista um significativo número de técnicas e modelos de aprendizado de máquina destinados à classificação de padrões em imagens, a maioria deles não é capaz de indicar o motivo que levou o modelo a tomar uma decisão. Isso pode se tornar um problema em muitas áreas como na Medicina, à medida que o médico precisa saber o motivo da predição do modelo antes de tomar uma decisão importante. No caso do diagnóstico do câncer de mama, em que a predição correta e em tempo adequado pode diminuir bastante a taxa de mortalidade da doença, essas habilidades podem ser um fator crítico para a escolha do método a ser empregado.

O câncer de mama é uma doença de alto risco que afeta uma grande quantidade de mulheres. No Brasil, a doença representou cerca de 30.1% dos casos de câncer em mulheres em 2022 (INCA, 2021). Já é sabido que a alta taxa de mortalidade da doença se deve ao seu diagnóstico tardio, o que eleva a importância das técnicas usadas para sua identificação precoce. A mamografia é a principal técnica usada para detecção da doença e consiste em exames de raio X com o objetivo de localizar possíveis nódulos e calcificações. Um problema nesse método é a detecção da doença no caso de mamas densas que podem esconder pequenas lesões. A escala mais usada para classificar a densidade mamária é a *Breast Image Reporting and Data System* (BI-RADS). Ela classifica a densidade em quatro grupos, I a IV, baseado na proporção de tecido gorduroso e fibroglandular da mama. As mamas dentro dos grupos I e II possuem menor risco de ocultarem lesões do que as caracterizadas como sendo dos grupos III e IV. A Figura 1 mostra exemplos dessas diferentes classes.

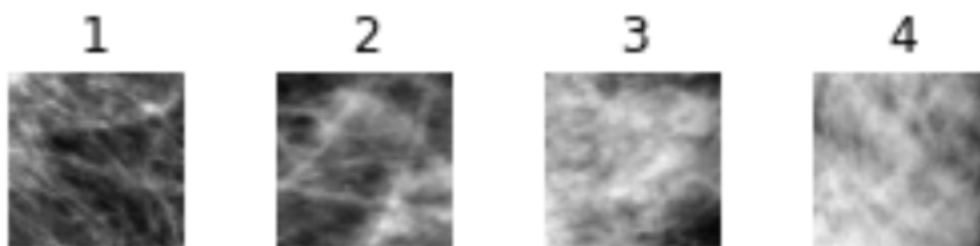


Figura 1. Exemplos de texturas mamográficas para cada classe BI-RADS

Uma limitação na classificação das imagens de raio X na escala BI-RADS é a falta de limites explícitos de densidade entre as classes. Dessa forma, além de proporcionar uma classificação acurada, um modelo de apoio ao diagnóstico computadorizado deve ser capaz de justificar os resultados de forma interpretável.

O conceito de interpretabilidade está intimamente relacionado à representação da informação visual através de variáveis não originalmente observáveis que constituem um espaço latente. Esse espaço é composto de um conjunto reduzido de variáveis obtidas através de técnicas de redução de dimensionalidade. Os *Variational Autoencoders* (VAEs) (Kingma e Welling, 2013) são arquiteturas de redes neurais profundas que conseguem definir um espaço latente a partir do qual é possível gerar imagens de saída com as características de imagens que foram usadas para seu treinamento. Logo, os VAEs são arquiteturas que devem ser consideradas quanto ao aspecto da interpretabilidade em modelos profundos.

Este artigo propõe uma nova arquitetura de VAE denominada *TexVar* capaz de representar e fundamentar a interpretação do seu espaço latente na classificação de densidades mamárias através de sua textura. O método procura recriar imagens mamográficas para cada classe BI-RADS no intuito de, através da manipulação do espaço latente, transformar uma imagem de uma determinada classe em outra classe, de modo a salientar suas características discriminantes.

2. Trabalhos relacionados

Variational Autoencoders são usados em diversas áreas que necessitam do uso de modelos generativos e recursos de interpretabilidade. VAEs podem ser usados simplesmente como geradores de novos dados para treinamento de um modelo de interpretação, como apresentado por Schockaert, Macher, e Schmitz (2020), que utilizam o *Local Interpretable Model-agnostic Explanations* (LIME) para treinar um modelo linear que se aproxima localmente do modelo caixa-preta. No trabalho, é proposto o VAE-LIME, que usa um VAE para aprender as características complexas do processo de alto-forno, objeto de estudo do artigo, a partir dos dados. Conseqüentemente, o VAE tem como objetivo gerar amostras artificiais para treinar um modelo interpretável local como substituto do modelo caixa-preta.

