

Atividades de Pesquisa e Ensino

Visão Computacional e Aprendizado de Máquina em Imagens Não-Naturais

Prof. Hugo Neves de Oliveira
hugo.n.oliveira@ufv.br

Universidade Federal de Viçosa

22/06/2023



Agenda

- 1 Introdução
- 2 Open World Segmentation
- 3 Self-Supervised Learning
- 4 Meta-Learning para Few-shot Weakly-supervised Segmentation
- 5 Atividades de Ensino

Visão Computacional Clássica vs. Deep

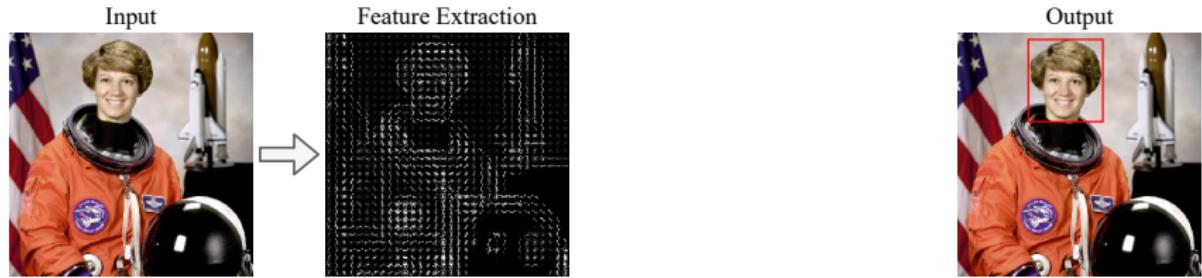
Input



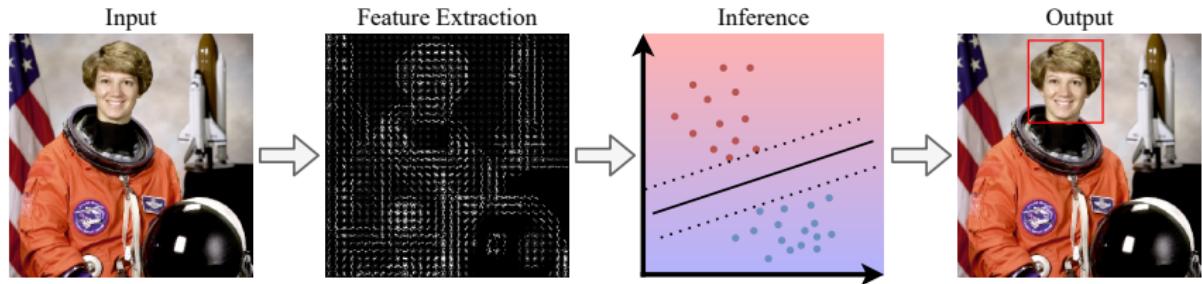
Output



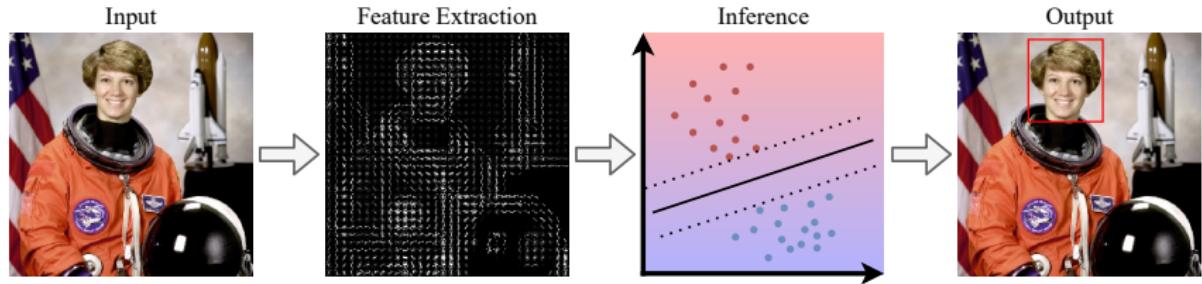
Visão Computacional Clássica vs. Deep



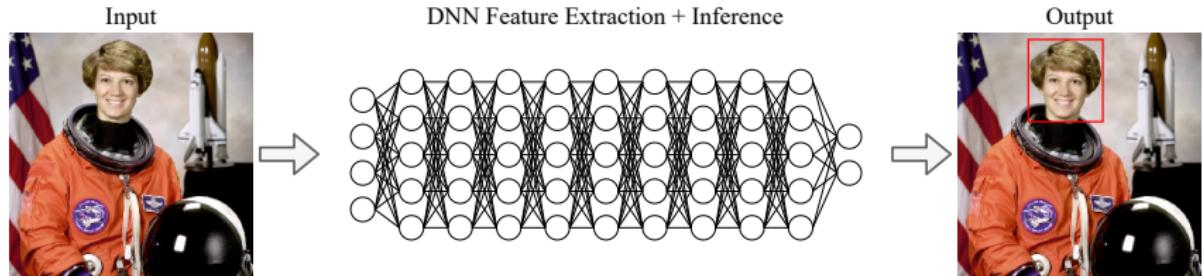
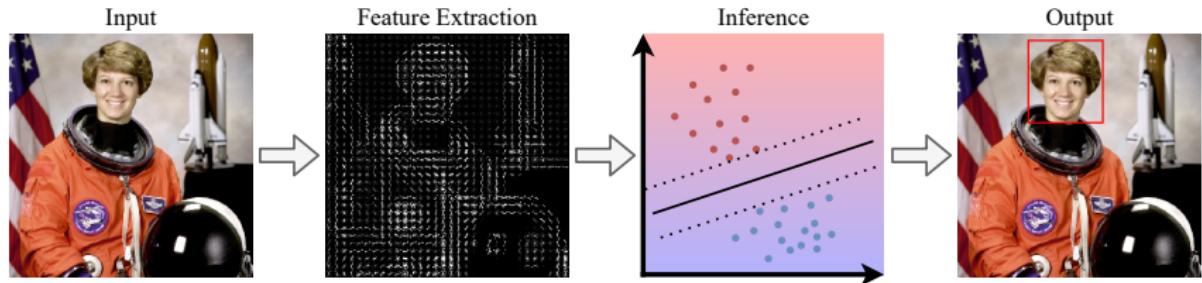
Visão Computacional Clássica vs. Deep



Visão Computacional Clássica vs. Deep



Visão Computacional Clássica vs. Deep



Limitações de Abordagens de Deep Learning

- Alta dependência de **grandes quantidades de dados anotados**;
- Generalização limitada para tarefas relacionadas com *domain shift* maior;
- Modelos Deep para imagens normalmente são desenhados com imagens naturais (RGB) em mente, muitas vezes falhando quando aplicados em outros domínios como imagens radiológicas ou multispectrais.

Pesquisas em Andamento

Visão Computacional Aplicada a Imagens
Não-Naturais e Dados Fracamente Anotados

Pesquisas em Andamento

Visão Computacional Aplicada a Imagens Não-Naturais e Dados Fracamente Anotados

Pesquisas visando mitigar os problemas causados por essas limitações do Deep Learning ao longo dos últimos anos:

Pesquisas em Andamento

Visão Computacional Aplicada a Imagens Não-Naturais e Dados Fracamente Anotados

Pesquisas visando mitigar os problemas causados por essas limitações do Deep Learning ao longo dos últimos anos:

- Aumento da robustez das Redes Neurais com o uso de **Open Set Recognition** [1]–[6];

Pesquisas em Andamento

Visão Computacional Aplicada a Imagens Não-Naturais e Dados Fracamente Anotados

Pesquisas visando mitigar os problemas causados por essas limitações do Deep Learning ao longo dos últimos anos:

- Aumento da robustez das Redes Neurais com o uso de **Open Set Recognition** [1]–[6];
- Geração de modelos mais generalistas utilizando abordagens como **Adaptação de Domínio** [7]–[9] e **Self-Supervised Learning** [10];

Pesquisas em Andamento

Visão Computacional Aplicada a Imagens Não-Naturais e Dados Fracamente Anotados

Pesquisas visando mitigar os problemas causados por essas limitações do Deep Learning ao longo dos últimos anos:

- Aumento da robustez das Redes Neurais com o uso de **Open Set Recognition** [1]–[6];
- Geração de modelos mais generalistas utilizando abordagens como **Adaptação de Domínio** [7]–[9] e **Self-Supervised Learning** [10];
- Aprendizado a partir de poucas imagens e a partir de anotações fracas por meio de **Meta-Learning** [11]–[13].

Projetos com Agências de Fomento

- **[FAPEMIG Universal] LittleBigData: Reconhecimento de padrões em grandes bases de imagens utilizando pequenos conjuntos de dados anotados (APQ-00449-17):**
 - Laboratório/Grupo de Pesquisa: PATREO/UFMG
 - Tema: Utilizando *big data* para reconhecimento de padrões em conjuntos de dados pequenos
 - Vínculo: Colaborador
- **[FAPESP] Deep learning e representações intermediárias para análise de imagens pediátricas (2020/06744-5):**
 - Laboratório/Grupo de Pesquisa: Creativision/IME/USP
 - Tema: *Deep Learning* para análise de imagens de radiologia pediátrica em cenários de *small data*
 - Vínculo: Colaborador

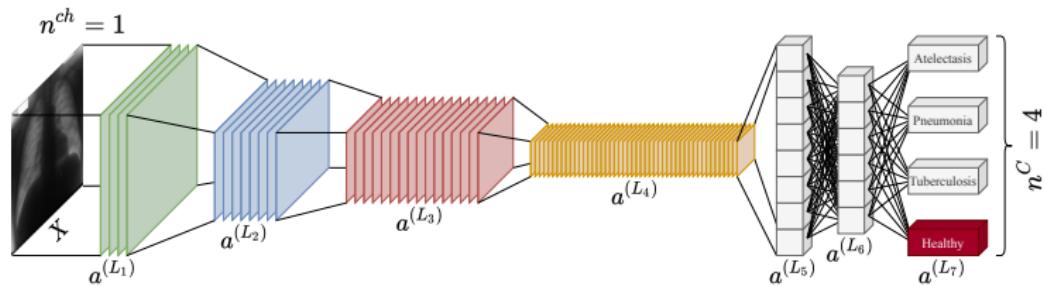
Projetos com Agências de Fomento

- **[Instituto Serrapilheira] Dense Labeling of Remote Sensing Images in the Wild:**
 - Laboratório/Grupo de Pesquisa: PATREO/UFMG
 - Tema: Investigar soluções para cenários de *open set* e *small-data* desafiadores no mapeamento de larga escala
 - Vínculo: Colaborador
- **[FUNDECT] Inteligência artificial e sensoriamento remoto aplicados no monitoramento do estoque de carbono e emissão de CO2 por incêndios:**
 - Tema: Uso de imagens de sensoriamento remoto para análise da preservação de reservas ambientais
 - Vínculo: Colaborador

