

## RESSALVA

Atendendo solicitação do autor, o texto completo desta dissertação será disponibilizado somente a partir de 18/12/2025.



UNIVERSIDADE ESTADUAL PAULISTA

“JÚLIO DE MESQUITA FILHO”

*INSTITUTO DE GEOCIÊNCIAS E CIÊNCIAS EXATAS*

Câmpus de Rio Claro

Programa de Pós-Graduação em Ciência da Computação

Deryk Willyan Biotto

**Transdução para Indução: Aprendizado Profundo de  
Representações em Cenários Não Supervisionados com  
Base em Informações de Listas Ranqueadas**

Rio Claro - SP

2025

UNIVERSIDADE ESTADUAL PAULISTA  
“Júlio de Mesquita Filho”  
Instituto de Geociências e Ciências Exatas  
Câmpus de Rio Claro

Deryk Willyan Biotto

**Transdução para Indução: Aprendizado Profundo de  
Representações em Cenários Não Supervisionados com  
Base em Informações de Listas Ranqueadas**

Dissertação de Mestrado apresentada ao Instituto de Geociências e Ciências Exatas do Câmpus de Rio Claro, da Universidade Estadual Paulista “Júlio de Mesquita Filho”, como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Denis Henrique Pinheiro Salvadeo

Rio Claro - SP

2025

B616t Biotto, Deryk Willyan  
Transdução para Indução: Aprendizado Profundo de Representações em Cenários Não Supervisionados com Base em Informações de Listas Ranqueadas / Deryk Willyan Biotto. -- Rio Claro, 2025  
58 p.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Instituto de Geociências e Ciências Exatas, Rio Claro  
Orientador: Denis Henrique Pinheiro Salvadeo

1. Aprendizado Métrico Profundo. 2. Recuperação de Imagens por Conteúdo. 3. Aprendizado Não Supervisionado. 4. Aprendizado Transdutivo. 5. Aprendizado Indutivo. I. Título.

UNIVERSIDADE ESTADUAL PAULISTA  
“Júlio de Mesquita Filho”  
Instituto de Geociências e Ciências Exatas  
Câmpus de Rio Claro

Deryk Willyan Biotto

**Transdução para Indução: Aprendizado Profundo de  
Representações em Cenários Não Supervisionados com  
Base em Informações de Listas Ranqueadas**

Dissertação de Mestrado apresentada ao Instituto de Geociências e Ciências Exatas do Câmpus de Rio Claro, da Universidade Estadual Paulista “Júlio de Mesquita Filho”, como parte dos requisitos para obtenção do título de Mestre em Ciência da Computação.

**Comissão Examinadora**

- Prof. Dr. Denis Henrique Pinheiro Salvadeo (Orientador)  
Instituto de Geociências e Ciências Exatas (IGCE)  
Universidade Estadual Paulista - UNESP
- Prof. Dr. Jurandy Gomes de Almeida Junior  
Universidade Federal de São Carlos - UFSCAR
- Prof. Dr. Danillo Roberto Pereira  
Faculdade de Ciências e Tecnologia (FCT)  
Universidade Estadual Paulista - UNESP

Conceito: Aprovado.

Rio Claro (SP), 18 de junho de 2025.

# Agradecimentos

Meu mais sincero agradecimento ao meu orientador, Denis Henrique Pinheiro Salvadeo, pela oportunidade e confiança em me orientar, pela paciência, amizade e incentivo.

Sou grato à equipe de pesquisa do PRISMA, especialmente aos colegas Emílio Bergamim Júnior, Davi Duarte de Paula, Jean Pierre Vargas e Vitor Privatto, pelas conversas produtivas e pelo compartilhamento de ideias que enriqueceram meu aprendizado.

Agradeço ainda à equipe de pesquisa SIAM, em especial aos pesquisadores Lucas Valem e Daniel Carlos Guimarães Pedronette pelas trocas enriquecedoras de ideias e experiências. Um agradecimento especial aos meus companheiros de mestrado, Gustavo Leticio e Vinicius Kawai, e aos colegas de pesquisa Bionda Rozin e Guilherme Jardim, pela amizade e apoio.

Registro ainda meu agradecimento ao pesquisador Sergio Antonio Cáceres Contreras, por ter sido o primeiro a me apresentar as redes neurais artificiais e seu imenso potencial, que têm sido a base e o principal elemento das minhas pesquisas, bem como a área na qual tenho me especializado.

Agradecimentos especiais à UNESP e ao professor Fabio Augusto Gomes Vieira Reis, líder do centro UNESPetro durante meu mestrado, por me permitir dedicar tempo e esforço a essa importante etapa da minha formação. Também expressei minha gratidão aos colegas de trabalho do UNESPetro Hermes, Sérgio Caetano, Ana, Márcia, Iata, José Maria, Dimas, Gabrielle, Pedro e Rafael, pela amizade, incentivo e apoio ao longo desta trajetória.

