

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO TECNOLÓGICO  
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA  
CIÊNCIAS DA COMPUTAÇÃO

Thiago Zimmermann Loureiro Chaves

**Uso de Visão Computacional com Redes Neurais Convolucionais para Classificação de  
Fungos**

Florianópolis  
2023

Thiago Zimmermann Loureiro Chaves  
**Uso de Visão Computacional com Redes Neurais Convolucionais para Classificação de Fungos**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de Bacharel em Ciências da Computação e aprovado em sua forma final pelo curso de Graduação em Ciências da Computação.

Florianópolis, 10 de Julho de 2023.

---

Prof. Lúcia Helena Martins Pacheco, Dr.  
Coordenadora do Curso

**Banca Examinadora:**

---

Prof. Aldo Von Wangenheim  
Orientador  
Universidade Federal de Santa Catarina

---

Prof. Elisandro Ricardo Drechsler dos Santos  
Avaliador  
Universidade Federal de Santa Catarina

---

Prof. Mauro Roisenberg  
Avaliador  
Universidade Federal de Santa Catarina

---

Prof. Antônio Carlos Sobieranski  
Avaliador  
Universidade Federal de Santa Catarina

Thiago Zimmermann Loureiro Chaves

## **Uso de Visão Computacional com Redes Neurais Convolucionais para Classificação de Fungos**

Trabalho de Conclusão de Curso submetido ao Curso de Graduação em Ciências da Computação do Centro Tecnológico da Universidade Federal de Santa Catarina como requisito para obtenção do título de Bacharel em Ciências da Computação.

Orientador: Prof. Aldo Von Wangenheim Co-orientador:

Florianópolis

2023

## RESUMO

Este Trabalho apresenta a aplicação de técnicas de visão computacional e Deep Learning com o objetivo de identificar e classificar espécies da família de fungos Hymenochaetaceae, uma iniciativa realizada sob a bolsa de estudo do projeto Protax, conduzido pelo Laboratório Mind Funga. O estudo utiliza tanto métodos tradicionais quanto os considerados estado da arte. Para o propósito da pesquisa, foram criados quatro conjuntos de dados (datasets), subdivididos em Train, Valid e Test. Cada dataset representava uma característica distinta do fungo, e todas as imagens utilizadas foram coletadas em campo e posteriormente tratadas em laboratório. A precisão obtida nos conjuntos de teste foram aproximadamente 100%, 100%, 100% e 94,12%.

**Palavras-chave:** Deep Learning, Visão Computacional, CNN, Vision Transformers, Classificação de Imagens, Macrofungos, Hymenochaetaceae

## ABSTRACT

This work presents the application of computer vision techniques and Deep Learning with the objective of identifying and classifying species from the Hymenochaetaceae family of fungi, an initiative implemented under the scholarship of the Protax project, conducted by the Mind Funga Laboratory. The study uses both traditional methods and those considered state of the art. For the purpose of the research, four datasets were created, subdivided into Train, Valid, and Test. Each dataset represented a distinct characteristic of the fungus, and all images used were collected in the field and subsequently processed in the laboratory. The accuracy obtained in the test sets were approximately 100%, 100%, 100%, and 94.12%.

**Palavras-chave:** Deep Learning, Computer Vision, CNN, Vision Transformers, Image Classification, Macrofungi, Hymenochaetaceae

## **AGRADECIMENTOS**

Agradeço aos familiares pelo constante apoio e auxílio emocional, que foram fundamentais para a realização deste trabalho. Aos amigos que tive o prazer de fazer na faculdade e que levarei para o resto da vida, em especial a Leonardo, Luiz, Nicolas, André e Samuel, agradeço pela ajuda emocional. Os nossos churrascos e o apoio em diversas matérias da faculdade foram essenciais para concluir este curso.

Agradeço também aos amigos que fiz e mantive durante a pandemia, em especial a Marco, Rodrigo, Marcelo, Roberto, João Vitor, Matheus, Eduardo, Binckas e Rhenzo. Vocês mantiveram minha sanidade mental nesse período conturbado.

Por fim, sou grato ao professor Aldo e à equipe do Mind Funga, em especial ao Prof. Ricardo, Genivaldo e Eloísa, que ajudaram ativamente na confecção deste trabalho.

Anything that could give rise to smarter-than-human intelligence — in the form of Artificial Intelligence, brain-computer interfaces, or neuroscience-based human intelligence enhancement — wins hands down beyond contest as doing the most to change the world. Nothing else is even in the same league (Eliezer Yudkowsky)

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>10</b>
<b>1.0.1</b>	<b>Contexto Específico</b>	<b>12</b>
1.1	OBJETIVOS	13
<b>1.1.1</b>	<b>Objetivo Geral</b>	<b>13</b>
<b>1.1.2</b>	<b>Objetivos Específicos</b>	<b>13</b>
<b>2</b>	<b>CONCEITOS RELACIONADOS</b>	<b>14</b>
2.1	O QUE É INTELIGÊNCIA ARTIFICIAL?	14
<b>2.1.1</b>	<b>Machine Learning</b>	<b>14</b>
<b>2.1.2</b>	<b>Processamento de Linguagem Natural (NLP)</b>	<b>14</b>
<b>2.1.3</b>	<b>Sistemas Especialistas</b>	<b>14</b>
<b>2.1.4</b>	<b>Visão Computacional</b>	<b>14</b>
2.1.4.1	<i>Reconhecimento/Classificação de Imagens</i>	15
2.1.4.2	<i>Detecção de Objetos</i>	15
2.1.4.3	<i>Segmentação Semântica</i>	16
2.2	CONVOLUTIONAL NEURAL NETWORK (CNN)	16
<b>2.2.1</b>	<b>Camada de Convolução (Convolutional Layer)</b>	<b>16</b>
<b>2.2.2</b>	<b>Camada de Pooling/Downsampling (Pooling Layer)</b>	<b>17</b>
<b>2.2.3</b>	<b>Camada Totalmente Conectada (Fully Connected Layer)</b>	<b>17</b>
<b>2.2.4</b>	<b>Pesos</b>	<b>17</b>
<b>2.2.5</b>	<b>Função de Ativação (Fully Connected Layer)</b>	<b>18</b>
<b>2.2.6</b>	<b>Função de Perda (Loss Function)</b>	<b>18</b>
<b>2.2.7</b>	<b>Data Augmentations</b>	<b>19</b>
2.3	MODELOS	19
<b>2.3.1</b>	<b>Lenet5</b>	<b>19</b>
<b>2.3.2</b>	<b>Alexnet</b>	<b>19</b>
<b>2.3.3</b>	<b>VGG16</b>	<b>20</b>
<b>2.3.4</b>	<b>Resnet</b>	<b>20</b>
<b>2.3.5</b>	<b>Efficientnet</b>	<b>21</b>
<b>2.3.6</b>	<b>Hrnet</b>	<b>21</b>
<b>2.3.7</b>	<b>Vision Transformers (ViT)</b>	<b>21</b>
2.3.7.1	<i>Small Patch/ Base Patch</i>	21
2.4	FASTAI	22
2.5	ESTRUTURA DE MACROFUNGOS	22
<b>2.5.1</b>	<b>Características das espécies classificadas em Hymenochaetaceae</b>	<b>23</b>
<b>3</b>	<b>ESTADO DA ARTE</b>	<b>25</b>
3.1	METODOLOGIA DA REVISÃO SISTEMÁTICA DE LITERATURA	25

3.1.1	<b>Bancos de dados de pesquisa</b> . . . . .	26
3.1.2	<b>Plataformas de buscas e como funcionam</b> . . . . .	26
3.2	<b>TRABALHOS RELACIONADOS</b> . . . . .	26
3.2.1	<b>Trabalhos Relacionados aos macrofungos</b> . . . . .	27
4	<b>DESENVOLVIMENTO</b> . . . . .	31
4.1	OBJETIVOS . . . . .	31
4.2	METODOLOGIA . . . . .	31
4.2.1	<b>Coleta de dados</b> . . . . .	31
4.2.2	<b>Especificações do Equipamento</b> . . . . .	31
4.2.3	<b>Datasets</b> . . . . .	32
4.2.3.1	<i>Esporos</i> . . . . .	33
4.2.3.2	<i>Himênio</i> . . . . .	33
4.2.3.3	<i>Poros</i> . . . . .	34
4.2.4	<b>Análise de dados</b> . . . . .	34
4.2.5	<b>Redes Escolhidas</b> . . . . .	35
4.3	RESULTADOS . . . . .	36
4.3.1	<b>Himênio</b> . . . . .	37
4.3.2	<b>Esporos</b> . . . . .	37
4.3.3	<b>Poros</b> . . . . .	38
5	<b>CONCLUSÕES</b> . . . . .	40
	<b>REFERÊNCIAS</b> . . . . .	41
	<b>APÊNDICE A – APÊNDICE</b> . . . . .	43
A.1	DATASET . . . . .	43
A.2	NOTEBOOKS . . . . .	43
	<b>APÊNDICE B – ARTIGO CIENTÍFICO</b> . . . . .	44

## LISTA DE FIGURAS

Figura 1 – Detecção de Cones - Projeto do Lapix . . . . .	15
Figura 2 – Detecção de Caminho Navegável - Projeto do Lapix . . . . .	16
Figura 3 – Estrutura Geral de uma Rede Neural Convolucional . . . . .	17
Figura 4 – Comparativo Funções de ativação . . . . .	18
Figura 5 – Estrutura Geral de uma Resnet . . . . .	20
Figura 6 – Desenho Macrofungo . . . . .	23
Figura 7 – Exemplo de imagens do Dataset . . . . .	24
Figura 8 – Metodologia de uma SLR . . . . .	25
Figura 9 – Artigos organizados por método de treinamento . . . . .	30
Figura 10 – Artigos organizados por função . . . . .	30
Figura 11 – Estrutura Dataset . . . . .	32
Figura 12 – Estereomicroscópio e Microscópio Respectivamente . . . . .	32
Figura 13 – Imagens de Esporos . . . . .	33
Figura 14 – Imagens de Himênio . . . . .	33
Figura 15 – Imagens de Poros 35x . . . . .	34
Figura 16 – Imagens de Poros 100x . . . . .	34
Figura 17 – Página do Wandb para os testes do Dataset Poros 35x . . . . .	35
Figura 18 – Train e Valid Loss Experimentos Poros 35x . . . . .	37
Figura 19 – Comparativo Redes Neurais Himênio 1000x de Zoom . . . . .	38
Figura 20 – Comparativo Redes Neurais Esporos 1000x de Zoom . . . . .	38
Figura 21 – Comparativo Redes Neurais Poros 35x de Zoom . . . . .	39
Figura 22 – Comparativo Redes Neurais Poros 100x de Zoom . . . . .	39
Figura 23 – Comparativo Geral dos Experimentos . . . . .	39

## 1 INTRODUÇÃO

Visão Computacional ou visão robótica é um ramo da inteligência artificial que baseia-se em um conjunto de técnicas de programação e modelos matemáticos cujo objetivo é extrair conhecimento de tanto mídias estáticas como: textos, imagens; quanto mídias dinâmicas como: vídeos e áudios. Portanto, com esse conhecimento o computador/modelo consegue compreender e automatizar tarefas análogas aos quais o sistema visual e nervoso realizam.

Essa tecnologia surgiu ao final da década de 60 através do estudo de percepção visual e neurofisiológica, testando como os neurônios reagiriam a diferentes estímulos. Resultado desse estudo foi a constatação de dois cientistas quanto ao funcionamento dos neurônios, onde os mesmos detectam padrões simples, como bordas por exemplo, e repassam a informação adiante formando formas cada vez mais complexas (HUBEL; WIESEL, 1968). Nesse sentido, após esse evento foi criado um certo otimismo, que no entanto, foi seguido de uma grande decepção devido à complexidade e falta de tecnologia na época, acarretando o que é chamado agora de inverno da IA. No entanto, nas décadas seguintes a tecnologia avançou a ponto dessa discussão voltar à tona, gerando dois marcos importantes tanto para Visão Computacional quanto para a IA que são:

- A invenção dos filtros e consequentemente as Redes Neurais Convolucionais (CNN) por Yann LeCun (LECUN et al., 1998).
- A criação de uma Rede Neural Profunda chamada Alexnet da Universidade de Toronto, cujas características revolucionaram a competição *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012)

O que antes era uma grande problema, reconhecer imagens de dígitos escritos a mão, Yann LeCun resolveu em 1998 através da utilização um algoritmo de backpropagation para aprender os coeficientes de convolução. O resultado foi que seu aprendizado foi realizado de forma automática, além de ultrapassar em eficiência os coeficientes calculados manualmente. Esse novo método se tornou a fundação para visão computacional moderna.