Uma técnica muito utilizada na concepção de VAEs é a perturbação que altera o valor de alguma característica da entrada gerando dados modificados. Utkin, Drobintsev, Kovalev, e Konstantinov (2021) fazem uso dessa técnica na interpretabilidade de modelos caixa-preta. O trabalho propõe um modelo de interpretabilidade de uma rede neural profunda através de um *autoencoder* (AE) simples e um VAE. O AE simples tenta reconstruir os exemplos, enquanto o variacional produz novas entradas para o primeiro, fazendo uso da técnica de perturbação e mudando alguma característica para ver como a rede se comporta. Uzunova, Ehrhardt, Kepp, e Handels (2019) também utilizam um VAE para explicar como classificadores caixa-preta diferenciam imagens patológicas e saudáveis. O VAE tem o papel de gerar perturbações na região patológica da imagem tentando alterar a previsão do classificador, conseguindo explicar quais regiões da imagem são determinantes para considerá-la como não saudável. A modalidade de imagem usada nos experimentos é a Tomografia de coerência óptica de retina (OCT) e imagens de lesão cerebral.

A utilização do espaço latente para interpretar as características fundamentais contidas nas imagens de uma base de doenças em plantas é explorada por Habaragamuwa, Oishi, e Tanaka (2021). O trabalho utiliza um modelo para interpretar as variações dos recursos mais importantes, visualizando as imagens geradas que correspondem às variações desses recursos. Liu et al. (2020) apresentam um método para gerar mapas de atenção visual a partir do espaço latente aprendido e também demonstrar que tais explicações de atenção servem mais do que para apenas explicar as previsões do VAE. O artigo mostra como esses mapas de atenção podem ser usados para localizar anomalias em imagens.

Wei e Mahmood (2021) resumem de forma abrangente os principais conceitos de VAEs e suas aplicações em algumas áreas como na análise de imagens médicas. O artigo aborda outras utilizações de VAEs além da interpretabilidade e apresenta ideias e temas de discussão para que os modelos possam ser usados como fontes de tomadas de decisões seguras.

3. Modelos generativos

Em contraste com modelos discriminativos, que recebem recursos, como entrada, e tentam classificá-los, modelos generativos tentam prever os recursos a partir de sua classe. Atualmente, as principais arquiteturas pertencentes ao conjunto de modelos generativos são as redes generativas adversárias (GANs) e os VAEs, ambas capazes de produzir conteúdo original.

As GANs são redes adversárias na qual 2 redes competem uma com a outra, o gerador e o discriminador. Enquanto o gerador tenta produzir imagens semelhantes às reais, a partir de seu espaço latente, o discriminador recebe as imagens reais misturadas às produzidas pelo gerador e tenta distingui-las. Dessa forma, uma rede ajuda a outra a evoluir, fazendo com que o gerador consiga gerar conteúdo novo parecido com o real.

Por sua vez, os VAEs são arquiteturas de redes neurais profundas que realizam a redução de dimensionalidade gerando um espaço latente regularizado e depois geram um novo conteúdo com as características dos recursos que foram usados para o treinamento. Uma VAE é formada por um codificador, que recebe a imagem de origem e a representa no seu espaço latente, e pelo decodificador, que recebe um vetor de números aleatórios para a geração de novas imagens. Ao forçar a entrada a passar por camadas com número reduzido de neurônios, o codificador precisa aprender a eliminar redundâncias e representar a informação em um espaço de dimensionalidade reduzida.

3.1. Redução de dimensionalidade

A redução de dimensionalidade é o processo de representar um dado com uma quantidade menor de atributos. Uma imagem é representada pelos valores de cada pixel e reduzir sua dimensionalidade implica em representá-la com menor quantidade de informação, remover redundâncias ou combinar seus valores, entre outras opções. O codificador do VAE faz uso da redução de dimensionalidade para tentar reduzir ao máximo os atributos que representam os dados sem perder a informação sobre seu conteúdo. A construção de um AE consiste em definir o melhor par de codificador e decodificador através de um processo de otimização iterativo. A cada iteração, o modelo é treinado propagando-se dados através do codificador e tentando-se reproduzi-los pelo decodificador. A saída decodificada é comparada com os dados originais de entrada e o erro é retropropagado na arquitetura para atualizar os pesos das conexões.

3.2. Variational AutoEncoders

Os VAEs são semelhantemente aos *autoencoders* e também possuem um codificador e um decodificador. Entretanto, eles se diferem quanto à capacidade de gerarem dados novos. Os VAEs são modelos generativos que, para conseguirem realizar essa tarefa, necessitam ter o espaço latente o mais regular possível. Com o espaço latente regular, é possível definir valores para cada ponto dentro do espaço latente e decodificá-los gerando uma nova imagem. A regularização do espaço latente é garantida de forma explícita através de um termo regularizador utilizado na função de perda. Essa função possui um termo de reconstrução para tornar o esquema de codificação-decodificação mais eficiente e um termo regularizador para regularizar a organização do espaço latente. Este último termo é expresso através da divergência de Kullback-Leibler (KL), que mantém diversos vetores latentes gerados e evita ter representações muito diferentes para dados semelhantes. Em um VAE, as variáveis do espaço latente apresentam distribuição Gaussiana, mas com uma componente aleatória para evitar a geração dos mesmos conjuntos de dados de saída.