# Resumo

Recentemente, o uso de aprendizado profundo em cenários supervisionados consolidou-se significativamente. No entanto, há um crescente interesse em explorar métodos de aprendizado não supervisionado. Abordagens transdutivas são promissoras para aprender relações contextuais ricas em cenários não supervisionados, mas enfrentam desafios ao lidar com grandes quantidades de dados. O objetivo central deste estudo é investigar a viabilidade de um modelo indutivo baseado em Redes Perceptron Multicamadas (MLP) aprender a partir de listas ranqueadas geradas por métodos transdutivos em cenários não supervisionados. Foi proposto uma metodologia chamada IRL (*Inductive Ranking Learning*), que utiliza técnicas para aprender similaridades e dissimilaridades de pares obtidos de listas ranqueadas obtidas com métodos transdutivos. A técnica consiste em ponderar os elementos mais relevantes e irrelevantes para o cálculo do erro dos prováveis pares positivos e negativos, de acordo com a posição que o elemento está na lista ranqueada do seu respectivo par, permitindo o aprendizado sem rótulos. A abordagem proposta possibilita a utilização de técnicas transdutivas para treinar modelos indutivos, promovendo a escalabilidade e o uso dessas técnicas em diversas tarefas, representando um grande potencial para o aprendizado não supervisionado.

**Palavras-chave:** Aprendizado Métrico Profundo, Recuperação de Imagens por Conteúdo, Aprendizado Não Supervisionado, Aprendizado Transdutivo, Aprendizado Indutivo.

# Abstract

Recently, the use of deep learning in supervised scenarios has significantly consolidated. However, there is a growing interest in exploring unsupervised learning methods. Transductive approaches are promising for learning rich contextual relationships in unsupervised scenarios but face challenges when dealing with large amounts of data. The main objective of this study is to investigate the feasibility of an inductive model based on Multilayer Perceptron (MLP) networks learning from ranked lists generated by transductive methods in unsupervised scenarios. A methodology called IRL (Inductive Ranking Learning) is proposed, which utilizes techniques to learn similarities and dissimilarities of pairs obtained from ranked lists by transductive methods. The technique involves weighting the most relevant and irrelevant elements to calculate the error of likely positive and negative pairs, according to the element's position in the ranked list of its respective pair, enabling label-free learning. The proposed approach facilitates the use of transductive techniques to train inductive models, promoting scalability and the use of these techniques in various tasks, representing great potential for unsupervised learning.

**Keywords:** Deep Metric learning, Content-Based Image Retrieval, Unsupervised Learning, Transductive Learning, Inductive Learning.

# Lista de ilustrações

Figura 1 – Exemplo do uso de aprendizado métrico . . . . .	16
Figura 2 – Perceptron simples e perceptron com várias camadas ocultas. . . . .	17
Figura 3 – <i>triplet loss</i> responsável por maximizar a distância entre o âncora e o negativo. . . . .	18
Figura 4 – Resultado de re-ranqueamento . . . . .	21
Figura 5 – Exemplo de uma abordagem auto-supervisionada. . . . .	22
Figura 6 – Comparação de abordagens transdutivas e indutivas. . . . .	24
Figura 7 – Fluxo geral da abordagem proposta . . . . .	32
Figura 8 – Amostras de imagens do conjunto de dados Corel5K com gráficos de distâncias das listas ordenadas, usando características extraídas com o modelo ConvNeXt. . . . .	34
Figura 9 – Gráfico do vetor obtido a partir das distâncias médias das listas ordenadas do conjunto de dados Corel5k. . . . .	35
Figura 10 – Exemplo de uma abordagem transdutiva tradicional comparada ao modelo indutivo final do método proposto. . . . .	39
Figura 11 – Curvas Geradas pelas Diferentes Estratégias de Normalização para Construção da Escala de Ponderação . . . . .	44
Figura 12 – Distribuição com t-SNE das Representações Aprendidas com Diferentes Estratégias de Normalização no Conjunto Corel5K . . . . .	44
Figura 13 – Escalas de Referência Geradas por Diferentes Medidas de Tendência Central no Conjunto Corel5k . . . . .	46
Figura 14 – Distribuição com t-SNE mostrando o impacto do método IRL com LHR no conjunto de dados Corel5k . . . . .	47

# Lista de tabelas

Tabela 1 – Resultados das Características DINOv2 . . . . .	41
Tabela 2 – Resultados das Características ConvNeXt . . . . .	42
Tabela 3 – Resultados das Estratégias de Normalização com o Conjunto Corel5K .	44
Tabela 4 – Comparativo de Medidas de Tendência Central na Construção da Escala de Referência no Conjunto Corel5k . . . . .	46

# Lista de abreviaturas e siglas

CBIR	<i>Content-Based Image Retrieval</i>
CNN	<i>Convolutional Neural Networks</i>
FC	<i>fully connected</i>
FFN	<i>feed-forward network</i>
ML	<i>Machine learning</i>
MLP	<i>Multilayer Perceptron</i>
DML	<i>Deep Metric Learning</i>
SSL	<i>Self-Supervised Learning</i>
IRL	<i>Inductive Ranking Learning</i>

# Lista de símbolos

$\mathcal{L}_c$	Função Contrastive Loss.
$\mathcal{L}_{\text{triplet}}$	Função Triplet Loss.
$\mathcal{L}_{\text{cr}}$	Função Contrastive Ranking Loss.

