



UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

CRISTIANO NÉIA DE OLIVEIRA BASSANI

**REDES NEURAS CONVOLUCIONAIS SEMI-SUPERVISIONADAS APLICADAS A
MUDANÇA DE DOMÍNIO**

DISSERTAÇÃO DE MESTRADO

CORNÉLIO PROCÓPIO
2022

CRISTIANO NÉIA DE OLIVEIRA BASSANI

**REDES NEURAS CONVOLUCIONAIS SEMI-SUPERVISIONADAS
APLICADAS A MUDANÇA DE DOMÍNIO**

**SEMI-SUPERVISED CONVOLUTIONAL NEURAL NETWORKS
APPLIED TO DOMAIN CHANGE**

Dissertação de Mestrado, apresentado ao Programa de Pós-Graduação em Informática da Universidade Tecnológica Federal do Paraná – UTFPR, como requisito para a obtenção do título de Mestre em Informática.

Orientador: Prof. Dr. Pedro Henrique Bugatti

CORNÉLIO PROCÓPIO

2022



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



**Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Cornélio Procópio**



CRISTIANO NEIA DE OLIVEIRA BASSANI

REDES NEURAS CONVOLUCIONAIS SEMI-SUPERVISIONADAS APLICADAS A MUDANÇA DE DOMÍNIO.

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Informática da Universidade Tecnológica Federal do Paraná (UTFPR).
Área de concentração: Computação Aplicada.

Data de aprovação: 16 de Agosto de 2022

Dr. Pedro Henrique Bugatti, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Alexandre Rossi Paschoal, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Marcelo Ponciano Da Silva, Doutorado - Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 16/08/2022.

Dedico este trabalho aos meus pais, Fernando e Valéria, aos meus irmãos André e Maria da Glória, minha companheira Sara, meus avós Maria e Djalma, minha avó Maria Aparecida e meu tio Marco, à meu falecido avô Juracy, meus amigos e irmãos Felício e Paulo, e a todos que contribuíram para esta caminhada.

AGRADECIMENTOS

Para se iniciar um projeto muitos fatores influenciam nas decisões de uma pessoa, intimamente sabemos que somos capazes de alcançar os nossos objetivos, sejam eles quais forem. Ao longo da jornada sempre nos depararemos com situações difíceis e complicadas, sejam elas profissionais ou pessoais. Para cada situação dessa Deus coloca anjos que nos guiam, dão forças e nos ajudam a não desvirtuar ou desistir do nosso caminho.

Pensando nisso, agradeço primeiramente à Deus, não somente por ter colocado todos estes anjos que aqui citarei em minha vida, mas também por ter me dado a vida, me concedido a sabedoria e me iluminado em todos os momentos desde quando nasci até o presente momento.

Agradeço aos meus pais, Valéria e Fernando, pois me proporcionaram absolutamente tudo para eu estar aqui, desde me conceder a vida, até a educação que me proporcionaram, e se hoje sou alguém, um homem formado, é tudo graças a vocês, que em hipótese alguma mediram esforços para fazer a mim e meus irmãos felizes. Eu sempre digo a mim mesmo que se eu tiver metade do caráter de vocês, já serei uma pessoa infinitamente bondosa e responsável, saibam que amo muito vocês e que aqui está o meu mais profundo Obrigado!

A minha companheira Sara, por sempre ter estado ao meu lado em todas as ocasiões, sejam elas tristes ou felizes, ruins ou boas. A cada decisão minha você me apoiou, muitas vezes nem mesmo concordando com o que eu decidia, mas ainda assim você me apoiava e nunca restringia. Você sempre me ajudou com sua sabedoria, e assim também farei contigo sempre, você é meu porto seguro e eu serei o seu. Te amo muito e obrigado por tudo!

Aos meus irmãos André e Maria da Glória, por me apoiarem e, por muitas vezes me aguentarem, as vezes eu passo dos limites com as brincadeiras, mas sempre é com objetivo de fazer vocês sorrirem, e não se enganem, vocês me ajudaram muito ao longo da jornada, André me dando suporte em diversos trabalhos das disciplinas ou conversando sobre várias coisas legais de madrugada, Maria da Glória me emprestando algumas coisas ou fazendo alguns favores enquanto eu estava ocupado trabalhando ou estudando, vocês foram um suporte para mim, eu amo muito vocês, muito obrigado meus irmãos!

Agradeço à todos meus avós, Maria, Djalma, Maria Aparecida e Juracy, por todo carinho, dedicação, atenção e paciência que tiveram sempre comigo. Peço perdão por muitas vezes estar ausente e não conseguir ter muito contato, mas saibam que amo demais

absolutamente todos vocês do fundo da alma, muito obrigado!

Por fim, agradeço imensamente a todos os meu amigos, Felício e Paulo, todos meus amigos que ganhei por toda vida e à família Unimed, em especial os amigos do setor de Contas Médicas, todos vocês são incríveis e foram imprescindíveis para este projeto do qual estou realizando, pois existem momentos em que não conseguimos estar na presença dos familiares, momentos do dia-a-dia, dos quais vocês estão sempre lá, seja pra dar um sorriso, uma palavra de alegria, uma risada, um apoio, ou até mesmo uma chamada de atenção. Saibam que todos vocês estão eternamente comigo em minhas palavras e orações, muito obrigado à todos!

Não ganhe o mundo e perca sua alma, sabedoria é melhor que prata e ouro. (Bob Marley)

RESUMO

BASSANI, Cristiano Néia de Oliveira. REDES NEURAIIS CONVOLUCIONAIS SEMI-SUPERVISIONADAS APLICADAS A MUDANÇA DE DOMÍNIO. 43 f. Dissertação – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

Dentre as muitas dificuldades observadas dentro de áreas afins da inteligência computacional, algumas se sobressaem devido a sua similaridade de ocorrência na grande maioria dos casos. Dentro desta área, e em específico dentro do processo de treinamentos de algoritmos de aprendizagem de máquina, um empecilho recorrente é a obtenção de amostras suficientes para realizar o treinamento dos algoritmos, bem como o gasto temporal e custoso de rotulação manual de amostras. Assim, com o intuito de mitigar tal problema, se propõe o uso de técnicas de aprendizado de máquina semi-supervisionado aplicado ao contexto de mudança de domínios em imagens atrelado a redes neurais convolucionais. Para tanto, foram realizados experimentos utilizando duas abordagens semi-supervisionadas (Self-Training e Co-Training) em conjunto redes neurais por convolução. Além disso, os experimentos foram realizados em 3 conjuntos de imagens públicos. Em suma, as contribuições do presente trabalho se apresentam como o desenvolvimento de diferentes arquiteturas de aprendizagem de máquina Semi-Supervisionada, a aplicação de Fine-Tuning e classificação de bases sobre diferentes domínios, os ganhos computacionais provenientes da utilização das técnicas Semi-Supervisionadas, as aplicabilidades do processo de Transfer Learning dentro da Aprendizagem de Máquina e por fim, os resultados obtidos das métricas de avaliação aplicadas sobre os diferentes datasets utilizados.

Palavras-chave: Aprendizado Semi-Supervisionado, Aprendizado de Máquina, Redes Neurais Covolucionais, Visão Computacional, Transferência de Aprendizado.

ABSTRACT

BASSANI, Cristiano Néia de Oliveira. SEMI-SUPERVISED CONVOLUTIONAL NEURAL NETWORKS APPLIED TO DOMAIN CHANGE. 43 f. Dissertação – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

Among the many difficulties observed within related areas of computational intelligence, some stand out due to their similarity of occurrence in the vast majority of cases. Within this area, and specifically within the training process of machine learning algorithms, a recurrent obstacle is obtaining adequate ones to carry out the training of algorithms, as well as the time and costly expense of manual rotation. Thus, in order to mitigate this problem, it is proposed the use of semi-supervised machine learning techniques applied to the context of changing domains in images linked to convolutional neural networks. For that, we used two semi-supervised strategies (self-training and co-training), as well as the use of three public image datasets. In short, the contributions of the present work are presented as the development of different architectures of Semi-Supervised machine learning, the application of Fine-Tuning and classification of bases on different domains, the computational gains from the use of Semi-Supervised, the applicability of the Transfer Learning process within Machine Learning and finally, the results obtained from the evaluation metrics applied to the different datasets used.

Keywords: Semi-Supervised Learning, Machine Learning, Convolutional Neural Network, Computational Vision, Transfer Learning

LISTA DE FIGURAS

FIGURA 1	– Modelo de aprendizado de máquina semi-supervisionado proposto por (AMORIM et al., 2019).	19
FIGURA 2	– Modelo de aprendizado de máquina semi-supervisionado Self-training proposto por (LING et al., 2018)	20
FIGURA 3	– Modelo de aprendizado de máquina semi-supervisionado Co-training proposto por (CHEN et al., 2019)	20
FIGURA 4	– Modelo descritivo do funcionamento de Rede Neural Convolutacional apresentado por (KIRANYAZ et al., 2021).	21
FIGURA 5	– Modelo esquemático do processo de Transfer Learning.	22
FIGURA 6	– Pseudo-rotulagem no processo de Self-Training usando transfer learning.	24
FIGURA 7	– Pseudo-rotulagem no processo de Co-Training usando transfer learning.	24
FIGURA 8	– Retreinamento a partir das amostras pseudo-rotuladas.	25
FIGURA 9	– Arquitetura Self-Training sobre diferentes domínios.	26
FIGURA 10	– Classes do conjunto de dados Papsmear. (a) carcinoma. (b) light. (c) moderate. (d) normal columnar. (e) normal intermediate. (f) severe dysplastic. (g) normal superficial	27
FIGURA 11	– Classes do conjunto de dados MAMMOSET. (a) malignant mass. (b) benign mass. (c) malignant calcification. (d) benign calcification	27
FIGURA 12	– Classes do conjunto de dados WHOI-Plankton. (a) Cerataulina. (b) Chaetoceros. (c) Ciliate-mix. (d) Corethron. (e) Cyliandrotheca. (f) Dactyliosolen. (g) Dino30. (h) Dinobryon. (i) Guinardia delicatula. (j) Guinardia striata. (k) Leptocylindrus. (l) Pseudonitzschia. (m) Rhizosolenia. (n) Thalassiosira.	27

LISTA DE TABELAS

TABELA 1	– Resultados obtidos por cada abordagem (Co-Training e Self-Training) para cada porcentagem do conjunto rotulado (10%, 20%, 30%, 40%, 50% e 100%) considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Papsmear	30
TABELA 2	– Resultados obtidos por cada abordagem (Co-Training e Self-Training) para cada porcentagem do conjunto rotulado (10%, 20%, 30%, 40%, 50% e 100%) considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Mammoset	31
TABELA 3	– Resultados obtidos por cada abordagem (Co-Training e Self-Training) para cada porcentagem do conjunto rotulado (10%, 20%, 30%, 40%, 50% e 100%) considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de WHOI-Plankton	31
TABELA 4	– Resultados obtidos por cada abordagem (Supervisionado, Co-training e Self-training) para cada classe considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Papsmear ..	32
TABELA 5	– Resultados obtidos por cada abordagem (Supervisionado, Co-training e Self-training) para cada classe considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Mammoset .	32
TABELA 6	– Resultados obtidos por cada abordagem (Supervisionado, Co-training e Self-training) para cada classe considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de WHOI-Plankton	33
TABELA 7	– Resultados gerais referente aos melhores resultados obtidos por percentual de dados de amostras rotuladas de treino por subdivisões de Fine-Tuning do Experimento Macro 1.	35
TABELA 8	– Resultados gerais referente aos melhores resultados obtidos por percentual de dados de amostras rotuladas de treino por subdivisões de Fine-Tuning do Experimento Macro 2.	35
TABELA 9	– Resultados específicos das métricas de avaliação geradas a partir da base de predição do Dataset Mammoset de divisão de treino de 20% de amostras rotuladas do Dataset WHOI-Plankton no Experimento Macro 1.	36
TABELA 10	– Resultados específicos das métricas de avaliação geradas a partir da base de predição do Dataset WHOI-Plankton de divisão de treino de 20% de amostras rotuladas do Dataset Papsmear no Experimento Macro 1.	37

LISTA DE SIGLAS

CNN - Convolutional Neural Network(Redes Neurais Convolucionais);

RGB - Red, Green, Blue(Vermelho, Verde, Azul; Padrão de cores de imagens);

OPF - Optimum Path Forest(Floresta de caminhos ótimos);

DDSM - Digital Dataset for Screening Mammography(Conjunto de dados digitais para triagem de mamografia);

SUMÁRIO

1	INTRODUÇÃO	13
1.1	PROBLEMATIZAÇÃO	14
1.2	MOTIVAÇÃO	15
1.3	OBJETIVOS	16
1.3.1	Objetivos Específicos	16
1.4	ORGANIZAÇÃO	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.0.1	Aprendizagem de Máquina	17
2.0.2	Aprendizado Supervisionado	17
2.0.3	Aprendizado Semi-Supervisionado	18
2.0.4	Self-Training e Co-Training	19
2.0.5	Redes Neurais Convolucionais	20
2.0.6	Transfer Learning	21
2.0.7	Fine-Tuning	22
3	METODOLOGIA PROPOSTA	23
3.1	DESCRIÇÃO DAS BASES DE IMAGENS	26
3.1.1	Papsmear	27
3.1.2	MAMMOSET (DDSM)	28
3.1.3	WHOI-Plankton	28
3.2	EXPERIMENTOS	28
3.3	RESULTADOS	30
3.3.1	Resultados com Domínio Fixo	30
3.3.2	Resultados com mudança de Domínio	32
4	CONCLUSÃO	39
4.0.1	Publicações	40
4.0.2	Trabalhos Futuros	40
	REFERÊNCIAS	41

1 INTRODUÇÃO

Com o advento do uso de aplicações de aprendizado de máquina e inteligência computacional, a procura por melhores ferramentas tecnológicas para diversos fins, sejam acadêmicos ou profissionais, tem aumentado consideravelmente. Dentro da área de aprendizado de máquina existem algumas vertentes das quais os pesquisadores buscam explorar e obter maiores conhecimentos, gerando contribuições para a comunidade científica. Dentre tais vertentes, existem três amplamente conhecidas, que são aplicações de aprendizagem de máquina Supervisionada, Não Supervisionada e Semi-supervisionada, conforme (KOTSIANTIS, 2007).