O outro marco foi quando a CNN Alexnet através da competição ILSVRC em Setembro de 2012, apresentou um erro de 15,3%, mais de 10% a menos do valor que se mantinha estagnado na competição. O paper lançado por Alex Krizhevsky e os demais mostrou que a profundidade e o número de camadas eram essenciais para atingir uma boa performance, que embora cara em poder de processamento e consumo de energia, seria possível devida a utilização de GPU's para treinamento.

A função da visão computacional por si só seria entender e extrair informações de imagens digitais com o objetivo de criar informações para compreender o contexto que está envolvido criando uma possível decisão. Uma das tarefas mais típicas são:

- Reconhecimento

- Detecção de objetos
- Identificação
- Detecção/Segmentação
- Análise de movimento
  - Egomotion (Determinar a rotação de uma sequência de imagens)
  - Rastreamento
  - Optical flow (Identificar para cada ponto da imagem o movimento em relação ao plano)
- Reconstrução de imagens

Ao analisar essa tecnologia de uma maneira mais tangível, ela comumente é aplicada na medicina no processamento de imagens médicas exames como microscopia, ultrassonografia, tomografia, onde a extração de informações tem função de realizar diagnósticos de pacientes. Um exemplo da utilização dessa ferramenta é um estudo realizado na *University of Central Florida in Orlando* pelo especialista de visão computacional Ulas Bagci, onde relata dois problemas, a crescente demanda de tomografias, causando uma sobrecarga aos médicos, consequentemente gerando mais erros e o fato que cerca de 35% dos nódulos do pulmão não são detectados na visualização inicial do médico. Nesse sentido, a inteligência artificial poderia ajudar a detectar erros profissionais ou ajudar com a demanda fazendo uma primeira avaliação. Um exemplo semelhante a isso ocorre no artigo (ZHANG, 2022)

Outro exemplo de aplicação, é na agricultura moderna através do mapeamento com uso de drone a fim de localizar pragas ou uma possível falta de água. Exemplo disso é a fazenda em Sunnyville, Califórnia onde são utilizados robôs para mapear ervas daninhas resistentes a herbicida. Além disso, a *Research Institute for the Semi-Arid Tropics (ICRISAT)* está usando a Microsoft Cortana e uma IA treinada com dados do solo, água, previsão do tempo e tempo de colheita para enviar dados como melhor período para plantar aos 176 fazendeiros cadastrados por meio de um aplicativo, que segundo eles é de fácil interpretação. Um exemplo semelhante a esse trabalho demonstrado é o artigo (PALLOTTINO et al., 2019). Além dessas ainda há inúmeras funções como:

- Carros Autônomos
- Análise do tráfego de veículos
- Monitoramento de buracos e problemas na estrada
- Inspeção de qualidade na indústria
- Leitura de textos e códigos de barra

- Manutenção preditiva
- Controle de estoque de loja e supermercados

### 1.0.1 Contexto Específico

Nesse contexto, o trabalho visa a utilização dessa ferramenta que é a visão computacional e Deep Learning, com objetivo da identificação e classificação das espécies da família de fungos Hymenochaetaceae como parte de uma bolsa de estudo no projeto Protax: Entendendo Hymenochaetaceae (basidiomycota, fungi): Da resolução de complexos taxonômicos à confirmação da classificação filogenética de gêneros, administrado pelo professor Elisandro Ricardo Drechsler dos Santos do Laboratório Mind Funga. A Família de fungos Hymenochaetaceae consiste em macrofungos com coloração marrom, cuja função ecológica geralmente é a degradação de madeira, no entanto, também podem ser parasitas. Embora a família de macrofungos Hymenochaetaceae seja reconhecida internacionalmente, a maioria dos trabalhos carecem de representantes Neotrópicos, ou seja da região da América Latina, devido a isso são espécies de difícil classificação e existem muitas discordâncias taxonômicas, necessitando uma melhor re-amostragem e delimitação. Portanto, o projeto Protax do laboratório tem o objetivo: Resolver complexos taxonômicos, ou seja, melhor delimitar táxons até então mal interpretados, com grande amplitude morfológica e geográfica, que poderiam revelar uma diversidade ainda desconhecida. Além disso, gerar melhor entendimento da diversidade e história evolutiva da família. Nesse contexto, o laboratório Mind Funga forneceu suporte ao trabalho de TCC criando e operando bancos de dados de tanto estruturas microscópicas, utilizando instrumentos profissionais como o microscópio LEICA, quanto estruturas macroscópicas, obtidas através de saídas de campo dos integrantes do laboratório. Portanto, esse trabalho visa incorporar uma parte desse grande projeto através da análise de aspectos microscópicos através de uma rede neural para cada característica como:

- Himênio
- Esporos
- Poros
  - 35x Zoom
  - 100x Zoom

A fim de escolher qual a melhor rede neural para essas diversas situações foram realizadas pesquisas voltadas ao estado da arte de redes neurais e diversos testes na base de dados. Com a criação dessas redes, foi possível identificar e classificar as espécies através da técnica de Inteligência Artificial, a classificação de imagens. Ao final dessa etapa, realizamos análises probabilísticas e estatísticas para identificar quais as características mais informativas para delimitar gêneros e espécies nesse grupo de fungos.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo Geral

Esse trabalho tem como objetivo geral utilizar de visão computacional e Deep Learning aplicada no projeto do Protax, a fim de identificar as espécies e possíveis características determinantes para classificação da "família" Hymenochaetaceae, que é de difícil categorização, inclusive para profissionais da área

### 1.1.2 Objetivos Específicos

- Analisar a literatura de visão computacional.
- Analisar a literatura de exemplos semelhantes ao projeto
- Realizar diversos testes com diferentes modelos.
- Analisar o estado da arte.
- Escrita de um artigo científico para apresentação a uma revista nacional ou local.

## 2 CONCEITOS RELACIONADOS

### 2.1 O QUE É INTELIGÊNCIA ARTIFICIAL?

Inteligência Artificial ou IA é um ramo da ciência da computação, definido como conjunto de técnicas e métodos que almejam imitar a inteligência humana, seja no processo de raciocínio ou em decisões e realizar tarefas que usam desses para resolver problemas. Essa tecnologia possui diversos ramos como:

- Machine Learning
- Processamento de Linguagem Natural (NLP)
- Sistemas Especialistas
- Visão Computacional

#### 2.1.1 Machine Learning

Aprendizado de máquina, ou Machine Learning, é um subcampo da Inteligência Artificial que se concentra no desenvolvimento de algoritmos e modelos estatísticos que permitem que o computador melhore o desempenho de uma tarefa específica através da experiência gerada pelos dados.

#### 2.1.2 Processamento de Linguagem Natural (NLP)

O Processamento de Linguagem Natural é um campo da inteligência artificial que se concentra na interação entre computadores e a linguagem humana. Os principais objetivos do NLP é a compreensão do significado de um texto, extração de informações relevantes, tradução entre idiomas, resumir textos, entre outros.

#### 2.1.3 Sistemas Especialistas

Os sistemas especialistas é um campo da inteligência artificial cujo objetivo é imitar a habilidade de tomada de decisões de um especialista humano dentro de um campo específico de conhecimento. Eles são projetados para resolver problemas complexos por meio de raciocínio sobre conhecimento, representado principalmente como regras de "se-então".

#### 2.1.4 Visão Computacional

Visão Computacional ou visão robótica é um ramo da inteligência artificial que baseia-se em um conjunto de técnicas de programação e modelos matemáticos cujo aplicação é extrair

conhecimento de mídias digitais, como imagens ou objetos multidimensionais a fim de compreender o contexto com base no conhecimento que adquiriu. Possui diversos usos como:

- Classificação de Imagens
- Detecção de Objetos
- Segmentação Semântica

#### 2.1.4.1 Reconhecimento/Classificação de Imagens

É o processo de prever uma classe ou uma label específica através da categorização de grupos de pixels e vetores no contexto de uma imagem.

Existem dois tipos de classificação:

- Não Supervisionada

É um método totalmente automatizado, onde o algoritmo analisa e agrupa em conjuntos através da descoberta de padrões implícitos sem a necessidade de intervenção humana. Um exemplo de algoritmo famoso desse método é o K-means.

- Supervisionada

É um método que utiliza amostras classificadas anteriormente como base para treinar o classificador, assim, com base no que aprendeu pode generalizar e identificar novos dados.

#### 2.1.4.2 Detecção de Objetos

Detecção de objetos é uma área em visão computacional que tem como objetivo identificar e localizar objetos específicos dentro de imagens ou vídeos. Diferente da classificação de imagens a detecção de objetos é capaz de identificar vários objetos em uma única imagem e proporcionar uma delimitação de cada objeto na forma de uma *bounding box*, que demarca a área específica onde o objeto está localizado.



Figura 1 – Detecção de Cones - Projeto do Lapix

### 2.1.4.3 Segmentação Semântica

A Segmentação Semântica é uma área da visão computacional na qual o objetivo é categorizar cada pixel de uma imagem em uma classe ou objeto. O objetivo é produzir um mapa de segmentação pixel a pixel de uma imagem, onde cada pixel é atribuído a uma classe ou objeto específico.



Figura 2 – Detecção de Caminho Navegável - Projeto do Lapix

## 2.2 CONVOLUTIONAL NEURAL NETWORK (CNN)

Uma Rede Neural Convolutiva é uma rede feedforward hierárquica de multicamadas onde cada camada, usando o banco de kernels convolucionais, realiza múltiplas transformações (LeCun et al 2010). Nesse sentido, a rede recebe imagens como input e que através de pesos e vieses é capaz de identificar um objeto ou até mesmo diferenciá-lo de outro.

A arquitetura das CNN é baseada no padrão de neurônios do cérebro humano e inspirado na organização do córtex visual. Os neurônios respondem a estímulos externos e repassam a informação adiante para os neurônios vizinhos.

Essa método consegue capturar referências espaciais e temporais através da incorporação de filtros que reduzem o número de parâmetros e reutiliza pesos já treinados.

Geralmente essa arquitetura possuem 3 tipos de camadas e uma função de ativação que serão explicados adiante.

### 2.2.1 Camada de Convolução (Convolutional Layer)

A camada de convolução possui um grupo de kernels/filtros e parâmetros que serão aprendidos durante o treinamento. Cada filtro é convolucionado com o input, com a função de criar um mapa de ativação (feature map), para mapear destaques ou padrões no espaço dos inputs. Seu input é um tensor com o formato (Número de inputs, altura do input, largura do input, canais do input) e retorna uma feature map com o formato (Número de inputs, altura do feature map, largura do feature map, canais do feature map).

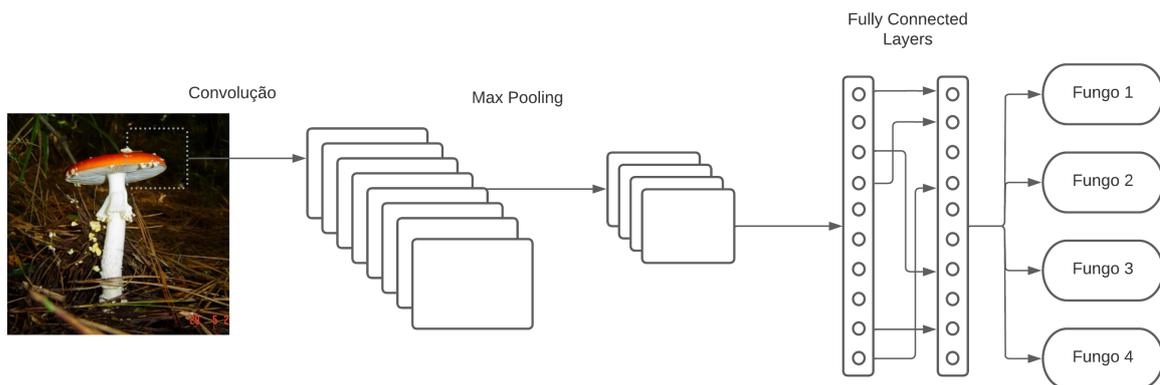


Figura 3 – Estrutura Geral de uma Rede Neural Convolucional

### 2.2.2 Camada de Pooling/Downsampling (Pooling Layer)

A camada de Pooling tem como função diminuir as dimensões do tensor, combinando resultados parecidos em clusters de neurônios com o objetivo de diminuir a complexidade das operações e ajudar no aumento de generalização diminuindo o overfitting. Possuem diversos tipos de pooling como:

- Max Pooling
- Average Pooling
- L2 Pooling
- Overlapping Pooling
- Spatial Pyramid Pooling

### 2.2.3 Camada Totalmente Conectada (Fully Connected Layer)

É uma camada que geralmente é usada no final das redes neurais após o feature map estiver planejado, com o objetivo de classificação. Essa camada é um MLP (Multilayer Perceptron) que possui uma camada oculta, função de ativação e uma loss function. Nesse sentido, através dessas ferramentas será classificado a imagem através do feature map.