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>15</b>
<b>2.1</b>	<b>Aprendizado de Máquina</b>	<b>15</b>
2.1.1	Aprendizado Métrico	15
2.1.2	Aprendizado Profundo	16
2.1.3	Aprendizado Métrico Profundo	17
<b>2.2</b>	<b>Aprendizado Não Supervisionado</b>	<b>19</b>
2.2.1	Baseados em Agrupamento	19
2.2.2	Redução de Dimensionalidade	20
2.2.3	Baseados em Recuperação de Imagens pelo Conteúdo	20
2.2.4	Auto-Supervisionado	21
2.2.4.1	Aprendizado Contrastivo	22
2.2.4.2	Professor-Estudante	23
<b>2.3</b>	<b>Aprendizado Transdutivo e Indutivo</b>	<b>24</b>
<b>3</b>	<b>TRABALHOS RELACIONADOS</b>	<b>26</b>
<b>3.1</b>	<b>Abordagens Híbridas Transdutivas com Indutivas</b>	<b>26</b>
<b>3.2</b>	<b>Ranqueamento com Aprendizado Profundo</b>	<b>27</b>
<b>3.3</b>	<b>Considerações Adicionais e Reflexão</b>	<b>28</b>
<b>4</b>	<b>MÉTODO PROPOSTO</b>	<b>31</b>
<b>4.1</b>	<b>Extração de Características</b>	<b>32</b>
<b>4.2</b>	<b>Listas Ordenadas e Escalas de Distância</b>	<b>33</b>
<b>4.3</b>	<b>Escala Normalizada para Referência de Ponderação</b>	<b>34</b>
<b>4.4</b>	<b>Ranqueamento com Inferência Transdutiva</b>	<b>36</b>
<b>4.5</b>	<b>Treinamento Indutivo com <i>Contrastive Ranking Loss</i></b>	<b>37</b>
<b>5</b>	<b>EXPERIMENTOS E RESULTADOS</b>	<b>40</b>
<b>5.1</b>	<b>Conjuntos de Dados</b>	<b>40</b>
<b>5.2</b>	<b>Protocolo Experimental</b>	<b>40</b>
<b>5.3</b>	<b>Resultados</b>	<b>41</b>
<b>5.4</b>	<b>Comparação com outras Estratégias para Ponderação</b>	<b>43</b>
5.4.1	Diferentes Funções de Normalização	43
5.4.2	Variantes Baseadas nas Distâncias das Listas Ranqueadas	45
<b>5.5</b>	<b>Análise e Discussão</b>	<b>46</b>

<b>6</b>	<b>CONCLUSÃO</b> . . . . .	<b>49</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>51</b>

# 1 Introdução

Recentemente, observou-se uma consolidação significativa do uso de aprendizado profundo (*deep learning*, DL) em cenários de aprendizado supervisionado (KRIZHEVSKY; SUTSKEVER; HINTON, 2012; SIMONYAN; ZISSERMAN, 2015; HE et al., 2016; DO-SOVITSKIY et al., 2021; LIU et al., 2021). Na conclusão do influente artigo de LeCun, Bengio e Hinton (2015), os autores destacaram a expectativa de que surgisse um interesse crescente por estudos na área de aprendizado não supervisionado. Uma das abordagens promissoras nesse sentido é o aprendizado auto-supervisionado, no qual diversas técnicas são empregadas para que modelos de redes neurais aprendam representações profundas a partir de grandes quantidades de dados não rotulados, criando seus próprios rótulos para o treinamento, aumentando assim a capacidade de generalização (CHEN et al., 2020; HE et al., 2020; GRILL et al., 2020; CHEN; HE, 2021; CARON et al., 2021).

Paralelamente, abordagens como técnicas de re-ranqueamento responsáveis por melhorar a precisão de listas ranqueadas, emergem como soluções eficazes para cenários não supervisionados e semi-supervisionados (VALEM; PEDRONETTE; ALMEIDA, 2018; LETICIO et al., 2023; VALEM; PEDRONETTE; LATECKI, 2023b). Tais técnicas são particularmente valorizadas pela capacidade de aprender informações contextuais dentro de um conjunto de dados, sendo frequentemente aplicadas em tarefas de recuperação de imagens baseadas em conteúdo (CBIR) (BAI et al., 2009; WEBBER; MOFFAT; ZOBEL, 2010; PEDRONETTE; TORRES, 2013; BAI; BAI, 2016). Além disso, estudos recentes têm aproveitado essas técnicas em tarefas de agrupamento, ou *clustering* (ROZIN et al., 2021; LOPES; PEDRONETTE, 2023). A principal vantagem desse tipo de abordagem reside na sua capacidade de capturar não apenas padrões dos exemplares, mas também suas relações contextuais (PEDRONETTE; TORRES, 2012; BAI; BAI, 2016).

No entanto, apesar do sucesso obtido com o uso do re-ranqueamento em agrupamento, o processo integral dessas abordagens permanece sendo transdutivo. Isso implica que, ao surgirem novos dados, é necessário repetir o processo de re-ranqueamento sobre todo o conjunto de dados, o que não permite o aproveitamento de aprendizados anteriores para novos dados. Isso pode se tornar um problema ainda maior em conjuntos de dados grandes e que estão sempre aumentando.

Este estudo investiga a capacidade de um modelo indutivo baseado em Redes Perceptron Multicamadas (MLP) (BAUM, 1988; CHEN et al., 2023) aprender a partir de listas ranqueadas geradas por métodos transdutivos, com ênfase especial nas técnicas de re-ranqueamento. O objetivo central é capacitar o modelo MLP a gerar representações de alta dimensão por meio de vetores mais discriminativos, aproveitando as relações contextuais

aprendidas com re-ranqueamento em cenários não supervisionados. Isso possibilita a realização de ranqueamento diretamente através dos vetores gerados pelo modelo treinado, agilizando tarefas de recuperação de imagens e tornando possível escalar essas tarefas para conjuntos de dados maiores. Além disso, esses vetores podem ser empregados em tarefas de classificação e clusterização, maximizando o uso da representação contextual adquirida após o treinamento indutivo com base nos resultados de um método transdutivo. O trabalho também apresenta um método desenvolvido denominado IRL (*Inductive Ranking Learning*) para esse processo de aprendizado indutivo, com resultados que indicam ganhos relevantes em tarefas de recuperação de imagens, superando modelos indutivos pré-treinados utilizados para extração de características e, em alguns casos, aproximando-se do desempenho obtido por abordagens transdutivas.