4. Materiais

Para a concepção de um modelo de VAE adequado à caracterização das texturas mamográficas, foram implementados modelos com diferentes arquiteturas aplicados a bases de dados de texturas de naturezas diversas.

O principal conjunto de dados utilizado neste estudo é a base de imagens mamográficas IRMA Oliveira, Gueld, Araújo, Ott, e Deserno (2008) que consiste em 5024 imagens de mamografias classificadas na escala BI-RADS. As imagens são recortes de mamografias de mamas esquerdas e direitas, nas posições Médio Lateral Oblíqua (MLO) e Crânio Caudal (CC). Cada imagem possui 128×128 pixels em tons de cinza. As 5024 imagens estão divididas igualmente entre as 4 classes BI-RADS, tendo 1256 para cada uma (Figura 1).

O segundo conjunto de dados utilizado neste trabalho foi o VisTex (Picard et al., 1995). Esse dataset foi necessário para validação visual das arquiteturas e do modelo de interpretabilidade, por possuir imagens cuja diferenciação de textura entre classes é mais perceptível. O conjunto de dados VisTex consiste em imagens de tamanhos variados de texturas diversas, das quais foram escolhidas imagens de cascas de árvores, tijolos, folhas e água. As amostras foram reformatadas para resolução de 128×128 pixels. A quantidade de imagens de cada classe nesse diretório público varia em torno de 10 imagens apenas. Dessa forma, foi realizado um aumento de dados através de transformações de rotação e espelhamento, bem como da adição de ruído, resultando em 600 imagens para cada classe. A Figura 2 mostra exemplos de imagens escolhidas dentre as 4 classes selecionadas do conjunto de dados.

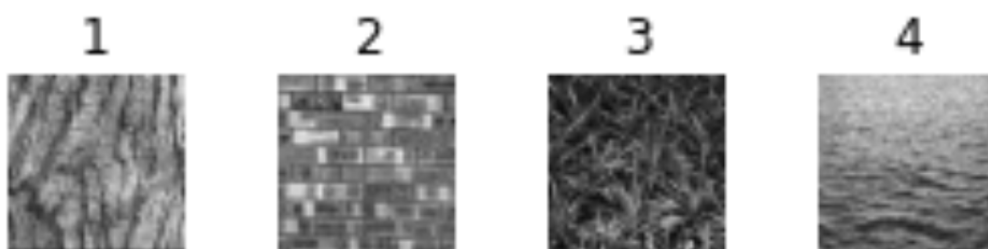


Figura 2. Exemplos de imagens de textura de casca de árvore (1), tijolos (2), folhas (3) e água (4) do conjunto VisTex.

Ambas as bases tiveram suas imagens separadas de forma aleatória em treino e teste, porém mantendo o balanceamento entre as classes, com a proporção de 70% e 30% respectivamente. Para garantir que em todos os métodos e modelos testados fossem usados as mesmas imagens para treino e teste, elas foram armazenadas em pastas específicas após a separação.

5. Metodologia

O TexVar é um autoencoder variacional projetado para ser eficiente com relação ao seu custo computacional e utilização de memória, mas com espaço latente capaz de representar de forma eficaz os elementos de textura presentes em imagens mamográficas. Sua concepção se baseia nas arquiteturas de VAEs clássicas, compostas de camadas convolucionais e densas, que inicialmente realizam o *downsampling* para a extração das variáveis em um espaço latente reduzido e em seguida realizam o *upsampling* para a produção das imagens de saída. A escolha dos hiper-parâmetros foi incremental, sendo o número de camadas convolucionais e densas, bem como o número de unidades em cada uma, variados de forma a se melhorar a qualidade visual das imagens reconstruídas. O tamanho

do espaço latente foi sucessivamente aumentado até atingir o número de 32 variáveis, capazes de reproduzir com maior fidelidade as texturas dos 2 conjuntos de dados utilizados nos experimentos. Dessa forma, a arquitetura final é composta por 2 camadas convolucionais com 128 e 96 filtros e 4 camadas densas totalmente conectadas no codificador, espaço latente com tamanho 32 e o decodificador com 4 camadas densas e 2 camadas convolucionais, um reflexo do codificador (Figura 3).