Para um algoritmo ser de aprendizagem Supervisionada, a base de dados estudada deve conter rótulos nas amostras, bem como características que auxiliam o algoritmo a reconhecer os padrões das amostras de acordo com as características de suas classes, sendo que, o mesmo elucida que algoritmos de aprendizagem de máquina Não Supervisionada possuem amostras com suas respectivas características, porém sem classes ou rótulos pré-existentes. No entanto a aprendizagem de máquina Semi-Supervisionada, pode ser vista como uma mescla das abordagens supracitadas, que faz uso de amostras rotuladas e não rotuladas, pois segundo (ENGELLEN; HOOS, 2020), a aprendizagem Semi-Supervisionada está conceitualmente situada entre a aprendizagem Supervisionada e Não Supervisionada, pois permite o uso de uma grande quantidade de amostras não rotuladas serem utilizadas com poucas amostras rotuladas.

Quando se tem o intuito de iniciar uma nova pesquisa abrangendo o uso de aplicações de aprendizado de máquina, se faz necessário o entendimento e a delimitação do problema a ser resolvido, pois para cada situação, uma abordagem pode se sobressair perante a outra, e o contrário pode ser válido em outra situação distinta. Em se tratando de visão computacional, pode-se citar alguns exemplos em que tais aprendizados de máquina foram utilizados, como no caso de (Ienco et al., 2020) que utilizou aprendizagem de máquina Supervisionada para realizar o mapeamento de ilhas a partir de imagens de satélite, ou a utilização de aprendizagem Não Supervisionada para auxiliar profissionais no processo de caracterização de tumores pulmonares e pancreáticos, como fez (Hussein et al., 2019). Por fim um exemplo de como (Choi; Song, 2020)

utilizou a aprendizagem Semi-Supervisionada, que no caso aqui citado, concentrou sua pesquisa na utilização da aprendizagem Semi-Supervisionada para identificar e reconhecer movimentos e emoções humanas a partir de rede neurais convolucionais. Vale ressaltar que as pesquisas supracitadas fizeram o uso de Redes Neurais Convolucionais (CNN-Convolutional Neural Network), que segundo (Kido et al., 2018) não requerem necessariamente que as características das imagens analisadas sejam extraídas a partir de um extrator de características, pois em outras palavras, as Redes Neurais Convolucionais realizam tal extração.

1.1 PROBLEMATIZAÇÃO

Tomando por base afirmações de que o uso de aprendizagem de máquina Supervisionada requer um processo custoso, que consome muito tempo quanto à rotulação de amostras, conforme (AMORIM et al., 2019), e que a aprendizagem Não Supervisionada foca no processo de clusterização de uma base de dados de amostras não rotuladas segundo (Hussein et al., 2019), a abordagem Semi-Supervisionada desperta o interesse dos pesquisadores para explorar suas funcionalidades e capacidades de classificação ou clusterização. Assim, porque tal interesse neste processo de aprendizagem de máquina? Para responder esta pergunta, é válido retroceder um pouco no que fora explicitado anteriormente, do qual a aprendizagem de máquina Semi-Supervisionada nada mais é que a mescla da abordagem Supervisionada e Não Supervisionada, agregando as características de rotulação e não rotulação de amostras ao mesmo tempo no processo de classificação.

Muitas perguntas surgem quando se adentra neste aprendizado de máquina, e ainda existem muitas destas sem respostas, pois existe uma vastidão de aplicabilidades a serem exploradas na mesma. Alguns pesquisadores procuram desenvolver arquiteturas que visam alcançar melhores resultados em métricas de avaliação para um dado problema, como o caso do próprio (AMORIM et al., 2019), que desenvolveu o algoritmo OPF-Semi, um algoritmo baseado em Floresta de caminhos ótimos (OPF - Optimum Path Forest) em junção com a abordagem Semi-supervisionada. O mesmo alcançou resultados relevantes para tal situação.

Outros autores como (T.R. et al., 2019) fazem uso do processo de Transfer Learning, ou Transferência de aprendizado por assim dizer, para montarem suas arquiteturas e re-treinar estas redes pré-existentes adaptando-as ao problema em questão. Levando em consideração que muitos trabalhos procuram re-treinar arquiteturas existentes para se adaptar ao problema proposto, surge a seguinte questão: Seria possível a classificação de amostras de um problema a partir de uma rede treinada previamente sob outro problema diferente? Pois, quando se utiliza aprendizagem de máquina Semi-Supervisionada, um dos objetivos pelo seu uso

é a capacidade de agregar características diferentes se comparando ao engessamento que a abordagem Supervisionada restringe às características mais marcantes das classes da base de dados treinada. Pensando de tal forma, mesmo que seja necessário que as classes previamente treinadas possuam certa similaridade, seria o algoritmo capaz de prever as classes de uma nova base distinta? Tal questão será abordada neste trabalho, e para elucidar o que seria essa similaridade de bases de dados, pode-se considerar por exemplo, que o algoritmo fora previamente treinado utilizando classes de tumores pulmonares, mas que consiga prever classes de problemas ósseos detectáveis a partir de imagens de raio-x, ou até mesmo identificar espécies de animais de uma determinada região do globo terrestre, com o algoritmo previamente treinado sob uma base de dados de animais distintos de uma região diferente.

1.2 MOTIVAÇÃO

A urgência na obtenção de resultados de maneira rápida e objetiva que o ramo profissional das grandes empresas exigem dos profissionais nos dias de hoje, tem aumentado consideravelmente devido a alta competitividade da concorrência de mercado. Quando se pensa em obter tais resultados de maneira rápida e efetiva, é necessário encontrar ou desenvolver abordagens que busquem encurtar caminhos ou trazer maior qualidade e confiabilidade para o processo desenvolvido ou a matéria prima produzida.

Pensando nisso, as aplicações de aprendizagem de máquina ou inteligência computacional, tem estado em evidência no processo de auxílio ou até mesmo efetivamente no processo de análise dos determinados problemas e as situações que as empresas que as fazem uso se propõe. Porém, para tanto, algumas destas aplicações utilizadas demandam custos tanto financeiro quanto de tempo gasto.

Quando se pensa em visão computacional, da qual muitas vezes se utiliza Redes Neurais Convolucionais em processos de reconhecimento de padrões, classificações, generalizações ou agrupamentos, não somente o treinamento dos algoritmos demanda um certo custo computacional, como também influencia diretamente no processo manual de rotulação que algumas aprendizagens de máquina exigem.

É válido destacar que cada problema tem sua resolução melhor realizada utilizando um determinado tipo de aprendizagem de máquina, e dada esta constatação, quando a aplicação utiliza redes neurais convolucionais, seja por abordagem supervisionada ou semi-supervisionada, é possível inferir que o processo de rotulação de amostras e treinamento dos algoritmos, por sua vez, pode ser relativamente maçante e custoso.

1.3 OBJETIVOS

O trabalho proposto teve como intuito analisar o comportamento de algoritmos de aprendizagem de máquina semi-supervisionadas em associação às redes neurais convolucionais, treinadas a partir do processo de transfer learning em diferentes contextos de imagens.

1.3.1 OBJETIVOS ESPECÍFICOS

Como objetivos específicos, visou-se observar e analisar a capacidade de classificação e generalização de bases de dados distintas em algoritmos previamente treinados sobre outras bases de dados, havendo um processo de refinamento de pesos treinando utilizando a técnica de Fine-Tuning. Além disso, foi realizada a análise de desempenho de tal processo a partir das métricas de avaliação (i.e. acurácia, revocação, F1-Score, precisão) e suas respectivas significâncias estatísticas.

1.4 ORGANIZAÇÃO

A dissertação foi composta por alguns tópicos, dos quais dentre são respectivamente a introdução como capítulo inicial, que tem por foco introduzir o tema abordado, explicitar a motivação pela qual fora escolhido o tema, bem como estruturar e objetivar os argumentos que auxiliem na constatação dos argumentos demonstrados.

O Capítulo 2 visou contextualizar as aplicações de aprendizagens de máquina, os processos e ferramentas que dão suporte para as abordagens utilizadas, bem como demonstrando a metodologia que foi utilizada, expondo a modelagem e funcionamento das arquiteturas geradas para a realização dos experimentos e cumprimentos das tarefas propostas.

O Capítulo 3 possui o intuito de descrever detalhadamente os passos processuais do trabalho, desde a metodologia desenvolvida e proposta para o desenvolvimento do algoritmo, até a descrição dos experimentos realizados e o resultados obtidos.

Por fim, o Capítulo 4 encerra as análises e discussões acerca do trabalho desenvolvido, realizando a explanação do contexto geral do trabalho por meio do tópico de conclusão.

2 FUNDAMENTAÇÃO TEÓRICA

A seção aqui descrita, tem intuito de introduzir e apontar os conceitos teóricos dos quais a presente pesquisa se baseou e fez uso de propostas desenvolvidas por outros autores presentes na literatura, como forma de auxílio ao desenvolvimento do trabalho.

2.0.1 APRENDIZAGEM DE MÁQUINA

Dentre os ramos existentes dentro de Inteligência Computacional, a Aprendizagem de Máquina ou Machine Learning, é a que mais se destaca por sua aplicabilidade em diversos problemas e situações devido seu alto poder de reconhecimento de padrões, classificação, generalizações e agrupamentos.

Conforme (QIU et al., 2016) esta área da Inteligência Computacional foca principalmente nas bases teóricas, performance e propriedades de aprendizagens dos algoritmos. Sua alta aplicabilidade em diversas áreas é justificada por sua multidisciplinaridade, auxiliando e impactando diferentes domínios da ciência, segundo aponta (RUDIN; WAGSTAFF, 2013).

Explorando a aprendizagem de máquina, pode-se observar que a mesma se subdivide em três diferentes paradigmas principais, que correspondem respectivamente à aprendizagem supervisionada, aprendizagem não-supervisionada e aprendizagem semi-supervisionada.

2.0.2 APRENDIZADO SUPERVISIONADO

Como dito anteriormente, a aprendizagem de máquina se subdivide em três tipos, dentre os quais está a aprendizagem Supervisionada, que já fora e ainda é a mais explorada juntamente com a abordagem Não Supervisionada. Para expor a maneira de funcionamento de tal abordagem de aprendizado de máquina, é válido verificar como funciona de maneira geral e como corresponde quando um algoritmo é desenvolvido utilizando tal abordagem, e para isso é necessário observar a disposição dos dados, que conforme (FONTANA, 2018), para se

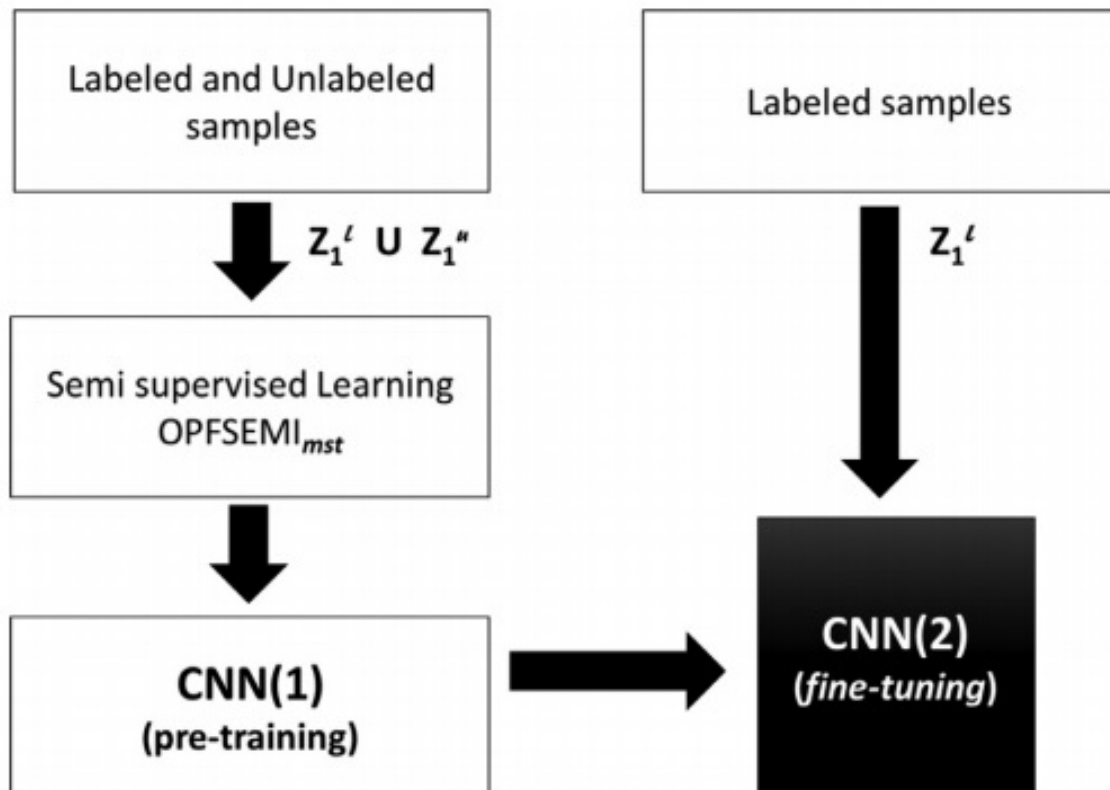
considerar um algoritmo de abordagem Supervisionada, o mesmo deve utilizar para a realização de treinamentos, amostras com dados rotulados.

2.0.3 APRENDIZADO SEMI-SUPERVISIONADO

O aprendizado de máquina Semi-Supervisionado, diferente do aprendizado Supervisionado e do aprendizado Não-Supervisionado, os quais são costumeiramente aplicados em muitos problemas previamente definidos envolvendo, respectivamente, bases rotuladas e não-rotuladas, não possui uma estrutura unicamente solidificada como as outras abordagens, e faz uso de amostras rotuladas e não rotuladas de forma simultânea.