### 2.2.4 Pesos

Pesos são parâmetros aprendíveis cuja função é controlar o sinal entre neurônios, determinando o quanto de influência cada característica do input impactará no output.

## 2.2.5 Função de Ativação (Fully Connected Layer)

É uma função cujo objetivo é ajudar a aprender padrões intrínsecos e sua escolha pode afetar no aceleramento ou desaceleramento do tempo de treinamento. Possuem diversas funções famosas para essa função como:

- Sigmoid
- Tangente Hiperbólica
- SWISH
- ReLU e variações

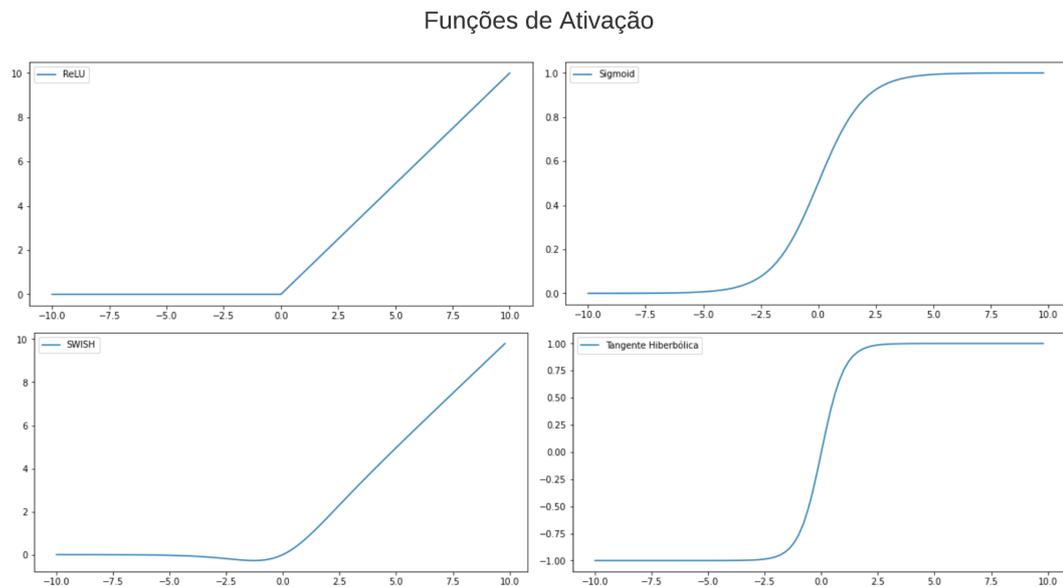


Figura 4 – Comparativo Funções de ativação

No entanto, a função mais utilizada é a ReLU (Rectified Linear Units) devido ao fato de ajudarem a resolver o "Vanishing Gradient Problem".

## 2.2.6 Função de Perda (Loss Function)

A função de perda é uma função cujo objetivo é verificar a qualidade de seu modelo avaliado através da comparação do valor esperado com o valor do output. Assim como a função de ativação possuem diversos tipos:

- Cross-entropy

- Exponential Loss
- Mean Square Error (MSE)
- Mean Absolute Error (MAE)
- ...

### 2.2.7 Data Augmentations

São conjuntos de técnicas para aumentar o conjunto de dados com cópias levemente modificadas do conjunto de dados. Essa tecnologia age como regulador e ajuda a diminuir overfitting. Exemplos dessas técnicas são:

- Rotação
- Ampliação
- Recorte
- Flips (Horizontal e Vertical)
- Brilho

## 2.3 MODELOS

### 2.3.1 Lenet5

LeNet-5 é uma arquitetura de rede neural convolucional projetada por Yann LeCun e publicada no paper (LECUN et al., 1998) em 1998. LeNet-5 foi uma das primeiras redes neurais convolucionais e contribuiu para os fundamentos do campo da aprendizagem profunda. A LeNet-5 foi projetada com intuito de reconhecer dígitos manuscritos e foi amplamente utilizada para tarefas de processamento de imagem. A arquitetura inclui duas camadas convolucionais, duas camadas de pooling e três camadas totalmente conectadas.

### 2.3.2 Alexnet

A AlexNet é uma arquitetura de rede neural convolucional que foi apresentada por Alex Krizhevsky no NIPS 2012. Ela representou um marco na evolução das redes neurais convolucionais, pois venceu com uma grande margem o desafio de classificação de imagens ImageNet LSVRC-2010, demonstrando o poder das CNNs para tarefas de classificação de imagens (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). Foi a primeira a utilizar GPUs para treinamento, com intuito de acelerar o processo de aprendizado.

### 2.3.3 VGG16

A VGG16 é uma arquitetura de rede neural convolucional que foi apresentada por Karen Simonyan em seu artigo (SIMONYAN; ZISSERMAN, 2014), publicado em 2014. A VGG16 se destacou devido à sua profundidade, consistindo de 16 camadas com pesos, e demonstrou um desempenho impressionante pra época no desafio ImageNet 2014.

### 2.3.4 Resnet

Resnet ou Residual Neural Network foi um modelo de Rede neural criado pelo Autor Kaiming He e publicado no artigo (HE et al., 2015). Esse modelo foi o primeiro que utilizou com sucesso redes neurais feedforward muito profundas, a ideia essencial foi introduzir um conceito chamado "Identity shortcut connection" cuja função é pular uma ou mais camadas. Outro conceito utilizado "Pre-activation variant of residual block", na qual os gradientes podem fluir por atalhos para qualquer outra camada anterior. Esse modelo foi criado em 2015 e ainda possui enorme relevância e ao longo dos anos foram criadas inúmeras versões que variam tanto número de camadas quanto a arquitetura.

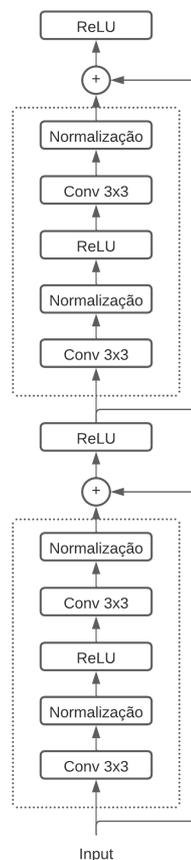


Figura 5 – Estrutura Geral de uma Resnet

### 2.3.5 Efficientnet

EfficientNet é uma família de arquiteturas de redes neurais profundas proposta no paper (TAN; LE, 2019) projetada para maximizar a eficiência computacional e a precisão do modelo. Isso é alcançado através de um método de escalonamento composto por largura, profundidade, resolução e o uso de blocos de construção eficientes chamados MBConv. Portanto, o resultado é uma arquitetura altamente eficiente que pode ser facilmente adaptada para diferentes tamanhos e tarefas.

### 2.3.6 Hrnet

Hrnet (High-Resolution Network) é uma arquitetura de rede neural convolucional proposta no paper (WANG et al., 2019) feita com o objetivo de melhorar a precisão do entendimento do corpo humano e suas poses, no entanto, também é utilizada para outros objetivos devido a sua grande eficiência. A Hrnet, diferente da maioria das redes neurais, mantém a resolução mais alta durante todo o processo de computação devido a múltiplos caminhos de processamento paralelos.

### 2.3.7 Vision Transformers (ViT)

Vision Transformers(DOSOVITSKIY et al., 2020) é uma aplicação do modelo Transformers, que já eram utilizados em NLP, em visão computacional. Transformers são um modelo de Deep Learning com aplicação em tradução e sumarização, designados para processar dados de forma sequencial lendo todo o input de uma vez só. Em sua arquitetura, a imagem é inicialmente dividida em pequenos patches, cada um dos quais é achatado em um vetor e processado por uma camada de embedding. Após isso adiciona-se um token especial de classificação na sequência de vetores e a sequência resultante é inserida em uma série de blocos Transformer, que aplicam uma atenção de múltiplas cabeças para entender as interações entre os patches. A previsão final do modelo é gerada pelo processamento do token de classificação através da função softmax, produzindo as probabilidades de classe.

#### 2.3.7.1 *Small Patch/ Base Patch*

"Small Patch"e "Base Patch"são duas variantes do Vision Transformer. A principal diferença entre elas está no tamanho do modelo (número de parâmetros) e na quantidade de dados necessários para treiná-las adequadamente. "Small Patch"tem menos parâmetros e é um modelo menor em comparação ao outro. Devido ao seu tamanho, ele geralmente é mais rápido para treinar. No entanto, ele também pode ter menos capacidade para aprender representações complexas.

## 2.4 FASTAI

Fastai é uma biblioteca de aprendizado profundo de código aberto feita por Jeremy Howard e Sylvain Gugger construída sobre o framework PyTorch (HOWARD; GUGGER, 2020). O Fastai foi criado no intuito de facilitar o processo de criação e treinamento de redes neurais, oferecendo uma interface de alto nível e simplificada. A biblioteca tem sido altamente utilizada em visão computacional, processamento de linguagem natural, entre outras. Além disso, a Fastai oferece uma abordagem de Transfer Learning, que permite ao usuário utilizar modelos pré-treinados para acelerar treinamento.

## 2.5 ESTRUTURA DE MACROFUNGOS

Macrofungos são fungos que se caracterizam por produzirem estruturas reprodutivas visíveis a olho nu. A maioria desses macrofungos pertencem as ordens de Basidiomycota e Ascomycota, possuindo geralmente a função de parasitas, decompositoras ou micorrízicas.

- Hifas

Hifas são unidades básicas (equivalentes a células) da estrutura uma vez que estão sempre presentes nos esporomas (estruturas reprodutivas macroscópicas) e somas (estruturas amorfas e não reprodutivas), além disso se transformam em todos os outros tipos (RYVARDEN, 2004). Sua função é garantir a nutrição, crescimento e reprodução dos fungos.

- Píleo

Píleo é o nome técnico para o "chapéu de um fungo", estrutura reprodutiva onde são produzidos os esporos.

- Himenóforo

Superfície inferior dos esporomas, onde são produzidos os esporos. Majoritariamente, podem ser lisos, com lamelas (cogumelos em geral; de esporomas frágeis) ou porosos (orelhas de pau; de esporomas resistentes) (KIRK et al., 2009)

- Himênio

Porção do himenóforo produtora de esporos, onde as hifas podem desenvolver basídios e ascos, hifas modificadas para produção dos esporos nos filos Basidiomycota e Ascomycota, respectivamente. (KIRK et al., 2009)

- Esporos

Um termo geral para a unidade reprodutiva em fungos, especializada em disseminação. São produtos de reprodução sexual ou assexual (KIRK et al., 2009)

- Setas

Estrutura estéril, no formato de uma seta, de diferentes tamanhos e formatos, com origem no himênio (KIRK et al., 2009)

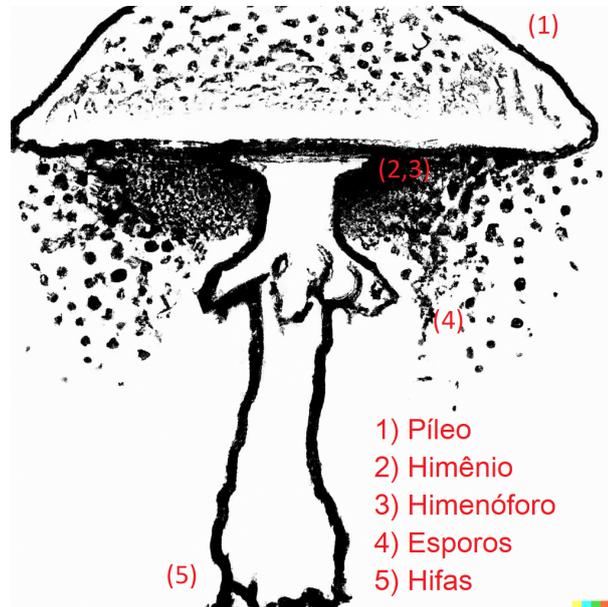


Figura 6 – Desenho Macrofungo

### 2.5.1 Características das espécies classificadas em Hymenochaetaceae

**Píleo:** O formato dos píleos podem variar de ressupinados (totalmente aderidos ao substrato), efuso-reflexos (parte aderida e parte se projetando), estipitado (projeção a partir de um estipe (uma haste reta)) a pileado (projeção sem um estipe).