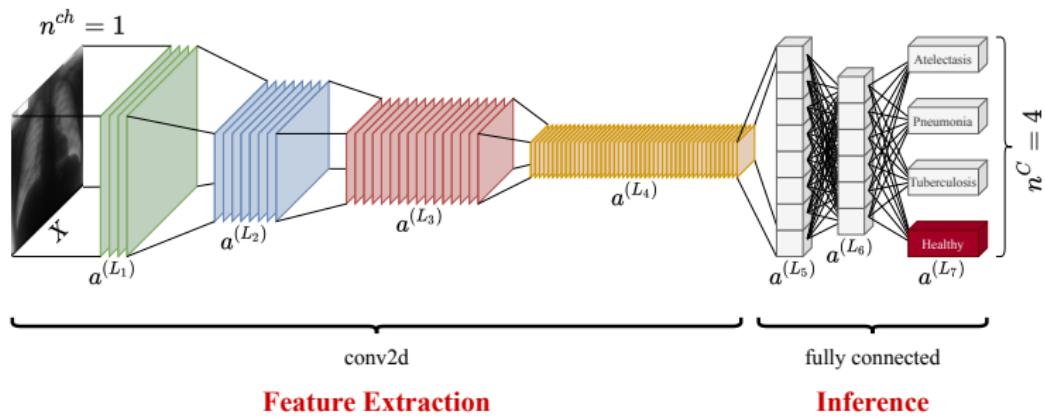
Agenda

- 1 Introdução
- 2 Open World Segmentation
- 3 Self-Supervised Learning
- 4 Meta-Learning para Few-shot Weakly-supervised Segmentation
- 5 Atividades de Ensino

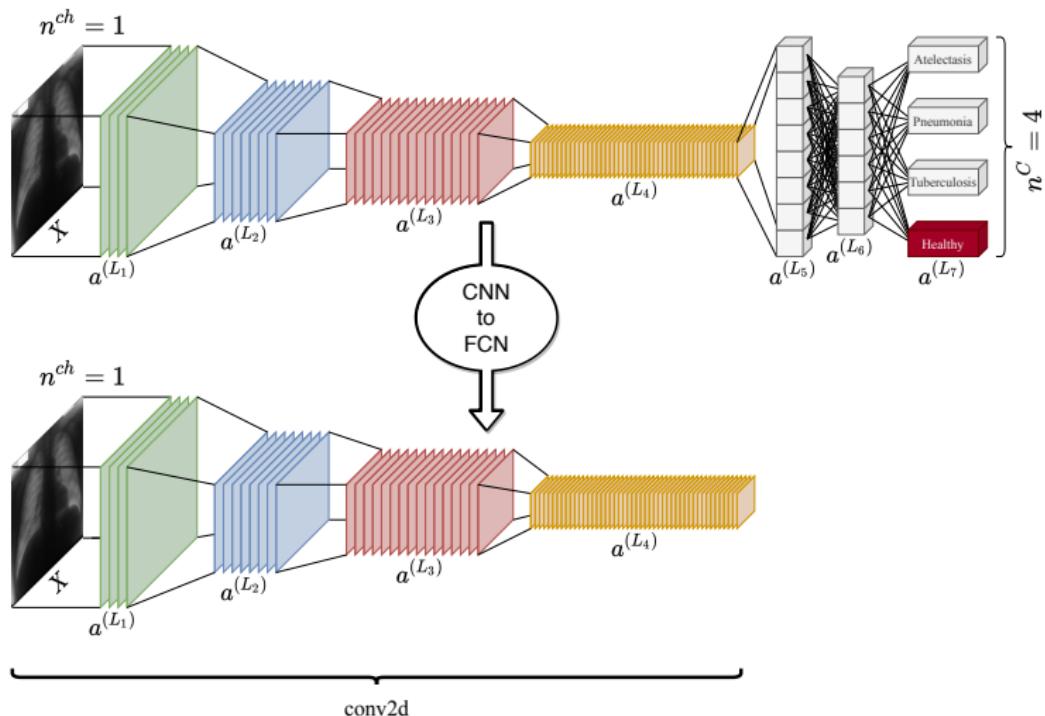
CNN para Classificação de Imagens



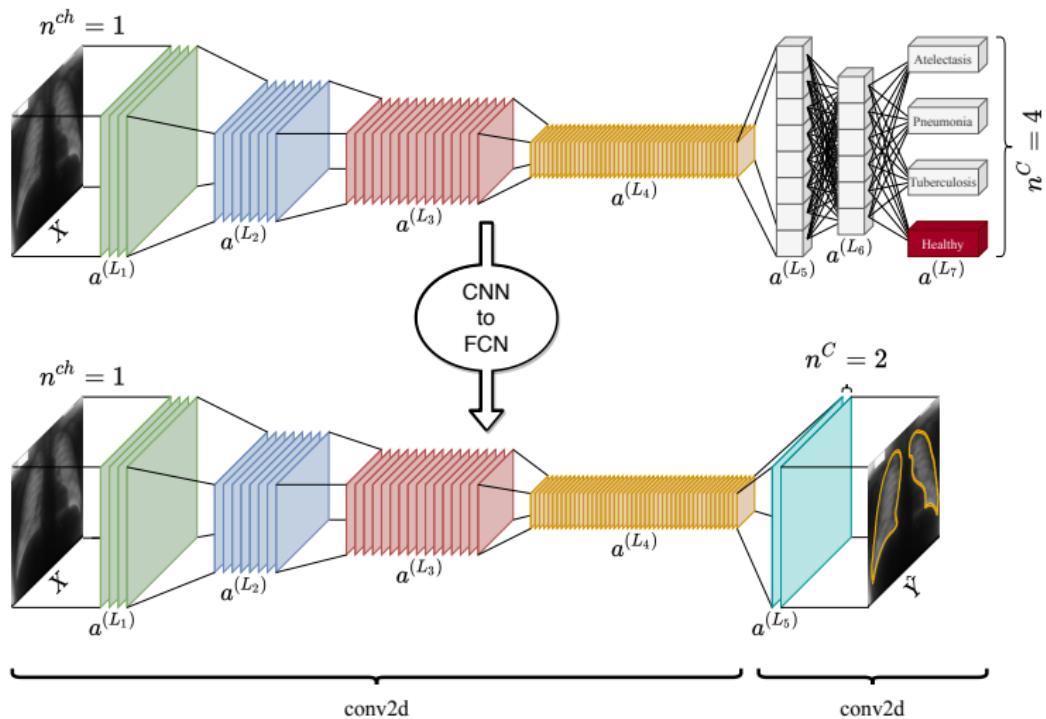
CNN para Classificação de Imagens



Transformando uma CNN numa FCN



Transformando uma CNN numa FCN



Limitações do Treino Closed Set

Reconhecimento de Imagens Closed Set

Modelos de Deep Learning tradicionais são treinadas assumindo total conhecimento sobre o conjunto das classes de um certo domínio. Esses modelos são chamados de modelos de conjunto fechado de classes (*Closed Set*) e podem apresentar uma baixa robustez quando inseridos numa tarefa do mundo real.

Limitações do Treino Closed Set

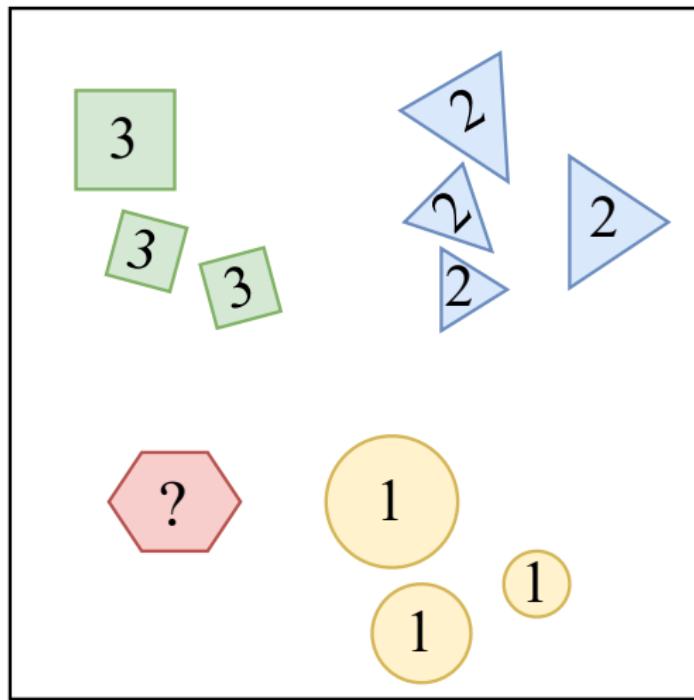
Reconhecimento de Imagens Closed Set

Modelos de Deep Learning tradicionais são treinadas assumindo total conhecimento sobre o conjunto das classes de um certo domínio. Esses modelos são chamados de modelos de conjunto fechado de classes (*Closed Set*) e podem apresentar uma baixa robustez quando inseridos numa tarefa do mundo real.

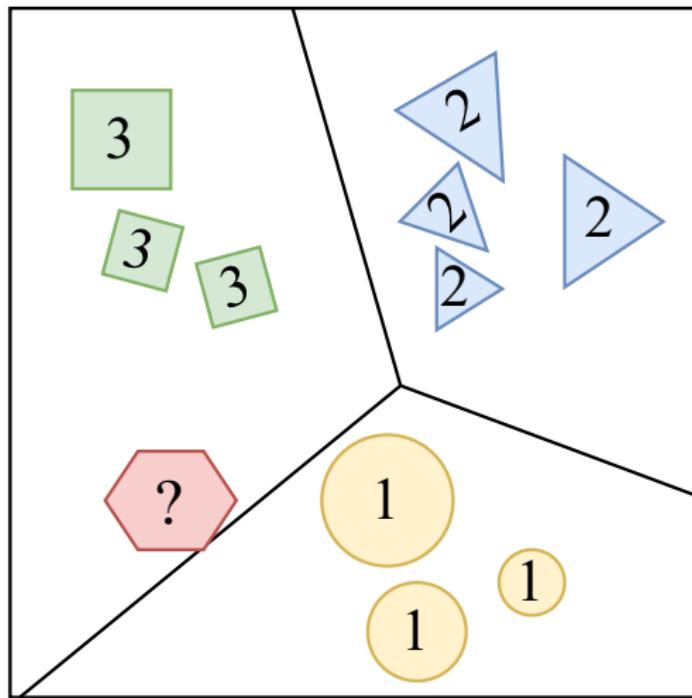
Solução

Reconhecimento Open Set

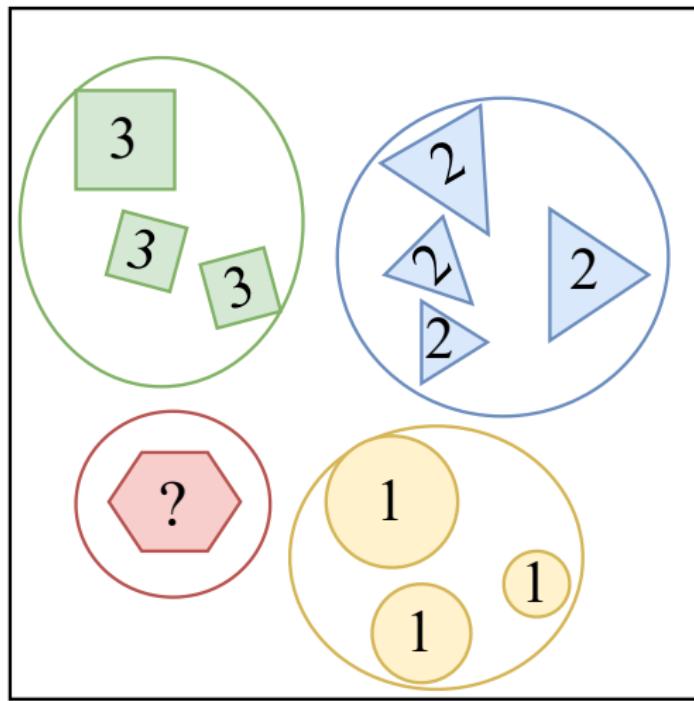
Distribuição de Classes Num Problema Exemplo