Conforme (ENGELLEN; HOOS, 2020), a mesma situa-se entre a abordagem Supervisionada e Não-Supervisionada. Assim como as demais, esta aprendizagem pode ser utilizada em diversas áreas, desde aplicações de robótica, até processos de big data e visão computacional. Segundo (AMORIM et al., 2019), que em seu trabalho fez uso de aprendizagem Semi-Supervisionada em processos de Redes Neurais Convolucionais, os autores expõem que a necessidade do uso deste artifício surge do custo temporal e de recursos computacionais nos processos de rotulação manual de amostras. Portanto, em (AMORIM et al., 2019) é utilizada a abordagem Self-Training, que é uma das abordagens de aprendizado Semi-Supervisionado. Já em (CHEN et al., 2019) utilizou-se de outra vertente da abordagem Semi-Supervisionada, respectivamente fora aplicado o algoritmo Co-Training.

Figura 1: Modelo de aprendizado de máquina semi-supervisionado proposto por (AMORIM et al., 2019).



A Figura 1 exemplifica a abordagem semi-supervisionada proposta por (AMORIM et al., 2019), que faz uso do algoritmo Floresta de Caminhos Ótimos, mesclando amostras rotuladas com amostras não rotuladas e submetendo-as em uma arquitetura CNN.

2.0.4 SELF-TRAINING E CO-TRAINING

Adentrando no processo de aprendizagem de máquina Semi-Supervisionado, é possível observar que o mesmo se subdivide em várias vertentes, abrindo uma infinidade de opções para se explorar e utilizar de auxílio para as situações e processos a serem realizados. Duas destas vertentes se referem às arquiteturas ou algoritmos que realizam treinamentos e possibilitam a classificação e agrupamentos de padrões, das quais são as arquiteturas Self-Training e Co-Training.

No que se refere a arquitetura Self-Training, a mesma é uma das abordagens mais básicas e conhecidas do aprendizado Semi-Supervisionado, a mesma faz uso do processo de pseudo-rotulação de amostras não rotuladas para realizar os treinamentos e classificações. Já o Co-Training também faz uso do processo de pseudo-rotulação, porém o mesmo utiliza múltiplos algoritmos sobre mesmas bases para realizar validação cruzada no processo de rotulação, o qual

faz uso no processo de treinamento apenas amostras igualmente pseudo-rotuladas(ENGELEN; HOOS, 2020).

Figura 2: Modelo de aprendizado de máquina semi-supervisionado Self-training proposto por (LING et al., 2018)

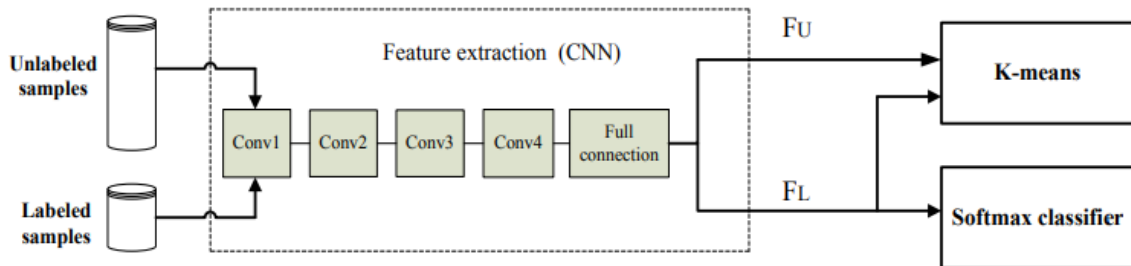
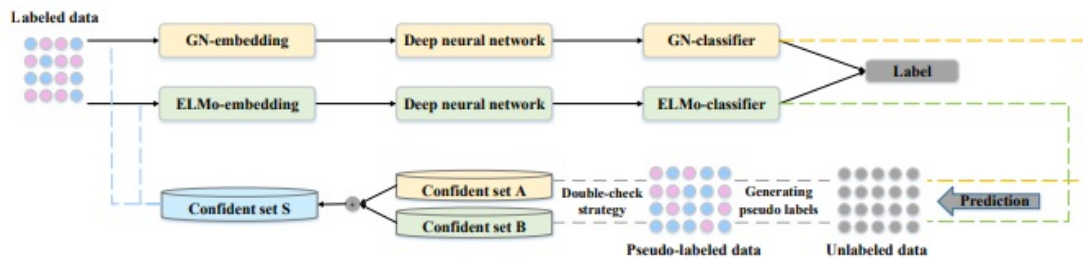


Figura 3: Modelo de aprendizado de máquina semi-supervisionado Co-training proposto por (CHEN et al., 2019)



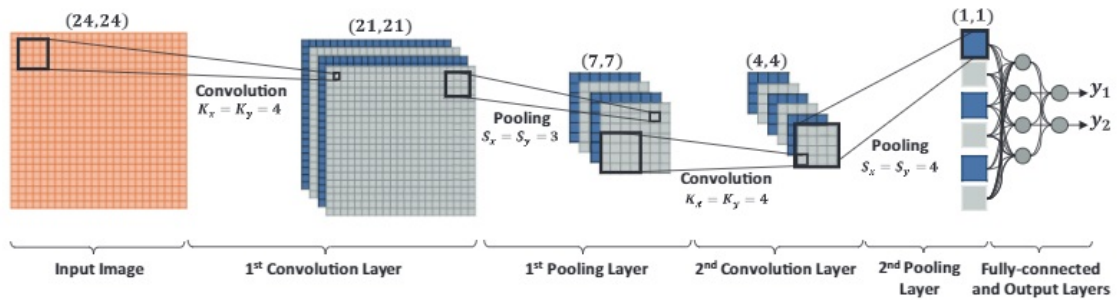
As Figuras 2 e 3 referenciam o modelo Self-training proposto por (LING et al., 2018), do qual partiu do uso simultâneo de amostras rotuladas e não rotuladas para gerar o treinamento da abordagem escolhida, bem como o modelo de abordagem Co-training proposto por (CHEN et al., 2019), que faz uso de um processo de validação de amostras pseudo-rotuladas a partir de dois algoritmos distintos, e que posteriormente aplica as amostras pseudo-rotuladas no processo de treinamento de uma rede final.

2.0.5 REDES NEURASIS CONVOLUCIONAIS

Conforme supracitado, os trabalhos realizados por (AMORIM et al., 2019) e (CHEN et al., 2019) fizeram uso das Redes Neurais Convolucionais como abordagens de extração de características e pré-processamento, assim os mesmos fizeram uso de bases de dados de imagens para solucionar as questões por eles propostas. As Redes Neurais Convolucionais, conhecidas também pela sigla CNN ou *Convolutional Neural Networks*, são camadas de pré-processamentos de imagens que realizam a extração de características em imagens e posteriormente submetem às camadas densas de Multi Layers Perceptron, ou Perceptrons

Multi Camadas, segundo (KIRANYAZ et al., 2021), para realizarem o refinamento de tais características e por fim gerar a classificação final da amostra.

Figura 4: Modelo descritivo do funcionamento de Rede Neural Convolucional apresentado por (KIRANYAZ et al., 2021).



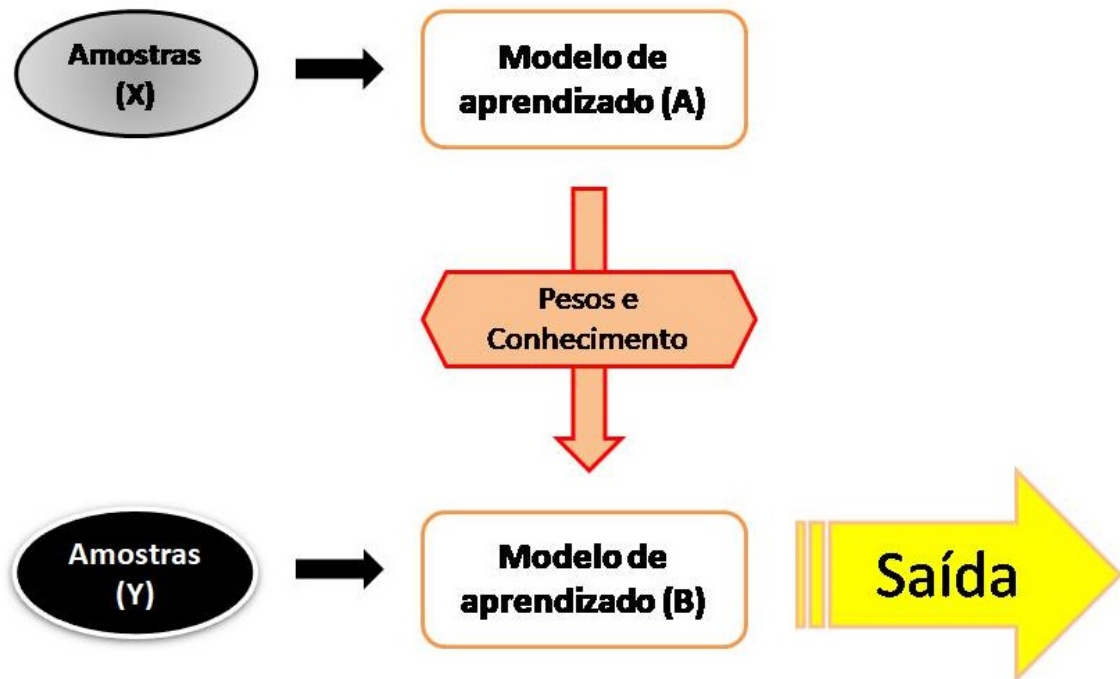
A Figura 4 ilustra o funcionamento das Redes Neurais Convolucionais, desde a amostra de entrada, que percorre as camadas de convolução e pooling realizando o pré-processamento, até ingressar na camada densa e de saída.

2.0.6 TRANSFER LEARNING

Um conceito que recentemente vem ganhando destaque pelos pesquisadores, o qual segundo (WEISS et al., 2016) apareceu em mais de 700 papers nos últimos tempos, e que frequentemente tem trazido benefícios pelo seu uso é o Transfer Learning, ou Transferência de Aprendizado.

A Transferência de Aprendizado consiste em reutilizar pesos previamente treinados sobre uma determinada arquitetura de aprendizado aplicando os mesmos no processo de treinamento de outro domínio abordado, pois conforme (WEISS et al., 2016) experiências obtidas em situações podem ser replicadas em situações semelhantes e quando se tem quantidades limitadas de dados relativos à nova questão a ser solucionada, bem como devido ao fato do processo de rotulação de amostras ser custoso quanto a tempo e desempenho computacional, convém aplicar um pré-treinamento a partir de dados semelhantes provenientes fazendo uso do chamado Transfer Learning.

Figura 5: Modelo esquemático do processo de Transfer Learning.



A Figura 5 exemplifica visualmente o funcionamento do processo de Transfer Learning, cuja finalidade tem por propagar o conhecimento gerado no treinamento de um modelo de aprendizado que pode ser replicado em outro modelo.

2.0.7 FINE-TUNING

Na literatura, muitos projetos científicos e tecnológicos em alguns momentos acabam necessitando de técnicas que auxiliem na obtenção de melhores resultados, muitas ferramentas e técnicas são utilizadas para tal, e uma delas é o chamado Fine-Tuning, ou como aponta (FRIEDERICH, 2017), seria este o “ajuste fino”. O termo é utilizado de maneira generalizada para adequar o trabalho a situação apresentada, refinando os detalhes dos parâmetros utilizados, e para tal, os autores promovem melhorias em suas pesquisas a partir destas técnicas de Fine-Tuning.

Para comprovar a fala anterior, um exemplo de melhoria feita por pesquisadores da literatura, pode se citar o caso de (CETINIC et al., 2018), que utilizou diferentes arquiteturas de CNN de maneira complementar para realizar o treinamento e teste de suas bases de dados.

3 METODOLOGIA PROPOSTA

Para a realização deste trabalho, inicialmente foram propostas três arquiteturas iniciais de aprendizagem de máquina, duas delas semi-supervisionadas, uma no formato de composição Self-Training e outra no formato de composição Co-Training, bem como uma outra na composição Supervisionada, das quais geraram os resultados iniciais para a verificação das métricas de avaliação.

Os experimentos destas arquiteturas iniciais foram realizados da seguinte maneira: associou-se o paradigma semi-supervisionado com CNNs para classificar imagens de 3 conjuntos de dados públicos. Para isso, foram utilizadas a Xception e a ResNet152_V2, que são duas arquiteturas CNN diferentes. Foram usados os mesmos hiperparâmetros para ambas as arquiteturas (por exemplo, taxas de aprendizado, otimizador, entre outros), redimensionamento de imagens, regularização e normalização de dados. Primeiro, considerou-se para a abordagem Self-Training e para a abordagem Supervisionada a utilização da arquitetura Xception, do qual fora definido a partir de testes iniciais, dos quais geraram melhores resultados dentre as aplicações da biblioteca Keras, a partir dos pesos previamente carregados do ImageNet. O Self-Training seleciona uma parte das amostras rotuladas para treinar a CNN. Em seguida, um modelo de aprendizado é gerado para classificar as amostras não rotuladas do conjunto de treinamento. Cada uma dessas amostras, rotuladas com pseudo-rótulos, são então agregadas incrementalmente ao conjunto rotulado para treinar novamente a CNN. As Figuras 6 e 7 mostram o pipeline das abordagens de Self-Training e Co-training propostas. A Figura 8 ilustra o retreinamento considerando as amostras pseudo-rotuladas. Assim ilustra o caminho que as amostras pseudo-rotuladas seguem depois do processo de pseudo-rotulação, das quais são submetidas ao modelo de CNN cujos pesos são previamente gerados e carregados no processo anterior de pseudo-rotulação. Nesta etapa, as amostras pseudo-rotuladas são inseridas no treinamento, propagando características e gerando novos pesos para promover a classificação de novas amostras

Em relação à abordagem de Co-Training, ela faz uso de duas CNNs diferentes(ou seja, Xception e ResNet152 V2) para gerar os pseudo-rótulos juntamente com o processo de

Figura 6: Pseudo-rotulagem no processo de Self-Training usando transfer learning.