**Superfície do píleo:** Podem variar de marrom amareladas a preto. Apresenta zonações concêntricas rasas ou sulcadas, rachaduras rasas, profundos a formato de escamas perpendiculares as zonações e/ou nas zonações.

**Himenóforo:** Os himenóforos podem variar de liso a poroides (visão perpendicular da projeção de cilindros).

**Esporos:** Os esporos são globosos a cilíndricos, de parede fina a grossa. Geralmente são lisos, mas podem ser ornamentados. Algumas espécies possuem esporos com reações a substâncias como KOH, Melzer ou Cotton Blue.

**Hifas:** Possuem hifas generativas (paredes finas, septos regulares e crescimento ilimitado) ou hifas generativas e esqueléticas (paredes grossas, crescimento limitado e não possuem septos).

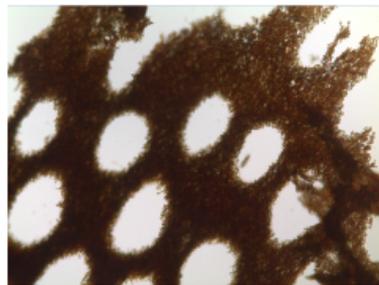
**Setas:** Possuem setas himeniais ou tramais, curtas a longas, com o ápice reto a curvado.

Nesse sentido, essas características destacadas poderiam ser evidenciadas e entendidas pelas redes neurais modeladas durante esse trabalho.

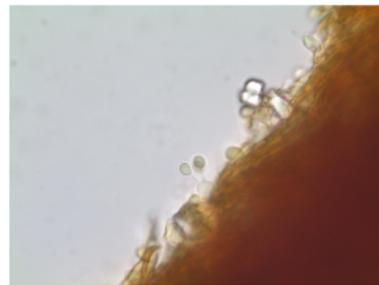
Caracteres a serem testados:

- Poros
  - Em 35x
  - Em 100x
- Himênio
  - Em 1000x
- Esporos
  - Em 1000x

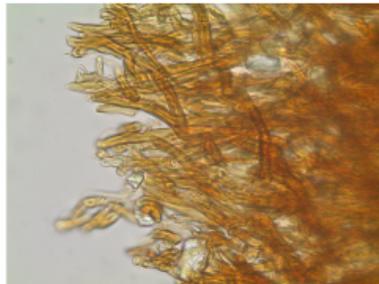
### Exemplos de Imagens do Dataset



Poros em 100x



Himênio



Hifas



Esporos em Melzer

Ex: Fomitiporella

Figura 7 – Exemplo de imagens do Dataset

### 3 ESTADO DA ARTE

Neste capítulo, busca-se fazer um revisão sistemática de literatura baseada no Estado da arte dos modelos e técnicas da classificação de imagens com redes neurais utilizadas em macrofungos.

#### 3.1 METODOLOGIA DA REVISÃO SISTEMÁTICA DE LITERATURA

Uma revisão sistemática de literatura (SLR) é um estudo ou trabalho que tem como objetivo reunir materiais semelhantes de vários autores através de procuras por palavras-chave e assim realizar uma análise estatística com a função de responder uma questão principal fazendo um estudo crítico. Esse método foi criado por Kitchenham e comumente é utilizado em Ciência da computação e outras áreas científicas.

Após definirmos o protocolo de pesquisa foi descrito os passos necessários para replicar os arquivos dessa lista. Como início será utilizado uma busca geral através de palavras chaves em diversos sites de divulgação científica e realizado um agrupamento de artigos científicos que posteriormente alguns serão excluídos através de critérios pré estabelecidos, por meio de uma leitura do artigo e uma verificação se a temática se adéqua ao TCC atual.

Será realizada uma revisão sistemática de literatura (SLR) cujo objetivo será verificar o estado da arte geral da classificação de imagens e Redes Neurais Convolucionais (CNN) contextualizada com tema do trabalho, em estruturas microscópicas de macrofungos.

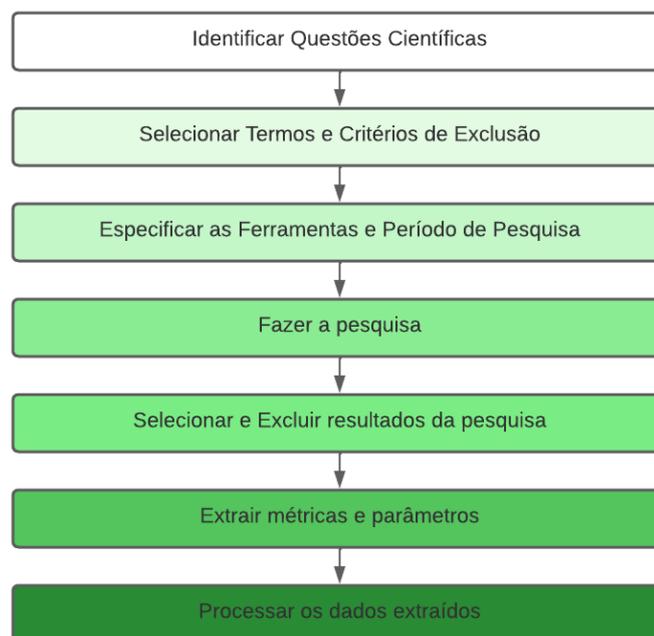


Figura 8 – Metodologia de uma SLR

### 3.1.1 Bancos de dados de pesquisa

Foram considerado Jornais, Revistas e conferências do ano 2018 até 2022 para a pesquisa contextualizada das estruturas microscópicas do fungos. Todo trabalho coletado tem sua data e suas informações a fim de facilitar o rastreio das fontes caso necessário. Foi utilizados os seguintes DataBases de acesso aberto:

- IEEE: Institute of Electrical and Electronics Engineers.
- SD: ScienceDirect Digital library.
- SP: Springer.
- GS: Google Scholar.

### 3.1.2 Plataformas de buscas e como funcionam

Essas databases provem mecanismos de busca configuradas com palavras-chave e expressões lógicas para inclusão e exclusão de artigos. A procura da plataforma se baseia no título, palavras-chave e o abstract dos artigos.

Tabela 1 – Banco de dados e strings de busca usados para obter o resultado preliminar.

#	Base	Initial Search String
1	IEEE	((Deep Learning OR Microscopic) AND (Fungi OR Protozoa OR Micology)) OR ((Deep convolutional neural network OR MICROSCOPIC) AND (fungi OR protozoa OR micology))
2	SD	("Deep Learning") OR ("Machine Learning") OR ("CNN") AND ("Microscopic) OR ("Mushroom") OR ("Micology"))
3	SP	(mushroom OR micology OR fungi OR funga) AND (Deep Learning OR Machine Learning OR CNN)
4	GS	(mushroom OR micology OR fungi OR funga OR microscopic) AND (Deep Learning OR Machine Learning OR CNN)

## 3.2 TRABALHOS RELACIONADOS

Dentro dessa seção serão abordados pesquisas do estado da arte no contexto de aplicações de visão computacional e redes neurais em macrofungos e seu o estado da arte vinculado a classificação de imagem.

### 3.2.1 Trabalhos Relacionados aos macrofungos

Diversas técnicas de Machine Learning e estatística foram utilizadas para classificar e identificar os fungos ao longo dos últimos 5 anos. A fim de identificar esses e métodos e artigos foi realizado uma SLR (Systematic Literature Review) baseado na metodologia de Kitchenham, muito utilizado em áreas científicas. Nessa Revisão Sistemática de Literatura foi verificado uma grande diversidade de modelos e técnicas de redes neurais e visão computacional como: Bag of Words, SVM (Support Vector Machine), Modelos criados pelos autores, Modelos pré treinados de Resnet, InceptionV3, entre outras e tecnologias atuais como Vision Transformers (ViT) (DOSOVITSKIY et al., 2020).

Por exemplo (ZIELIŃSKI et al., 2020), esse trabalho tinha como objetivo a identificação de fungos microscópicos a fim de fazer o diagnóstico de infecções fúngicas. Os autores propõem a utilização de modelos pré treinados de redes neurais associados a um algoritmo de Deep Bag of Words e classificar os resultados através de SVM em um dataset formado de 180 imagens de fungos cultivados em ágar a 37°C, com uma resolução 3600 x 5760. O resultado chegou a uma acurácia de 93.9%.

No artigo (GAIKWAD; RUMMA; HANGARGE, 2022) os autores propuseram a análise de fungos que afetam plantações de manga e maçã, atividade importante para agricultura da Índia, mas a um nível microscópico, através modelos pré treinados de AlexNet e SqueezeNet. Em relação ao treinamento e o dataset foram utilizadas 78 imagens com split entre Train e Val de (80,20) e treinados 25 épocas com um learning rate de 0.0001 chegando a uma acurácia de 93,8% e 87,5% respectivamente.

Outro exemplo de trabalho com fungos na agricultura é o (MARCOS; RODOVALHO; BACKES, 2019), onde é utilizado redes neurais para identificar ferrugem-do-café, uma doença da planta do café causada pelo fungo *Hemileia vastatrix*, problema que pode causar perdas de até 45% da produção. O dataset era composto de 159 imagens adquiridas por uma câmera digital Sony Cyber-shot na resolução de 2340x4160, no entanto, o número de imagens foi acrescido por *data augmentations*. Nesse contexto, utilizaram um arquitetura própria de rede neural, contendo apenas duas camadas convolucionais, seguidos ReLU e max pooling adquirindo um média do coeficiente Dice de 0.79

$$D = 2 * \frac{|A \cap B|}{|A| + |B|} \quad (3.1)$$

Sendo A e B duas imagens binárias e D o quão similar são as duas

Outro exemplo é o paper (PICEK et al., 2022) cujo trabalho de classificação de fungos está relacionado ao seu aplicativo FungiVision, que segundo eles aumentou o interesse do público em geral aos fungos. Para o dataset utilizaram o Atlas of Danish Fungi, base de dados que possui centenas de milhares de dados com 8300 espécies diferentes. Em relação ao treinamento foi utilizado a nova técnica de visão computacional, os Vision Transformers, superando mode-

los famosos como Inception e Resnets, chegando a 91.68% de top3 accuracy e 80.45% de top1 accuracy.

No artigo (HAO et al., 2019) é realizado um trabalho de identificação de fungos presentes no teste rotina de Leucorrhea usados em exames ginecológicas, importante evidência para detecção de vaginite. Foram utilizadas 1000 imagens, sendo 3640 amostras positivas e 6360 negativas tiradas com o microscópio Olympus CX31. Para o treinamento foram utilizados uma versão modificada do modelo LeNet5 e k-fold cross-validation com um k=10, dividindo 90% treinamento e 10% teste e como resultado final, cujos resultados chegaram a 93,26 % de acurácia.

Outro exemplo é o artigo (PROMMAKHOT; SRINONCHAT, 2020) onde novamente visão computacional é usado para investigar doenças relacionadas a fungos, no entanto, esse trabalho é baseado em micoses. O dataset foi baseado em 1500 imagens RGB, 100x100 px com um split de 80:20. Em relação ao classificador foi utilizado uma rede neural criada pelo autor, composta de 16 camadas criadas em Tensorflow e Keras com seguintes características: 1 camada de entrada, 4 camadas convolucionais 3x3 e 5x5, 4 camadas de max pooling 2x2, 4 camadas de ReLU, 2 camadas totalmente conectadas e uma camada de saída. Os resultados obtidos foram uma acurácia de 98,03%.