Classificação Closed Set



Classificação Open Set

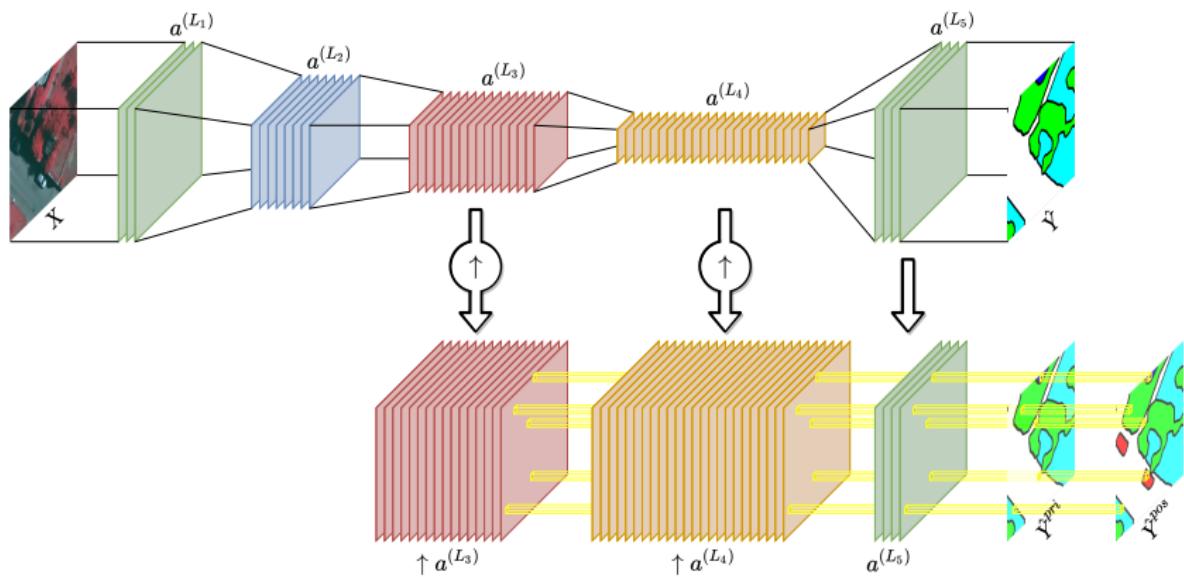


Estendendo Reconhecimento Open Set para Segmentação Open Set

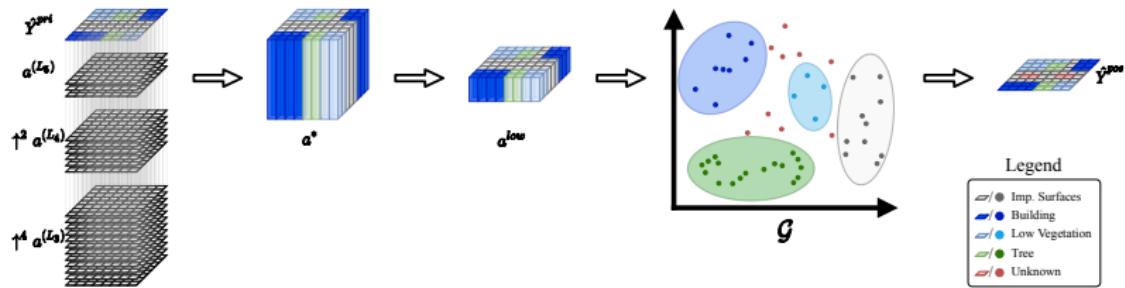
Segmentação Open Set

Vários problemas de segmentação de objetos são inherentemente Open Set, precisando lidar com amostras de classes desconhecidas durante a sua implementação. Um que viemos explorando particularmente bem é a segmentação Open Set de áreas de plantação a partir de imagens de SAR [4], [6].

OSS via Open Principal Component Scoring (OpenPCS)



OSS via Open Principal Component Scoring (OpenPCS)



Além de Open Set Segmentation

Futuro do OSS

Além de identificar automaticamente amostras de classes desconhecidas, seria ideal se esses modelos fossem capazes de se atualizar para aprender novas classes que forem sendo apresentadas já durante sua implantação num problema real.

Além de Open Set Segmentation

Futuro do OSS

Além de identificar automaticamente amostras de classes desconhecidas, seria ideal se esses modelos fossem capazes de se atualizar para aprender novas classes que forem sendo apresentadas já durante sua implantação num problema real.

Segmentação de Mundo Aberto

A Área de *Open World Recognition* (OWR) trata esse problema de aprendizado contínuo no contexto de open set para problemas de classificação. Esse projeto de pesquisa visa aplicar esse mesmo conceito a segmentação de imagens de satélite [2], [4], [6], atualizando modelos de segmentação continuamente à medida que novas amostras de classes desconhecidas forem aparecendo. Essa área se chama *Open World Segmentation* (OWS).

Aprendendo a Partir de Poucas Amostras Rotuladas

OWS

No contexto de Open World Segmentation, portanto, algoritmos capazes de aprender a partir de poucos dados rotulados das classes desconhecidas que são aparecendo durante a implementação de um modelo são cruciais. Duas vertentes principais de algoritmos de *Few-Shot Learning* vêm emergindo:

- *Meta-Learning*
- *Self-Supervised Learning*

Agenda

- 1 Introdução
- 2 Open World Segmentation
- 3 Self-Supervised Learning
- 4 Meta-Learning para Few-shot Weakly-supervised Segmentation
- 5 Atividades de Ensino

Aprendizado de Representações

Aprendendo Representações Úteis Sem o Uso de Rótulos

A área de pesquisa de Self-Supervised Learning (SSL) vem emergindo como uma abordagem robusta para o aprendizado de características em Redes Neurais Profundas. Contrariamente ao Aprendizado Supervisionado tradicional para esse tipo de modelo, SSL não utiliza amostras rotuladas para realizar o pré-treino de modelos de Deep Learning.

Aprendizado de Representações

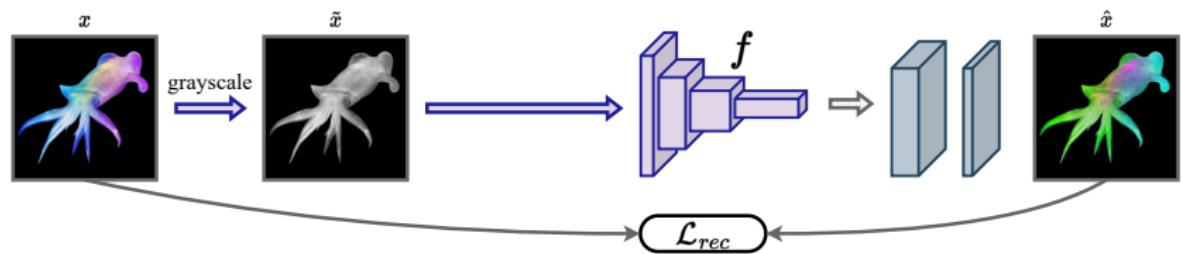
Aprendendo Representações Úteis Sem o Uso de Rótulos

A área de pesquisa de Self-Supervised Learning (SSL) vem emergindo como uma abordagem robusta para o aprendizado de características em Redes Neurais Profundas. Contrariamente ao Aprendizado Supervisionado tradicional para esse tipo de modelo, SSL não utiliza amostras rotuladas para realizar o pré-treino de modelos de Deep Learning.

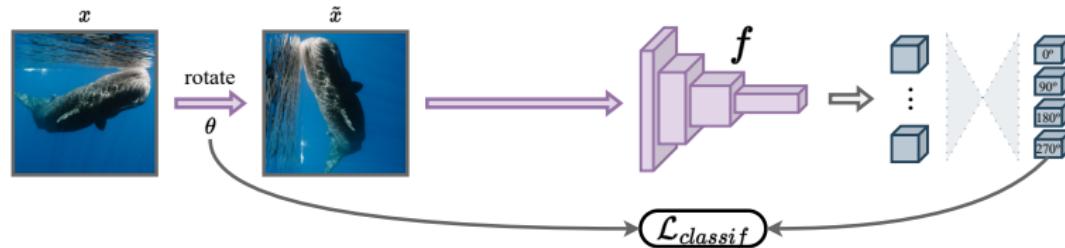
Aprendendo Representações Úteis Sem o Uso de Rótulos

SSL faz uso de **tarefas de pretexto** (*pretext tasks*) que não necessitam de rótulos gerados por humanos. Por exemplo, os algoritmos tradicionais de Deep SSL são treinados para predizer **pseudo-rótulos** [14] que podem ser obtidos diretamente de modificações nos dados.