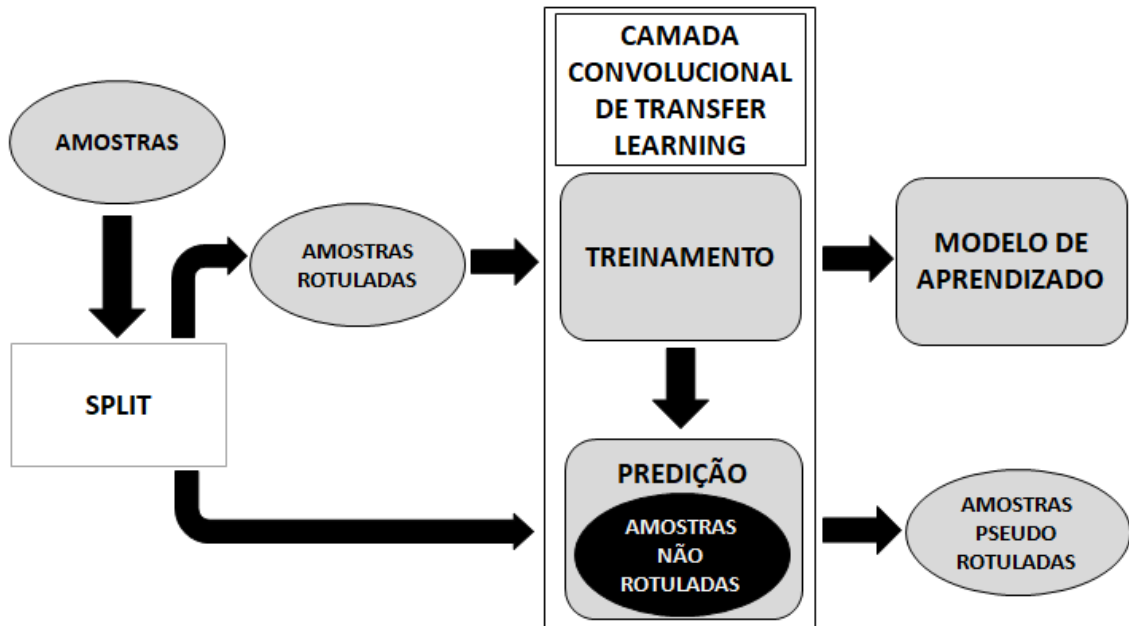
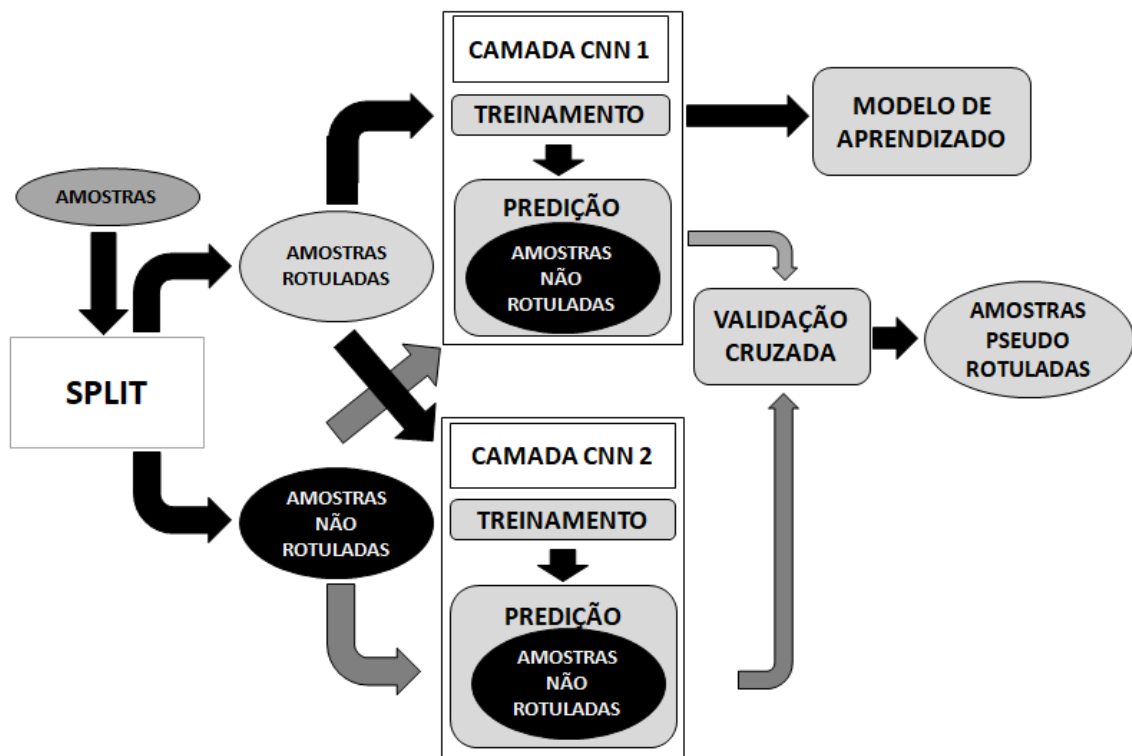


Figura 7: Pseudo-rotulagem no processo de Co-Training usando transfer learning.



validação cruzada dos rótulos previstos por ambas as CNNs. Nesta situação, usou-se apenas as amostras rotuladas igualmente entre as duas arquiteturas. Esta segunda abordagem semi-supervisionada é diferente da primeira. Neste caso, as mesmas amostras são submetidas a duas CNNs simultaneamente. Em seguida, são gerados dois conjuntos de pseudo-rótulos, que passam por uma validação mútua para atribuir um pseudo-rótulo para uma determinada amostra

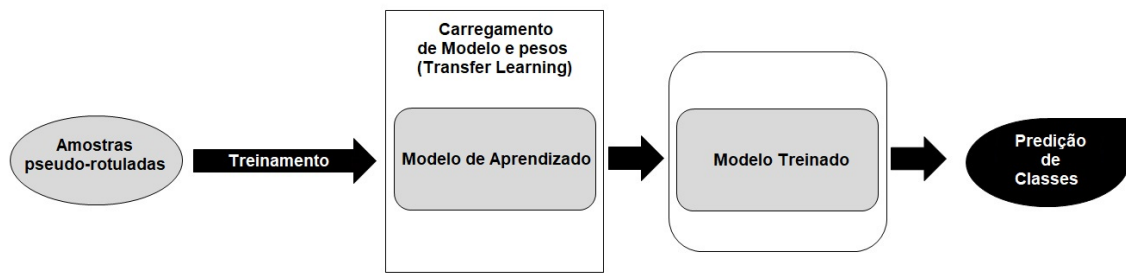


Figura 8: Retreinamento a partir das amostras pseudo-rotuladas.

não rotulada. O bloco de validação cruzada funciona da seguinte forma: se os pseudo-rótulos (saídas de ambas as arquiteturas) forem iguais, a amostra é inserida no conjunto de amostras pseudo-rotuladas. Caso contrário, a amostra não entra em tal um conjunto. No processo de validação cruzada, são descartadas quantidades de amostras como é possível inferir de acordo com a acurácia simultânea das duas arquiteturas, levando em consideração que são três diferentes bases de dados utilizadas e que serão expostas posteriormente, são descartadas aproximadamente entre 55% a 63% das amostras na base Papsmear, aproximadamente 63% e 69% das amostras da base MAMMOSET e por fim aproximadamente 13% e 18% na base de dados WHOI-Plankton. Terminada a pseudo-rotulagem, as seguintes redes, anteriormente treinadas são recarregadas e o novo conjunto de treinamento (conjunto rotulado mais o pseudo-rótulo) é submetido no processo de treinamento da CNN.

É importante notar que, inicialmente, considerando ambas as abordagens, todas as CNNs foram inicializadas com pesos (Transfer-Learning) da ImageNet(pré-treinados). Neste caso, a Figura 7 mostra o processo de pseudo-rotulagem realizado pela abordagem Co-Training. O processo de retreinamento é idêntico ao da Self-Training (Figura 8).

Para avaliar ambas as abordagens propostas, primeiro dividiu-se cada conjunto de dados de imagem em um conjunto de treinamento e de teste, considerando 80% e 20%, respectivamente. Então, o conjunto de treinamento foi dividido em conjuntos rotulados e não rotulados. Para melhor analisar as abordagens propostas foram consideradas diferentes divisões de conjuntos rotulados (10%, 20%, 30%, 40% e 50%). Essa mesma metodologia foi aplicada 10 vezes para gerar conjuntos mutuamente exclusivos e fornecer análises estatísticas para o processo de avaliação.

Após definidos os experimentos supracitados, foram aplicadas as abordagens de aprendizagem de máquina Semi-Supervisionada baseada no modelo Self-Training e Co-training. Para a realização dos experimentos e obtenção de resultados, foram utilizadas três bases públicas como datasets de treinamento, sendo elas Papsmear, WHOI-Plankton e MAMMOSET (ver Seção 3.1).

Partindo para a abordagem proposta, o algoritmo foi desenvolvido de maneira a utilizar diferentes particionamentos das bases de imagens empregadas, subdividindo em imagens não rotuladas, imagens rotuladas e imagens de teste. Seguindo para o pipeline do método proposto, a mesma utilizou uma rede convolucional e usou inicialmente somente as imagens previamente rotuladas. Posteriormente, as imagens não rotuladas foram selecionadas e com base no processo de classificação, foram atribuídos pseudo-rótulos à estas amostras referenciadas.

Tendo concretizada a pseudo-rotulação, a abordagem foi submetida a estas novas amostras pseudo-rotuladas novamente ao processo de treinamento, tomando por base a mesma rede convolucional. Por fim, após este treinamento, a abordagem foi submetida a dois testes: o primeiro se refere ao processo de classificação e generalização de uma base com as mesmas classes às da base de dados inicialmente treinada; o segundo, e principal experimento foi a aplicação do processo de generalização e classificação de bases distintas à abordagem recém treinada, verificando sua capacidade de obter assertividade das classes preditas quando treinada com uma base divergente da base de teste, como é possível ver na Figura 9 o seu funcionamento, ressaltando que para este segundo caso utilizou-se apenas a abordagem de Self-training.

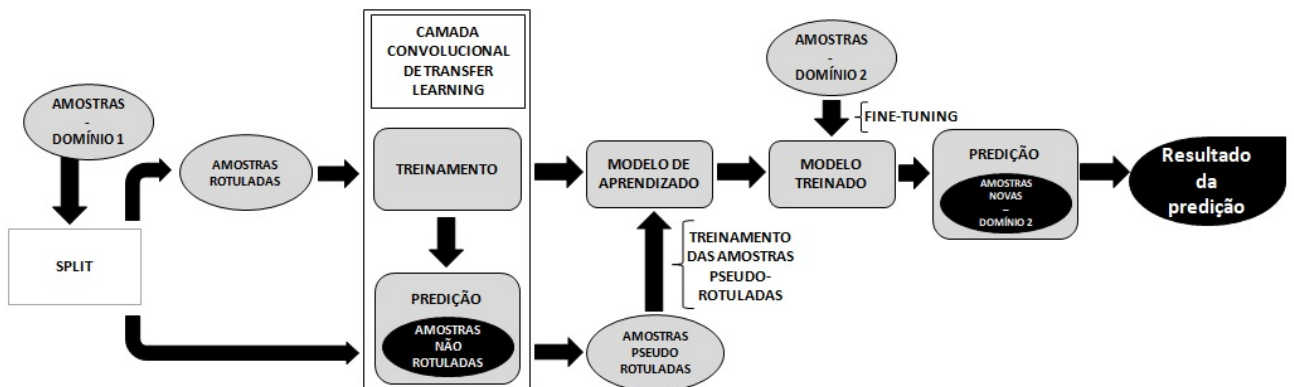


Figura 9: Arquitetura Self-Training sobre diferentes domínios.

3.1 DESCRIÇÃO DAS BASES DE IMAGENS

Para a realização do presente trabalho foram definidos 3 datasets. Para tanto, foram selecionados os datasets públicos Papsmear, MAMMOSET e WHOI-Plankton, descritos na presente seção.

A base Papsmear corresponde a imagens cujas amostras referem-se a células saudáveis e cancerosas de exames de papanicolau do cólon do útero. Já, o dataset MAMMOSET é uma base de imagens que referem-se a exames radiológicos utilizados para detecção de câncer na mama referentes a achados envolvendo calcificação ou massa. Por fim, o dataset WHOI-

Plankton é um acervo de imagens de diferentes espécies de plânctons, as quais podem gerar diferentes impactos em ecossistemas marinhos.

3.1.1 PAPSMEAR

O Dataset Papsmear é uma base de dados de imagens composta por imagens de lâminas cervicais de útero, a qual possibilita a detecção de estágios pré-cancerígenos do cólon do útero, conforme (KOSS, 1993).

As Figuras 10 de (a)-(g) representam exemplos de imagens de cada uma das classes referentes aos estágios pré-cancerígenos do colo de útero do dataset papsmear. As imagens do respectivo dataset possuem dimensões que variam de 53x129 pixels a 252x28 pixels, bem como quantidades de amostras desbalanceadas de classe para classe. As imagens encontram-se no espaço de cores RGB com profundidade de 8 bits. Nesta base de dados, foram utilizadas 50 amostras por classe de treinamento e 20 amostras por classe na base de testes, totalizando 490 amostras.