Um exemplo similar ao trabalho que será realizado nesse TCC é o (BILLONES et al., 2020) onde embora não seja a mesma espécie trabalhada, o conceito por trás é o mesmo. Devido a dificuldade de diferenciar diferentes espécies do gênero *Aspergillus* a olho nu, utilizam de redes neurais em estruturas microscópicas para a tentativa de sua diferenciação, no caso do autor são utilizados 9 classes de espécies de *Aspergillus* com aproximadamente 4545 imagens. Como classificador foi utilizado uma CNN própria cuja estrutura era: uma layer de input com três camadas de convolução 2D, com a função de ativação ReLU e uma camada núcleo que se baseia em uma camada totalmente conectada com 64 filtros com a função de ativação softmax e dropout rate de 0.2. O modelo foi treinado por 20 épocas e gerando uma acurácia de 94.2%.

Outra aplicação relevante a esse assunto é o trabalho de (TAHIR et al., 2018) que utiliza de redes neurais para classificar esporos microscópicos de diferentes fungos. No dataset foram utilizados 40 mil imagens de 5 diferentes classes: *Penicillioides*, *Restrictus*, *Versicolor*, *Cladosporium* e *Eurotium* e classificando através de uma arquitetura própria cuja estrutura são: 1 camada de input, 3 vezes um grupo de camadas de convolução, max pooling e ReLU seguido por duas camadas de FC e um softmax. Após vários testes o autor citou que o melhor learning rate que acharam foi 0.01 e embora não cite o batch size e o número de épocas, apenas afirmou que são características dependentes que devem ser escolhidas em combinação. Como resultado conseguiram 94.8% de acurácia.

Um exemplo de uma junção de visão computacional e um tema muito atual é o (RAHMAN et al., 2022) onde os autores propõem um sistema conjunto de IoT e aprendizado de máquina baseado em fazendas produtoras de cogumelos, que muitas vezes sofrem dificuldade em identificar possíveis cogumelos venenosos. O modelo de trabalho tem três focos: Sistema de monitoramento inteligente, Automação da fazenda e Classificação de cogumelos utilizando

aprendizado de máquina, o que focaremos nesse resumo. A aquisição de imagens foi realizada através de um raspberry pi e utilizam de um dataset disponível no Kaggle. Em relação ao treinamento foram utilizadas 6 técnicas diferentes: Árvores de decisão, Regressão Logística, K-nearest neighbor, SVM, Redes Bayesianas e Random Forrest, atingindo uma acurácia de 100% , 92,78%, 100%, 93,37%, 90,93% e 100% respectivamente.

Uma técnica nova é descrita em (DONG et al., 2021) onde os autores utilizam de Two-dimensional correlation spectroscopy (2DCOS) para discriminação de espécies vinculado a Deep learning. Essa técnica é versátil e possui a capacidade de extrair informação de estímulos físicos e químicos, além de adquirir descrições detalhadas de espectros síncronos e assíncronos. Para esse trabalho foi selecionado regiões características das 2054 boletos coletados e produzido 2DCOS imagens espectrais através de espectrômetro FT-MIR. Nesse contexto, para classificar os dados foi aplicado uma Resnet com 20 camadas, chegando em 99,76% de acurácia.

Assim como no artigo (PROMMAKHOT; SRINONCHAT, 2020) o paper (GAO et al., 2021) trata de fungos em nível microscópico para identificação de doenças relacionadas. Nesse trabalho foram coletados amostras de diferentes partes do corpo (Pele, Unhas, Cabelo) de diferentes espécies, resultando um total de 292 amostras e 1092 imagens, todas coletadas de uma câmera sensível a UV (XC-EU). O método de visão computacional usado foi um treinamento de uma Resnet50 com pesos treinado com a Imagenet. Os resultados foram de 99,5% para pele, 95,2% para as unhas e 60% para o cabelo.

O artigo (SOPO; HAJATI; GHEISARI, 2021) consiste em um trabalho da Victoria University Sydney, que almeja classificar um dataset DeFungi Dataset, que se baseia em 2500 imagens em 5 classes. Nessa pesquisa foi utilizada uma Resnet50 pré treinada com ImageNet e K-fold cross validation testing, gerando um resultado de 73,2% de acurácia sem transfer learning e 85,04% com transfer learning, resultados relativamente baixos comparados a todos os trabalhos apresentados anteriormente.

O trabalho (LIU et al., 2020) aborda um problema bem atual, que é ceratopatia (Doença não inflamatória no olho) causada por infecção de fungos, cuja ascensão é devido ao abuso de hormônios, antibióticos e uso de lentes de contato. Esse trabalho utiliza de dois algoritmos para pré processamento, o primeiro é o contrast stretching utilizado para realçar nervos, esporos, lesões e retirar o ruído das imagens. O segundo é o algoritmo de fusão de imagens baseado em HMF, cuja função é fundir imagens de multi-ângulos de acordo com certos pesos e regras definidas antes do pré processamento. Em relação ao dataset, foram criadas a partir de um microscópio Heidelberg HRT-3, gerando 1213 imagens de tamanho 384x384. O treinamento foi realizado com os modelos de rede Alexnet e VGG16 com uma acurácia de 99.95 e 99.98% respectivamente.

Assim como o artigo (MARCOS; RODOVALHO; BACKES, 2019) e (GAIKWAD; RUMMA; HANGARGE, 2022) o paper (PRZYBYŁ et al., 2020) utiliza de redes neurais para agricultura, mais especificamente em sementes para plantação de Colza, uma planta rica em óleo vegetal. Nesse estudo foram extraídos 520 imagens 2572x1936, 32 bit de sementes isoladas e observadas por microscópio e pré processadas através de Matlab e o Pid System utilizados

para filtrar as imagens. Em relação aos métodos utilizados, foram realizados experimentos com MLP (Multilayer Perceptron), RBF (Radial basis function network) e CNN. Foi constatado que o MLP e a CNN possuíram resultados iguais, com uma acurácia de 89%, no entanto, a MLP ganhou de ambos nos casos de teste com uma acurácia de 90%.

Artigos Selecionados	SVM	Bag of Words	Alexnet	Rede Própria	Lenet5	Vision Transformers	Resnet	Árvores de decisão	MLP
(ZIELINSKI et al., 2020)	X	X							
(GAIKWAD; RUMMA; HANGARGE, 2022)			X						
(MARCOS; RODOVALHO; BACKES, 2019)				X					
(PICEK et al., 2022)						X			
(HAO et al., 2019)					X				
(PROMMAKHOT; SRINONCHAT, 2020)				X					
(BILLONES et al., 2020)				X					
(TAHIR et al., 2018)				X					
(RAHMAN et al., 2022)	X							X	X
(DONG et al., 2021)							X		
(GAO et al., 2021)							X		
(SOPO; HAJATI; GHEISARI, 2021)							X		
(LIU et al., 2020)			X						
(PRZYBYŁ et al., 2020)				X					X

Figura 9 – Artigos organizados por método de treinamento

Artigos Selecionados	Infecções	Plantação	Aplicativo	Diferenciar Espécies
(ZIELINSKI et al., 2020)	X			
(GAIKWAD; RUMMA; HANGARGE, 2022)		X		
(MARCOS; RODOVALHO; BACKES, 2019)		X		
(PICEK et al., 2022)			X	
(HAO et al., 2019)	X			
(PROMMAKHOT; SRINONCHAT, 2020)	X			
(BILLONES et al., 2020)				X
(TAHIR et al., 2018)				X
(RAHMAN et al., 2022)		X		
(DONG et al., 2021)				X
(GAO et al., 2021)	X			
(SOPO; HAJATI; GHEISARI, 2021)				X
(LIU et al., 2020)	X			
(PRZYBYŁ et al., 2020)		X		

Figura 10 – Artigos organizados por função

## 4 DESENVOLVIMENTO

### 4.1 OBJETIVOS

- Utilizar os métodos observados no estado da arte de classificação de imagens
- Definir a melhor estratégia e rede neural para o problema em questão
- Definir qual é a estrutura mais propícia para a identificação entre classes

### 4.2 METODOLOGIA

Nesse seção será descrita a metodologia desse trabalho de conclusão de curso, demonstrando como foi coletado as amostras para os datasets analisados, e como eles foram tratados. Além disso, será descrito os métodos e bibliotecas que foram utilizadas para os experimentos em questão.

#### 4.2.1 Coleta de dados

Os fungos aos quais as fotos pertencem foram coletados em diversos locais do Brasil por membros do grupo de pesquisa MIND Funga no período de confecção do projeto Protax, entre 2020 e 2022. Através dessa coleta, foram selecionadas três espécies da Família Hymenochaetaceae: *Fulvifomes squamosus*, *Phylloporia clariceae* e *Phellinotus piptadeniae*.

As fotos foram tiradas a partir de equipamentos da marca LEICA, tanto o Esteromicroscópio quanto o Microscópio óptico, junto com suas câmeras acopladas. Após a escolha da posição, aumento e luz das lâminas nesses equipamentos, as fotos são capturadas pelo software da própria marca que é compatível com as câmeras dos equipamentos. Antes da captura das imagens, foram montadas lâminas microscópicas com cortes finos dos poros e dos tubos dos fungos em reagente de Melzer (solução aquosa de Iodeto de potássio, iodo e hidrato de cloral).

#### 4.2.2 Especificações do Equipamento

Esteromicroscópio: LEICA EZ4D com câmera acoplada.

Software Esteromicroscópio: Leica Application Suite Version 3.4.0 2016 (LAS EZ)

Microscópio Óptico: LEICA DM 2500

Câmera do Micro: LEICA DFC295

Software Microscópio: Leica Application Suite Version 3.8.0 2011 (LAS V3.8)

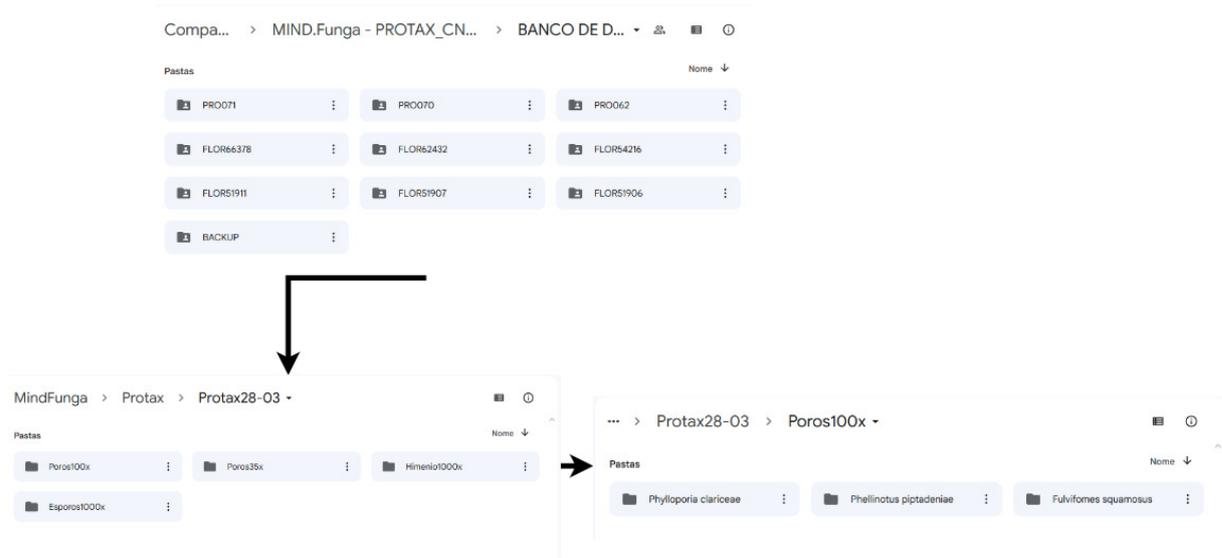


Figura 11 – Estrutura Dataset



Figura 12 – Estereomicroscópio e Microscópio Respectivamente

### 4.2.3 Datasets

As fotografias foram obtidas a partir de três exemplares distintos de cada uma das três espécies diferentes, totalizando 10 imagens por espécime. Assim, selecionamos três estruturas e quatro conjuntos de dados, que incluem:

- Esporos
- Himênio
- Poros 35x de Zoom
- Poros 100x de Zoom

Nesse contexto, foram separadas 92 imagens para as estruturas Himênio e Esporos e 90 imagens para os Poros em zoom de 35x,100x. Abaixo estão representados os 4 datasets através da espécie *Fulvifomes squamosus*:

#### 4.2.3.1 Esporos



Figura 13 – Imagens de Esporos

#### 4.2.3.2 Himênio

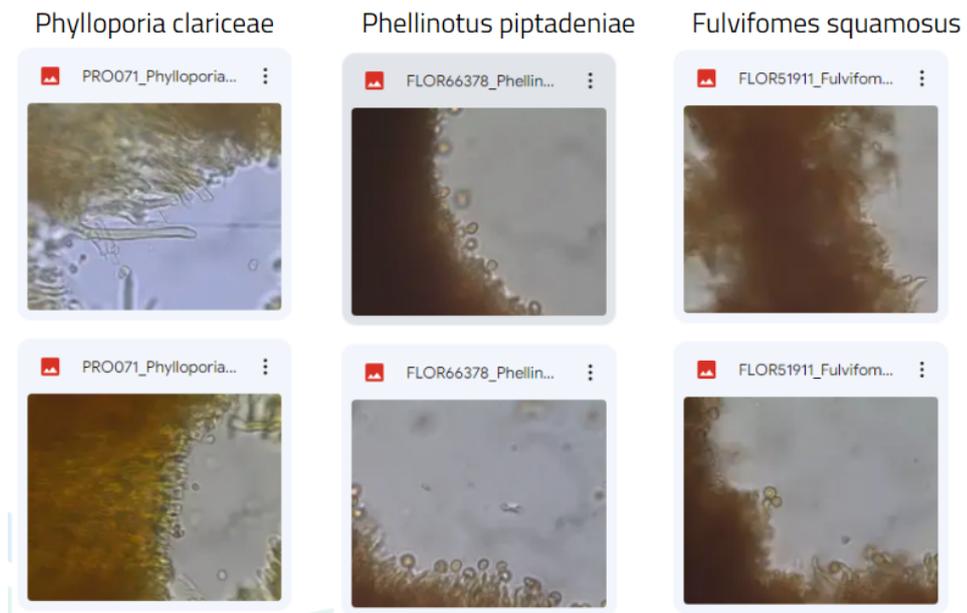


Figura 14 – Imagens de Himênio

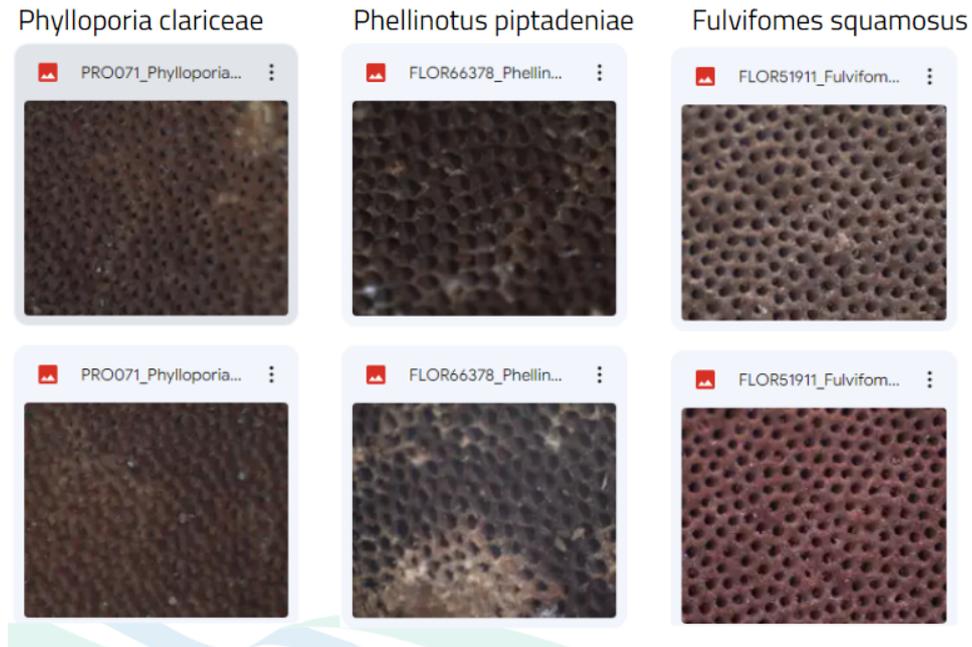


Figura 15 – Imagens de Poros 35x



Figura 16 – Imagens de Poros 100x

#### 4.2.3.3 Poros

#### 4.2.4 Análise de dados

As imagens dos fungos foram organizadas e enviadas separadas por espécimes, totalizando três exemplares para cada espécie estudada. Dentro de cada pasta referente a um espécime, havia quatro pastas distintas, cada uma representando uma característica desse organismo. Nesse contexto, os dados foram então divididos em quatro conjuntos de dados refe-

rentes as estruturas, sendo cada conjunto composto pelas três espécies analisadas. Essa divisão foi necessária devido ao formato de dados necessário para treinar as redes neurais. Após esse processo, foi necessário dividir o cada dataset através de scripts em subdatasets: Train, Valid e Test com proporção de 70%, 20% e 10% respectivamente.

Para esse trabalho de conclusão de curso foram utilizadas 4 bibliotecas para treinar as redes neurais sendo elas:

- Fastai
- Torch Vision
- Timm
- Wandb

Nesse contexto, o Fastai foi a plataforma base para o projeto, uma vez que lá estão definidas as funções básicas para o treinamento como: definir o dataset, batch-size, aplicar transformadas e fit-one-cycle. No entanto, as bibliotecas Torch Vision e Timm foram utilizadas para obter modelos de redes neurais não implementadas originalmente na plataforma como a HRNet, Efficientnet e os Vision Transformers. Além disso, foi utilizada uma quarta biblioteca chamada Wandb, cujo objetivo é fazer o plot e o monitoramento de métricas de distintas redes neurais. Nessa plataforma é possível ver a comparação de Acurácia, Train-Loss, Val-loss e entre outras dos diferentes modelos.

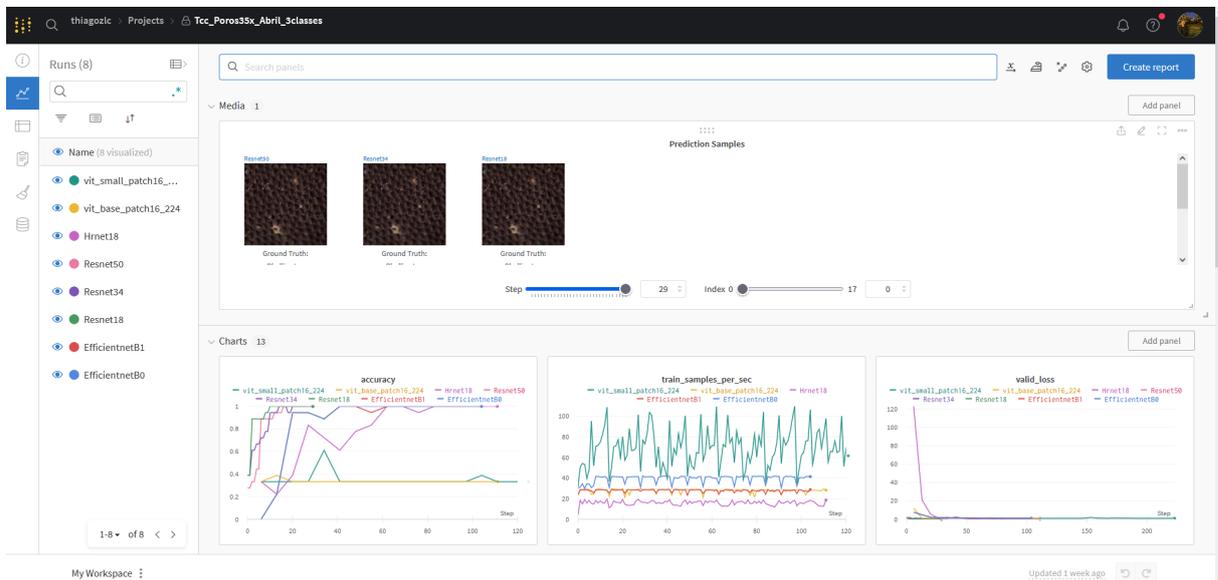


Figura 17 – Página do Wandb para os testes do Dataset Poros 35x

#### 4.2.5 Redes Escolhidas

Foram escolhidas 3 Resnets, 2 Efficientnets, 1 Hrnet, 2 Vision Transformers escolhidos em função de verificar testes em diferentes arquiteturas e profundidades.

- Resnet18
- Resnet34
- Resnet50
- Efficientnet B0
- Efficientnet B1
- Hrnet18
- VIT Small Patch
- VIT Base Patch

Optamos pelas ResNets em virtude de sua ampla tradição e aplicação na maioria dos problemas de classificação de imagens e segmentação semântica. Já as EfficientNets foram selecionadas considerando a relevância atual do modelo e a comparação de acurácia apresentada no artigo (TAN; LE, 2019), onde a EfficientNet B0 é comparável à ResNet50 e a B1 à ResNet101. Com base em experiências prévias em projetos similares, concluímos que não seria necessário empregar redes mais densas.

A HRNet foi escolhida por ser um modelo contemporâneo, que habitualmente demonstrou desempenhos superiores em projetos semelhantes conduzidos no laboratório Lapix. No que se refere aos Vision Transformers, optamos por sua utilização, devido ao amplo e recente reconhecimento de sua eficiência. Nessa contexto, selecionamos dois modelos, um básico e outro intermediário, considerando as limitações de tamanho do nosso conjunto de dados.

### 4.3 RESULTADOS

Neste estudo, foram realizados testes em três estruturas de fungos (Himênio, Esporos e Poros 35x e 100x de zoom) utilizando a biblioteca Fastai para o treino das redes neurais e a biblioteca wandb para registrar e organizar as métricas e resultados. Nesse contexto, foram utilizadas oito redes neurais diferentes, incluindo três modelos de Resnet, dois modelos de Efficientnet, uma Hrnet e dois modelos de Vision Transformers, visando comparar e avaliar a eficácia de cada arquitetura de rede na tarefa proposta. Essa abordagem permitiu explorar várias técnicas de Deep Learning e contribuiu para uma melhor compreensão das propriedades dos fungos estudados. Para a realização deste estudo, todas as redes neurais foram treinadas durante 15 épocas. Durante o treinamento, foram utilizadas taxas de aprendizagem (learning rates) que variaram entre  $1e-3$  e  $1e-7$ . A escolha desses valores foi motivada pela ausência de variação significativa ou pela percepção de um platô nas funções de perda na validação (Validation Loss) e perda no treinamento (Train Loss). Estes fenômenos foram observados na maioria das redes que passaram pelo processo de treinamento.