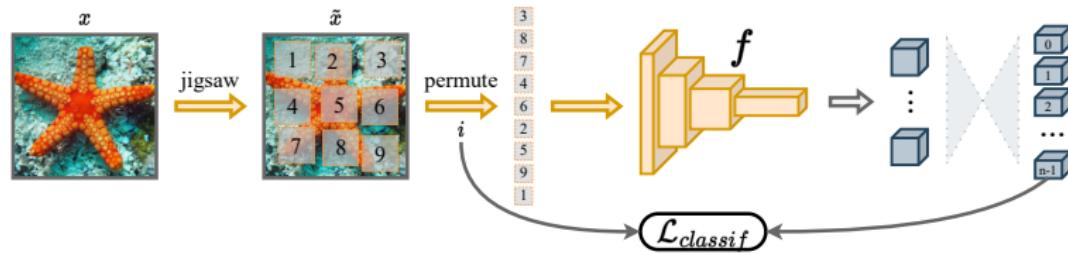
SSL com Tarefas *Handcrafted*



SSL com Tarefas *Handcrafted*



SSL com Tarefas *Handcrafted*



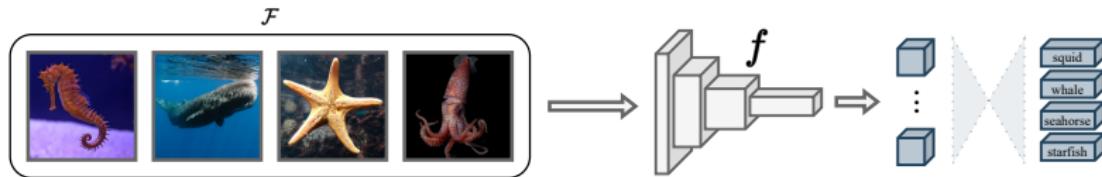
SSL com Tarefas *Handcrafted*

 \mathcal{F} 

SSL com Tarefas *Handcrafted*



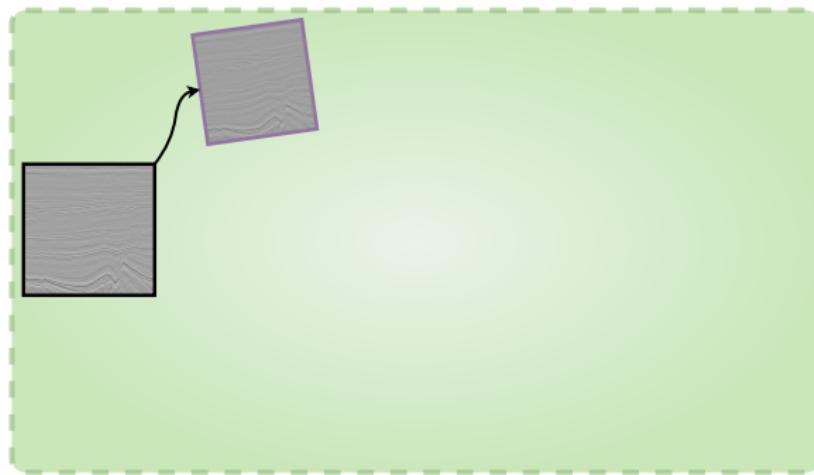
SSL com Tarefas *Handcrafted*



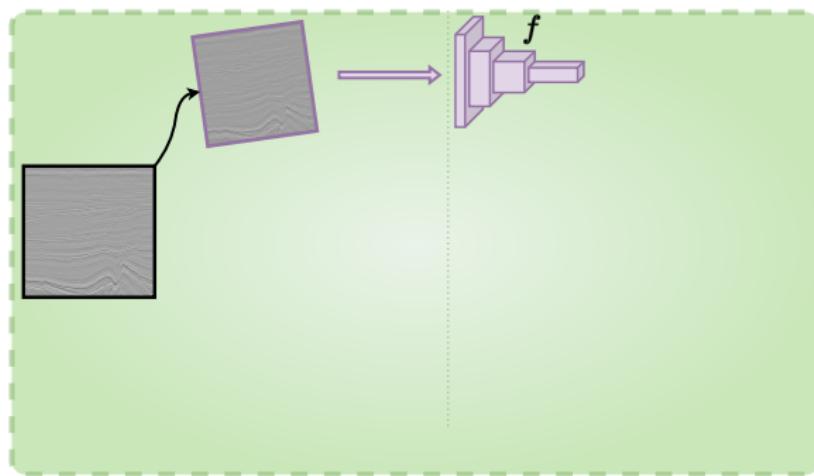
SSL para Imagens Sísmicas



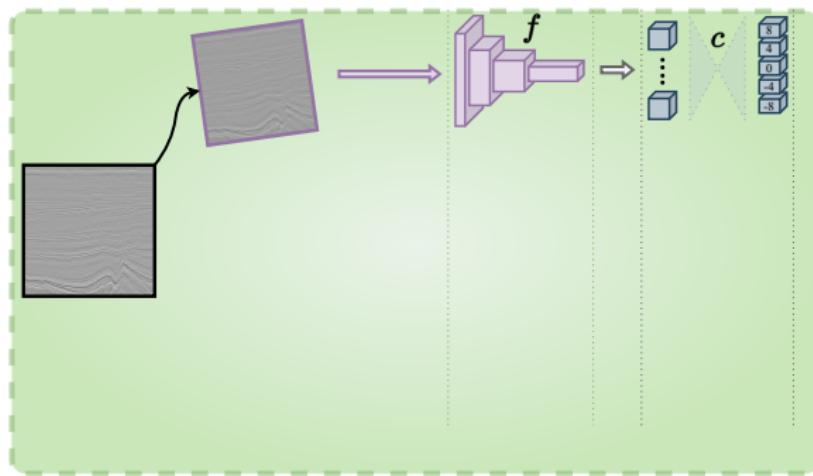
SSL para Imagens Sísmicas



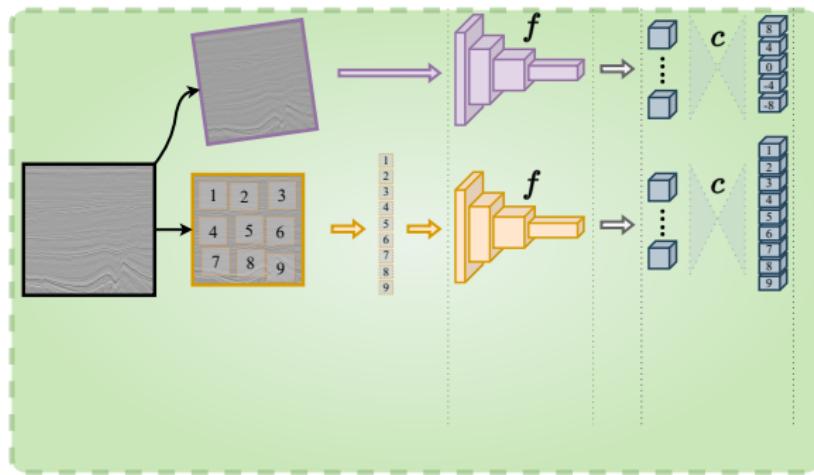
SSL para Imagens Sísmicas



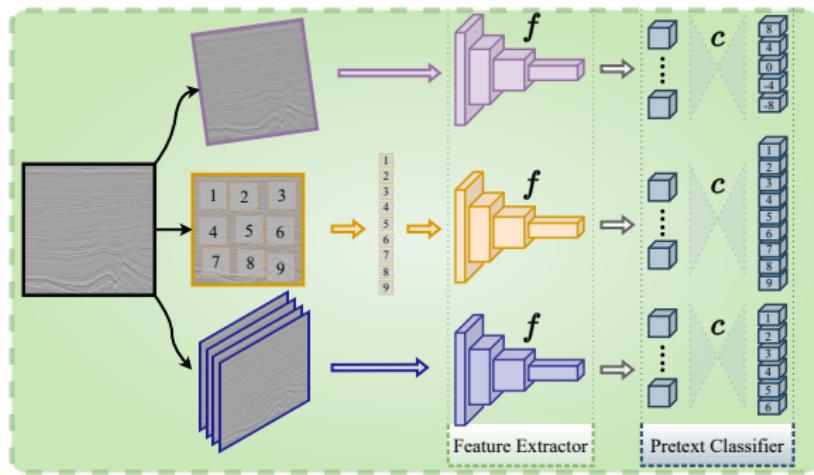
SSL para Imagens Sísmicas



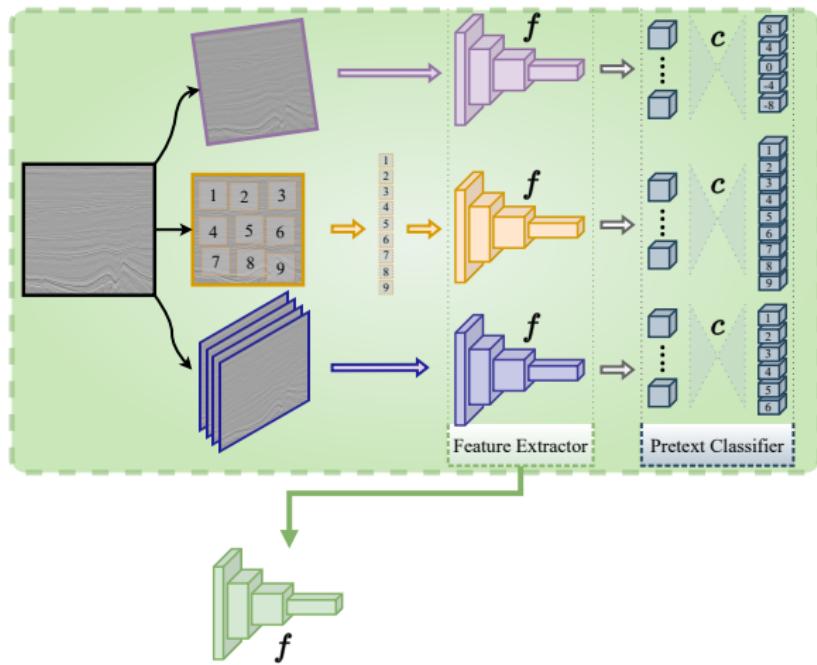
SSL para Imagens Sísmicas



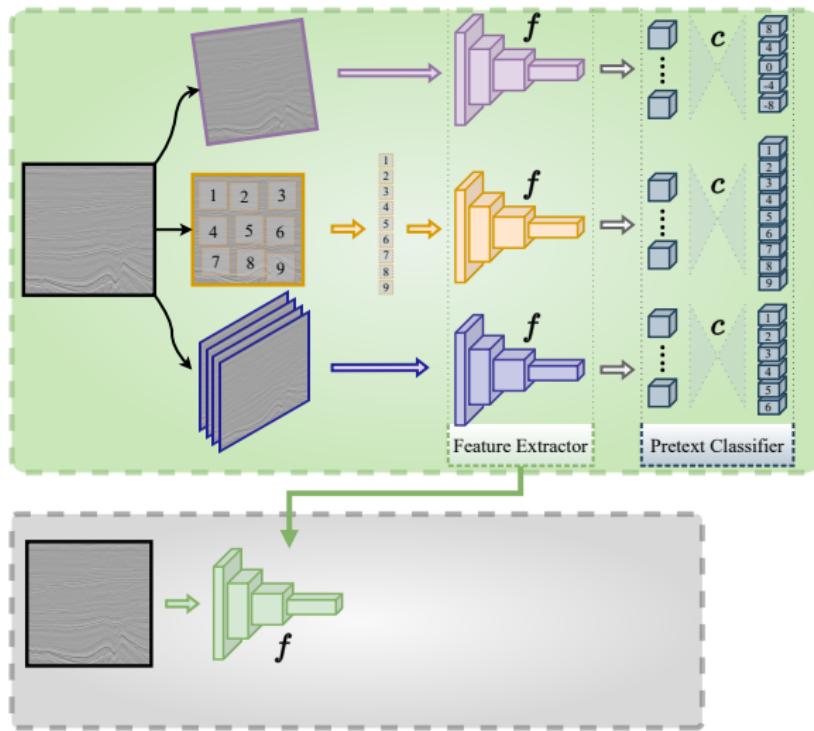
SSL para Imagens Sísmicas



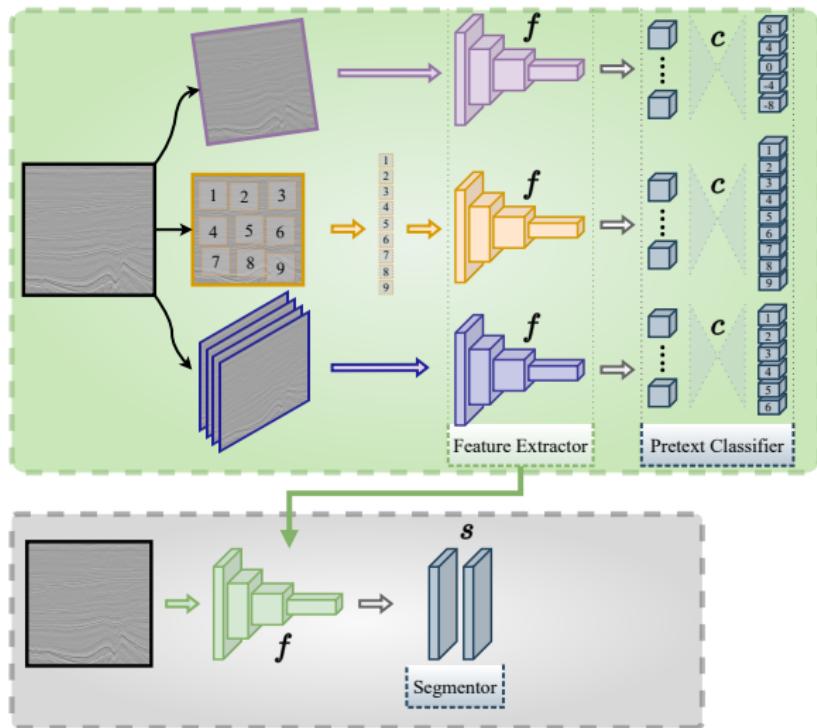
SSL para Imagens Sísmicas



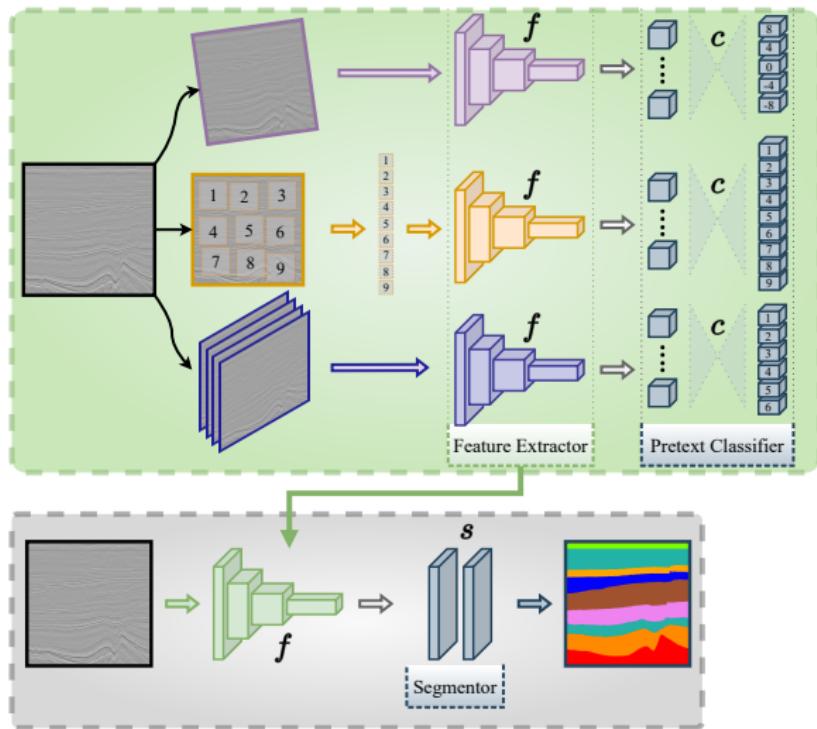
SSL para Imagens Sísmicas



SSL para Imagens Sísmicas



SSL para Imagens Sísmicas



Tendências Atuais de SSL

SSL Moderno para Imagens Não-Naturais

Além de tarefas de pretexto especificamente desenhadas para domínios não-RGB específicos, esse projeto também prevê o uso de técnica mais modernas da literatura de SSL, como SSL utilizando **funções de perda contrastivas** [15]–[17], o uso de redes Transformers e *Self-Attention* para pré-treino auto-supervisionado [18] e o uso de tarefas de pretexto de baseadas em *masking*, como nos **Masked AutoEncoders** [19].