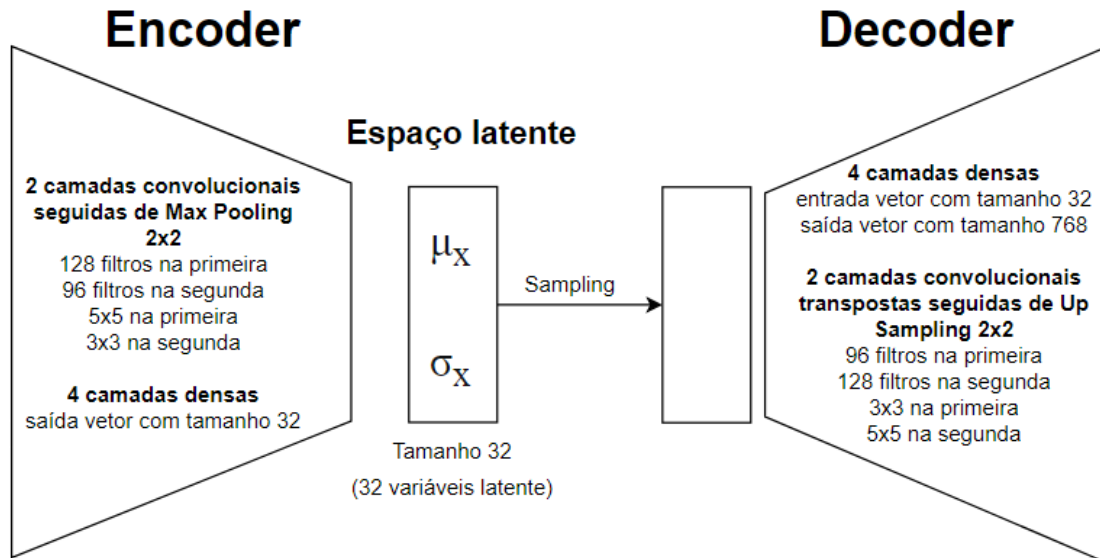


Figura 3. Arquitetura do TexVar.

O modelo de interpretabilidade proposto foi baseado no método de visualização do espaço latente apresentado por Kingma e Welling (2013), em que é utilizada a base MNIST e a arquitetura de um VAE contendo apenas 2 variáveis no espaço latente. Por se tratar de um espaço bidimensional, os valores das variáveis podem ser facilmente visualizados em um gráfico de dispersão. A arquitetura TexVar, no entanto, se difere por possuir 32 variáveis em seu espaço latente. Neste caso, a visualização deve ser feita através da seleção de 2 variáveis por vez, o que pode ser uma escolha difícil e gerar visualizações incompletas ou distorcidas. Não obstante, foram realizadas visualizações onde uma imagem de uma classe tem variáveis latentes manipuladas em direção a uma imagem de outra classe. Isso permite observar a maneira com a qual uma determinada variável altera a imagem, podendo interpretá-la como a caracterização de um elemento de textura.

A visualização do efeito das variáveis latentes na transformação da textura de uma classe em outra foi feita com o auxílio das médias e da matriz de covariância \mathbf{S} das variáveis. Foi calculada a distância de Mahalanobis entre os vetores com as médias para as duas classes, \mathbf{x} e \mathbf{y} , da forma:

$$D = \sqrt{(\mathbf{x} - \mathbf{y})^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{y})}. \quad (1)$$

Em seguida, escolhem-se as variáveis com maior valor de distância, por serem as mais capazes de diferenciar as classes.

6. Resultados

Os experimentos realizados para este estudo foram executados no ambiente Colab (Google). Os principais *frameworks* e bibliotecas usados foram o Tensorflow juntamente com o Keras para construção de toda arquitetura do VAE e o Matplotlib para geração das imagens e resultados.

Após o treinamento do VAE, a etapa de testes consistiu em submeter cada imagem do conjunto de teste ao codificador para gerar sua representação no espaço latente e produzir a imagem de saída reconstruída. As Figuras 4 e 5 apresentam exemplos de imagens de entrada e saída, respectivamente para as classes da base VisTex e BI-RADS. Observa-se que em ambos os datasets, o TexVar produziu imagens mais suavizadas que as originais, com menor definição de bordas, mas com características suficientes para a sua diferenciação. No caso das texturas mamográficas, observa-se a presença de elementos cada vez mais densos, quando as imagens passam da classe I para a IV.

