

Comparação de extratores de características de textura para o diagnóstico de COVID-19 em imagens médicas

José Maria Clementino Junior ¹

¹Instituto De Ciências Matemáticas e de Computação - Universidade de São Paulo (ICMC -USP)
São Carlos SP – Brazil

juniorclementino@usp.br

Abstract. *The discovery of knowledge when applied in the context of medical images through techniques that involve image processing is of great importance to assist the medical professional in decision making. However, a large part of the methods focuses only on quantitative results. In this way, the whole process of functioning or presenting results is carried out by algorithms called black boxes that do not prioritize explainability. However, there are processing techniques, such as feature extraction, which are interested in extracting characteristic points of the image (color, texture, and shape). In addition to this representation providing good quantitative results, it also provides the professional with the entire results process. In this context, this final work of the course aimed to analyze the results provided by different feature extractors and by the combination of different multidimensional projections through interactive visual techniques. Six machine learning algorithms were used to perform the quantitative validation of the combinations performed for diagnosis*

Resumo. *A descoberta de conhecimento quando aplicada no contexto de imagens médicas por meio de técnicas que envolvem processamento de imagens é de grande importância para auxiliar o profissional médico na tomada de decisões. No entanto, grande parte dos métodos focam apenas em resultados quantitativos, desta forma todo processo de funcionamento ou apresentação dos resultados são feitas por algoritmos denominados caixas-pretas que não priorizam a explicabilidade. No entanto, existem técnicas de processamento, como a extração de características, que tem interesse de extrair pontos característicos da imagem (cor, textura e forma) além desta representação fornecer bons resultados quantitativos também disponibiliza ao profissional todo o processo de como foram obtidos os resultados. Neste contexto, este trabalho final da disciplina, teve como objetivo analisar os resultados providos por diferentes extratores de características e pela combinação de diversas projeções multidimensionais por meio de técnicas visuais interativas. Para o diagnóstico foram utilizados seis algoritmos de aprendizado de máquina para realizar a validação quantitativa da combinações realizadas.*

1. Introdução

O uso de extratores de características são de grande importância na área de visão computacional, em específico, no processamento de imagens que tem como objetivo extrair informações de regiões ou um ponto interesse, que podem ser previamente selecionados.

Uma das principais áreas que pode favorecer-se com as técnicas de extração é a medicina, pois podem atuar em diversas etapas do processo de descoberta de conhecimento, tais como o pré-processamento (anotações, segmentação) até as tarefas de aprendizado. Essa abordagem pode ser de forma automática, como também de maneira semiautomática com a interação do profissional médico com objetivo de enriquecer e validar os resultados obtidos pelo algoritmos.

Neste trabalho, surgiu a oportunidade de explorar o domínio de imagens de raio X-de pulmão para predição de diagnóstico de COVID-19 em pacientes. Sabe-se por meio da literatura [Peng et al. 2010, Backes et al. 2010] que um dos melhores extratores para este tipo de imagem são os extratores de textura. São diversas as variedades de extratores de textura presente na literatura, as variações e combinações trazem uma dificuldade de analisar qual o melhor extrator para determinado tipo de imagem e diagnóstico. Sendo assim, fez-se necessário a aplicações de técnicas visuais interativas no processo de tomada de decisão, proporcionando ao especialista médico o comportamento e os resultados obtidos por cada extrator combinado com as projeções e os algoritmos de classificação.

O trabalho desenvolvido está organizado nas seguintes seções: na Seção 2 é detalhada metodologia e o detalhamento de cada etapa realizada. E na Seção 3 detalha a discussão dos resultados e as considerações relevantes do trabalho.

2. Visão Geral do Trabalho

2.1. Data Set

A primeira base de dados disponibilizada pelo Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (HC FMRP/USP) apresenta imagens no formato DICOM que não sofreram nenhum tipo de alteração em relação à resolução e ao conteúdo interno da imagem. A base possui no total 890 imagens divididas em 3 classes, sendo elas 199 imagens de COVID, 381 Normal e 310 Intersticiais-não-covid, se tratando de uma base não estruturada onde suas imagens se diferem entre 419 resoluções diferentes. Para o estudo de caso foram utilizadas 580 imagens referente à soma entre a classe COVID e Normal, nas quais as imagens foram convertidas em extensão PNG e a resolução reduzida em 50% para ser realizado as extrações de características.

A segunda base de dados é pública, fornecida pela Kaggle (Link). A versão utilizada se trata de uma base de imagens padronizada na qual possui no total 21.164 imagens dividida em 4 classes sendo elas 3615 COVID, 10192 Normal, 1345 Viral-pneumonia e 6012 Lung-Opacity, todas no formato JPG com resolução de 299 x 299 pixels. Para o estudo de caso foram utilizadas 250 imagens da classe COVID e 250 imagens da classe Normal escolhidas de forma aleatória, perfazendo um total de 500 imagens.