Tendências Atuais de SSL

SSL Moderno para Imagens Não-Naturais

Além de tarefas de pretexto especificamente desenhadas para domínios não-RGB específicos, esse projeto também prevê o uso de técnica mais modernas da literatura de SSL, como SSL utilizando **funções de perda contrastivas** [15]–[17], o uso de redes Transformers e *Self-Attention* para pré-treino auto-supervisionado [18] e o uso de tarefas de pretexto de baseadas em *masking*, como nos **Masked AutoEncoders** [19].

Portando para Imagens Não-Naturais

A simples utilização ingênua desses modelos em imagens não-naturais pode resultar em resultados espúrios ou num aprendizado de representações pouco útil na prática. Logo, uma área que ainda possui um gap de pesquisa em SSL é relacionado a como realizar essa portabilidade de forma mais eficaz para domínios diferentes do RGB (i.e. imagens médicas, de satélite, sísmicas, etc).

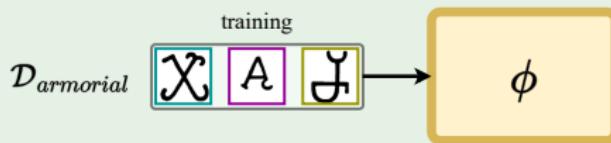
Agenda

- 1 Introdução
- 2 Open World Segmentation
- 3 Self-Supervised Learning
- 4 Meta-Learning para Few-shot Weakly-supervised Segmentation
- 5 Atividades de Ensino