A motivação para este estudo foi responder a seguinte pergunta: É possível usar informações contextuais extraídas por métodos transdutivos como base para treinar modelos indutivos em tarefas de aprendizado não supervisionado? Espera-se que este estudo tenha o potencial de contribuir significativamente para o aprendizado de máquina em cenários não supervisionados. Ao possibilitar que modelos indutivos se beneficiem do conhecimento adquirido por métodos transdutivos, aumenta-se a relevância dos métodos transdutivos no cenário atual, elevando-os a um novo patamar. Além disso, contribui para que modelos indutivos aprendam com informações contextuais, aprimorando assim seu desempenho. Durante a elaboração deste trabalho, foi desenvolvido um artigo científico intitulado "Transduction to Induction: Self-Supervised Representation Learning Based on Ranking Information", atualmente submetido ao periódico *Neurocomputing* e em processo de decisão final após revisão por pares. Esse artigo apresenta resultados e ideias centrais exploradas nesta dissertação.

Os capítulos subsequentes deste trabalho estão organizados de forma a proporcionar um entendimento detalhado e progressivo sobre os temas abordados. No Capítulo 2, são apresentados os fundamentos teóricos que embasam este estudo, com destaque para conceitos essenciais de aprendizado de máquina, aprendizado métrico profundo e aprendizado não supervisionado, além de uma revisão sobre métodos transdutivos e indutivos. No Capítulo 3, são discutidos os trabalhos relacionados, com ênfase em abordagens híbridas entre transdução e indução, bem como técnicas de ranqueamento e aprendizado profundo. O Capítulo 4 detalha o método proposto, explicando as etapas de extração de características, construção de listas ordenadas, geração de escala normalizada e o treinamento indutivo utilizando a função *Contrastive Ranking Loss*. O Capítulo 5 apresenta a avaliação experimental, incluindo os conjuntos de dados utilizados, o protocolo experimental adotado e os resultados obtidos. Por fim, são discutidas as implicações dos resultados, seguidas das conclusões e possibilidades de trabalhos futuros no Capítulo 6.

## 6 Conclusão

A motivação inicial deste estudo foi investigar se informações contextuais extraídas por métodos transdutivos poderiam ser utilizadas como base para o treinamento de modelos indutivos em tarefas de aprendizado não supervisionado. Com os resultados deste estudo, é possível sugerir que métodos transdutivos podem, de fato, ser usados como base para modelos indutivos aprenderem representações mais discriminativas a partir de listas ranqueadas, aproveitando a rica informação contextual que essas listas fornecem. Até este estudo, nenhuma abordagem similar foi encontrada.

Importante destacar que o objetivo deste trabalho não foi propor um método mais competitivo que abordagens transdutivas tradicionais, tampouco competir com técnicas híbridas existentes. Em vez disso, a proposta consistiu em investigar um novo espaço de problema, no qual modelos indutivos possam se beneficiar das relações contextuais obtidas pelas listas ranqueadas geradas por métodos transdutivos. A proposta visa abrir caminho para abordagens alternativas que exploram essa interação, com foco na escalabilidade, generalização e reaproveitamento de aprendizado transdutivo em cenários não supervisionados.

O principal benefício deste trabalho se aplica às tarefas de Recuperação de Imagens Baseada em Conteúdo (CBIR). Ao demonstrar que redes neurais podem aprender a gerar representações que resultam em listas ranqueadas com relevância comparável aos métodos transdutivos, esse avanço tem o potencial de permitir a escalabilidade desses modelos para conjuntos de dados maiores, sem a necessidade de aplicar repetidamente o método transdutivo quando surgem novas amostras. Essa capacidade pavimenta o caminho para futuros estudos explorarem a aplicação dessa abordagem em cenários de grande escala, visando validar sua eficácia e eficiência em ambientes com grandes volumes de dados. Além disso, este estudo buscou aumentar a visibilidade e a importância dos métodos transdutivos, posicionando-os como um recurso chave para o aprendizado de representações.

Em comparação com métodos de pesquisas anteriores, que empregam técnicas de ranqueamento baseadas em aprendizado profundo ou adotam abordagens híbridas transdutivas e indutivas, o método proposto se destaca por aplicar aprendizado profundo utilizando informações de listas ranqueadas em um cenário não supervisionado. Além disso, ele é compatível com qualquer metodologia transdutiva capaz de gerar listas ranqueadas, oferecendo uma flexibilidade única. Outra vantagem significativa do método desenvolvido neste estudo é a geração de representações em vetores de alta dimensão, que podem ser reutilizadas em diversas outras tarefas, ampliando sua aplicabilidade e utilidade em diferentes contextos de aprendizado de máquina.

Embora a metodologia desenvolvida apresente limitações, houve um progresso significativo no desempenho do modelo indutivo em comparação com as características originais obtidas com modelos pré-treinados. No entanto, os resultados ainda não atingem, em todos os casos, o nível de eficácia observado nos métodos transdutivos. Há espaço para explorar novas direções, que incluem: investigar critérios alternativos de ponderação de erro, conceber funções de perda mais eficazes e robustas a lotes maiores, adotar arquiteturas de redes neurais mais sofisticadas, visto que neste estudo foi empregada apenas uma MLP, e substituir a amostragem aleatória por estratégias de seleção de amostras mais eficientes. Este estudo pretende servir como ponto de partida para futuras investigações no espaço de problema explorado: a transdução para indução.