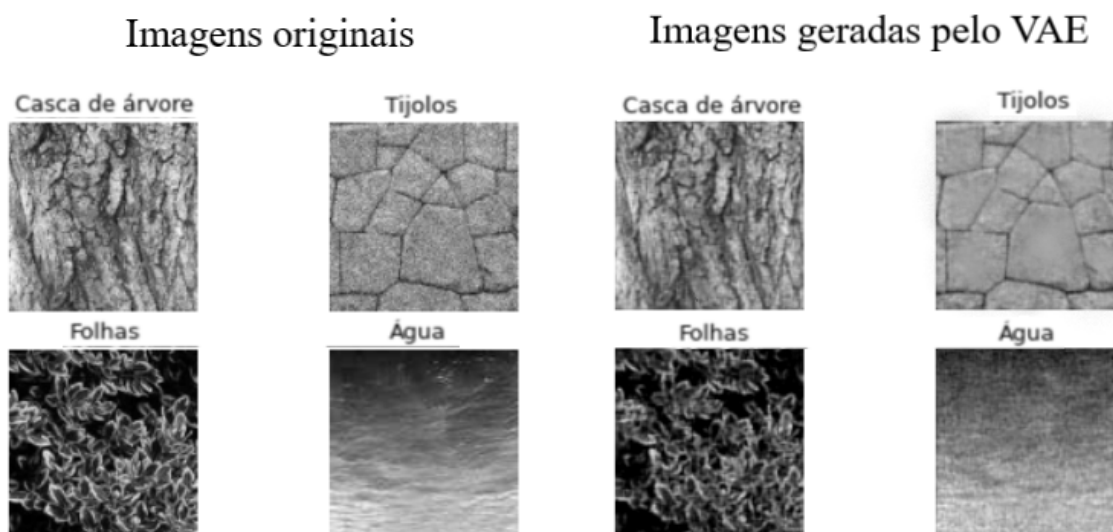


Figura 4. Exemplos de imagens geradas para cada classe VisTex com a arquitetura TexVar

A visualização e interpretação do significado das variáveis latentes foi feita variando-se o valor das variáveis de maior distância de Mahalanobis, enquanto as demais mantinham seus valores médios. Foram avaliadas as 5 variáveis com maior valor de distância, conforme relacionado nas Tabelas 1 e 2, respectivamente para as bases VisTex e BI-RADS. Apenas as 10 maiores distâncias são mostradas em ordem decrescente. Para a base BI-RADS, é possível perceber que apenas as 7 primeiras possuem valores maiores que 0, indicando que grande parte das variáveis não discriminam adequadamente as classes.

A Figura 6 apresenta gráficos de dispersão onde a distribuição das imagens quanto às variáveis do espaço latente podem ser observadas. Para as variáveis V7 e V2 da base VisTex, as amostras de classes se apresentam bem separadas, uma vez que as texturas correspondentes são bastante distintas. Já para a base BI-RADS, observa-se que as variáveis proporcionam menor capacidade de separação, tendo em vista que são classes cujas texturas fluem de forma suave da classe I para a IV. A melhor separação é obtida pela variável

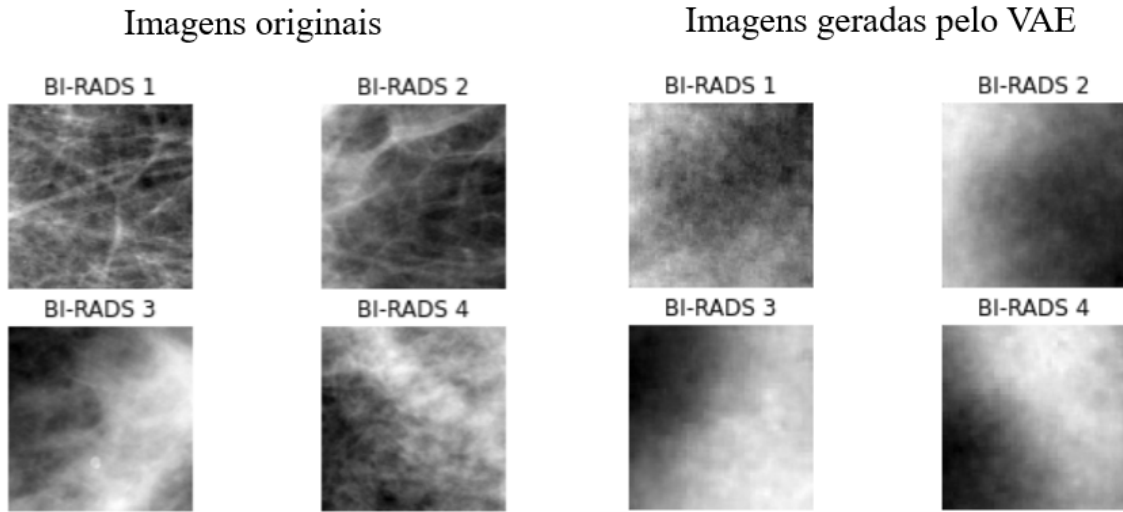


Figura 5. Exemplos de imagens geradas para cada classe BI-RADS com a arquitetura TexVar

V16 que posiciona a classe I na região superior do espaço e a classe IV na porção inferior, com as classes II e III na região intermediária.

| Variável | V2 | V24 | V1 | V7 | V13 | V11 | V30 | V16 | V20 | V4 |
|----------|------|------|------|------|------|------|------|------|------|------|
| DM | 8.66 | 6.32 | 6.12 | 5.79 | 3.89 | 3.19 | 2.50 | 2.29 | 2.10 | 1.73 |