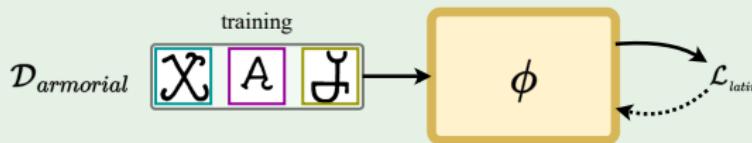
Aprendizado Supervisionado Tradicional



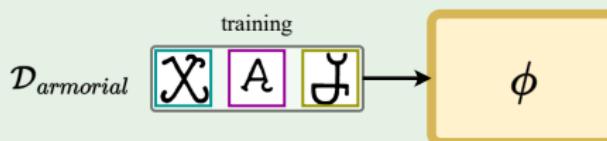
Aprendizado Supervisionado Tradicional



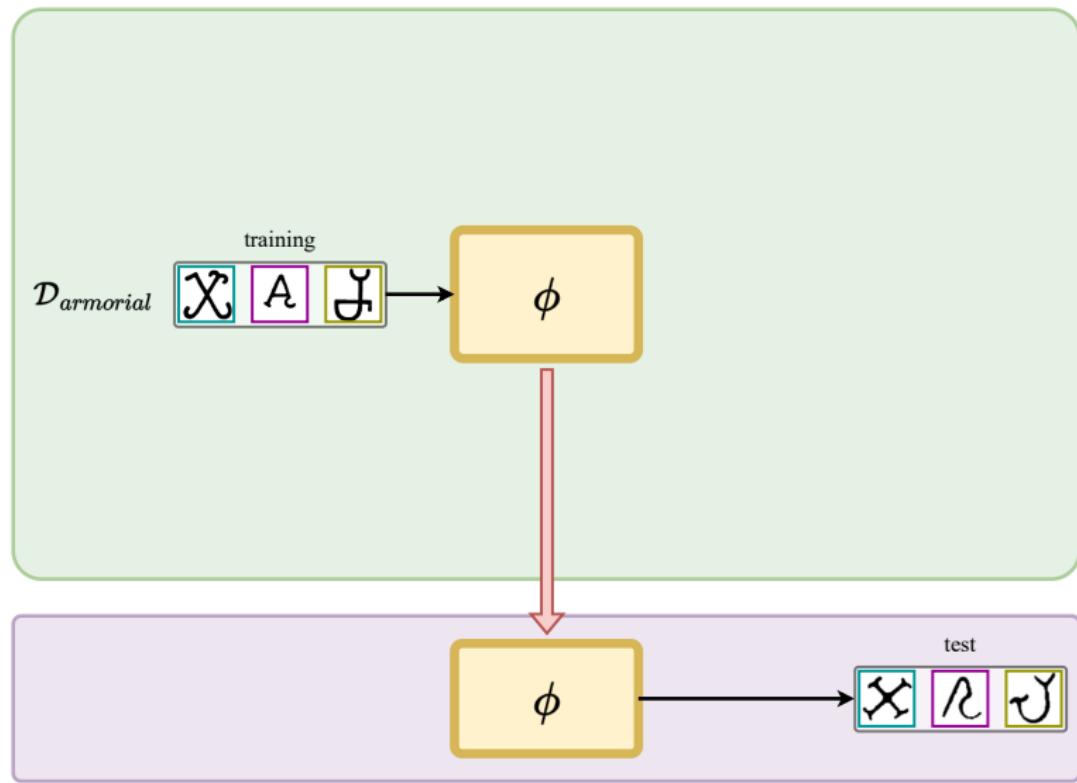
Aprendizado Supervisionado Tradicional



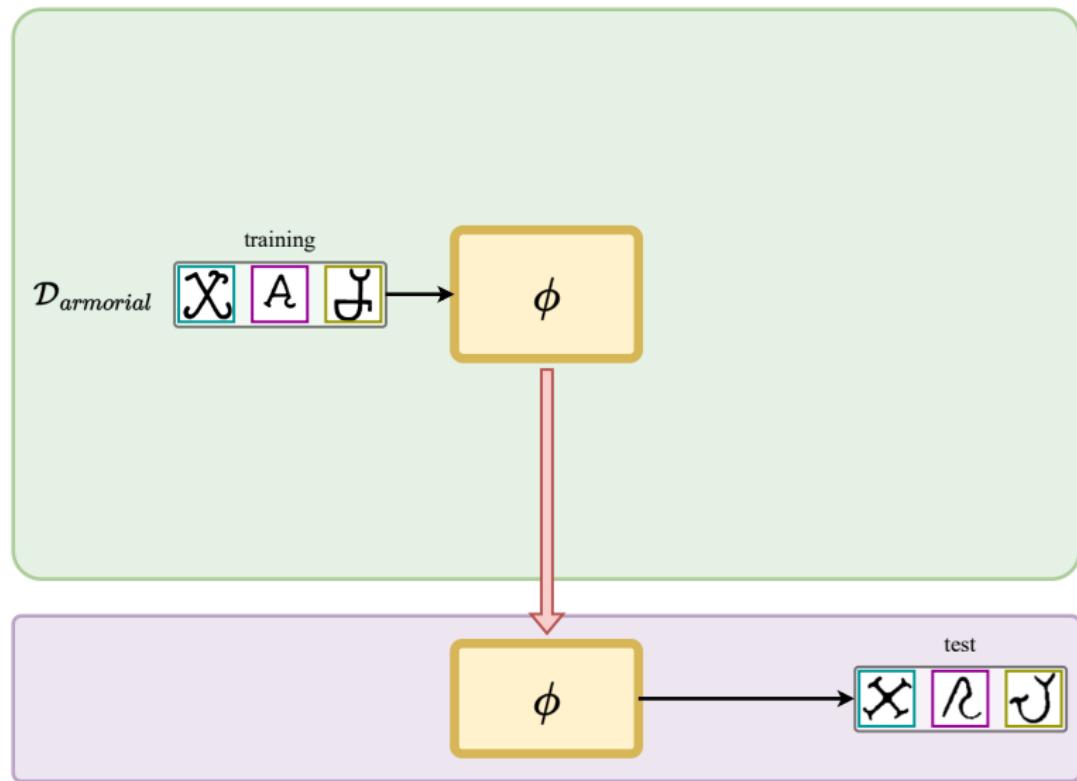
Aprendizado Supervisionado Tradicional



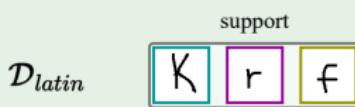
Aprendizado Supervisionado Tradicional



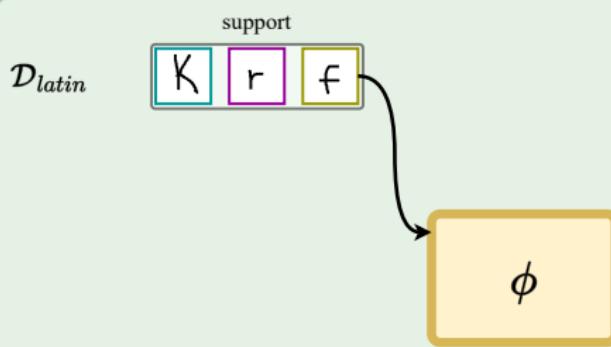
Aprendizado Supervisionado Tradicional



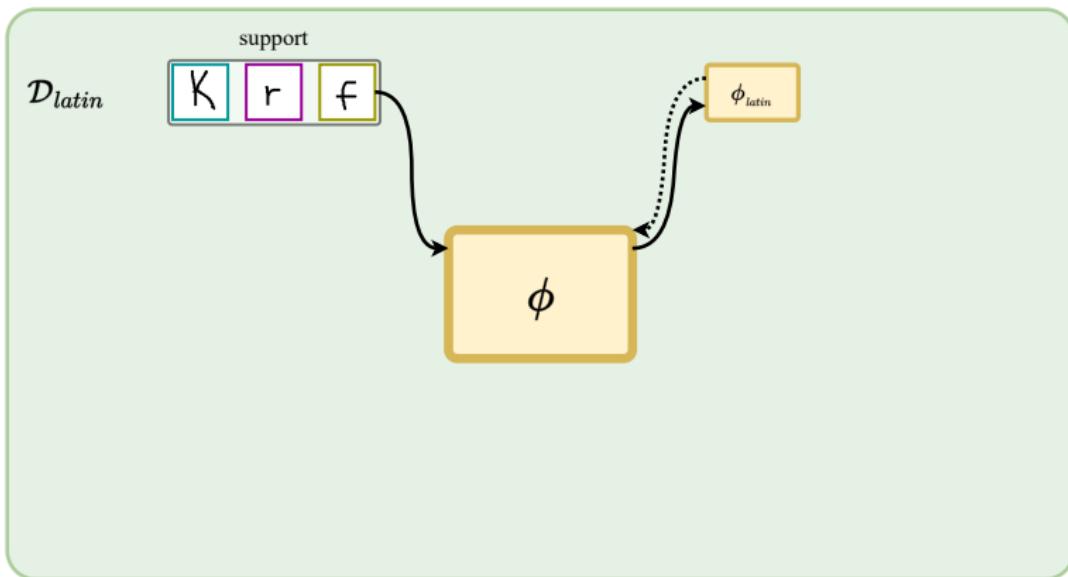
Meta-Learning e Aprendizado Episódico



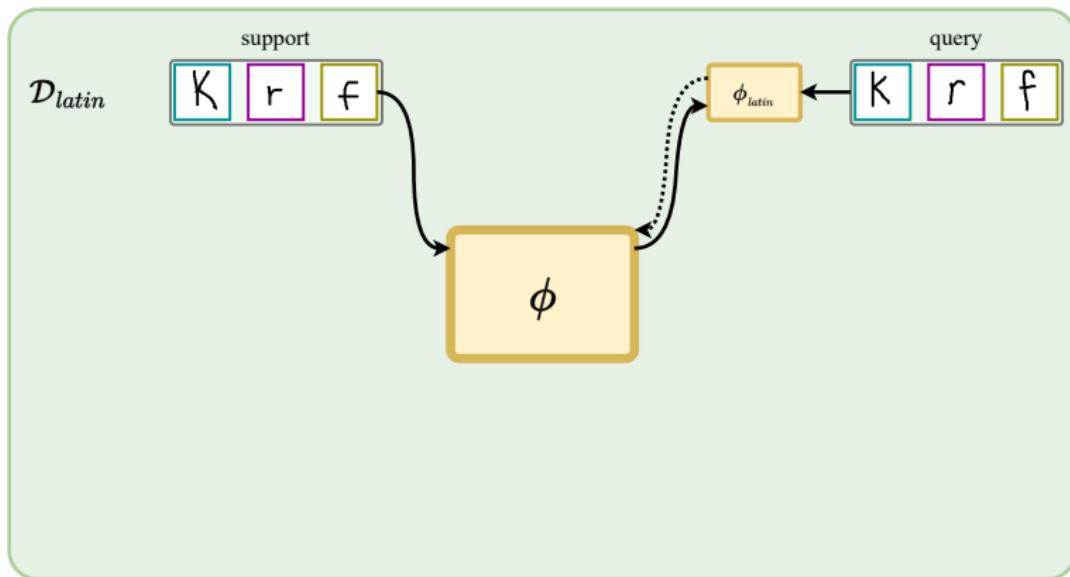
Meta-Learning e Aprendizado Episódico



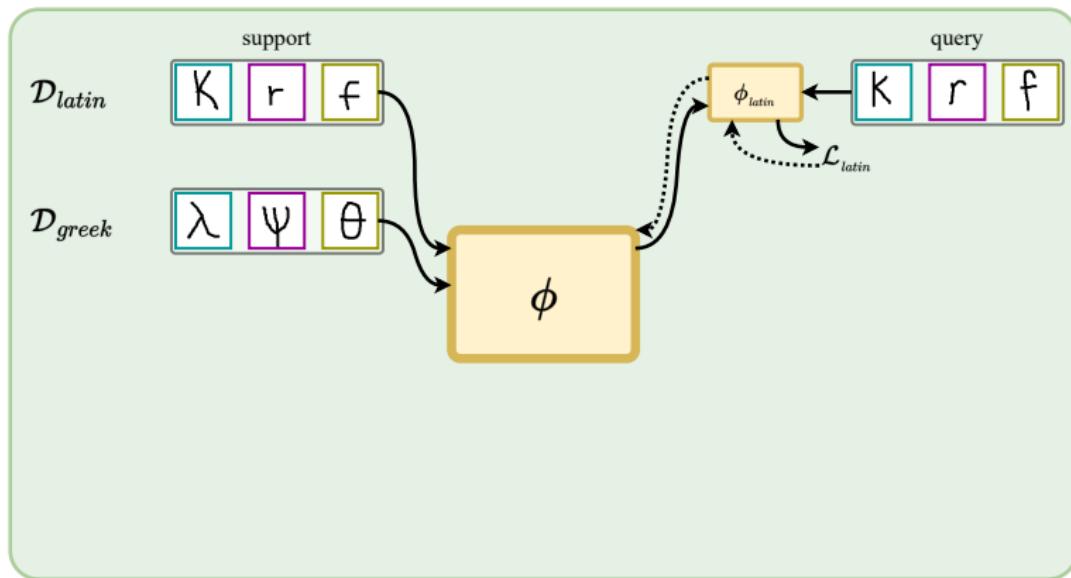
Meta-Learning e Aprendizado Episódico



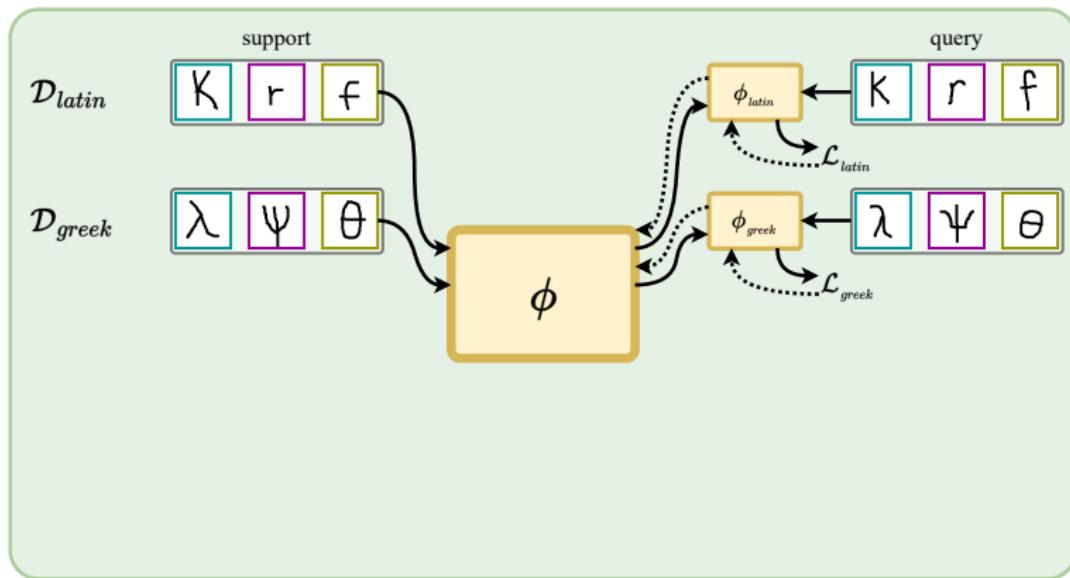
Meta-Learning e Aprendizado Episódico



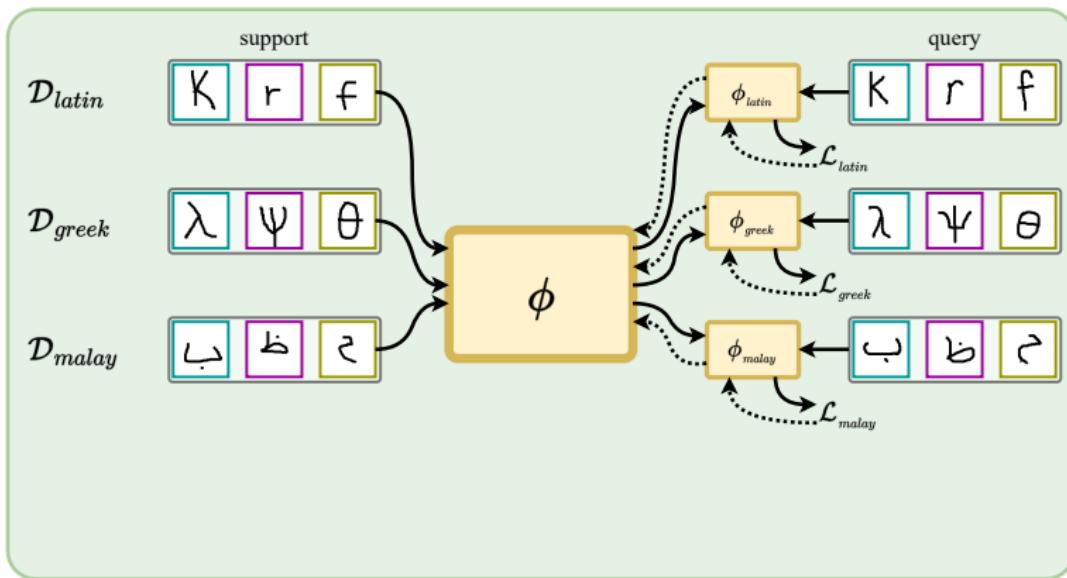
Meta-Learning e Aprendizado Episódico



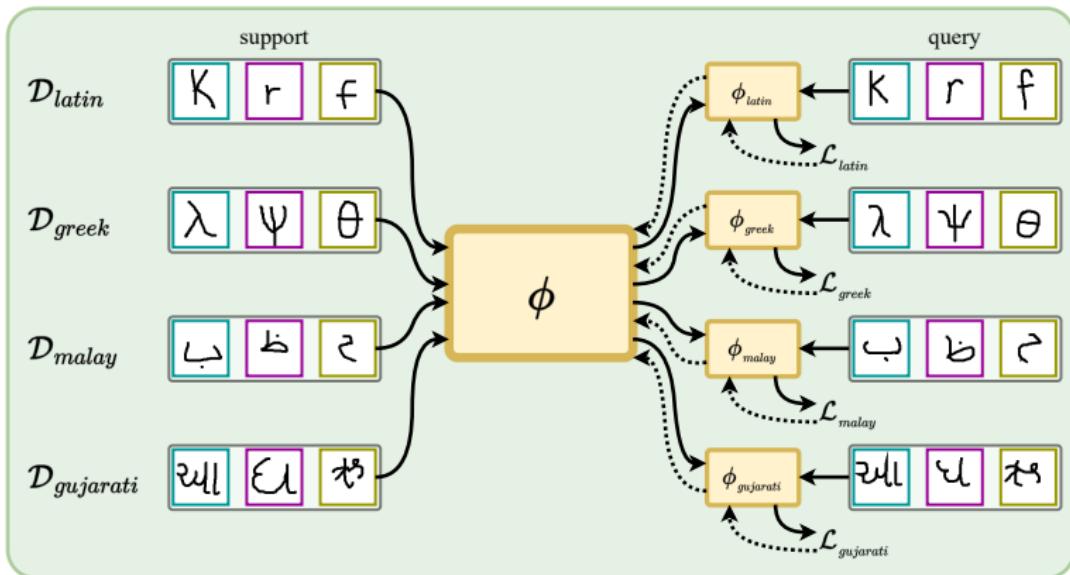
Meta-Learning e Aprendizado Episódico



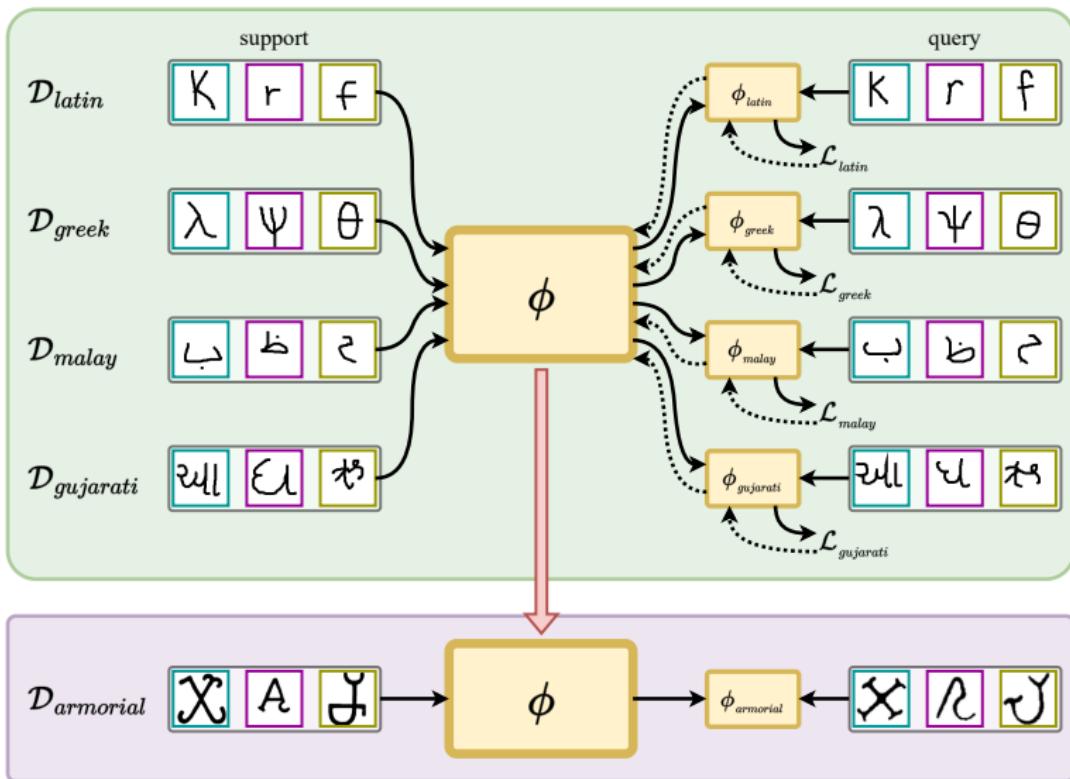
Meta-Learning e Aprendizado Episódico



Meta-Learning e Aprendizado Episódico