A combinação de métodos transdutivos com aprendizado indutivo apresenta um potencial significativo para o avanço do aprendizado de máquina em cenários não supervisionados. A expectativa com os resultados desse estudo é que o interesse por esse tipo de abordagem cresça e que mais estudos sejam conduzidos para explorar seus benefícios.

# Referências

- ALZUBAIDI, L. et al. Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *Journal of big Data*, Springer, v. 8, p. 1–74, 2021. Citado 3 vezes nas páginas 16, 17 e 24.
- BAI, S.; BAI, X. Sparse contextual activation for efficient visual re-ranking. *IEEE Transactions on Image Processing*, IEEE, v. 25, n. 3, p. 1056–1069, 2016. Citado na página 13.
- BAI, X. et al. Learning context-sensitive shape similarity by graph transduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 32, n. 5, p. 861–874, 2009. Citado na página 13.
- BAUM, E. B. On the capabilities of multilayer perceptrons. *Journal of complexity*, Elsevier, v. 4, n. 3, p. 193–215, 1988. Citado 2 vezes nas páginas 13 e 37.
- BROMLEY, J. et al. Signature verification using a "siamese" time delay neural network. *Advances in neural information processing systems*, v. 6, 1993. Citado na página 18.
- CARON, M. et al. Emerging properties in self-supervised vision transformers. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. [S.l.: s.n.], 2021. p. 9650–9660. Citado 2 vezes nas páginas 13 e 23.
- CHALVIDAL, M.; SERRE, T.; VANRULLEN, R. Learning functional transduction. *Advances in Neural Information Processing Systems*, v. 36, 2024. Citado 2 vezes nas páginas 24 e 26.
- CHEN, J. et al. Unsupervised clustering using graph transduction. In: IEEE. *2010 International Conference on Biomedical Engineering and Computer Science*. [S.l.], 2010. p. 1–4. Citado na página 25.
- CHEN, S. et al. Cyclemlp: a mlp-like architecture for dense visual predictions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 45, n. 12, p. 14284–14300, 2023. Citado 2 vezes nas páginas 13 e 37.
- CHEN, T.; CHEN, H. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE Transactions on Neural Networks*, IEEE, v. 6, n. 4, p. 911–917, 1995. Citado na página 31.
- CHEN, T. et al. A simple framework for contrastive learning of visual representations. In: PMLR. *International Conference on Machine Learning*. [S.l.], 2020. p. 1597–1607. Citado 2 vezes nas páginas 13 e 23.
- CHEN, X.; HE, K. Exploring simple siamese representation learning. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2021. p. 15750–15758. Citado 2 vezes nas páginas 13 e 23.

- CHOI, R. Y. et al. Introduction to machine learning, neural networks, and deep learning. *Translational vision science & technology*, The Association for Research in Vision and Ophthalmology, v. 9, n. 2, p. 14–14, 2020. Citado na página 16.
- CHOPRA, S.; HADSELL, R.; LECUN, Y. Learning a similarity metric discriminatively, with application to face verification. In: IEEE. *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*. [S.l.], 2005. v. 1, p. 539–546. Citado 2 vezes nas páginas 18 e 37.
- CIANO, G. et al. On inductive–transductive learning with graph neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 44, n. 2, p. 758–769, 2021. Citado na página 26.
- COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE transactions on information theory*, IEEE, v. 13, n. 1, p. 21–27, 1967. Citado na página 24.
- DAMRICH, S. et al. From t-SNE to UMAP with contrastive learning. In: *The Eleventh International Conference on Learning Representations*. [S.l.: s.n.], 2023. Citado na página 20.
- DAYAN, P.; SAHANI, M.; DEBACK, G. Unsupervised learning. *The MIT encyclopedia of the cognitive sciences*, MIT Press, p. 857–859, 1999. Citado na página 19.
- DOSOVITSKIY, A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. In: *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. [S.l.: s.n.], 2021. Citado na página 13.
- ESPADOTO, M. et al. Toward a quantitative survey of dimension reduction techniques. *IEEE transactions on visualization and computer graphics*, IEEE, v. 27, n. 3, p. 2153–2173, 2019. Citado na página 20.
- EZUGWU, A. E. et al. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence*, Elsevier, v. 110, p. 104743, 2022. Citado na página 19.
- FU, Z. et al. Deep metric learning with self-supervised ranking. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2021. v. 35, n. 2, p. 1370–1378. Citado na página 28.
- GAMMERMAN, A.; VOVK, V.; VAPNIK, V. Learning by transduction. In: *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI1998)*. [S.l.: s.n.], 1998. p. 148–155. Citado na página 24.
- GAO, R. et al. Zero-vae-gan: Generating unseen features for generalized and transductive zero-shot learning. *IEEE Transactions on Image Processing*, IEEE, v. 29, p. 3665–3680, 2020. Citado na página 26.
- GKELIOS, S. et al. Deep convolutional features for image retrieval. *Expert Systems with Applications*, Elsevier, v. 177, p. 114940, 2021. Citado 2 vezes nas páginas 29 e 32.
- GLOBERSON, A.; ROWEIS, S. Metric learning by collapsing classes. *Advances in Neural Information Processing Systems*, v. 18, 2005. Citado na página 16.