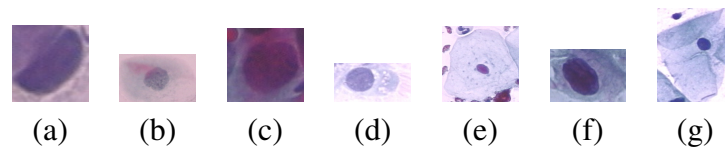


Figura 10: Classes do conjunto de dados Papsmear. (a) carcinoma. (b) light. (c) moderate. (d) normal columnar. (e) normal intermediate. (f) severe dysplastic. (g) normal superficial

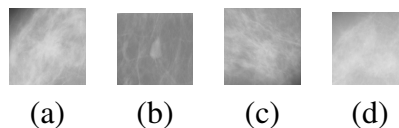


Figura 11: Classes do conjunto de dados MAMMOSET. (a) malignant mass. (b) benign mass. (c) malignant calcification. (d) benign calcification

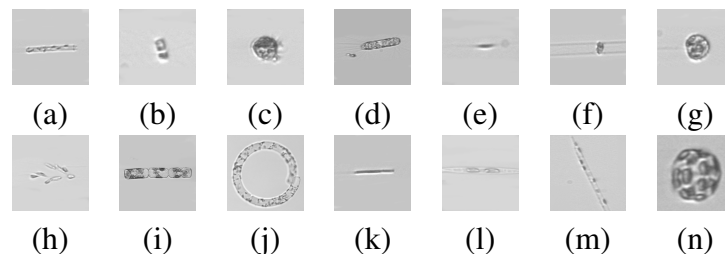


Figura 12: Classes do conjunto de dados WHOI-Plankton. (a) Cerataulina. (b) Chaetoceros. (c) Ciliate-mix. (d) Corethron. (e) Cyliandrotheca. (f) Dactyliosolen. (g) Dino30. (h) Dinobryon. (i) Guinardia delicatula. (j) Guinardia striata. (k) Leptocylindrus. (l) Pseudonitzschia. (m) Rhizosolenia. (n) Thalassiosira.

3.1.2 MAMMOSET (DDSM)

O dataset MAMMOSET é uma base de dados que possui imagens de regiões de interesse (*regions of interest* - ROIs) da mama provindas de exames radiológicos, utilizados para detectar pontos na imagem que possam indicar um futuro câncer ou até mesmo um câncer em estágio avançado, segundo (OLIVEIRA et al., 2017).

As Figuras 11(a)-(d) compreendem exemplos de imagens das classes do dataset MAMMOSET. Diferentemente do dataset papsmear, neste caso as amostras encontram-se em escalas de cinza, com dimensões igualitárias de 1000x1000 pixels e com quantidades balanceadas de 500 amostras por classe de treinamento e 100 amostras por classe de teste, totalizando 2400 imagens.

3.1.3 WHOI-PLANKTON

De acordo com (ORENSTEIN et al., 2015), o dataset WHOI-Plankton é uma base de dados que possui larga quantidade de imagens de uma vasta quantidade de espécies oceânicas de planctons, utilizada usualmente para o processo de classificação de espécies por pesquisadores, para que possam catalogá-las.

Por fim, o dataset WHOI-Plankton apresenta-se como o de maior cardinalidade a ser utilizado. Devido sua alta quantidade de classes e amostras, bem como por limitações de processamento total do dataset, foram selecionadas randomicamente apenas quatorze classes de diferentes espécies de plâncton, vistos nas Figuras 12(a)-(n). Tais amostras compreendem imagens em escalas de cinza, com dimensões variadas. Além disso, trata-se de uma dataset com quantidades irregulares de amostras por classes. Para tanto, as quantidades foram regularizadas da seguinte maneira, 100 amostras por classe na base de treinamento e 20 amostras por classe na base de testes, totalizando 1680 amostras.

3.2 EXPERIMENTOS

Para os experimentos iniciais foram comparadas as abordagens semi-supervisionadas baseadas em Self-training e a Co-training, atreladas a CNNs, contra a abordagem supervisionada tradicional.

Além disso, para o desenvolvimento científico deste trabalho, houve a necessidade da realização de dois tipos de experimentos aqui intitulados Experimentos Macro, que são subdivididos a cada passo do algoritmo, sendo estes 3 passos ao todo. Porém, ao final se

totalizam por trinta diferentes variações destes dois Experimentos Macro, e para cada um dos três datasets. As cinco primeiras variações dos experimentos consistem na subdivisão de amostras rotuladas treinadas no primeiro passo, seguindo por 50% de dados rotulados da primeira base de dados e indo respectivamente para 40%, 30%, 20% e 10% de dados rotulados.

Ambos Experimentos Macro possuem similaridades entre si, porém o que será divergente é a última parte do processo. Para começar a explicar sobre tais experimentos, partimos do passo inicial, onde a abordagem recebe as amostras da base de dados inicial, por exemplo utilizando a base Mammaset. Feito isso, a abordagem realiza os treinamentos e gera pseudo-rótulos de amostras desta mesma base, ou seja, pseudo-rotula amostras não rotuladas da base de dados Mammaset.

O segundo passo começa a partir deste ponto, em que as imagens pseudo-rotuladas da primeira base de dados são utilizadas como fontes de dados para o treinamento e geração dos pesos que serão propagados através da abordagem, porém neste passo não existe a predição de novas amostras não rotuladas desta base de dados, apenas o salvamento dos pesos de treinamento.

Por fim, o terceiro e último passo deste experimento, parte do pressuposto de carregar os pesos gerados no passo anterior na arquitetura CNN. Diferentemente dos passos iniciais, este passo recebe amostras de domínio diferente, ou seja amostras de outro conjunto de imagens diferente do mammoset por exemplo, agora é utilizado Papsmear ou WHOI-Plankton. Neste ponto o experimento é segmentado, subdividindo o processo macro do experimento em 3 vertentes. Para entender o funcionamento das 3 vertentes, é válido sinalizar o funcionamento do terceiro e último passo, que por meio das amostras do novo domínio a ser treinado, gera um processo de Fine-Tuning dos pesos carregados do passo anterior, por meio do treinamento de amostras rotuladas desta nova base. As três vertentes de experimentos gerada consistem na utilização de três diferentes porcentagens de amostras para a realização do Fine-Tuning, das quais são respectivamente 100% de amostras rotuladas do novo domínio, 20% de amostras rotuladas e 10% de amostras rotuladas desta base de dados. De maneira geral, vale ressaltar que a partir deste ponto, após o processo de Fine-Tuning, é gerada a predição de amostras não rotuladas deste domínio, obtendo os valores das métricas de avaliação de acurácia, precisão, revocação e F1-Score

3.3 RESULTADOS

3.3.1 RESULTADOS COM DOMÍNIO FIXO

Para analisar as abordagens iniciais propostas (Self-Training e Co-Training agregados às CNNs) fora elaborado um comparativo com o paradigma supervisionado, considerando os três conjuntos de dados mencionados anteriormente.

A Tabela 1 mostra os resultados obtidos considerando o conjunto de dados do Papsmear. De acordo com os experimentos, todas as abordagens apresentaram melhores resultados (todas as métricas), considerando 50% das amostras rotuladas quando comparadas com a supervisionada (100% das amostras rotuladas). Por exemplo, a abordagem Co-training CNN obteve um ganho de até 15% de acurácia.

Tabela 1: Resultados obtidos por cada abordagem (Co-Training e Self-Training) para cada porcentagem do conjunto rotulado (10%, 20%, 30%, 40%, 50% e 100%) considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Papsmear

Labeled	Approach	Acurácia	Precisão	Revocação	F1-Score
10%	Co-Training	37.71%±0.255	42.03%±0.258	39.74%±0.267	38.69%±0.253
	Self-Training	33.43%±0.317	41.51%±0.365	37.29%±0.351	33.11%±0.297
20%	Co-Training	38.86%±0.262	40.91%±0.253	40.91%±0.276	39.09%±0.248
	Self-Training	33.43%±0.305	37.57%±0.319	35.26%±0.321	32.43%±0.282
30%	Co-Training	38.14%±0.280	38.71%±0.262	40.20%±0.294	38.34%±0.270
	Self-Training	35.86%±0.314	38.89%±0.315	37.77%±0.329	34.91%±0.294
40%	Co-Training	40.14%±0.273	41.54%±0.262	42.26%±0.287	40.00%±0.253
	Self-Training	37.14%±0.350	40.49%±0.364	40.17%±0.379	35.83%±0.341
50%	Co-Training	45.00%±0.276	44.83%±0.261	47.37%±0.291	44.83%±0.273
	Self-Training	39.86%±0.345	44.43%±0.332	42.00%±0.363	38.46%±0.306
100%	Supervisionado	38.86%±0.290	42.03%±0.311	44.00%±0.322	40.66%±0.303

Esse mesmo comportamento foi observado ao analisar os conjuntos de dados Mammoset e WHOI-Plankton (Tabelas 2 e 3). Por exemplo, considerando os resultados obtidos a partir do dataset Mammoset (Tabela 2), a abordagem Co-training CNN obteve um ganho de até 19% de acurácia com apenas 40% de amostras rotuladas, quando comparadas com o supervisionado. Em relação à precisão e ao F1-Score, os ganhos foram ainda maiores (45% e 38%, respectivamente). A arquitetura Self-Training também apresentou melhores resultados que o paradigma supervisionado.

Analisando os resultados do conjunto de dados WHOI-Plankton, o Self-training alcançou os melhores resultados (88,71% de acerto) com 40% de amostras rotuladas, enquanto o supervisionado obteve uma acurácia de 84,86%.

Para demonstrar os melhores resultados obtidos por estas abordagens, é possível observar nas Tabelas 4, 5 e 6, respectivamente, todas as métricas para cada classe do conjuntos de dados Papsmear, Mammoset e WHOI-Plankton, considerando os melhores casos.

Tabela 2: Resultados obtidos por cada abordagem (Co-Training e Self-Training) para cada porcentagem do conjunto rotulado (10%, 20%, 30%, 40%, 50% e 100%) considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Mammoset

Labeled	Approach	Acurácia	Precisão	Revocação	F1-Score
10%	Co-Training	31.56%±0.115	33.70%±0.113	33.25%±0.121	32.55%±0.104
	Self-Training	30.03%±0.330	28.85%±0.286	31.55%±0.348	24.60%±0.267
20%	Co-Training	34.75%±0.118	36.50%±0.106	36.50%±0.125	35.88%±0.108
	Self-Training	26.52%±0.328	14.45%±0.240	27.95%±0.346	17.80%±0.252
30%	Co-Training	35.83%±0.121	39.10%±0.132	37.95%±0.128	37.60%±0.114
	Self-Training	35.08%±0.305	36.05%±0.281	36.90%±0.321	31.10%±0.239
40%	Co-Training	37.55%±0.128	39.60%±0.113	39.40%±0.135	38.65%±0.110
	Self-Training	36.64%±0.352	33.10%±0.167	38.50%±0.371	31.45%±0.216
50%	Co-Training	36.59%±0.166	39.15%±0.137	38.50%±0.175	37.30%±0.134
	Self-Training	37.00%±0.319	38.70%±0.215	38.85%±0.336	32.60%±0.200
100%	Supervisionado	30.49%±0.365	21.80%±0.251	33.75%±0.405	23.95%±0.265

Tabela 3: Resultados obtidos por cada abordagem (Co-Training e Self-Training) para cada porcentagem do conjunto rotulado (10%, 20%, 30%, 40%, 50% e 100%) considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de WHOI-Plankton

Labeled	Approach	Acurácia	Precisão	Revocação	F1-Score
10%	Co-Training	82.43%±0.124	87.70%±0.109	86.73%±0.131	86.49%±0.100
	Self-Training	78.93%±0.172	85.84%±0.137	83.09%±0.181	82.79%±0.131
20%	Co-Training	84.00%±0.115	88.67%±0.086	88.43%±0.122	88.09%±0.091
	Self-Training	85.71%±0.113	90.87%±0.085	90.17%±0.119	89.94%±0.084
30%	Co-Training	85.29%±0.106	90.39%±0.076	89.70%±0.113	89.49%±0.073
	Self-Training	87.79%±0.095	93.29%±0.082	92.36%±0.100	92.21%±0.069
40%	Co-Training	86.29%±0.106	91.31%±0.067	90.76%±0.112	90.50%±0.068
	Self-Training	88.71%±0.095	94.17%±0.073	93.36%±0.100	93.13%±0.065
50%	Co-Training	87.00%±0.102	91.91%±0.068	91.53%±0.108	91.29%±0.072
	Self-Training	88.71%±0.110	94.01%±0.065	93.39%±0.116	92.97%±0.076
100%	Supervisionado	84.86%±0.082	94.70%±0.060	94.17%±0.090	94.01%±0.060

Por exemplo, a Tabela 4 mostra os resultados para cada classe considerando a abordagem supervisionada o Co-Training com 50% de amostras rotuladas e o Self-Training também com 50% de amostras rotuladas (ou seja, melhor caso da Tabela 1) para a base Papsmear. Analisando os resultados, pode-se observar que a CNN Co-Training apresentou melhores resultados para as classes 1 e 3 a 6 em comparação com a supervisionada. Por exemplo, para a classe 1 o ganho de precisão foi 2,8 vezes melhor.

Considerando a Tabela 5, base MAMMOSET, a CNN de Co-Training com 40% de amostras rotuladas apresenta melhores resultados para as classes 0, 1 e 3. Vale ressaltar que a CNN de Self-Training (50% de amostras rotuladas amostras) obteve para a classe 0 uma precisão 4,4 vezes melhor que a supervisionada. Por fim, ao analisar a Tabela 6, base WHOI-Plancton, as abordagens apresentaram os melhores resultados para quase todas as classes. Esses resultados levam ao passo seguinte do presente trabalho, do qual gera um ensemble na abordagem semi-supervisionada Self-Training.