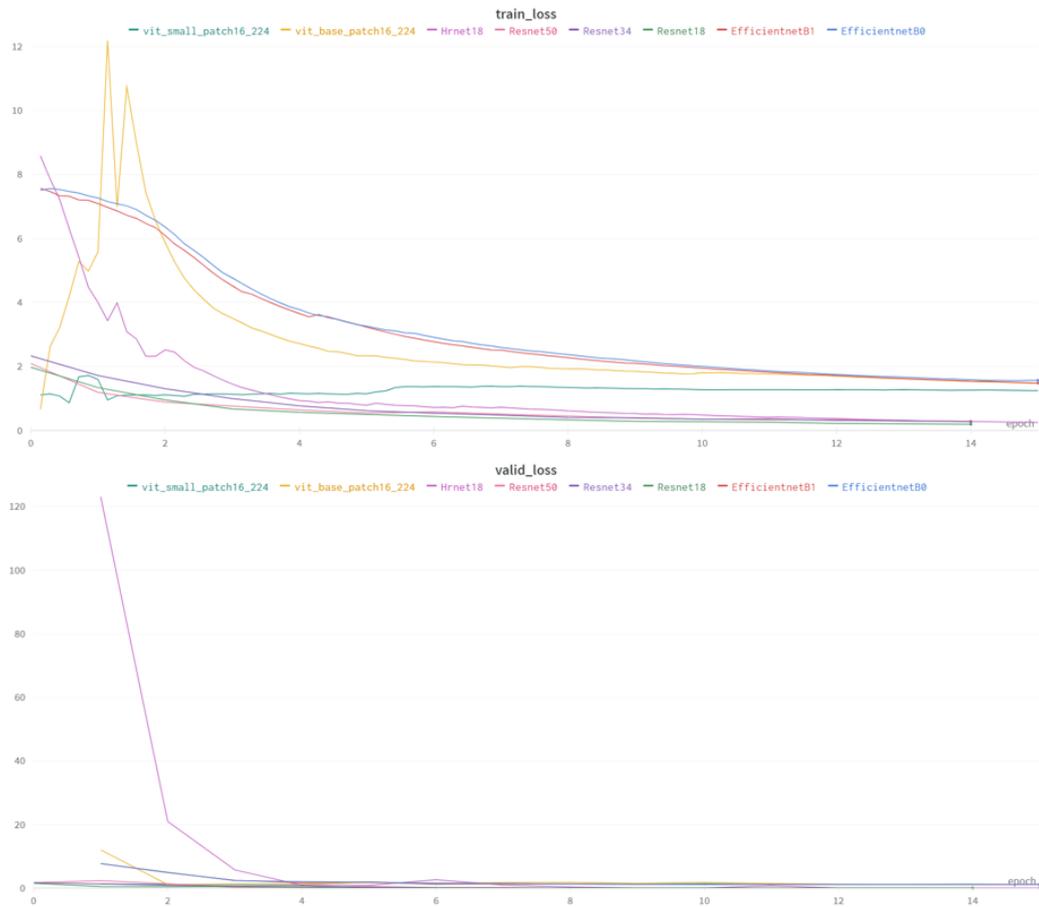


Figura 18 – Train e Valid Loss Experimentos Poros 35x

### 4.3.1 Himênio

Para essa estrutura em específico as melhores redes neurais sob a métrica de acurácia foram a Resnet50, Efficientnet B1 e a Hrnet, todas empatadas com 100% de acurácia no conjunto de teste. No entanto, a Resnet 50 e Hrnet possuíram Train e Valid Loss consideravelmente menores com 0,45 / 0,047 e 0,23 / 0,065 respectivamente. Outro detalhe importante foi que foi a estrutura de fungo com melhor sucesso entre os Vision Transformers, que não performaram bem em geral, gerando 72% de acurácia no small-patch e 61% no base-patch.

### 4.3.2 Esporos

Para essa estrutura a rede que mais se destacou foi Resnet34 com 94,12% de acurácia e 0,27 de Valid-loss, seguido da Efficientnet B0 e Hrnet com 88,24% e 82,35% de acurácia respectivamente. Além disso, essa foi a estrutura com piores resultados para os Vision Transformers, chegando a apenas 29% de acurácia.

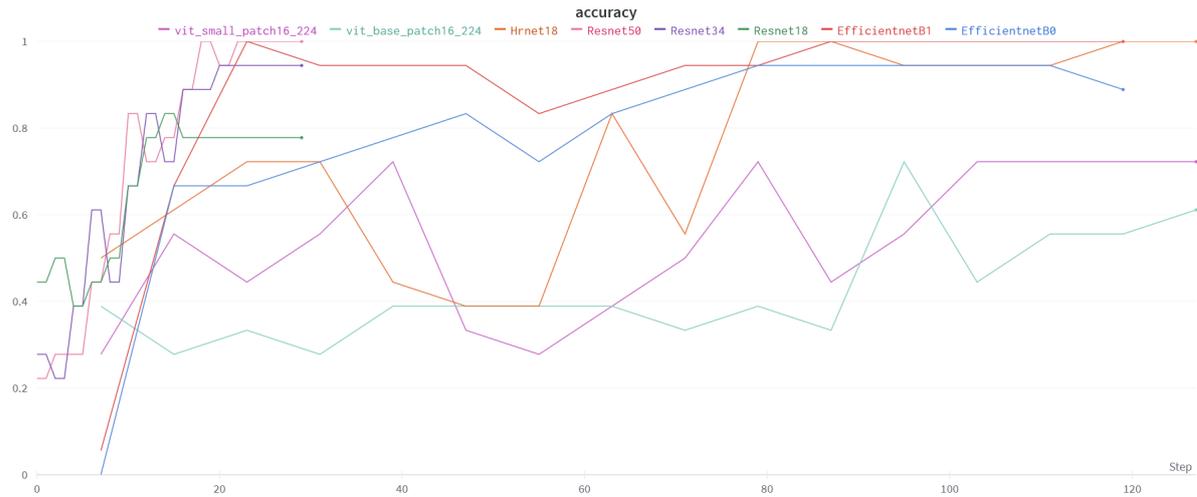


Figura 19 – Comparativo Redes Neurais Himênio 1000x de Zoom



Figura 20 – Comparativo Redes Neurais Esporos 1000x de Zoom

### 4.3.3 Poros

Para essa estrutura foram utilizadas duas versões de dataset, um utilizando 35x de zoom e outro com 100x. Os dois experimentos tiveram resultados semelhantes, onde a maioria das redes conseguiram chegar a 100% de acurácia, sendo elas: Resnet18, Resnet34, Resnet50, Efficientnet B0, Efficientnet B1, Hrnet.

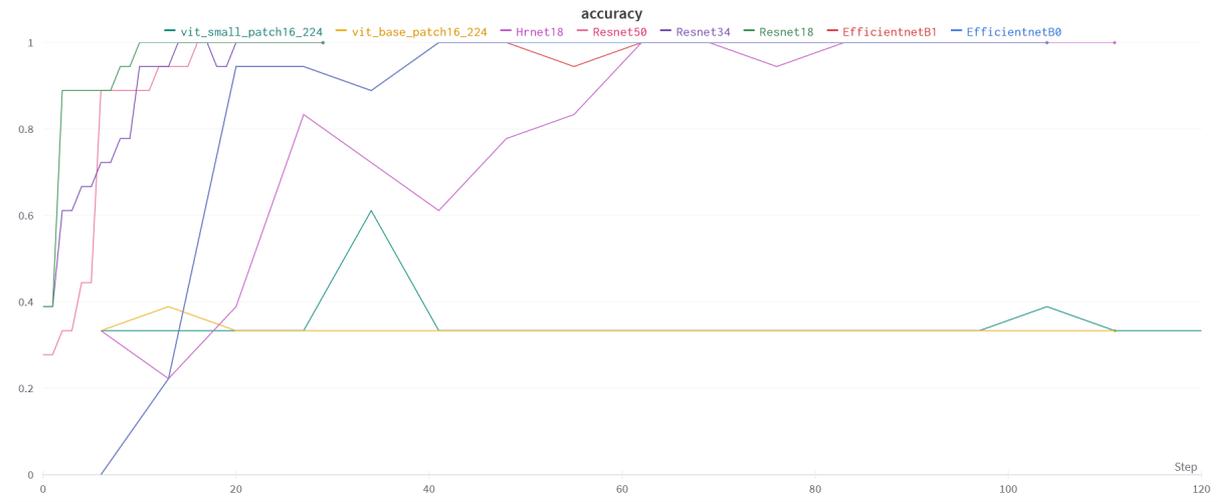


Figura 21 – Comparativo Redes Neurais Poros 35x de Zoom

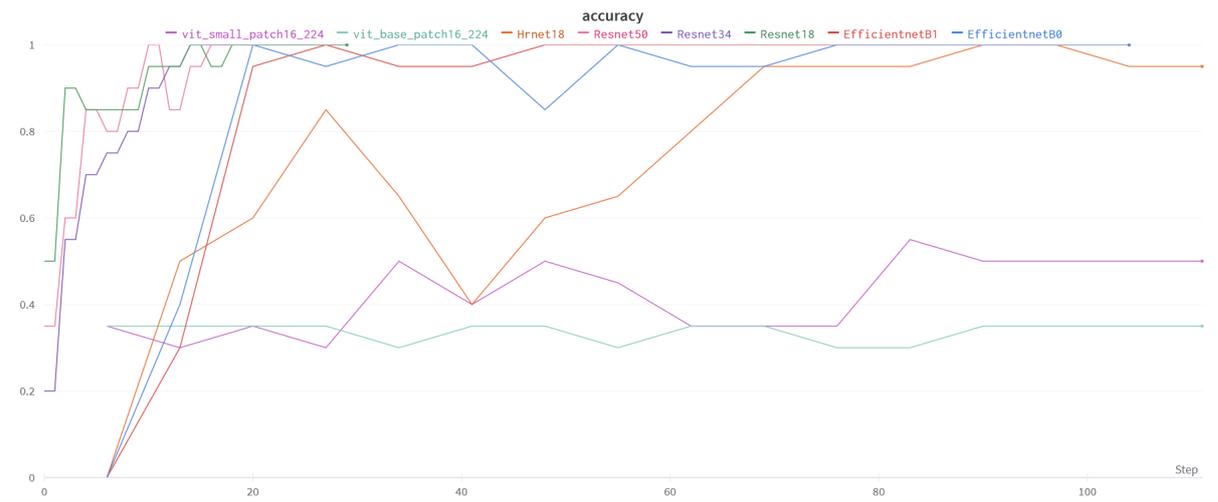


Figura 22 – Comparativo Redes Neurais Poros 100x de Zoom

Acurácia	Esporos 1000x	Himênio 1000x	Poros 35x	Poros 100x
Resnet18	76,47%	77,78%	100%	100%
Resnet34	94,12%	94,44%	100%	100%
Resnet50	76,47%	100%	100%	100%
EfficientnetB0	88,24%	88,89%	100%	100%
EfficientnetB1	76,47%	100%	100%	100%
Hrnet18	82,35%	100%	100%	95%
VIT Small Patch	29,41%	72,22%	33,33%	50%
VIT Base Patch	29,41%	61,11%	33,33%	35%

Figura 23 – Comparativo Geral dos Experimentos

## 5 CONCLUSÕES

Em conclusão, os resultados obtidos nos quatro experimentos demonstram que as arquiteturas de redes neurais convolucionais, como ResNets, EfficientNets e HRNet, apresentaram desempenhos semelhantes e igualmente satisfatórios na análise e processamento do conjunto de dados. Por outro lado, os Vision Transformers, que têm sido aclamados por sua eficiência em diversos cenários, não atingiram as expectativas para este tipo específico de dados. Os resultados sugerem que essas arquiteturas podem não ser a melhor opção para todos os problemas. Uma hipótese para a falha no trabalho em questão seria a falta de dados, uma vez que o artigo (DOSOVITSKIY et al., 2020), dos próprios criadores, afirma que, apesar da eficiência do modelo ser alta, esse modelo requer muito mais dados para treinamento do que uma rede tradicional.

Além disso, um dos objetivos do projeto era identificar quais possíveis estruturas dos fungos poderiam facilitar a identificação de uma espécie das demais e tudo indica que a melhor estrutura para esse problema seria os Poros, seja 35 ou 100x. Além de gerarem resultados com 100% de acurácia para um dataset de 90 imagens, foram gerados ótimos resultados em todas Resnets, todas EfficientNets e Hrnet testadas.

## REFERÊNCIAS

- BILLONES, R. K. C. et al. Aspergillus species fungi identification using microscopic scale images. In: **2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)**. [S.l.: s.n.], 2020. p. 1–5.
- DONG, J.-E. et al. Deep learning for species identification of bolete mushrooms with two-dimensional correlation spectral (2dcos) images. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 249, p. 119211, 2021. ISSN 1386-1425. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1386142520311902>.
- DOSOVITSKIY, A. et al. **An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale**. arXiv, 2020. Disponível em: <https://arxiv.org/abs/2010.11929>.
- GAIKWAD, S.; RUMMA, S.; HANGARGE, M. Classification of microscopic images of fungi using deep learning models. **SAMRIDDHI A Journal of Physical Sciences Engineering and Technology**, Vol 13, p. 105–110, 04 2022.
- GAO, W. et al. The design and application of an automated microscope developed based on deep learning for fungal detection in dermatology. **Mycoses**, v. 64, n. 3, p. 245–251, 2021. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/myc.13209>.
- HAO, R. et al. Automatic detection of fungi in microscopic leucorrhea images based on convolutional neural network and morphological method. In: **2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)**. [S.l.: s.n.], 2019. p. 2491–2494.
- HE, K. et al. Deep residual learning for image recognition. **CoRR**, abs/1512.03385, 2015. Disponível em: <http://arxiv.org/abs/1512.03385>.
- HOWARD, J.; GUGGER, S. fastai: A layered api for deep learning. **arXiv preprint arXiv:2002.04688**, 2020.
- HUBEL, D. H.; WIESEL, T. N. Receptive fields and functional architecture of monkey striate cortex. **The Journal of Physiology**, Wiley, v. 195, n. 1, p. 215–243, 1968.
- KIRK, P. et al. Dictionary of the fungi. **Mycol Res**, v. 113, p. 908–910, 01 2009.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: **Advances in neural information processing systems**. [S.l.: s.n.], 2012. p. 1097–1105.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, IEEE, v. 86, n. 11, p. 2278–2324, 1998.
- LIU, Z. et al. Automatic diagnosis of fungal keratitis using data augmentation and image fusion with deep convolutional neural network. **Computer Methods and Programs in Biomedicine**, v. 187, p. 105019, 2020. ISSN 0169-2607. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169260719308673>.