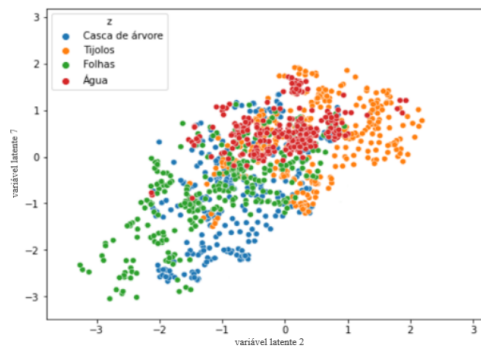
Tabela 1. Distâncias de Mahalanobis (DM) entre as classes Tronco de árvore e Tijolo para 10 variáveis mais discriminantes do espaço latente.

| Variável | V16 | V3 | V5 | V29 | V26 | V10 | V11 | V28 | V7 | V22 |
|----------|-------|------|------|------|------|------|------|------|------|------|
| DM | 13.15 | 9.31 | 6.81 | 6.03 | 4.44 | 3.83 | 3.78 | 0.45 | 0.22 | 0.22 |

Tabela 2. Distância de Mahalanobis (DM) entre as classes BI-RADS I e BI-RADS IV para as 10 variáveis mais discriminantes do espaço latente.

As variáveis com maiores valores de distância foram selecionadas para a geração de mapas de sequências de imagens que possibilitassem a interpretação das variáveis do espaço latente. Como esperado, nos testes com a base VisTex, as variáveis obtidas no espaço latente variaram de forma mais acentuada entre as classes. A Figura 7 exemplifica a transição de uma imagem da classe Casca de Árvore para a classe Tijolos a partir da variável latente com maior distância de Mahalanobis (V2). Observa-se que a primeira textura se descaracteriza gradativamente, passando a uma textura de menor granularidade, completamente distintas das 2, que depois se transforma na segunda textura. Por serem muito distintas, é difícil descrever verbalmente o significado da variável latente que deu origem à transformação. Já no caso da base BI-RADS, a variável latente manipulada (V16) causa uma transformação mais suave, levando gradativamente a textura da mama gordurosa à textura da mama densa, como mostrado na Figura 8). Outra característica interessante dessa transformação é que o canto superior direito da imagem, que é mais escuro na imagem da classe I permanece mais escuro na imagem da classe

Exemplo de variáveis na base VisTex



Exemplos de variáveis na base BI-RADS

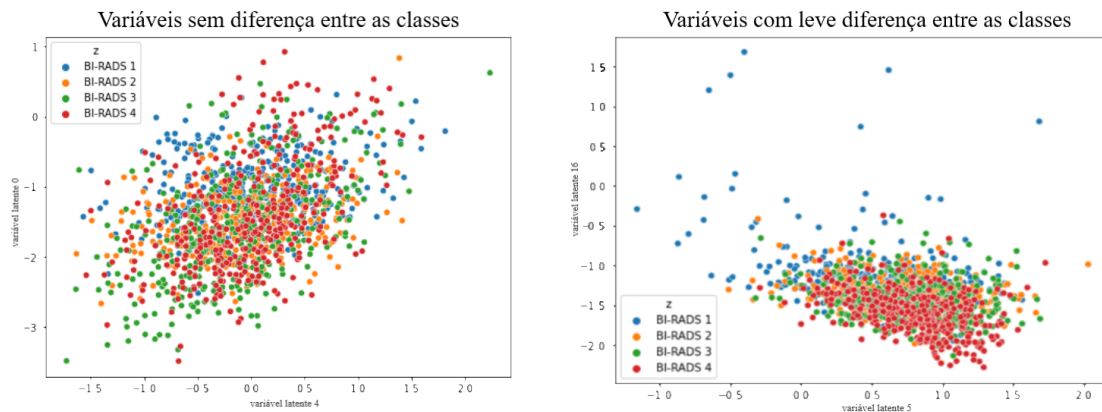


Figura 6. Visualização de variáveis do espaço latente. A imagem superior mostra a distribuição de classes para as variáveis V7 e V2 da base VisTex onde as classes estão bem separadas no espaço. A imagem da segunda linha à esquerda mostra um par de variáveis da base BI-RADS (variáveis V0 e V4) para as quais a separação das classes é baixa. A imagem da segunda linha à direita mostra a separação das classes através das variáveis V5 e V16. Neste caso a classe I ocupa a região superior do espaço enquanto a classe IV ocupa a porção inferior, com as classes II e III na região intermediária.

IV, indicando que a distribuição de tons de cinza não foi particularmente afetada pela transformação. Dessa forma, observa-se que a variável latente conseguiu abstrair especificamente a caracterização da densidade mamária de forma independente, a despeito da faixa de tons de cinza presentes na imagem.