Tabela 4: Resultados obtidos por cada abordagem (Supervisionado, Co-training e Self-training) para cada classe considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Papsmear

Class	Metrics	Supervisionado	Co-Training 50%	Self-Training 50%
0	Acurácia	13.00%±0.178	5.00%±0.035	14.00%±0.204
	Precisão	10.20%±0.143	19.00%±0.122	12.00%±0.098
	Revocação	14.40%±0.198	5.20%±0.039	14.80%±0.217
	F1-Score	12.00%±0.166	8.20%±0.057	12.20%±0.135
1	Acurácia	16.00%±0.219	45.00%±0.145	34.00%±0.245
	Precisão	23.80%±0.223	35.40%±0.048	48.20%±0.158
	Revocação	31.00%±0.298	47.40%±0.154	35.80%±0.258
	F1-Score	26.40%±0.241	40.00%±0.073	36.40%±0.139
2	Acurácia	25.00%±0.145	20.00%±0.117	19.00%±0.129
	Precisão	26.60%±0.080	26.40%±0.178	48.40%±0.307
	Revocação	26.80%±0.174	21.00%±0.124	20.00%±0.136
	F1-Score	23.00%±0.052	23.00%±0.139	23.60%±0.114
3	Acurácia	37.00%±0.256	47.00%±0.044	14.00%±0.171
	Precisão	22.00%±0.020	39.60%±0.062	8.20%±0.096
	Revocação	30.20%±0.268	49.60%±0.050	14.80%±0.181
	F1-Score	21.20%±0.104	44.00%±0.051	10.20%±0.120
4	Acurácia	70.00%±0.252	83.00%±0.067	92.00%±0.044
	Precisão	73.80%±0.178	75.00%±0.067	67.00%±0.178
	Revocação	80.00%±0.294	87.40%±0.072	96.80%±0.048
	F1-Score	76.00%±0.241	80.60%±0.065	77.80%±0.150
5	Acurácia	74.00%±0.147	77.00%±0.057	70.00%±0.183
	Precisão	92.40%±0.107	88.20%±0.068	95.20%±0.074
	Revocação	79.00%±0.146	81.00%±0.057	73.80%±0.191
	F1-Score	84.00%±0.099	84.00%±0.019	81.00%±0.098
6	Acurácia	37.00%±0.130	38.00%±0.103	36.00%±0.386
	Precisão	45.40%±0.108	30.20%±0.053	32.00%±0.247
	Revocação	46.60%±0.201	40.00%±0.110	38.00%±0.407
	F1-Score	42.00%±0.102	34.00%±0.067	28.00%±0.230

Tabela 5: Resultados obtidos por cada abordagem (Supervisionado, Co-training e Self-training) para cada classe considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de Mammaset

Class	Metrics	Supervisionado	Co-Training 40%	Self-Training 50%
0	Acurácia	14.20%±0.317	29.40%±0.068	62.60%±0.377
	Precisão	12.80%±0.286	44.80%±0.019	46.40%±0.162
	Revocação	15.80%±0.353	30.80%±0.072	66.00%±0.398
	F1-Score	14.20%±0.318	36.40%±0.050	43.60%±0.249
1	Acurácia	15.74%±0.161	26.38%±0.055	16.60%±0.166
	Precisão	24.20%±0.230	21.60%±0.038	22.00%±0.182
	Revocação	17.20%±0.177	27.60%±0.056	17.20%±0.174
	F1-Score	20.00%±0.198	23.80%±0.037	19.20%±0.176
2	Acurácia	46.60%±0.442	40.60%±0.064	35.60%±0.347
	Precisão	24.80%±0.269	44.00%±0.054	31.60%±0.315
	Revocação	51.60%±0.490	42.60%±0.064	37.40%±0.365
	F1-Score	31.20%±0.296	43.20%±0.050	34.00%±0.335
3	Acurácia	45.40%±0.437	53.80%±0.091	33.20%±0.302
	Precisão	25.40%±0.284	48.00%±0.040	54.80%±0.389
	Revocação	50.40%±0.486	56.60%±0.099	34.80%±0.317
	F1-Score	30.40%±0.288	51.20%±0.025	33.60%±0.307

3.3.2 RESULTADOS COM MUDANÇA DE DOMÍNIO

Para analisar a metodologia de predição de classes sobre diferentes domínios, bem como verificar a efetividade das abordagens aplicadas, se faz necessário visualizar ponto-a-ponto os resultados obtidos a partir das métricas de avaliação desta arquitetura aprimorada.

Tabela 6: Resultados obtidos por cada abordagem (Supervisionado, Co-training e Self-training) para cada classe considerando cada métrica (Acurácia, Precisão, Revocação e F1-Score) para o conjunto de dados de WHOI-Plankton

Class	Metrics	Supervisionado	Co-Training 50%	Self-Training 40%
0	Acurácia	61.00%±0.108	60.00%±0.061	64.00%±0.102
	Precisão	100.00%±0.000	90.40%±0.059	100.00%±0.000
	Revocação	67.80%±0.117	63.20%±0.065	67.40%±0.107
	F1-Score	80.20%±0.089	74.00%±0.035	79.80%±0.076
1	Acurácia	89.00%±0.022	78.00%±0.075	93.00%±0.044
	Precisão	100.00%±0.000	90.20%±0.092	96.20%±0.060
	Revocação	98.80%±0.027	81.80%±0.080	97.80%±0.049
	F1-Score	99.40%±0.013	85.60%±0.072	96.80%±0.032
2	Acurácia	84.00%±0.022	89.00%±0.022	89.00%±0.022
	Precisão	96.40%±0.033	94.00%±0.064	93.20%±0.053
	Revocação	93.00%±0.022	93.80%±0.027	93.80%±0.026
	F1-Score	94.60%±0.025	93.60%±0.038	93.20%±0.021
3	Acurácia	84.00%±0.022	93.00%±0.027	81.00%±0.082
	Precisão	95.40%±0.047	94.00%±0.042	93.40%±0.049
	Revocação	93.00%±0.022	98.00%±0.027	85.20%±0.084
	F1-Score	94.20%±0.018	95.80%±0.011	88.80%±0.048
4	Acurácia	90.00%±0.000	95.00%±0.000	95.00%±0.000
	Precisão	100.00%±0.000	94.00%±0.042	100.00%±0.000
	Revocação	100.00%±0.000	100.00%±0.000	100.00%±0.000
	F1-Score	100.00%±0.000	96.80%±0.020	100.00%±0.000
5	Acurácia	90.00%±0.000	95.00%±0.000	95.00%±0.000
	Precisão	98.00%±0.027	97.00%±0.045	100.00%±0.000
	Revocação	100.00%±0.000	100.00%±0.000	100.00%±0.000
	F1-Score	98.80%±0.016	98.40%±0.023	100.00%±0.000
6	Acurácia	90.00%±0.000	93.00%±0.027	95.00%±0.000
	Precisão	97.20%±0.063	94.20%±0.062	98.00%±0.027
	Revocação	100.00%±0.000	98.00%±0.027	100.00%±0.000
	F1-Score	98.40%±0.036	96.00%±0.042	98.80%±0.016
7	Acurácia	88.00%±0.027	88.00%±0.057	92.00%±0.044
	Precisão	82.40%±0.051	85.60%±0.047	77.60%±0.058
	Revocação	97.60%±0.033	92.60%±0.062	96.80%±0.048
	F1-Score	89.40%±0.030	88.80%±0.034	86.00%±0.041
8	Acurácia	90.00%±0.000	95.00%±0.000	95.00%±0.000
	Precisão	94.00%±0.042	98.00%±0.027	89.80%±0.065
	Revocação	100.00%±0.000	100.00%±0.000	100.00%±0.000
	F1-Score	96.80%±0.020	98.80%±0.016	94.40%±0.037
9	Acurácia	85.00%±0.035	84.00%±0.074	91.00±0.054
	Precisão	89.00%±0.070	88.60%±0.097	96.00%±0.041
	Revocação	94.20%±0.039	88.20%±0.078	95.60%±0.060
	F1-Score	91.40%±0.038	88.00%±0.044	95.40%±0.015
10	Acurácia	90.00%±0.000	95.00%±0.000	91.00%±0.065
	Precisão	98.00%±0.027	89.80%±0.064	87.20%±0.096
	Revocação	100.00%±0.000	100.00%±0.000	95.80%±0.069
	F1-Score	98.80%±0.016	94.60%±0.036	90.40%±0.040
11	Acurácia	83.00%±0.027	86.00%±0.022	84.00%±0.108
	Precisão	93.20%±0.045	92.20%±0.029	97.80%±0.030
	Revocação	92.00%±0.027	90.20%±0.027	88.40%±0.116
	F1-Score	92.60%±0.013	91.40%±0.025	92.40%±0.066
12	Acurácia	85.00%±0.035	85.00%±0.050	89.00%±0.065
	Precisão	89.40%±0.032	85.00%±0.097	94.60%±0.052
	Revocação	94.20%±0.039	89.40%±0.055	93.60%±0.070
	F1-Score	91.80%±0.022	86.80%±0.050	94.00%±0.053
13	Acurácia	79.00%±0.054	82.00%±0.097	88.00%±0.027
	Precisão	92.80%±0.027	93.80%±0.065	94.60%±0.005
	Revocação	87.80%±0.059	86.20%±0.101	92.60%±0.032
	F1-Score	89.80%±0.045	89.40%±0.047	93.80%±0.016

Iniciando a explanação dos seguintes experimentos, parte-se do entendimento da lógica de funcionamento dos mesmos, para o Experimento Macro 1 apenas nas cinco camadas iniciais foram retreinadas e tiveram seus pesos ajustados pelo processo de Fine-Tuning, mantendo o

treinamento inicial das camadas mais profundas, nesta situação foram retreinadas duas camadas convolucionais e duas camadas de pooling apenas. No caso do Experimento Macro 2, o algoritmo passou por um retreinamento em sua totalidade, gerando e ajustando os pesos camada por camada da arquitetura de Rede Neural Convolucional, ou seja, neste caso o Fine-Tuning fora realizado em todas as camadas de convolução, de pooling e densas, de maneira resumida, para o experimento Macro 1 as 5 camadas mais profundas arquitetura Xception utilizadas foram retreinadas, que são duas camadas de convolução, duas de pooling e 1 densa e para o experimento Macro 2, todas as camadas da Xception foram retreinadas. Para tanto, pode e deve-se observar tanto de maneira mais ampla para os resultados mais generalizados, analisando tanto a nível de arquitetura ou Experimento Macro, quanto por maiores especificidades, vendo a nível de classes ou domínio aplicado. Com isso, neste trabalho é possível observar os melhores valores macro obtidos nos experimentos apresentados nas Tabelas 7-8, sendo respectivamente a Tabela 7 o Experimento Macro 1 e a Tabela 8 referente aos melhores resultados do Experimento Macro 2.

Por questões de facilidade na leitura e interpretação dos experimentos da presente seção, foi definida a seguinte nomenclatura: porcentagem de amostras rotuladas% base_de_treino/base_de_predição. Por exemplo, a configuração 10% Pap/Mam significa que primeiramente a abordagem foi treinada com 10% de amostras rotuladas e, posteriormente, tal abordagem foi aplicada para a predição da base MAMMOSET. Dessa forma, caracterizando a mudança de domínio na análise. Além de tais análises, foram também realizadas análises referentes ao grau de fine-tuning da arquitetura, ou seja, após o treinamento em um domínio diverso foram aplicadas diferentes porcentagens (100%, 20% e 10%) de amostras do domínio alvo para calibrar a arquitetura. Para os experimentos de mudança de domínio foi aplicada o Self-training como abordagem semi-supervisionada.

Tomando por base os resultados macros apresentados, é possível visualizar que, quando se utilizou as camadas mais profundas dentro do Experimento Macro 2 (Tabela 8), os resultados sofreram perda significativa de desempenho em todas as variantes do experimento. A exemplo disso, verificando o resultado mais significativo e discrepante entre o Experimento Macro 1 (Tabela 7) e o Experimento Macro 2 (Tabela 8), pode-se observar a seguinte situação: no caso em que a base inicial de treinamento e pseudo-rotulação foi realizada com o Dataset WHOI-Plankton e o Dataset de Fine-Tuning e predição foi a Mammoset, os melhores resultados do Experimento Macro 1 estiveram na subdivisão de treinamento e pseudo-rotulação de 20% de dados rotulados no treinamento, onde obteve-se acurácia de $57.12\% \pm 0.1551$ na subdivisão de 100% de amostras no processo de Fine-Tuning, $25.23\% \pm 0.1120$ na subdivisão de 20% de amostras no processo de Fine-Tuning, e por fim $30.47\% \pm 0.1554$ na subdivisão de 10%

Tabela 7: Resultados gerais referente aos melhores resultados obtidos por percentual de dados de amostras rotuladas de treino por subdivisões de Fine-Tuning do Experimento Macro 1.