- GOLDBERGER, J. et al. Neighbourhood components analysis. *Advances in Neural Information Processing Systems*, v. 17, 2004. Citado na página 16.
- GRILL, J.-B. et al. Bootstrap your own latent—a new approach to self-supervised learning. *Advances in neural information processing systems*, v. 33, p. 21271–21284, 2020. Citado 2 vezes nas páginas 13 e 23.
- GRIRA, N.; CRUCIANU, M.; BOUJEMAA, N. Unsupervised and semi-supervised clustering: a brief survey. *A review of machine learning techniques for processing multimedia content*, Citeseer, v. 1, n. 2004, p. 9–16, 2004. Citado na página 19.
- GUI, J. et al. A survey on self-supervised learning: Algorithms, applications, and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1–20, 2024. Citado 2 vezes nas páginas 22 e 23.
- HADSELL, R.; CHOPRA, S.; LECUN, Y. Dimensionality reduction by learning an invariant mapping. In: IEEE. *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*. [S.l.], 2006. v. 2, p. 1735–1742. Citado 2 vezes nas páginas 18 e 37.
- HAMILTON, W.; YING, Z.; LESKOVEC, J. Inductive representation learning on large graphs. *Advances in neural information processing systems*, v. 30, 2017. Citado na página 27.
- HE, K. et al. Momentum contrast for unsupervised visual representation learning. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 9729–9738. Citado 2 vezes nas páginas 13 e 23.
- HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2016. p. 770–778. Citado 2 vezes nas páginas 13 e 17.
- HENAFF, O. Data-efficient image recognition with contrastive predictive coding. In: III, H. D.; SINGH, A. (Ed.). *Proceedings of the 37th International Conference on Machine Learning*. [S.l.]: PMLR, 2020. (Proceedings of Machine Learning Research, v. 119), p. 4182–4192. Citado na página 23.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural networks*, Elsevier, v. 2, n. 5, p. 359–366, 1989. Citado 2 vezes nas páginas 31 e 37.
- JIA, W. et al. Feature dimensionality reduction: a review. *Complex & Intelligent Systems*, Springer, v. 8, n. 3, p. 2663–2693, 2022. Citado na página 20.
- JIN, P. et al. Quantifying the generalization error in deep learning in terms of data distribution and neural network smoothness. *Neural Networks*, Elsevier, v. 130, p. 85–99, 2020. Citado na página 24.
- KARADERI, T. et al. Visual microfossil identification via deep metric learning. In: SPRINGER. *International Conference on Pattern Recognition and Artificial Intelligence*. [S.l.], 2022. p. 34–46. Citado na página 18.
- KAYA, M.; BILGE, H. Ş. Deep metric learning: A survey. *Symmetry*, MDPI, v. 11, n. 9, p. 1066, 2019. Citado na página 18.

- KHAN, M. M. R. et al. Adbscan: Adaptive density-based spatial clustering of applications with noise for identifying clusters with varying densities. In: IEEE. *2018 4th international conference on electrical engineering and information & communication technology (iCEEICT)*. [S.l.], 2018. p. 107–111. Citado na página 19.
- KHOSLA, A. et al. Novel dataset for fine-grained image categorization: Stanford dogs. In: *Proc. CVPR workshop on fine-grained visual categorization (FGVC)*. [S.l.: s.n.], 2011. v. 2, n. 1. Citado na página 40.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, v. 25, 2012. Citado 2 vezes nas páginas 13 e 17.
- KULIS, B. et al. Metric learning: A survey. *Foundations and Trends® in Machine Learning*, Now Publishers, Inc., v. 5, n. 4, p. 287–364, 2013. Citado na página 15.
- LACHAUD, G.; CONDE-CESPEDES, P.; TROCAN, M. Comparison between inductive and transductive learning in a real citation network using graph neural networks. In: IEEE. *2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. [S.l.], 2022. p. 534–540. Citado na página 25.
- LAO, G. et al. Three degree binary graph and shortest edge clustering for re-ranking in multi-feature image retrieval. *Journal of Visual Communication and Image Representation*, Elsevier, v. 80, p. 103282, 2021. Citado na página 29.
- LATECKI, L. J.; LAKAMPER, R.; ECKHARDT, T. Shape descriptors for non-rigid shapes with a single closed contour. In: IEEE. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*. [S.l.], 2000. v. 1, p. 424–429. Citado na página 21.
- LE-KHAC, P. H.; HEALY, G.; SMEATON, A. F. Contrastive representation learning: A framework and review. *IEEE Access*, IEEE, v. 8, p. 193907–193934, 2020. Citado na página 19.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Citado 2 vezes nas páginas 13 e 16.
- LECUN, Y. et al. Backpropagation applied to handwritten zip code recognition. *Neural computation*, MIT Press, v. 1, n. 4, p. 541–551, 1989. Citado na página 17.
- LETICIO, G. et al. pyudlf: A python framework for unsupervised distance learning tasks. In: *Proceedings of the 31st ACM International Conference on Multimedia*. [S.l.: s.n.], 2023. p. 9680–9684. Citado na página 13.
- LETICIO, G. R. et al. Manifold information through neighbor embedding projection for image retrieval. *Pattern Recognition Letters*, Elsevier, 2024. Citado na página 20.
- LIU, G.-H.; YANG, J.-Y. Content-based image retrieval using color difference histogram. *Pattern recognition*, Elsevier, v. 46, n. 1, p. 188–198, 2013. Citado na página 40.
- LIU, H. et al. Pay attention to mlps. *Advances in neural information processing systems*, v. 34, p. 9204–9215, 2021. Citado na página 37.