7. Conclusão

Este artigo apresentou a proposta de um autoencoder variacional onde as variáveis do espaço latente são capazes de representar de forma interpretável as características de texturas em imagens. A eficácia do modelo foi exemplificada no problema da representação de texturas na escala BI-RADS, que são usadas para a classificação do risco de câncer de mama.

Uma vez que as fronteiras entre as classes mamográficas são tênues e mal definidas por natureza, uma segunda base de texturas, a VisTex, foi utilizada para validação do

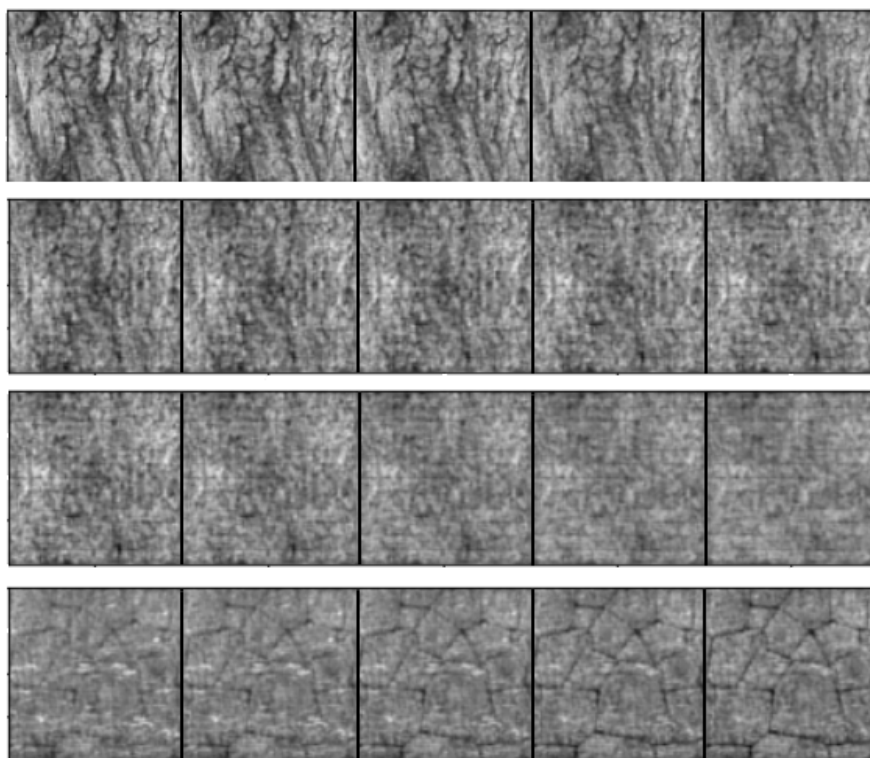


Figura 7. Transição de uma imagem da classe Casca de Árvore para a classe Tijolos, através da manipulação da variável latente V2.

modelo. O autoencoder foi capaz de recuperar grande parte das características das imagens de entrada, utilizando 32 variáveis latentes. As variáveis com maior discriminação entre classes foram investigadas com relação à sua capacidade de produzir imagens representativas de cada textura. O modelo mostrou ser capaz de produzir variáveis associadas à densidade mamária, gerando sequências de imagens que gradativamente caracterizaram a transição entre imagens de mamas gordurosas para densas.

Trabalhos futuros incluem a melhoria da arquitetura com objetivo de gerar imagens com melhor representação dos elementos de alta frequência, uma vez que a reconstrução das imagens pelo decodificador ainda gera imagens mais suavizadas que as originais. Um aspecto a ser investigado é a normalização das imagens de entrada, uma vez que a normalização de histogramas altera o contraste das imagens, o qual é uma das características de textura. Outra frente de investigação inclui critérios mais sistemáticos para o dimensionamento das variáveis latentes, ou formas de combinar seus efeitos, utilizando-se técnicas de análise multivariada como a Análise de Componentes Principais, o Implante Estocástico Vizinho Disperso ou outros métodos de redução de dimensionalidade.