% Treino/Predição	Fine-Tuning 100% Acc - Prec - Rec - F1	Fine-Tuning 20% Acc - Prec - Rec - F1	Fine-Tuning 10% Acc - Prec - Rec - F1
10% Pap/Mam	29.32% 30.60% 26.65% 16.10%	39.04% 41.85% 39.60% 34.40%	29.97% 32.60% 30.50% 25.50%
20% Pap/Mam	27.54% 25.45% 26.45% 18.10%	26.20% 27.40% 28.25% 21.35%	23.76% 30.35% 28.05% 21.75%
30% Pap/Mam	24.34% 22.00% 25.85% 16.10%	30.48% 31.15% 29.25% 22.45%	28.24% 26.60% 28.80% 21.25%
40% Pap/Mam	28.24% 21.20% 25.85% 17.20%	35.71% 37.15% 34.90% 30.60%	33.54% 39.30% 36.95% 32.30%
50% Pap/Mam	25.17% 15.70% 24.35% 13.25%	32.78% 36.90% 35.65% 30.55%	33.54% 37.65% 36.10% 32.60%
10% Who/Mam	24.09% 17.15% 26.55% 13.60%	27.59% 28.15% 28.70% 22.80%	34.11% 41.90% 36.60% 30.50%
20% Who/Mam	57.12% 51.95% 51.25% 48.95%	25.23% 28.10% 27.00% 22.20%	30.47% 30.10% 31.05% 26.50%
30% Who/Mam	40.51% 33.30% 37.75% 31.20%	38.14% 42.80% 39.75% 37.05%	38.01% 40.85% 38.10% 36.85%
40% Who/Mam	44.98% 41.75% 42.20% 39.15%	40.31% 42.20% 38.50% 35.30%	36.42% 37.75% 36.40% 34.60%
50% Who/Mam	37.83% 35.15% 35.40% 31.75%	39.04% 41.60% 39.85% 37.60%	30.15% 33.00% 31.80% 28.75%
10% Mam/Pap	32.69% 36.20% 32.69% 32.09%	15.55% 7.86% 15.57% 8.60%	12.22% 10.69% 14.46% 8.17%
20% Mam/Pap	32.06% 36.14% 32.09% 28.69%	14.12% 15.26% 14.14% 8.03%	14.92% 5.34% 14.91% 5.86%
30% Mam/Pap	31.90% 34.74% 31.94% 30.57%	16.19% 9.11% 16.17% 10.09%	15.71% 9.43% 15.71% 8.17%
40% Mam/Pap	23.65% 21.94% 23.69% 21.00%	13.01% 5.54% 13.03% 6.60%	16.50% 5.54% 16.49% 7.46%
50% Mam/Pap	35.39% 35.54% 35.46% 32.63%	17.14% 9.71% 17.17% 8.83%	15.55% 6.11% 15.50% 5.93%
10% Who/Pap	35.55% 36.91% 35.66% 33.94%	29.68% 29.86% 29.69% 27.26%	20.63% 21.46% 20.71% 16.49%
20% Who/Pap	39.20% 37.64% 37.54% 35.64%	31.11% 32.31% 31.20% 28.89%	21.11% 20.09% 21.17% 16.49%
30% Who/Pap	36.82% 38.14% 36.89% 34.51%	27.14% 28.49% 27.23% 24.37%	24.44% 24.23% 24.49% 19.54%
40% Who/Pap	35.87% 33.94% 35.91% 33.03%	32.38% 33.66% 32.37% 27.51%	22.53% 22.00% 22.66% 19.29%
50% Who/Pap	39.36% 37.49% 39.37% 37.03%	30.00% 28.74% 30.09% 27.54%	23.17% 22.86% 23.23% 21.00%
10% Pap/Who	92.46% 93.04% 92.33% 92.06%	54.20% 60.59% 54.26% 53.07%	35.39% 44.83% 35.47% 32.56%
20% Pap/Who	91.98% 92.70% 91.89% 91.53%	52.61% 57.67% 52.67% 50.53%	38.41% 45.84% 38.46% 35.86%
30% Pap/Who	91.66% 92.53% 91.57% 91.19%	58.01% 59.19% 58.01% 55.41%	41.98% 43.80% 42.01% 36.77%
40% Pap/Who	92.38% 92.96% 92.88% 92.36%	54.60% 55.89% 54.61% 51.80%	42.77% 51.23% 43.52% 41.09%
50% Pap/Who	92.30% 92.93% 92.14% 91.87%	53.33% 56.63% 53.34% 51.80%	39.68% 49.14% 39.69% 37.83%
10% Mam/Who	90.47% 91.16% 90.37% 90.14%	42.54% 45.41% 42.60% 39.99%	27.69% 31.01% 27.79% 24.36%
20% Mam/Who	90.95% 91.71% 90.84% 90.67%	46.30% 50.79% 45.91% 43.11%	33.01% 37.87% 33.06% 29.87%
30% Mam/Who	91.03% 91.96% 90.91% 90.66%	46.42% 48.04% 46.40% 43.17%	30.87% 31.97% 30.90% 27.33%
40% Mam/Who	90.87% 91.66% 90.76% 90.57%	41.98% 46.04% 42.03% 38.93%	31.35% 30.99% 31.36% 27.46%
50% Mam/Who	89.68% 91.03% 89.57% 89.41%	41.50% 45.99% 41.51% 38.70%	40.23% 40.43% 40.23% 35.84%

Tabela 8: Resultados gerais referente aos melhores resultados obtidos por percentual de dados de amostras rotuladas de treino por subdivisões de Fine-Tuning do Experimento Macro 2.

% Treino/Predição	Fine-Tuning 100% Acc - Prec - Rec - F1	Fine-Tuning 20% Acc - Prec - Rec - F1	Fine-Tuning 10% Acc - Prec - Rec - F1
10% Pap/Mam	23.59% 7.85% 25.85% 11.55%	23.32% 19.85% 28.60% 16.20%	26.29% 21.15% 26.50% 15.25%
20% Pap/Mam	22.55% 9.00% 24.95% 10.40%	16.54% 11.95% 24.15% 9.20%	26.45% 18.40% 25.25% 16.65%
30% Pap/Mam	24.97% 8.60% 24.70% 11.75%	20.61% 16.05% 25.60% 11.75%	22.77% 22.80% 25.10% 14.85%
40% Pap/Mam	22.26% 8.60% 24.25% 10.90%	21.40% 9.70% 25.00% 10.20%	24.47% 21.65% 25.55% 13.80%
50% Pap/Mam	21.21% 10.75% 25.05% 9.25%	16.80% 9.85% 24.75% 7.95%	21.69% 11.40% 25.00% 11.20%
10% Who/Mam	23.22% 6.45% 24.90% 10.15%	24.34% 18.30% 23.75% 12.95%	23.50% 9.90% 23.75% 11.55%
20% Who/Mam	22.48% 6.50% 24.95% 10.15%	25.87% 24.00% 27.05% 15.30%	22.32% 18.15% 24.45% 12.55%
30% Who/Mam	23.47% 11.55% 25.10% 10.55%	20.95% 18.50% 24.85% 11.25%	27.37% 19.55% 28.05% 16.30%
40% Who/Mam	20.34% 5.75% 25.00% 9.15%	23.57% 21.25% 29.15% 17.35%	22.74% 25.00% 25.45% 16.30%
50% Who/Mam	19.07% 16.60% 25.05% 8.50%	29.13% 21.30% 26.05% 15.80%	25.93% 25.20% 25.00% 17.10%
10% Mam/Pap	22.85% 19.14% 22.97% 16.40%	18.25% 8.09% 18.29% 8.43%	13.17% 3.03% 13.20% 3.83%
20% Mam/Pap	16.34% 14.11% 16.43% 11.71%	17.62% 13.00% 17.63% 8.86%	13.33% 3.43% 13.34% 4.80%
30% Mam/Pap	19.83% 12.89% 19.86% 13.34%	20.00% 12.51% 19.94% 12.03%	15.07% 4.80% 15.11% 6.91%
40% Mam/Pap	24.44% 22.31% 24.51% 20.11%	13.96% 3.66% 13.91% 5.54%	14.44% 7.49% 14.51% 5.60%
50% Mam/Pap	17.61% 19.17% 17.71% 14.71%	21.74% 21.79% 23.64% 16.46%	14.28% 2.03% 14.29% 3.60%
10% Who/Pap	36.50% 37.89% 36.54% 33.57%	27.45% 27.49% 27.51% 23.14%	15.87% 18.74% 15.97% 11.29%
20% Who/Pap	36.34% 37.31% 36.40% 35.06%	29.04% 32.97% 29.06% 26.11%	20.00% 22.63% 20.09% 16.60%
30% Who/Pap	35.71% 39.11% 35.77% 34.94%	24.60% 31.17% 24.66% 24.51%	18.72% 17.63% 18.80% 15.11%
40% Who/Pap	36.98% 38.57% 37.06% 35.77%	26.19% 26.66% 26.20% 23.34%	20.31% 19.74% 20.37% 17.83%
50% Who/Pap	33.96% 37.26% 34.03% 32.89%	27.29% 25.20% 27.31% 23.34%	21.26% 21.74% 21.37% 16.97%
10% Pap/Who	66.74% 64.94% 66.73% 62.76%	41.50% 43.59% 41.53% 37.17%	21.42% 19.67% 21.47% 15.56%
20% Pap/Who	72.77% 73.73% 72.70% 70.00%	34.36% 30.87% 34.37% 27.07%	23.17% 17.69% 23.17% 16.71%
30% Pap/Who	75.39% 77.46% 75.31% 73.24%	32.69% 29.27% 32.73% 26.40%	29.12% 26.01% 29.10% 22.37%
40% Pap/Who	80.55% 80.61% 80.53% 78.83%	36.90% 36.16% 36.91% 33.19%	22.06% 22.13% 22.06% 18.00%
50% Pap/Who	76.82% 76.71% 76.70% 74.01%	38.73% 37.23% 38.71% 34.49%	24.28% 23.71% 24.37% 18.60%
10% Mam/Who	45.15% 41.60% 45.19% 38.70%	15.95% 10.01% 15.99% 9.64%	16.11% 11.46% 15.11% 9.18%
20% Mam/Who	50.07% 48.64% 50.04% 44.60%	22.93% 16.44% 22.93% 15.90%	14.20% 7.50% 14.24% 6.09%
30% Mam/Who	45.71% 42.61% 45.73% 40.43%	18.17% 13.27% 18.24% 11.59%	11.34% 5.79% 11.34% 5.50%
40% Mam/Who	45.95% 41.53% 45.93% 40.99%	11.98% 5.73% 11.99% 6.09%	11.66% 5.81% 11.67% 5.37%
50% Mam/Who	65.95% 64.19% 65.90% 62.37%	20.15% 14.41% 20.20% 13.94%	11.74% 6.37% 11.77% 5.40%

de amostras no processo de Fine-Tuning. Já os melhores resultados obtidos no segundo Experimento Macro estão contidos na separação de 10% dos domínios Who/Mam, onde as acurácias alcançaram respectivamente nas divisões de 100%, 20% e 10% de Fine-Tuning os valores de $23.22\% \pm 0.4083$, $24.34\% \pm 0.2587$, e por fim $23.50\% \pm 0.3577$.

Considerando os experimentos nos quais a base de imagens Papsmear foi o domínio fonte, e a base de imagens WHOI-Plankton foi utilizada no Fine-Tuning e na predição, os melhores resultados obtidos foram os oriundos de 30% de dados rotulados no Experimento Macro 1 e 40% de dados rotulados no Experimento Macro 2. Os resultados de acurácia geral no Experimento Macro 1 alcançaram nas subdivisões de 100%, 20% e 10% de amostras rotuladas de treino, respectivamente aos seguintes números de acurácia e desvio padrão, o primeiro teve acurácia de $91.66\% \pm 0.0778$ de desvio padrão, $58.01\% \pm 0.2918$, e $41.98\% \pm 0.2875$. Já para o Experimento Macro 2 obteve-se acurácias de $80.55\% \pm 0.2103$ na subdivisão de 100% amostras de Fine-Tuning, $36.90\% \pm 0.3005$ na subdivisão de 20% de amostras de Fine-Tuning e por fim $22.06\% \pm 0.2519$.

Atentando-se para a demais métricas de avaliação, focando especificamente nestes dois resultados específicos, pode-se observar que com base nos valores de precisão, revocação e F1-Score, é possível notar que os valores não destoam entre si, e até mesmo se comparados aos valores de acurácia, os valores se mantem muito próximos, o que leva ao próximo ponto de exploração, que seria a verificação dos valores destas métricas de avaliação a nível de classes nas bases de dados de predição.

Iniciando tal análise é possível notar no primeiro caso, em que o a base de imagens de predição foi o Mammoset, as classes preditas tiveram resultados relativamente discrepantes entre si. Algumas classes foram melhor preditas que outras, a exemplo disso, pode-se notar que resultados de precisão foram melhores na classe 2 se comparada com a classe 1, com $67.60\% \pm 0.0934$, enquanto a classe 1 obteve $26.40\% \pm 0.0577$, conforme se pode visualizar na Tabela 9.

No segundo caso, em que o Dataset de predição foi o WHOI-Plankton, percebe-se que as métricas são mais equilibradas entre as classes, não há uma discrepância tão significativa entre elas, as mesmas se mantém dentro de um limiar de equilíbrio, como pode-se observar na Tabela 10.

Tabela 9: Resultados específicos das métricas de avaliação geradas a partir da base de predição do Dataset Mammoset de divisão de treino de 20% de amostras rotuladas do Dataset WHOI-Plankton no Experimento Macro 1.

% de Fine-Tuning	Classe	Acurácia	Precisão	Revocação	F1-Score
100%	0	$53.41\% \pm 0.0837$	$52.80\% \pm 0.0396$	$44.40\% \pm 0.2452$	$43.80\% \pm 0.2361$
	1	$40.74\% \pm 0.1317$	$26.40\% \pm 0.0577$	$45.20\% \pm 0.1835$	$31.40\% \pm 0.0371$
	2	$72.14\% \pm 0.0549$	$67.60\% \pm 0.0934$	$64.00\% \pm 0.0731$	$65.60\% \pm 0.0802$
	3	$62.21\% \pm 0.1422$	$61.00\% \pm 0.1995$	$51.40\% \pm 0.2461$	$55.00\% \pm 0.2432$
20%	0	$31.92\% \pm 0.1374$	$34.60\% \pm 0.0737$	$33.80\% \pm 0.2591$	$29.20\% \pm 0.0936$
	1	$23.42\% \pm 0.0918$	$14.80\% \pm 0.0444$	$38.20\% \pm 0.2463$	$19.40\% \pm 0.0921$
	2	$20.75\% \pm 0.1292$	$23.00\% \pm 0.1414$	$20.40\% \pm 0.1866$	$19.20\% \pm 0.1417$
	3	$24.89\% \pm 0.0841$	$40.00\% \pm 0.0822$	$15.60\% \pm 0.1055$	$21.00\% \pm 0.1093$
10%	0	$42.10\% \pm 0.1604$	$32.60\% \pm 0.0780$	$51.40\% \pm 0.2609$	$37.00\% \pm 0.0959$
	1	$27.23\% \pm 0.0720$	$19.40\% \pm 0.0537$	$34.40\% \pm 0.2023$	$22.60\% \pm 0.1152$
	2	$25.49\% \pm 0.2131$	$33.20\% \pm 0.2382$	$20.00\% \pm 0.2009$	$24.00\% \pm 0.2189$
	3	$27.10\% \pm 0.1255$	$35.20\% \pm 0.1402$	$18.40\% \pm 0.1397$	$22.40\% \pm 0.1443$

Tabela 10: Resultados específicos das métricas de avaliação geradas a partir da base de predição do Dataset WHOI-Plankton de divisão de treino de 20% de amostras rotuladas do Dataset Papsmear no Experimento Macro 1.