Para apresentação e discussão das próximas etapas é considerada apenas primeira base de dados do (HC FMRP/USP), porém é importante salientar que o processo para obtenção do conhecimento é o mesmo para ambas as bases.

2.2. Etapa 1 - Preparação dos Dados

Nesta etapa, três atividades principais foram realizadas: (i) Conversão dos dados, (ii) Extração das características, e (iii) Representação dos dados.

A **Conversão dos dados (i)** é referente a conversão das imagens que estavam no formato *Digital Imaging and Communications in Medicine* (DICOM) com diferentes resoluções para o formato PNG padronizando a resolução de todas as imagens para 299x299 pixels.

Extração de Características (ii): A extração de característica é um método de captura do conteúdo visual de uma imagem para indexação e recuperação em bases de dados. Este processo gera uma vetor de características que funciona como uma assinatura da imagem, esta informação é relevante para resolver a tarefa computacional relacionado a uma determinada aplicação [Haralick-1973]. Existem diversos padrões de extração de características, em relação ao domínio de imagens médicas do tipo raio-x pulmonar, a literatura tem mostrado que os extratores de textura têm tido as melhores precisões para medir similaridade entre imagens, alguns trabalhos que utilizam este tipo de extração como base são [X] [Y]. Os extratores utilizados neste trabalho são descritos a seguir:

- **First Order Statistical (FOS):** Os características baseadas em estatísticas de primeira ordem que caracterizam textura são utilizados para classificação de imagens. Nesta técnica, não é considerado a informação de textura relacionado a distribuição de intensidade da vizinhança de cada pixel da imagem, e sim uma taxa de ocorrência de cada intensidade de pixel em uma imagem.
- **Gray Level Co-occurrence Matrix (GLCM):**Proposto por [Haralick-1973], o GLCM é um método estatístico para descrição de texturas, este método gera uma matriz de ordem $(Ng \times Ng)$ onde é computado a probabilidade de que dois valores de intensidade de cinza estejam envolvidos por uma determinada relação espacial. Desta matriz, diferentes medidas estatísticas são extraídas para caracterizar a textura de uma imagem, em um total de quatorze medidas apresentadas por Haralick. Como parâmetros desta técnica, é esperado uma distância d e um ângulo θ . Este método de análise de textura se preocupa com a detecção de aspectos de uma imagem que são invariantes a rotação, para isso as GLCM são calculadas extraindo a intensidade de pixel nos ângulos 0° , 45° , 90° e 135° graus com valor de distância de 1 do pixel central.
- **Local Phase Quantization (LPQ):**Proposto por [Ahonen-2008], o LPQ é um descritor de textura feito para superar o LBP em aplicações onde imagens estão afetadas por opacidade ou iluminação não uniforme, tem como fundamento a invariância de desfoque do espectro da fase de Fourier. Similar ao LBP, para cada pixel da imagem na posição $w = (x,y)$, um dado é armazenado para representar a textura de pixels em uma vizinha N_w centrada em w , de tamanho $M \times M$ pixels. A quantização de fase utiliza o sinal das componentes real e imaginária da Transformada de Fourier $F_w(u)$, $u = (u,v)$, em no mínimo quatro valores de frequências próximos ao centro de origem, são elas $(0,a)$, (a,a) , $(a, 0)$ e $(a,-a)$. São gerados 8 bits, que assumem o valor 0 ou 1 dependendo de cada valor ser negativo ou não-negativo. Após este procedimento, os valores são concatenados em uma ordem arbitrária formando um valor inteiro de 8 bits que representa a textura da imagem.

Após realizado a extração das características, a etapa de **Representação dos dados (iii)** realiza a formatação dos dados obtidos pelos extratores para que possam ser entregues as técnicas de visualização. A tabela 1 sumariza formato dos dados e as dimensionalidades geradas pelos extratores.

Índice	Imagem	featureFos	featureGlcm	featureLpq	Category
0	nomeImagem.png	list[...]	list[...]	list[...]	Class
		16	28	255	

2.3. Etapa 2 - Visualização dos Dados

Conforme citado anteriormente, os extratores entregam diferentes números de dimensões (features), como por exemplo a *featureLpq* que contem 255 dimensões, impossibilitando analisar e compreender dimensões elevadas, pois a percepção humana funciona bem para até três dimensões apenas. No entanto, a redução da dimensionalidade pode facilitar a análise dos dados por meio de projeções multidimensionais. Essas técnicas são eficientes para projetar as instâncias de dados do espaço original em um espaço de dimensão inferior ($R^n \rightarrow R^2$), que pode ser representado por várias técnicas de visualização, para exemplo, uma nuvem de pontos onde cada ponto representa uma imagem e sua assinatura é dada por uma coordenada 2D (x, y). Assim, supondo que a projeção retém a relação de distância entre os elementos no espaço de alta dimensão, a vizinhança visual dos pontos também representa suas semelhanças. Com isso as atividades podem ser exemplificadas da seguinte forma: (i) projeção multidimensional, (ii) funções de distância e (iii) visualização.