Referências

- Habaragamuwa, H., Oishi, Y., e Tanaka, K. (2021). Achieving explainability for plant disease classification with disentangled variational autoencoders. *CoRR*, *abs/2102.03082*. Retrieved from <https://arxiv.org/abs/2102.03082>
- INCA. (2021, Set). *Estatísticas de câncer*. <https://www.inca.gov.br/numeros-de-cancer>. (acessado em: 15/10/2021)

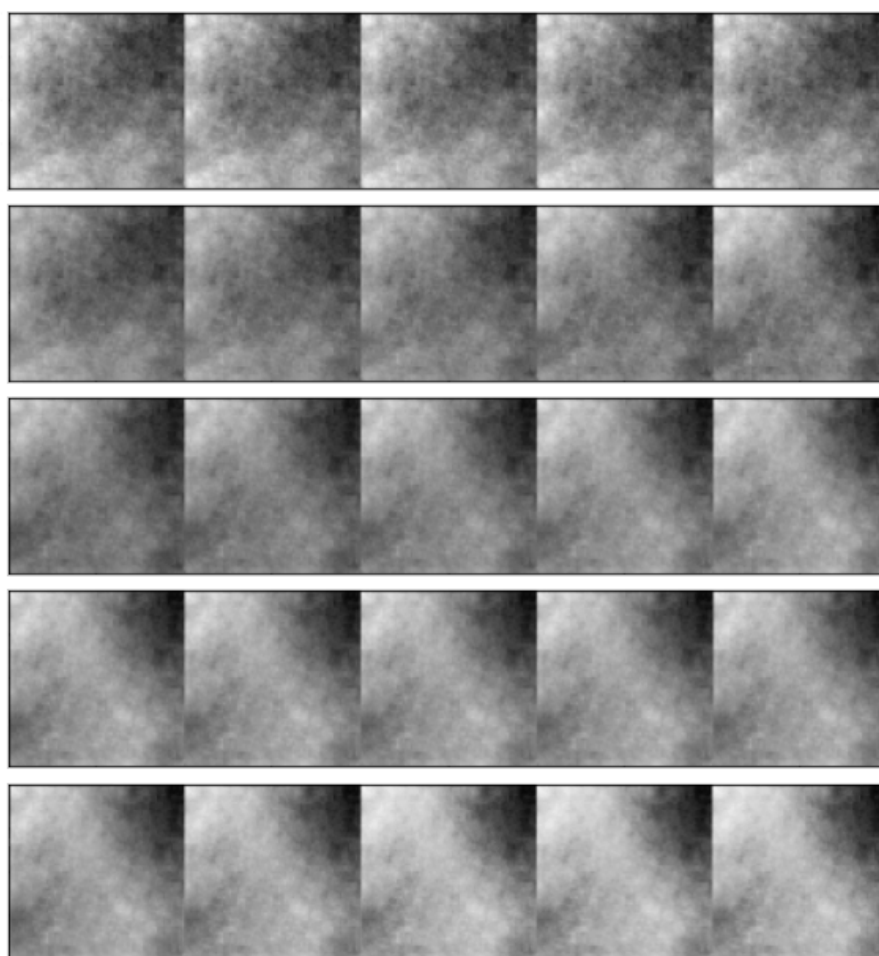


Figura 8. Transição de uma imagem da classe BI-RADS I para a classe BI-RADS IV, através da manipulação da variável latente V16.

- Kingma, D. P., e Welling, M. (2013). *Auto-encoding variational bayes*. arXiv. Retrieved from <https://arxiv.org/abs/1312.6114> doi: 10.48550/ARXIV.1312.6114
- Liu, W., Li, R., Zheng, M., Karanam, S., Wu, Z., Bhanu, B., ... Camps, O. (2020). Towards Visually Explaining Variational Autoencoders. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 8639–8648.
- Oliveira, J. E., Gueld, M. O., Araújo, A. d. A., Ott, B., e Deserno, T. M. (2008). Toward a standard reference database for computer-aided mammography. In *Medical imaging 2008: Computer-aided diagnosis* (Vol. 6915, p. 69151Y).
- Picard, R., Graczyk, C., Mann, S., Wachman, J., Picard, L., Negroponte, N., e Campbell, L. (1995, Mar). Retrieved from <https://vismod.media.mit.edu/pub/VisTex/>
- Schockaert, C., Macher, V., e Schmitz, A. (2020). VAE-LIME: Deep Generative Model Based Approach for Local Data-Driven Model Interpretability Applied to the Ironmaking Industry.
- Utkin, L., Drobintsev, P., Kovalev, M., e Konstantinov, A. (2021). Combining an auto-

encoder and a variational autoencoder for explaining the machine learning model predictions. *Conference of Open Innovation Association, FRUCT, 2021-Janua*. doi: 10.23919/FRUCT50888.2021.9347612

Uzunova, H., Ehrhardt, J., Kepp, T., e Handels, H. (2019). Interpretable explanations of black box classifiers applied on medical images by meaningful perturbations using variational autoencoders. In E. D. Angelini e B. A. Landman (Eds.), *Medical imaging 2019: Image processing* (Vol. 10949, pp. 264 – 271). SPIE. Retrieved from <https://doi.org/10.1117/12.2511964> doi: 10.1117/12.2511964

Wei, R., e Mahmood, A. (2021). Recent advances in variational autoencoders with representation learning for biomedical informatics: A survey. *IEEE Access*, 9, 4939-4956. doi: 10.1109/ACCESS.2020.3048309