% de Fine-Tuning	Classe	Acurácia	Precisão	Revocação	F1-Score
100%	0	72.01%±0.0274	98.00%±0.0447	51.20%±0.0268	67.20%±0.0249
	1	98.37%±0.0149	97.00%±0.0274	100.00%±0.0000	98.20%±0.0164
	2	94.01%±0.0194	94.20%±0.0390	94.00%±0.0000	94.20%±0.0179
	3	84.14%±0.0426	90.20%±0.0311	78.80%±0.0733	83.80%±0.0444
	4	100.00%±0.0000	100.00%±0.0000	100.00%±0.0000	100.00%±0.0000
	5	98.89%±0.0152	98.00%±0.0274	100.00%±0.0000	98.80%±0.0164
	6	99.39%±0.0137	100.00%±0.0000	98.80%±0.0268	99.40%±0.0134
	7	84.54%±0.0418	72.80%±0.0572	97.60%±0.0329	83.60%±0.0462
	8	91.81%±0.0197	84.40%±0.0358	100.00%±0.0000	91.40%±0.0219
	9	90.01%±0.0397	87.20%±0.0268	93.00%±0.0616	90.20%±0.0402
	10	96.02%±0.0431	92.60%±0.0823	100.00%±0.0000	95.80%±0.0460
	11	91.21%±0.0193	95.20%±0.0268	87.80%±0.0589	91.00%±0.0212
	12	92.41%±0.0139	92.00%±0.0274	93.00%±0.0224	92.60%±0.0134
13	90.54%±0.0324	93.80%±0.0045	87.80%±0.0589	90.40%±0.0329	
20%	0	46.59%±0.1009	52.20%±0.1613	40.80%±0.0753	45.20%±0.0963
	1	60.59%±0.1146	66.80%±0.1763	54.60%±0.1445	58.80%±0.1165
	2	84.85%±0.0653	90.20%±0.1238	79.00%±0.0809	83.80%±0.0743
	3	66.65%±0.0270	71.00%±0.0925	62.40%±0.1290	65.00%±0.0490
	4	88.25%±0.0534	78.60%±0.0786	97.80%±0.0492	86.80%±0.0482
	5	79.05%±0.0992	98.00%±0.0447	62.20%±0.1548	75.40%±0.1220
	6	0.00%±0.0000	0.00%±0.0000	0.00%±0.0000	0.00%±0.0000
	7	54.59%±0.1838	56.60%±0.0888	53.00%±0.2951	52.60%±0.1798
	8	57.52%±0.0721	31.60%±0.1045	93.40%±0.0984	46.00%±0.1012
	9	16.25%±0.0065	31.40%±0.0358	6.00%±0.0000	9.80%±0.0045
	10	15.96%±0.1638	13.60%±0.1301	17.80%±0.1887	15.20%±0.1519
	11	78.72%±0.0618	82.60%±0.1379	74.40%±0.0619	77.60%±0.0550
	12	77.19%±0.0971	77.40%±0.1139	77.80%±0.2184	74.80%±0.1270
13	85.99%±0.0597	78.60%±0.1379	93.00%±0.0616	84.80%±0.0756	
10%	0	30.24%±0.1064	29.40%±0.1246	30.00%±0.2009	27.40%±0.1176
	1	9.50%±0.0900	11.80%±0.1145	6.80%±0.0733	7.60%±0.0730
	2	48.64%±0.2323	64.00%±0.2666	35.60%±0.2556	42.40%±0.2617
	3	27.38%±0.1918	52.20%±0.4685	12.40%±0.1537	16.20%±0.1610
	4	72.64%±0.1161	53.20%±0.1612	94.40%±0.0550	66.40%±0.1258
	5	71.44%±0.2550	79.80%±0.2510	63.40%±0.2881	67.20%±0.2505
	6	63.98%±0.2692	58.40%±0.2150	69.00%±0.3718	60.00%±0.2947
	7	45.91%±0.2884	47.20%±0.2662	42.20%±0.3116	43.00%±0.2760
	8	30.38%±0.2867	57.20%±0.5253	12.20%±0.1470	19.00%±0.2189
	9	10.78%±0.0247	14.60%±0.0251	6.00%±0.0000	7.80%±0.0045
	10	0.00%±0.0000	0.00%±0.0000	0.00%±0.0000	0.00%±0.0000
	11	62.91%±0.1480	68.00%±0.1967	60.00%±0.2928	56.80%±0.1771
	12	55.18%±0.1771	35.40%±0.1588	78.60%±0.1476	47.60%±0.1721
13	58.78%±0.0699	42.00%±0.1239	77.60%±0.1565	52.80%±0.1182	

A partir deste ponto, é possível inferir que, para determinados Datasets, a arquitetura CNN que fora desenvolvida e utilizada se adapta melhor que em outros Datasets. Por exemplo, como no caso em que as amostras de predição foi realizada para a base de imagens WHOI-Plankton, onde houve um equilíbrio generalizado entre as métricas de avaliação de cada classe, pois os resultados ficaram muito próximos, mas também é válido se considerar que, para outros tipos de Datasets, no caso do Mammoset, o algoritmo se ajusta melhor a uma determinada classe, onde somente ela fora predita de maneira mais assertiva, tendo as demais classes do Dataset com resultados das métricas de avaliação piores que esta classe citada, ou até mesmo a classe que se sobressaiu perante as outras pode possuir amostras que as características mais relevantes existentes nela.

Por fim, a partir deste ponto, ressaltando a importância do processo de Fine-Tuning utilizado, é válido inferir que os resultados se portam de maneira positiva a níveis de classes, tendo em vista que a cada aumento de amostras de Fine-Tuning, os resultados obtidos nas métricas de avaliação apresentam relativas melhoras, evidenciando a importância de se realizar

o ajuste a base de dados a ser predita.

4 CONCLUSÃO

O presente trabalho visou a aplicação de técnicas semi-supervisionada agregadas a redes neurais por convolução. Para tal análise foram utilizados duas abordagens semi-supervisionadas denominadas Self-training e Co-training, bem como a abordagem Supervisionada. Além de tais análises experimentais do comportamento das abordagens no contexto das CNNs, foram também realizados experimentos com mudança de domínio no qual uma dada CNN foi treinada com imagens de um domínio fonte, por meio das abordagens de semi-supervisão, porém recalibradas e testadas em um domínio divergente. Para tanto, foram realizados experimentos com diversas porcentagens de rotulação de dados, bem como de calibração de tais arquiteturas CNNs com imagens do domínio alvo. A partir dos experimentos realizados e dos resultados obtidos pode-se argumentar que as técnicas semi-supervisionadas quando agregadas a CNNs apresentam resultados interessantes tanto do ponto de vista da eficácia como da eficiência, visto que diminui em larga escala a necessidade de dados rotulados para treinamento das arquiteturas, bem como a aplicação de Fine-Tuning sobre domínio diferente, possibilita a obtenção de resultados a partir de prévio treinamento sobre uma base que não seja a base alvo de classificação, fazendo com que não seja necessário uma base alvo com quantidades elevadas de dados, que geram elevado consumo energético e sobrecarga de memória de processamento. Além do fato apresentado, também vale ressaltar que as lógicas utilizadas nos desenvolvimentos dos algoritmos, bem como os próprios algoritmos servem como ferramentas para demais pesquisadores que possam e queiram fazer o uso de tal, contribuindo mais com a comunidade científica na evolução dos algoritmos de aprendizagem de máquina dentro de Visão Computacional.

Sob este argumento, e partindo do pressuposto do funcionamento da técnica de Transfer Learning, a verificação da aplicação de diferentes domínios no processo de classificação e generalização a partir de uma arquitetura previamente treinada com outro domínio, utilizando técnicas de aprendizagem de máquina semi-supervisionada possui resultados interessantes e promissores, pois, os mesmos se portam de maneira positiva, alcançando resultados próximos aos obtidos na abordagem supervisionada, sendo possível

de se treinar um algoritmo de aprendizagem de máquina para reconhecer e classificar outros padrões. Porém, para o método utilizado se faz necessário a utilização da técnica de Fine-Tuning, para que os pesos previamente treinados se ajustem ao domínio estudado, bem como as quantidades de amostras no Fine-Tuning também influenciarão nos resultados e métricas de avaliação do modelo, quanto maiores as quantidades de amostras ao processo de Fine-Tuning, melhor o ajustamento de pesos ao domínio a ser predito. Em algumas situações, conseguir uma quantidade grande de dados de uma determinada ocasião para poder submeter à um treinamento de aprendizagem de máquina acaba por ser um grande empecilho, a técnica apresentada demonstrou resultados interessantes e promissores, tanto utilizando as bases de treinamento inicial de um domínio diferente, quanto aplicando a técnica de refinamento de pesos Fine-Tuning.

4.0.1 PUBLICAÇÕES

- BASSANI, Cristiano Neia Oliveira; SAITO, Priscila Tiemi Maeda; BUGATTI, Pedro Henrique. Assessment of Semi-Supervised approaches applied to Convolutional Neural Networks. In: 21st International Conference on Artificial Intelligence and Soft Computing (ICAISC 2022), 2022,. 21st International Conference on Artificial Intelligence and Soft Computing (ICAISC 2022), 2022, Zakopane. Proceedings of the 21st ICAISC, 2022. v. 1. p. 1-10.
- BASSANI, Cristiano Neia Oliveira; SAITO, Priscila Tiemi Maeda; BUGATTI, Pedro Henrique. Avaliação de abordagens semi-supervisionadas aplicadas a redes neurais convolucionais. In: XVI Brazilian e-Science Workshop, 2022, Rio de Janeiro. Proceedings of the XVI Brazilian e-Science Workshop - CSBC, 2022. p. 9-16.

4.0.2 TRABALHOS FUTUROS

Como trabalhos futuros vislumbram-se os seguintes:

- análise de técnicas de adaptação de domínio junto a abordagens semi-supervisionadas;
- proposta de um arcabouço de agregação entre o paradigma semi-supervisionado e CNNs para aplicações com mudança de domínio;
- experimentos e testes com outros tipos de abordagens semi-supervisionadas.

REFERÊNCIAS

- AMORIM, W. et al. Semi-supervised learning with connectivity-driven convolutional neural networks. **Pattern Recognition Letters**, p. 16 – 22, 08 2019.
- CETINIC, E.; LIPIC, T.; GRGIC, S. Fine-tuning convolutional neural networks for fine art classification. **Expert Systems with Applications**, Elsevier, v. 114, p. 107–118, 2018.
- CHEN, J. et al. Co-training semi-supervised deep learning for sentiment classification of mooc forum posts. **Symmetry**, v. 12, p. 1 – 24, 12 2019.
- Choi, D. Y.; Song, B. C. Semi-supervised learning for continuous emotion recognition based on metric learning. **IEEE Access**, v. 8, p. 113443–113455, 2020.
- ENGELEN, J. E. V.; HOOS, H. H. A survey on semi-supervised learning. **Machine learning**, SPRINGER, v. 109, n. 2, p. 373–440, 2020.
- FONTANA, E. **Introdução aos Algoritmos de Aprendizagem Supervisionada**. [S.l.]: Universidade Federal do Paraná, 2018.
- FRIEDERICH, S. Fine-tuning. 2017.
- Hussein, S. et al. Lung and pancreatic tumor characterization in the deep learning era: Novel supervised and unsupervised learning approaches. **IEEE Transactions on Medical Imaging**, v. 38, n. 8, p. 1777–1787, 2019.
- Ienco, D. et al. Weakly supervised learning for land cover mapping of satellite image time series via attention-based cnn. **IEEE Access**, v. 8, p. 179547–179560, 2020.
- Kido, S.; Hirano, Y.; Hashimoto, N. Detection and classification of lung abnormalities by use of convolutional neural network (cnn) and regions with cnn features (r-cnn). In: **2018 International Workshop on Advanced Image Technology (IWAIT)**. [S.l.: s.n.], 2018. p. 1–4.
- KIRANYAZ, S. et al. 1d convolutional neural networks and applications: A survey. **Mechanical Systems and Signal Processing**, v. 151, p. 1 – 21, 2021.
- KOSS, L. G. Cervical (pap) smear: New directions. **Cancer**, v. 71, n. S4, p. 1406–1412, 1993.
- KOTSIANTIS, S. B. Supervised machine learning: A review of classification techniques. **Informatica (Slovenia)**, v. 31, n. 3, p. 249–268, 2007.
- LING, Z. et al. Semi-supervised learning via convolutional neural network for hyperspectral image classification. In: **2018 24th International Conference on Pattern Recognition (ICPR)**. [S.l.: s.n.], 2018. p. 1–6.
- OLIVEIRA, P. et al. Mammoset: An enhanced dataset of mammograms. In: . [S.l.: s.n.], 2017. p. 1 – 11.

ORENSTEIN, E. C. et al. **WHOI-Plankton- A Large Scale Fine Grained Visual Recognition Benchmark Dataset for Plankton Classification**. 2015.

QIU, J. et al. A survey of machine learning for big data processing. **EURASIP Journal on Advances in Signal Processing**, p. 28, 05 2016.

RUDIN, C.; WAGSTAFF, K. L. A survey of machine learning for big data processing. **Mach Learn** **95**, 1–9 (2014), p. 28, 11 2013.

T.R., S. et al. Hybrid xception model for human protein atlas image classification. In: **EEE 16th India Council International Conference (INDICON)**. [S.l.: s.n.], 2019. p. 1–4.

WEISS, K.; KHOSHGOFTAAR, T. M.; WANG, D. A survey of transfer learning. **Journal of Big Data**, p. 28, 05 2016.