Alguns dos métodos mais usados de **projeções multidimensionais** (i) são *t-Distributed Stochastic Neighbor Embedding* (t-SNE), *Multidimensional Scaling* (MDS) e *Principal Component Analysis* (PCA). O método t-SNE [Van der Maaten and Hinton 2008] visa minimizar a divergência entre os dados no espaço original e o espaço projetado visa encontrar um conjunto de coordenadas de baixa dimensão que representam as relações entre os dados originais [Ritz et al. 2020]. É uma variação do *Stochastic Neighbour Embedding* (SNE) que fornece otimização adicional, reduzindo sobreposições e *clusters* no centro das coordenadas. Ao preservar a vizinhança, gera melhores visualizações. *Multidimensional scaling* (MDS) [Borg and Groenen 2005] é uma ferramenta para analisar similaridade em um conjunto de elementos. Esta técnica exhibe a estrutura de dados, modelando os dados como distâncias entre pontos em um espaço geométrico, exibindo informações essenciais e suavizando o ruído dos dados. PCA [Jolliffe 2002, Jolliffe 1986] pode ser empregado como uma ferramenta de análise exploratória que reduz a dimensionalidade dos dados por meio da identificação de combinações lineares ortogonais, conhecidas como componentes principais (PC) [Wang et al. 2020]. Os componentes principais representam as dimensões com maior variabilidade [Armeni et al. 2019]. Em outras palavras, eles identificam as dimensões mais descritivas no espaço do recurso, aquelas que têm a maior capacidade de representar o conteúdo do conjunto de dados.

Funções de distâncias (ii) são utilizadas para medir a dissimilaridade ou similaridade entre dois elementos de um conjunto, ou seja, o quão menor for a distância entre dois elementos mais similares eles são. Os cálculos são realizados em um espaço métrico onde conceito de distância entre elementos pode ser obtidos, quando uma função satisfaz as propriedades do espaço métrico, trata-se de uma distância métrica [Zezula-2006]. Há uma grande diversidade de funções de distâncias fornecidas pela literatura, as mais utilizadas são as funções da família *Minkowski*, entre elas a distância *Manhattan*, Euclidiana *Chebyshev*. Durante o desenvolvimento do trabalho foram testados diferentes parâmetros

que podem oferecer informações interessantes durante a exploração visual das imagens. Assim, foram utilizadas diferentes funções de distância, como a Euclidiana, *Manhattan* e Cosseno. Diferentes técnicas de projeção multidimensional (PCA, t-SNE e MDS), com o intuito de observar relacionamentos entre as imagens em um ambiente 2D, visto que tais técnicas são capazes de capturar estes relacionamentos em alta dimensionalidade e preservá-los em baixa dimensionalidade.

Para representar visualmente as projeções, usamos uma representação *Scatter Plots*, onde os valores de X e Y são os resultados de cada algoritmo de projeção multidimensional. As distâncias entre as imagens são armazenadas em uma matriz de distância simétrica com zeros na diagonal principal. Eles são calculados usando diferentes funções de distância como uma entrada para a projeção multidimensional, que obteve em duas dimensões (2D) as coordenadas que representavam tais relações de distância no espaço original (multidimensional) de *features*. A decisão da escolha foi de acordo com os critérios apresentados em sala de aula e presentes no *from Data to Viz* ¹ seguindo a seguinte ordem do guia de representação: *Numeric* → *SeveralNumeric* → *NotOrdered* → *Correlogram* → *Scatterplot*.

Na Figura 1 apresenta todas as combinações de extratores (FOS, GLCM e LPQ) com as diferentes funções de distancia (*euclidian*, *cosine* e *manhattan*) para a técnica de projeção *Principal Component Analysis* (PCA).

Na Figura 2 apresenta todas as combinações de extratores (FOS, GLCM e LPQ) com as diferentes funções de distancia (*euclidian*, *cosine* e *manhattan*) para a técnica de projeção *Multidimensional Scaling* (MDS).

Na Figura 3 apresenta todas as combinações de extratores (FOS, GLCM e LPQ) com as diferentes funções de distancia (*euclidian*, *cosine* e *manhattan*) para a técnica de projeção *Principal Component Analysis* (PCA).

Ambas os resultados das figuras apresentadas, podem auxiliar em diferentes análises, tais como:

- Analisar as diferenças entre as funções de distância utilizadas.
- Analisar a relação entre a projeção e os futuros resultados da classificação.
- Analisar a segregação visual dos grupos e os seus relacionamentos.
- Analisar as diferentes Extratores utilizadas.
- Analisar os diferentes Datasets utilizados.