Meta-Learning e Aprendizado Episódico



Meta-Learning para Classificação

Métodos de Meta-Learning para Visão

A vasta maioria dos algoritmos de Meta-Learning para tarefas visuais são desenhados especificamente para **classificação few-shot** [20]–[23]. Apesar de alguns esforços recentes [24], [25], ainda há um grande **gap** de métodos de Meta-Learning para outras tarefas como **segmentação** e **detecção**.

Meta-Learning para Classificação

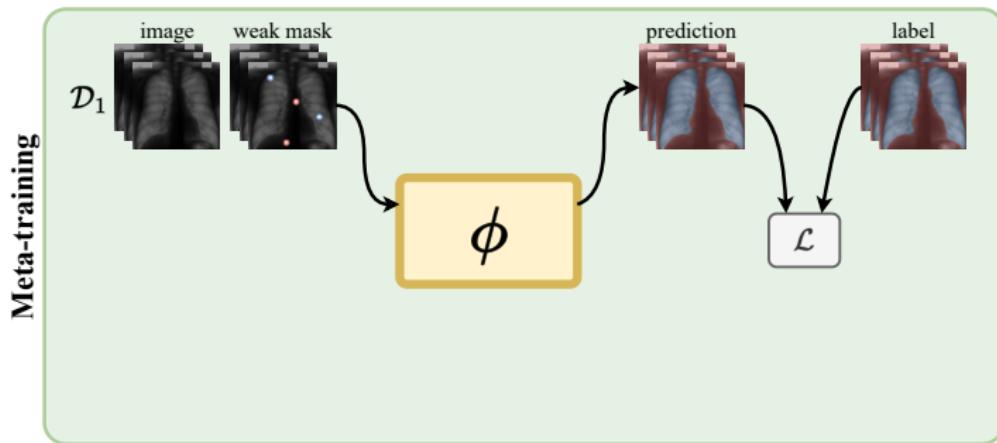
Métodos de Meta-Learning para Visão

A vasta maioria dos algoritmos de Meta-Learning para tarefas visuais são desenhados especificamente para **classificação few-shot** [20]–[23]. Apesar de alguns esforços recentes [24], [25], ainda há um grande **gap** de métodos de Meta-Learning para outras tarefas como **segmentação** e **detecção**.

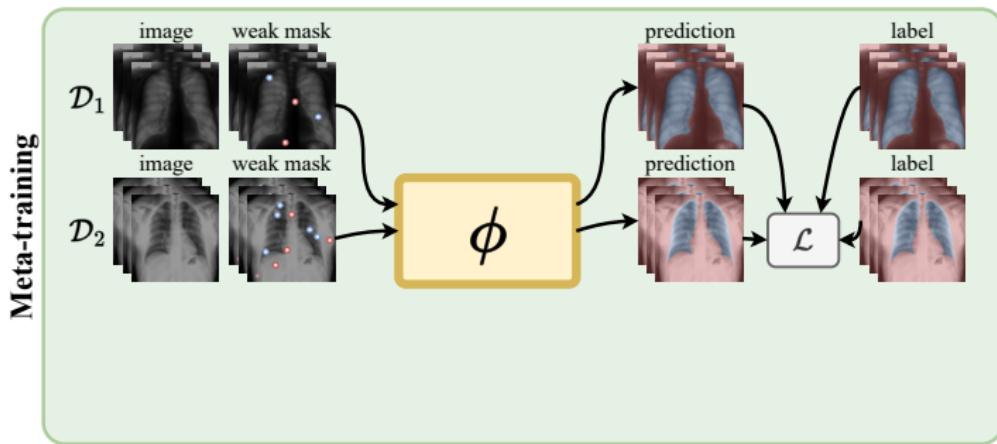
N

os últimos anos venho realizando pesquisas em colaboração com pesquisadores e alunos de outras instituições e publicando métodos capazes de realizar *Few-shot Weakly-supervised Segmentation* (FWS) utilizando Meta-Learning [11]–[13].