- LIU, Y. et al. Learning to propagate labels: Transductive propagation network for few-shot learning. In: *International Conference on Learning Representations*. [s.n.], 2019. Disponível em: <<https://openreview.net/forum?id=SyVuRiC5K7>>. Citado na página 26.
- LIU, Z. et al. Swin transformer: Hierarchical vision transformer using shifted windows. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2021. p. 10012–10022. Citado na página 13.
- LIU, Z. et al. A convnet for the 2020s. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2022. p. 11976–11986. Citado na página 40.
- LOPES, L. T.; PEDRONETTE, D. C. G. Self-supervised clustering based on manifold learning and graph convolutional networks. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. [S.l.: s.n.], 2023. p. 5634–5643. Citado na página 13.
- LOSHCHILOV, I.; HUTTER, F. Decoupled weight decay regularization. In: *International Conference on Learning Representations*. [S.l.: s.n.], 2019. Citado na página 41.
- LU, J.; HU, J.; ZHOU, J. Deep metric learning for visual understanding: An overview of recent advances. *IEEE Signal Processing Magazine*, IEEE, v. 34, n. 6, p. 76–84, 2017. Citado 2 vezes nas páginas 17 e 18.
- LU, L. et al. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, v. 3, p. 218–229, 03 2021. Citado na página 31.
- MAATEN, L. van der; HINTON, G. Visualizing data using t-sne. *Journal of Machine Learning Research*, v. 9, n. 86, p. 2579–2605, 2008. Citado na página 20.
- MAJI, S.; BOSE, S. Cbir using features derived by deep learning. *ACM/IMS Transactions on Data Science (TDS)*, ACM New York, NY, v. 2, n. 3, p. 1–24, 2021. Citado na página 29.
- MANNING, C. D. *Introduction to information retrieval*. [S.l.]: Syngress Publishing,, 2008. Citado na página 41.
- MCCLOSKEY, M.; COHEN, N. J. Catastrophic interference in connectionist networks: The sequential learning problem. In: *Psychology of learning and motivation*. [S.l.]: Elsevier, 1989. v. 24, p. 109–165. Citado na página 25.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, p. 115–133, 1943. Citado na página 17.
- MCINNES, L. et al. Umap: Uniform manifold approximation and projection. *Journal of Open Source Software*, The Open Journal, v. 3, n. 29, p. 861, 2018. Disponível em: <<https://doi.org/10.21105/joss.00861>>. Citado 3 vezes nas páginas 20, 36 e 41.
- MOHAMED, O. et al. Content-based image retrieval using convolutional neural networks. In: SPRINGER. *Lecture notes in real-time intelligent systems*. [S.l.], 2019. p. 463–476. Citado na página 29.

- MOHAN, D. D. et al. Deep metric learning for computer vision: A brief overview. *Handbook of Statistics*, Elsevier, v. 48, p. 59–79, 2023. Citado 2 vezes nas páginas 17 e 18.
- MOORE, B. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE transactions on automatic control*, IEEE, v. 26, n. 1, p. 17–32, 1981. Citado na página 20.
- MUNDHENK, T. N.; HO, D.; CHEN, B. Y. Improvements to context based self-supervised learning. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 9339–9348. Citado na página 22.
- OQUAB, M. et al. DINOv2: Learning robust visual features without supervision. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856. Citado na página 40.
- PARKHI, O. M. et al. Cats and dogs. In: IEEE. *2012 IEEE conference on computer vision and pattern recognition*. [S.l.], 2012. p. 3498–3505. Citado na página 40.
- PEDRONETTE, D. C. G.; ALMEIDA, J.; TORRES, R. d. S. A scalable re-ranking method for content-based image retrieval. *Information sciences*, Elsevier, v. 265, p. 91–104, 2014. Citado na página 29.
- PEDRONETTE, D. C. G.; TORRES, R. d. S. Exploiting contextual information for image re-ranking and rank aggregation. *International Journal of Multimedia Information Retrieval*, Springer, v. 1, p. 115–128, 2012. Citado na página 13.
- PEDRONETTE, D. C. G.; TORRES, R. d. S. Image re-ranking and rank aggregation based on similarity of ranked lists. *Pattern Recognition*, Elsevier, v. 46, n. 8, p. 2350–2360, 2013. Citado 3 vezes nas páginas 13, 21 e 27.
- PEDRONETTE, D. C. G. et al. Multimedia retrieval through unsupervised hypergraph-based manifold ranking. *IEEE Transactions on Image Processing*, IEEE, v. 28, n. 12, p. 5824–5838, 2019. Citado 3 vezes nas páginas 29, 36 e 41.
- RADENOVIĆ, F.; TOLIAS, G.; CHUM, O. Fine-tuning cnn image retrieval with no human annotation. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 41, n. 7, p. 1655–1668, 2018. Citado na página 20.
- RAI, P.; SINGH, S. A survey of clustering techniques. *International Journal of Computer Applications*, International Journal of Computer Applications, 244 5 th Avenue, # 1526, New . . . , v. 7, n. 12, p. 1–5, 2010. Citado na página 19.
- REHMAN, M. et al. Content based image retrieval: survey. *World Applied Sciences Journal*, v. 19, n. 3, p. 404–412, 2012. Citado na página 20.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, American Psychological Association, v. 65, n. 6, p. 386, 1958. Citado na página 17.
- ROZIN, B. et al. A rank-based framework through manifold learning for improved clustering tasks. *Information Sciences*, Elsevier, v. 580, p. 202–220, 2021. Citado na página 13.

- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations by error propagation, parallel distributed processing, explorations in the microstructure of cognition, ed. de rumelhart and j. mcclelland. vol. 1. 1986. *Biometrika*, v. 71, p. 599–607, 1986. Citado na página 17.
- SAXENA, A. et al. A review of clustering techniques and developments. *Neurocomputing*, Elsevier, v. 267, p. 664–681, 2017. Citado na página 19.
- SCHROFF, F.; KALENICHENKO, D.; PHILBIN, J. Facenet: A unified embedding for face recognition and clustering. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 815–823. Citado 2 vezes nas páginas 18 e 28.
- SHINDE, P. P.; SHAH, S. A review of machine learning and deep learning applications. In: IEEE. *2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA)*. [S.l.], 2018. p. 1–6. Citado na página 15.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. In: COMPUTATIONAL AND BIOLOGICAL LEARNING SOCIETY. *3rd International Conference on Learning Representations (ICLR 2015)*. [S.l.], 2015. p. 1–14. Citado 2 vezes nas páginas 13 e 17.
- TANG, C. et al. Sparse mlp for image recognition: Is self-attention really necessary? In: *Proceedings of the AAAI conference on artificial intelligence*. [S.l.: s.n.], 2022. v. 36, n. 2, p. 2344–2351. Citado na página 37.
- TOLSTIKHIN, I. O. et al. Mlp-mixer: An all-mlp architecture for vision. *Advances in neural information processing systems*, v. 34, p. 24261–24272, 2021. Citado na página 37.
- TORRES, R. da S.; FALCAO, A. X. Content-based image retrieval: theory and applications. *RITA*, v. 13, n. 2, p. 161–185, 2006. Citado na página 21.
- VALEM, L. P.; PEDRONETTE, D. C. G. An unsupervised distance learning framework for multimedia retrieval. In: *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval*. [S.l.: s.n.], 2017. p. 107–111. Citado na página 21.
- VALEM, L. P.; PEDRONETTE, D. C. G.; ALMEIDA, J. Unsupervised similarity learning through cartesian product of ranking references. *Pattern Recognition Letters*, v. 114, p. 41–52, 2018. Citado na página 13.
- VALEM, L. P.; PEDRONETTE, D. C. G.; LATECKI, L. J. Graph convolutional networks based on manifold learning for semi-supervised image classification. *Computer Vision and Image Understanding*, Elsevier, v. 227, p. 103618, 2023. Citado na página 21.
- VALEM, L. P.; PEDRONETTE, D. C. G.; LATECKI, L. J. Rank flow embedding for unsupervised and semi-supervised manifold learning. *IEEE Transactions on Image Processing*, IEEE, v. 32, p. 2811–2826, 2023. Citado 2 vezes nas páginas 13 e 29.
- VAPNIK, V. N.; VAPNIK, V. et al. Statistical learning theory. Wiley New York, 1998. Citado na página 24.
- VELIČKOVIĆ, P. et al. Graph attention networks. In: *International Conference on Learning Representations*. [S.l.: s.n.], 2018. Citado na página 27.

- WAN, J. et al. Deep learning for content-based image retrieval: A comprehensive study. In: *Proceedings of the 22nd ACM international conference on Multimedia*. [S.l.: s.n.], 2014. p. 157–166. Citado na página 32.
- WANG, J. et al. Learning fine-grained image similarity with deep ranking. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2014. p. 1386–1393. Citado na página 27.
- WANG, X. et al. Multi-similarity loss with general pair weighting for deep metric learning. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2019. p. 5022–5030. Citado na página 28.
- WANG, X. et al. Ranked list loss for deep metric learning. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2019. p. 5207–5216. Citado na página 28.
- WEBBER, W.; MOFFAT, A.; ZOBEL, J. A similarity measure for indefinite rankings. *ACM Transactions on Information Systems (TOIS)*, ACM New York, NY, USA, v. 28, n. 4, p. 1–38, 2010. Citado na página 13.
- WEINBERGER, K. Q.; SAUL, L. K. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, v. 10, n. 2, 2009. Citado na página 15.
- WU, C.-Y. et al. Sampling matters in deep embedding learning. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2017. p. 2840–2848. Citado 2 vezes nas páginas 27 e 28.
- WU, W. et al. Hashing-accelerated graph neural networks for link prediction. In: *Proceedings of the Web Conference 2021*. [S.l.: s.n.], 2021. p. 2910–2920. Citado na página 29.
- XIA, J. et al. Revisiting dimensionality reduction techniques for visual cluster analysis: An empirical study. *IEEE Transactions on Visualization and Computer Graphics*, IEEE, v. 28, n. 1, p. 529–539, 2021. Citado na página 19.
- XING, E. et al. Distance metric learning with application to clustering with side-information. *Advances in Neural Information Processing Systems*, v. 15, 2002. Citado na página 15.
- XU, R.; WUNSCH, D. Survey of clustering algorithms. *IEEE Transactions on neural networks*, IEEE, v. 16, n. 3, p. 645–678, 2005. Citado na página 19.
- YOU, J.; YING, R.; LESKOVEC, J. Position-aware graph neural networks. In: PMLR. *International conference on machine learning*. [S.l.], 2019. p. 7134–7143. Citado na página 27.
- ZHU, X.; GHAHRAMANI, Z. Learning from labeled and unlabeled data with label propagation. *ProQuest number: information to all users*, Citeseer, 2002. Citado na página 26.