Podendo levantar também algumas questões, para ser analisadas em parceria com especialistas do domínio:

- Qual a função de distância utilizar?
- Qual extrator de características (Feature) utilizar?
- Quais as combinações de técnicas utilizar em cada cenário?

¹<https://www.data-to-viz.com/>

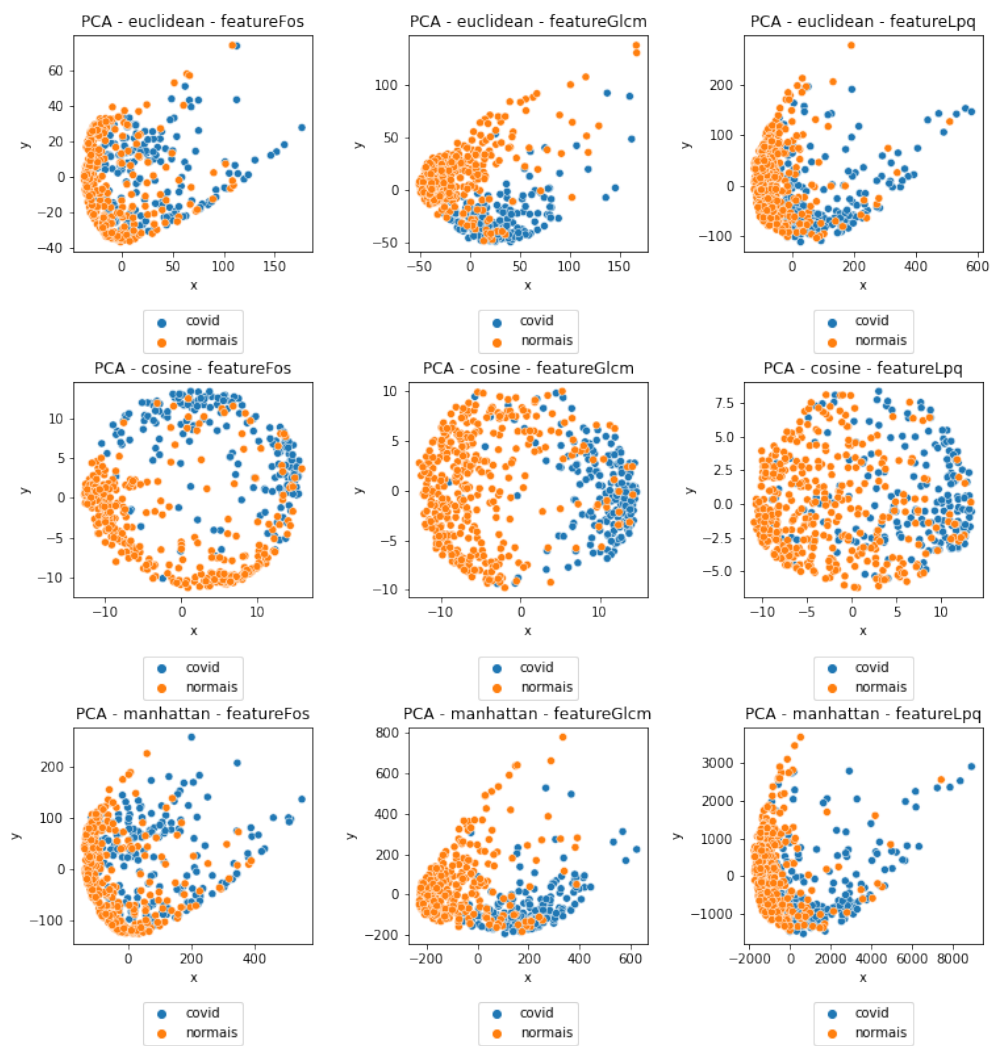


Figure 1. Projeção PCA: combinações entre os extratores e funções de distância

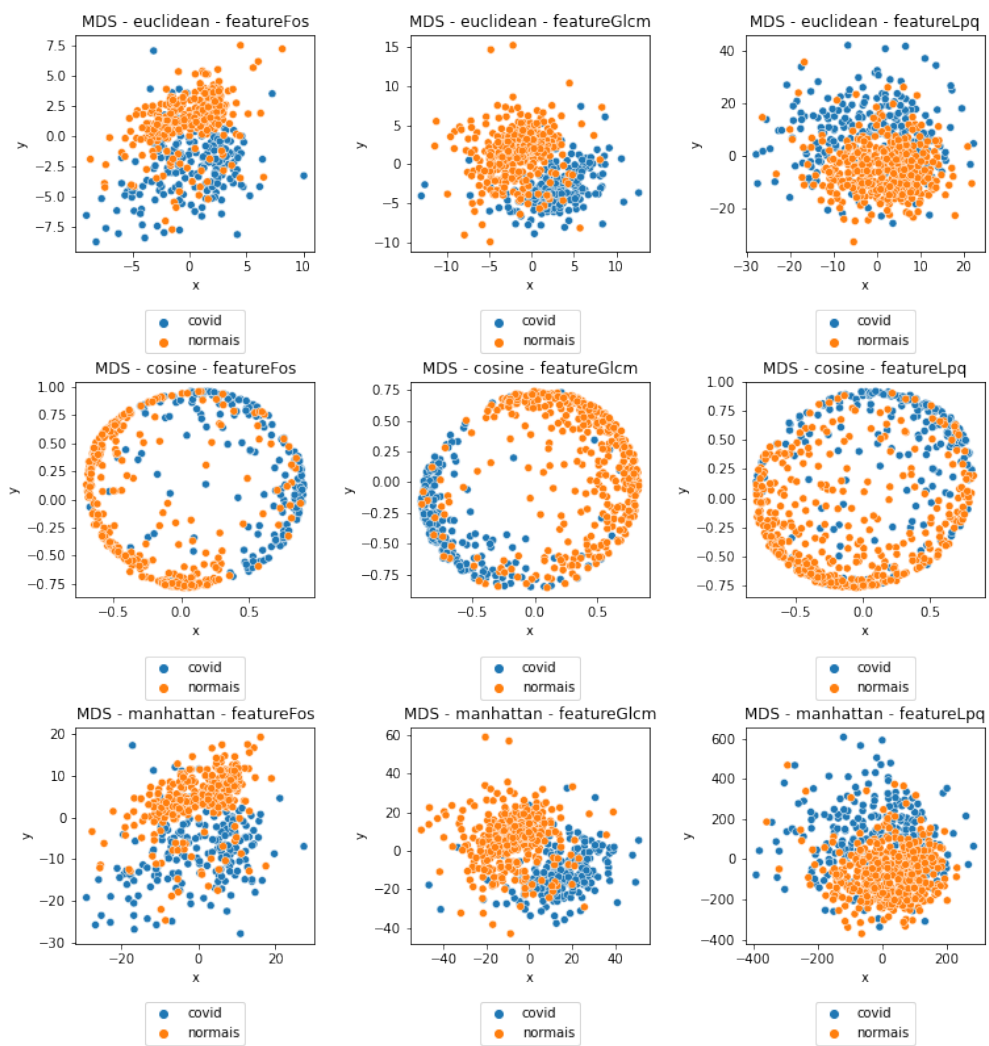


Figure 2. Projeção MDS: combinações entre os extratores e funções de distância

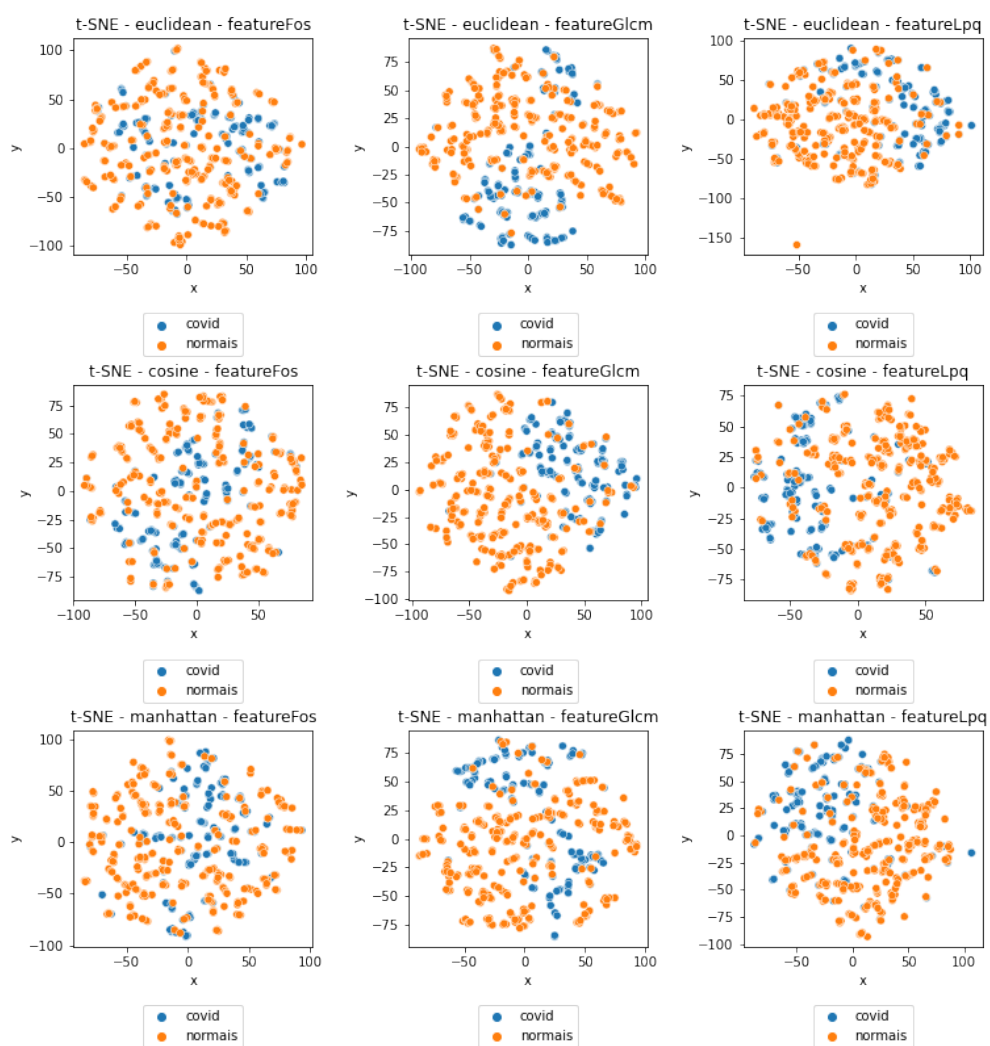


Figure 3. Projeção t-SNE: combinações entre os extratores e funções de distância

2.4. Etapa 3 - Visualização dos Dados Interativa

Com as diversas possibilidades de combinações de projeções multidimensionais, extratores de características e funções de distância somam 29 combinações possíveis. Apresentar todas essas informações de maneira estática dificultam a análise das informações. Com isso foi desenvolvido um sistema que permite a visualização interativa/dinâmica na qual o usuário possa escolher e visualizar as informações desejadas. O sistema interativo é dividido em 2 módulos principais: (1) Visualização e Exploração Imagem Única e (2) Visualização e Análise Quantitativa e Qualitativa das Projeções, que são detalhadas a seguir:

(1) Visualização e Exploração Imagem Única

- **1 (a) Histograma da Imagem:** Uma forma de representar uma imagem digital é por meio de um histograma, que pode ser definido como uma contagem/distribuição ou diagrama das frequências/estatísticas dos níveis de cinza do número de ("pixels"). Os histogramas podem ser utilizados para fins de reconhecimento de padrões. Algumas técnicas de processamento com histogramas: Ajuste de brilho e contraste; Equalização / Refinamento Imagem ;Segmentação; e Mapas de pseudocores. Na visualização permitimos que o usuário modifique o valor de **bins** do plot de maneira interativa. Os histogramas agregam dados numéricos em grupos de intervalos idênticos, denominados **bins** e exibem a frequência de valores dentro de cada bin. Proporcionando assim a visualização da distribuição da intensidade dos pixels na imagem.
- **1 (b) Visualização da Imagem:** é uma forma de visualizar imagens própria do *ipywidgets* em que no contexto de imagens de COVID é muito útil, pois possibilita de maneira interativa, zoom, demarcação e visualização de intensidade dos pixels. Podendo colaborar na visualização de fraturas, lesões caracterizadas covid entre outros achados importante para os especialista do domínio.

(2) **Visualização e Análise Quantitativa e Qualitativa das Projeções** A apresentação de todas as visualizações são alteradas automaticamente de acordo com a atribuições dos parâmetros selecionados pelo usuário, que são disponibilizados por meios do componente *Dropdown*.

- **2 (a) Data Frame:** Apresenta as informações tais como: nome da imagem, categoria (Covid ou Normal), tipo de distância, tipo de extrator de característica e o nome das projeções. A escolha pelo formato *DataFrame* foi a possibilidade do usuário visualizar a alterações dos parâmetros.
- **2 (b) Avaliação da Projeção: Silhouette :** Para avaliar as projeções fornecidas, foi utilizado a métrica de *Silhouette*, quanto mais próximo ao coeficiente (1) melhor é a distribuição resultante da projeção.
- **2 (c) Projeção Multidimensional:** é apresentado um *Scatter plot* das projeções em 2 dimensões denominadas (x,y) o gráfico apresenta a separação entre a categorias: Covid (cor roxo) e Normal (cor laranja) na visualização também é apresentada o nome de cada imagem (ponto) selecionado.
- **2 (d) Acurácia COVID-19 proveniente das projeções:** apresenta os resultados de acurácia de 8 classificadores *Nearest Neighbors*, *SVM Linear*, *SVM Polinomial grau 3*, *Decision Tree (gini)*, *Random Forest*, *Neural Net*, *AdaBoost*, *Naive Bayes* para as 27 combinações possíveis.