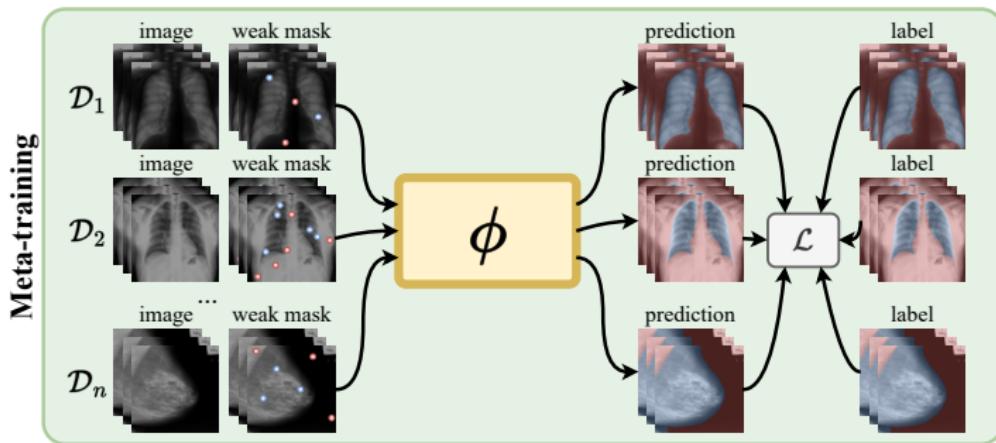
Meta-Learning para FWS



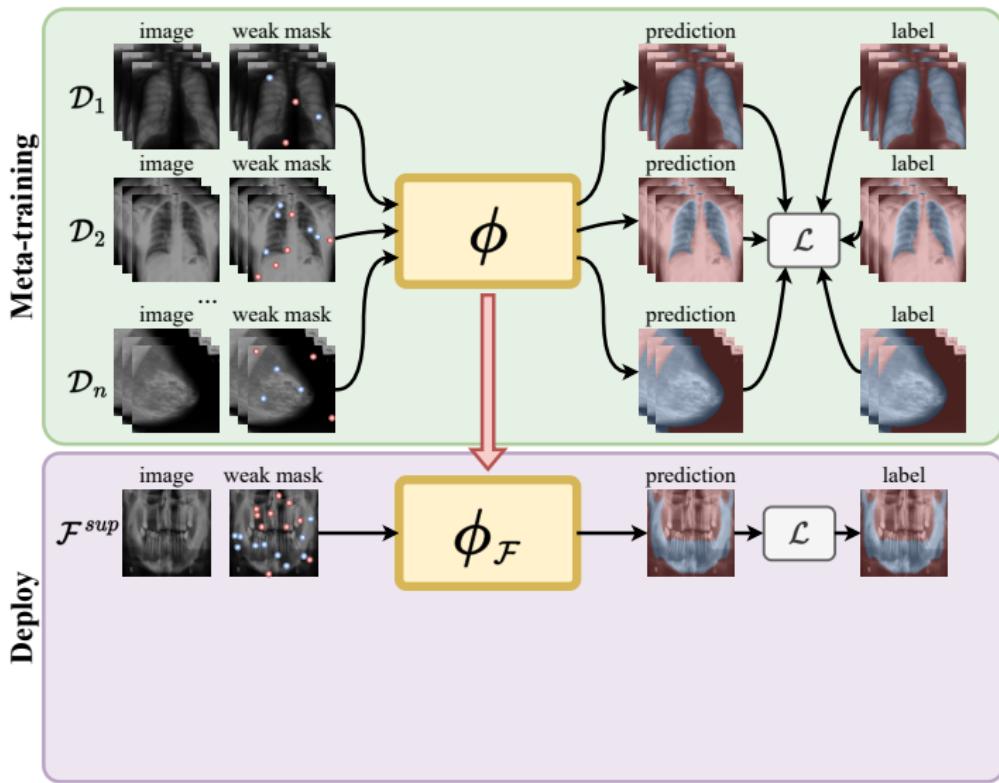
Meta-Learning para FWS



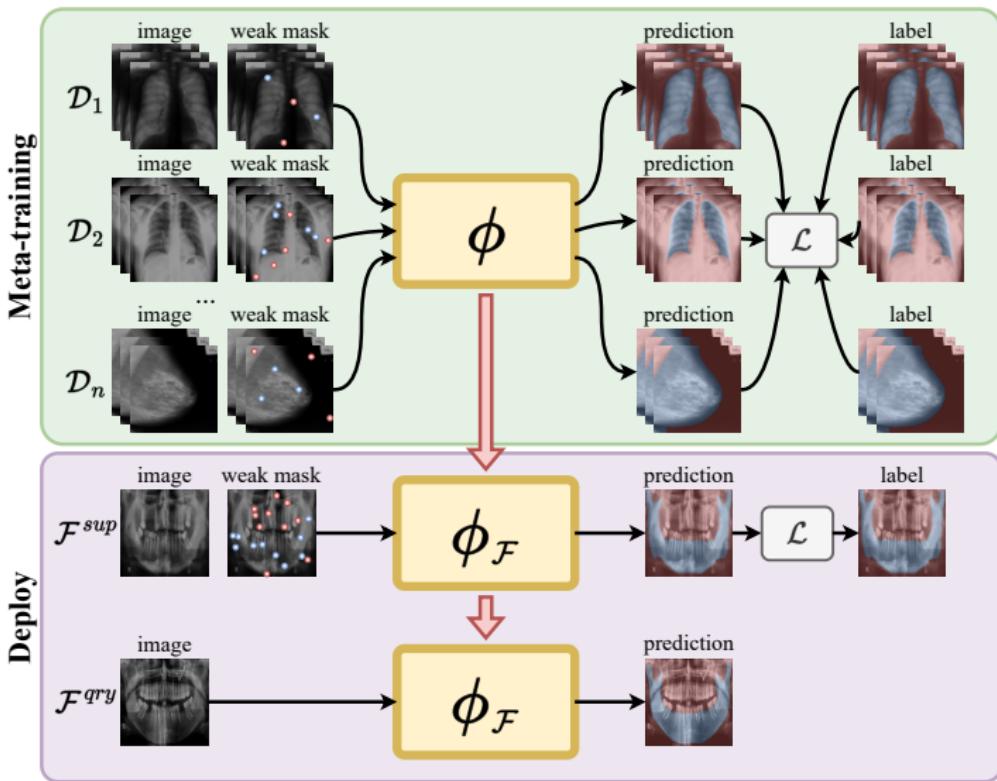
Meta-Learning para FWS



Meta-Learning para FWS



Meta-Learning para FWS



Limitação do Meta-Learning Atual

Necessidade de Dados Auxiliares Rotulados

Nos exemplos anteriores foi mostrado como se pode utilizar dados auxiliares relacionados à tarefa alvo para se atingir *few-shot learning*. Isso limita o uso desse tipo de algoritmo para domínios com **múltiplos datasets similares anotados**.

Limitação do Meta-Learning Atual

Necessidade de Dados Auxiliares Rotulados

Nos exemplos anteriores foi mostrado como se pode utilizar dados auxiliares relacionados à tarefa alvo para se atingir *few-shot learning*. Isso limita o uso desse tipo de algoritmo para domínios com **múltiplos datasets similares anotados**.

Trabalhos Futuros em Meta-Learning

Na etapa atual da minha pesquisa venho estudando como é possível se livrar da necessidade de rótulos nos dados auxiliares para pré-treino do meta-modelo. Experimentos exploratórios já executados mostram que isso pode ser atingido por meio de uma **fusão entre Meta-Learning e Self-Supervised Learning**.

Agenda

- 1 Introdução
- 2 Open World Segmentation
- 3 Self-Supervised Learning
- 4 Meta-Learning para Few-shot Weakly-supervised Segmentation
- 5 Atividades de Ensino

Disciplinas do Bacharelado em Ciência da Computação

- INF420 – **Inteligência Artificial I**;
- INF311 – **Programação para Dispositivos Móveis**;
- INF394 – **Processamento Digital de Imagens**;
- INF423 – **Inteligência Artificial II**.

Disciplinas do PPGCC

- INF620 – **Introdução à Inteligência Artificial e ao Aprendizado de Máquina;**
- INF623 – **Inteligência Artificial;**
- INF629 – **Aprendizado de Máquina;**
- INF721 – **Aprendizado em Redes Neurais Profundas.**

Referências I

- Caio CV da Silva, Keiller Nogueira, Hugo N Oliveira et al. "Towards open-set semantic segmentation of aerial images". Em: *Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)*. IEEE. 2020, pp. 16–21.
- Hugo Oliveira, Caio Silva, Gabriel LS Machado et al. "Fully Convolutional Open Set Segmentation". Em: *Machine Learning* (2021), pp. 1–52.
- Marcos Vendramini, Hugo Oliveira, Alexei Machado et al. "Opening Deep Neural Networks With Generative Models". Em: *International Conference on Image Processing (ICIP)*. IEEE. 2021, pp. 1314–1318.
- Jorge A Chamorro Martinez, Hugo Oliveira, Jefersson A dos Santos et al. "Open Set Semantic Segmentation for Multitemporal Crop Recognition". Em: *Geoscience and Remote Sensing Letters (GRSL)* 19 (2021), pp. 1–5.
- Ian Nunes, Matheus B Pereira, Hugo Oliveira et al. "Conditional Reconstruction for Open-Set Semantic Segmentation". Em: *International Conference on Image Processing (ICIP)*. IEEE. 2022, pp. 946–950.

Referências II

- Thiago Carvalho, Jorge A Chamorro, Hugo Oliveira et al. "Outlier Exposure for Open Set Crop Recognition from Multitemporal Image Sequences". Em: *accepted for publication in Geoscience and Remote Sensing Letters (GRSL)* (2023).
- Hugo Oliveira e Jefersson Alex dos Santos. "Deep Transfer Learning for Segmentation of Anatomical Structures in Chest Radiographs". Em: *SIBGRAPI*. IEEE. 2018, pp. 204–211.
- Hugo Oliveira, Edemir Ferreira e Jefersson Alex dos Santos. "Truly Generalizable Radiograph Segmentation with Conditional Domain Adaptation". Em: *IEEE Access* 8 (2020), pp. 84037–84062.
- Hugo Oliveira, Virginia Mota, Alexei MC Machado et al. "From 3D to 2D: Transferring Knowledge for Rib Segmentation in Chest X-Rays". Em: *Pattern Recognition Letters* 140 (2020), pp. 10–17.
- Bruno A A Monteiro, Hugo Oliveira e Jefersson A dos Santos. "Self-Supervised Learning for Seismic Image Segmentation From Few-Labeled Samples". Em: *Geoscience and Remote Sensing Letters (GRSL)* 19 (2022), pp. 1–5.

Referências III

- Pedro Henrique Targino Gama, Hugo Oliveira e Jefersson Alex dos Santos. "Weakly Supervised Medical Image Segmentation". Em: *SIBGRAPI* (2021).
- Pedro Henrique Targino Gama, Hugo Oliveira, José Marcato Junior et al. "Weakly Supervised Few-Shot Segmentation Via Meta-Learning". Em: *arXiv preprint arXiv:2109.01693* (2021).
- Hugo Oliveira, Pedro HT Gama, Jefersson A dos Santos et al. "Domain Generalization in Medical Image Segmentation via Meta-Learners". Em: *SIBGRAPI*. IEEE. 2022.
- Longlong Jing e Yingli Tian. "Self-Supervised Visual Feature Learning with Deep Neural Networks: A Survey". Em: *IEEE TPAMI* (2020).
- Ting Chen, Simon Kornblith, Mohammad Norouzi et al. "A Simple Framework for Contrastive Learning of Visual Representations". Em: *ICML*. PMLR. 2020, pp. 1597–1607.
- Mathilde Caron, Ishan Misra, Julien Mairal et al. "Unsupervised Learning of Visual Features by Contrasting Cluster Assignments". Em: *NeurIPS*. Vol. 33. 2020, pp. 9912–9924.

Referências IV

- Xinlei Chen e Kaiming He. "Exploring Simple Siamese Representation Learning". Em: *CVPR*. 2021, pp. 15750–15758.
- Xiao Liu, Fanjin Zhang, Zhenyu Hou et al. "Self-Supervised Learning: Generative or Contrastive". Em: *IEEE Transactions on Knowledge and Data Engineering* 35.1 (2021), pp. 857–876.
- Kaiming He, Xinlei Chen, Saining Xie et al. "Masked AutoEncoders are Scalable Vision Learners". Em: *CVPR*. 2022, pp. 16000–16009.
- Chelsea Finn, Pieter Abbeel e Sergey Levine. "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks". Em: *ICML*. PMLR. 2017, pp. 1126–1135.
- Jake Snell, Kevin Swersky e Richard Zemel. "Prototypical Networks for Few-Shot Learning". Em: *NIPS*. 2017, pp. 4080–4090.
- Luca Bertinetto, Joao F Henriques, Philip HS Torr et al. "Meta-learning with differentiable closed-form solvers". Em: *arXiv preprint arXiv:1805.08136* (2018).
- Kwonjoon Lee, Subhransu Maji, Avinash Ravichandran et al. "Meta-Learning with Differentiable Convex Optimization". Em: *CVPR*. 2019, pp. 10657–10665.

Referências V

- Kate Rakelly, Evan Shelhamer, Trevor Darrell et al. "Few-shot Segmentation Propagation with Guided Networks". Em: *arXiv preprint arXiv:1806.07373* (2018).
- Kaixin Wang, Jun Hao Liew, Yingtian Zou et al. "PANet: Few-Shot Image Semantic Segmentation with Prototype Alignment". Em: *CVPR*. 2019, pp. 9197–9206.