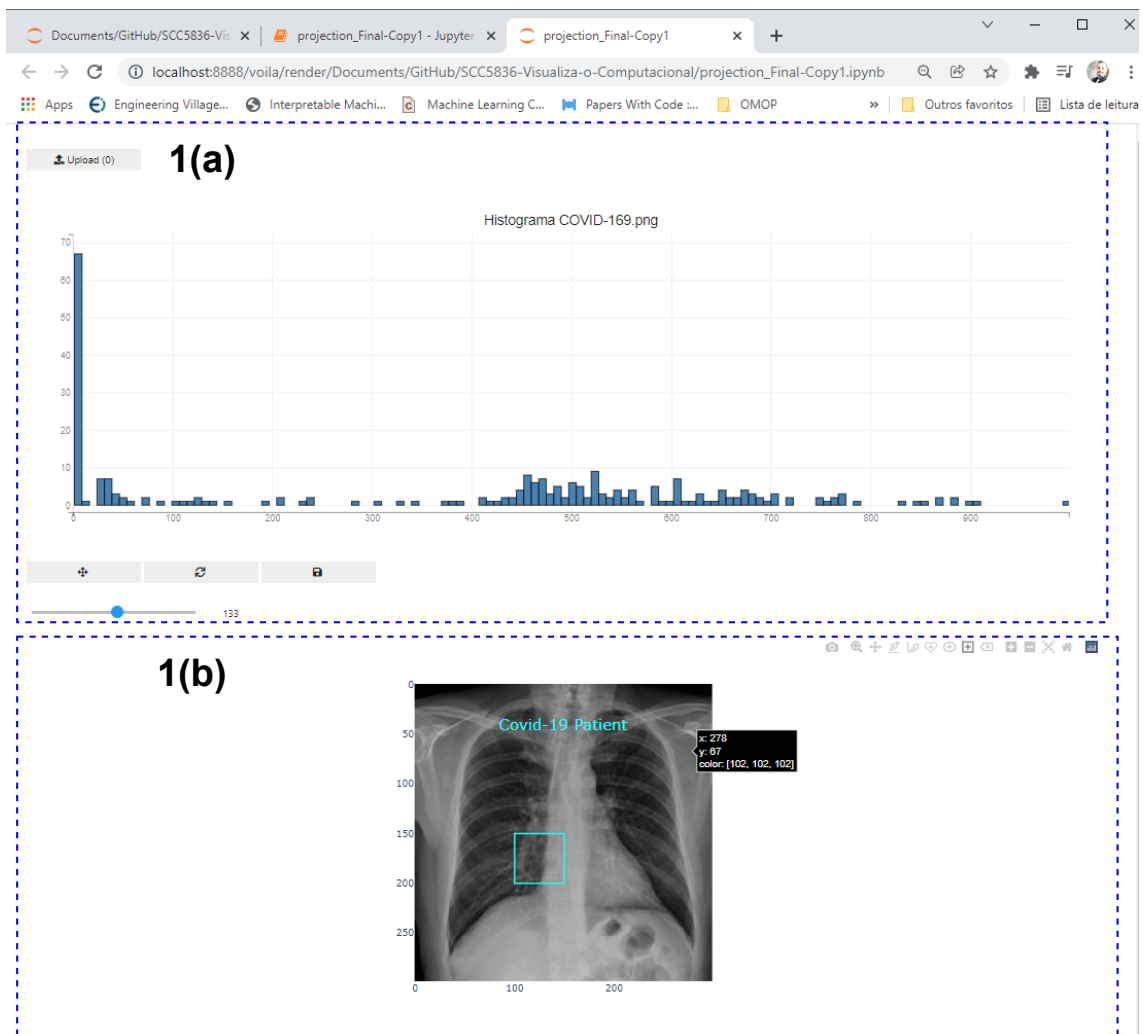


Figure 4. Visualização e Exploração Imagem Única

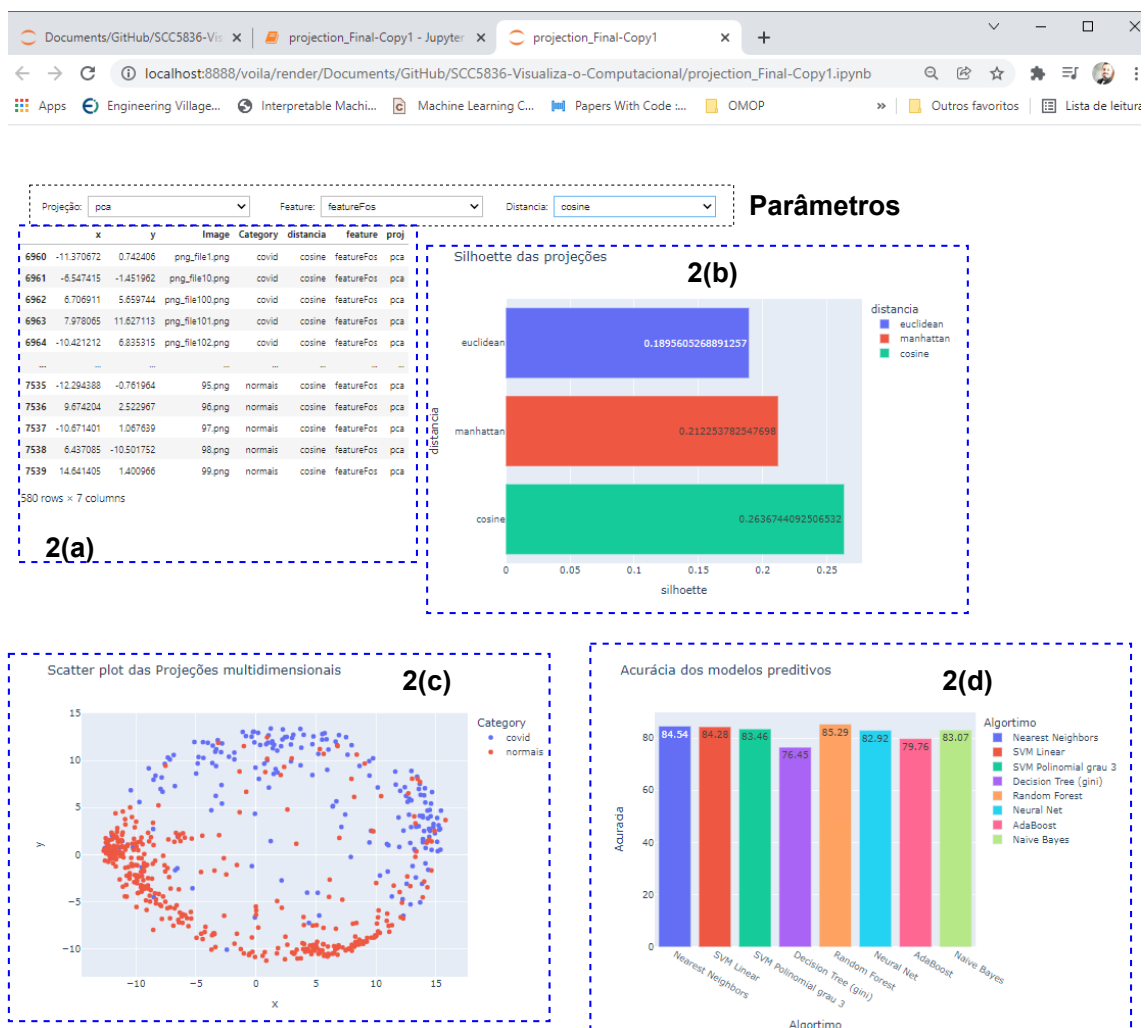


Figure 5. Visualização e Análise Quantitativa e Qualitativa das Projeções

3. Considerações

As técnicas interativas de visualização aplicadas, favorecem a compreensão dos especialistas de domínio, incorporando informações referentes (i) a exploração visual de uma única imagem, fornecendo a distribuição dos pixels e também uma ferramenta para auxiliar a visualização da imagem; (ii) analisar o comportamento de todas combinações de extratores de características, funções de distancias e projeções; (iii) proporciona também além de uma análise exploratória visual da distribuição das características, apresenta a técnica de validação das distribuições por meio da métrica da *Silhouette*; (iii) como uma forma de validação das projeções, é possível analisar os quantitativamente o desempenho dos algoritmos de classificação de acordo com as combinações utilizadas.

References

- Armeni, I., Sener, O., Zamir, A. R., Fischer, M., and Savarese, S. (2019). Systems and methods for performing three-dimensional semantic parsing of indoor spaces. US Patent 10,424,065.
- Backes, A. R., Casanova, D., and Bruno, O. M. (2010). A complex network-based approach for texture analysis. In *Iberoamerican Congress on Pattern Recognition*, pages 354–361. Springer.
- Borg, I. and Groenen, P. J. (2005). *Modern multidimensional scaling: Theory and applications*. Springer Science & Business Media.
- Jolliffe, I. (1986). *Principal Component Analysis*. Springer Verlag.
- Jolliffe, I. (2002). *Principal Component Analysis*. 2nd. ed. New York: Springer-Verlag.
- Peng, S.-H., Kim, D.-H., Lee, S.-L., and Lim, M.-K. (2010). Texture feature extraction based on a uniformity estimation method for local brightness and structure in chest ct images. *Computers in biology and medicine*, 40(11-12):931–942.
- Ritz, T., Baues, J., Krenkel, O., Schirmacher, P., and Longerich, T. (2020). Multispectral imaging to define morpho-molecular classes of human hcc. *Zeitschrift für Gastroenterologie*, 58(01):4–44.
- Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).
- Wang, Z. J., Wang, V. Y., Babarenda Gamage, T. P., Rajagopal, V., Cao, J. J., Nielsen, P. M., Bradley, C. P., Young, A. A., and Nash, M. P. (2020). Efficient estimation of load-free left ventricular geometry and passive myocardial properties using principal component analysis. *International Journal for Numerical Methods in Biomedical Engineering*, page e3313.