- MARCOS, A. P.; RODOVALHO, N. L. S.; BACKES, A. R. Coffee leaf rust detection using convolutional neural network. In: **2019 XV Workshop de Visão Computacional (WVC)**. [S.l.: s.n.], 2019. p. 38–42.
- PALLOTTINO, F. et al. Optoelectronic proximal sensing vehicle-mounted technologies in precision agriculture: A review. **Computers and Electronics in Agriculture**, v. 162, p. 859–873, 2019. ISSN 0168-1699. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169919304569>.
- PICEK, L. et al. Automatic fungi recognition: Deep learning meets mycology. **Sensors**, v. 22, n. 2, 2022. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/22/2/633>.
- PROMMAKHOT, A.; SRINONCHAT, J. Exploiting convolutional neural network for automatic fungus detection in microscope images. In: **2020 8th International Electrical Engineering Congress (iEECON)**. [S.l.: s.n.], 2020. p. 1–4.
- PRZYBYŁ, K. et al. Application of deep and machine learning using image analysis to detect fungal contamination of rapeseed. **Sensors**, v. 20, n. 24, 2020. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/20/24/7305>.
- RAHMAN, H. et al. Iot enabled mushroom farm automation with machine learning to classify toxic mushrooms in bangladesh. **Journal of Agriculture and Food Research**, v. 7, p. 100267, 2022. ISSN 2666-1543. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666154321001691>.
- RYVARDEN, L. **Neotropical polypores. Part 1 : Introduction , Ganodermataceae Hymenochaetaceae**. Oslo: Fungiflora, 2004. (Synopsis fungorum ; 19).
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.
- SOPO, C. J. P.; HAJATI, F.; GHEISARI, S. **DeFungi: Direct Mycological Examination of Microscopic Fungi Images**. arXiv, 2021. Disponível em: <https://arxiv.org/abs/2109.07322>.
- TAHIR, M. W. et al. A fungus spores dataset and a convolutional neural network based approach for fungus detection. **IEEE Transactions on NanoBioscience**, v. 17, n. 3, p. 281–290, 2018.
- TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. **International Conference on Machine Learning**, 2019.
- WANG, K. et al. Deep high-resolution representation learning for human pose estimation. **IEEE Conference on Computer Vision and Pattern Recognition**, 2019.
- ZHANG, B. J. Y. Lung nodule detectability of artificial intelligence-assisted ct image reading in lung cancer screening. **Journal**, 2022.
- ZIELIŃSKI, B. et al. Deep learning approach to describing and classifying fungi microscopic images. 2020. Disponível em: <https://openreview.net/forum?id=qbNMvyepH>.

## APÊNDICE A – APÊNDICE

### A.1 DATASET

<https://codigos.ufsc.br/thiago.zlc/dataset-protax>

### A.2 NOTEBOOKS

<https://github.com/ThiagoZLC/NotebooksTCC>

**APÊNDICE B – ARTIGO CIENTÍFICO**

# An Infrastructure for Cataloguing and Classifying Macrofungi using Convolutional Neural Networks <sup>1</sup>

Thiago Zimmermann Loureiro Chaves, Aldo Von Wangenheim, Joicymara Santos  
Xavier, Mahatma Titton, Elisando Drechsler

*Radarweg 29, Amsterdam*

*Elsevier Inc<sup>a,b</sup>, Global Customer Service<sup>b,2\*</sup>*

<sup>a</sup> *1600 John F Kennedy Boulevard, Philadelphia*

<sup>b</sup> *360 Park Avenue South, New York*

---

## Abstract

This work presents the application of computer vision and Deep Learning techniques with the objective of identifying and classifying macrofungi species, an initiative led by the Mind Funga Laboratory and funded by the PRONEM project. The study uses both traditional methods and those considered state-of-the-art. For the purpose of the research, a dataset was developed with more than 17 thousand images and 500 species, subdivided into three data sets: Train, Valid, and Test. All the images used were collected in the field or submitted by contributors and subsequently evaluated by experts. The accuracy obtained in the test sets was approximately 95.23%.

*Keywords:* Deep Learning, Computer Vision CNN, Image Classification, Macrofungi

2010 MSC: 00-01,99-00

---

<sup>1</sup> This is a footnote associated with the title.

<sup>2\*</sup> This is to indicate the corresponding author.

Email address: support@elsevier.com (Global Customer Service)

URL: www.elsevier.com (Elsevier Inc)

## 1. Introduction

Fungi are ubiquitous organisms that play vital roles in virtually all terrestrial ecosystems functioning. Their diversity is represented by the term Funga, equivalently to Fauna and Flora, which represent the diversity of animals and plants respectively (Kuhar et al., 2018). The Fungi kingdom is one of the most diverse groups, with mainly accepted estimates that range from ca. 1.5 to 5.1 million species (Hawksworth, 1991; Blackwell, 2011). A more recent estimation falls between 2.2 to 3.8 million species. Nevertheless, only a fraction of the Funga is known, as just ca. 10% of these species are described (Hawksworth & Lücking, 2017). Fungi act as key nutrient recyclers and form essential mutualistic relationships with several groups (Kendrick, 2011). Therefore, their ecological importance contrasts with this vast knowledge gap, being crucial to describe the Funga to fully understand the fungal community's role in complex ecosystems processes.

The slow process of species discovery and knowledge build worsens the Funga neglected scenario. It can take decades to build well-founded knowledge about some plant taxa (Goodwin et al. 2020) that arguably receives more attention than fungi. Hence, there is an urge to find ways to add up efforts and quicken the process of describing the Funga. The lack of knowledge about the Funga reflects the negative relationships between the general public and fungi, which can be considered mycophobic. Part of this repulsion to fungi could be attributed to their role in decaying food and causing diseases (Irga et al. 2018), coupled with the lack of awareness about their ecological importance. The Funga has its significance downplayed even in higher education, with mycology frequently not being treated as a course, but as a section in related grades (Grube et al. 2017).

As the Funga is neglected and plays an essential array of ecosystem processes, more efforts must be made to fill the knowledge gap in fungal diversity and make it available to the general public, especially with innovative techniques that speed up the process..

In this context, this work aims to use this tool, which is computer vision and Deep Learning, with the goal of identifying and classifying macrofungi species. Moreover, with these trained neural networks, we aspire to launch an application for both the general population and specialists, mainly to involve the community better with the fungi and remove this aversion that is becoming unfortunately common.

### *1.1 Objectives*

The objectives of this project are: (i) Analyze the literature on computer vision; (ii) Analyze the literature on examples similar to the project, (iii) Conduct a variety of tests with different models; (iv) Analyze the state of the art in image classification; (v) Develop neural networks to use in a future app

## 2. State of the Art

In this paper we roughly divide the analyzed scientific papers, from the point of view of the image processing technology used, into three general categories:(a) classic CV approaches (CCV), (b) deep learning (DL) approaches, and (c) hybrid CV approaches (HCV), where DL and CCV are employed together.

### 2.1. Background: classic computer vision vs. deep learning

Modern computer vision approaches can be roughly divided into classic computer vision (CCV) and deep learning based computer vision (DL). We will explain both concepts briefly here.

**Classic CV:** The general principle behind the CCV paradigm is that an image interpretation task starts from meaningless pixels and moves stepwise towards a meaningful representation of that image's content. This is performed as a pipeline of transforms in different levels of abstraction, whereas the first transforms in this pipeline are transforms from images into new and simplified images ( $I \rightarrow I$ ), such as noise filters, border detectors or segmentation algorithms. Later, when the abstraction level grows, these transforms are performed from images into models

( $I \rightarrow M$ ) that represent descriptions of specific elements or the content of these images, such as segment color and shape parameters or texture descriptors. Further on, these models are transformed into other, more abstract models ( $M \rightarrow M$ ) that describe the meaning or classification of those objects [22]. These transform steps can be applied on three different image processing domains: (i) value, when only context-less individual pixel values are being considered, such as in histograms or thresholds; (ii) space, when pixels are considered spatially, in the context of their neighborhoods, such as in border detectors, segmentation methods or general convolutions; and (iii) frequency, when pixel variation patterns along the image are taken into consideration, such as in Fourier or Wavelet transforms [5]. One characteristic of CCV is that these pipelines are extremely problem- and image type-specific and strongly parameter-dependent. This means that a processing pipeline that works well for a given type of image content, e.g. classification of tree crowns, will not work for agricultural weed detection. This parameter sensitivity also means that, e.g., a segmentation algorithm that works well for a set of images, will not work as well with another set of images of the same kind, but that shows different lightness and contrast characteristics. This made CCV solutions challenging to develop, requiring deep mathematical knowledge of the individual methods in order to allow an adequate composition of a pipeline for a specific problem. These characteristics also made CCV solutions less robust and highly image-quality-dependent, which has for a long time limited the popularization of CV software for natural history.

**Deep Learning and Convolutional Neural Networks:** The advent of DL techniques broke the CCV paradigm: DL convolutional neural networks (CNN) can learn sequences of convolution operations that represent image transforms in all three image domains and all levels of abstraction [6]. Due to their deep sequential structure, DL CNNs can also learn hierarchical transformation and representation sequences [6]. A DL application is able to, in one single step, start from pixels and end with a representation of the meaning of a given image. DL also simplified the development process of CV applications: the main steps of a CV solution can now be trained. However, DL solutions have the disadvantage of being black-box processes: the convolution operations performed by the network are *learned* and then coded as *data* in the structure of a given network, and not separate processes that can be isolated and individually analyzed.

A graphic with an overview of both paradigms is shown in Figure 3 on [7]. For this work, we considered CCV approaches all those that did not use any DL technique in its processing pipeline.

## 2.2. Scientific Works

A systematic literature review (SLR) will be conducted with the goal of verifying the general state of the art of image classification and Convolutional Neural Networks (CNN) contextualized with the subject of the work, in structures of macrofungi. Journals, Magazines, and conferences from 2018 to 2022 were considered for the contextualized research of the structures of fungi. All collected work has its date and information in order to facilitate the tracking of sources if necessary. The open access databases that were utilized for this research include the Institute of Electrical and

Electronics Engineers (IEEE), ScienceDirect Digital Library (SD), Springer (SP), and Google Scholar (GS).

Numerous Machine Learning and statistical techniques have been used to classify and identify fungi over the past 5 years. In order to identify these methods and papers, an SLR (Systematic Literature Review) was performed based on Kitchenham's methodology, widely used in scientific areas. In this Systematic Literature Review, a great diversity of models and techniques of neural networks and computer vision were found, such as: Bag of Words, SVM (Support Vector Machine), Models created by the authors, Pre-trained models of Resnet, InceptionV3, among others and current technologies like Vision Transformers (ViT).

After the literature review, 31 articles were collected but after a careful selection based on theme and technology used, they were filtered down to 14 articles. This process resulted in a refinement of the list down to 14 articles. These chosen articles primarily focus on four subject areas: fungal infections or diseases, fungal infestations in crops, the development of a fungus classification application, and the identification of different fungus species.

Firstly, speaking of the works related to human fungal infections. The article [19] deals with the identification of the type of fungus related to Candidiasis, such as *Candida Albicans*. In this article, pre-trained neural networks of Resnet and Alexnet were used, along with Bag of Words methods, achieving an accuracy of 93.9%. In the paper [20], a work is carried out to identify fungi present in routine Leucorrhoea tests used in gynecological examinations, important evidence for the detection of vaginitis. For the training, a modified version of the LeNet5 model and k-fold cross-validation were used, generating a result of 93.26% accuracy. Another example where computer vision is used is the article [21], where it is used to investigate diseases related to

fungi, more specifically, skin mycoses. For this work, a network with an architecture created by the laboratory was used and generated a result of 98.03% accuracy. Just like in the article [21] the paper [22] deals with microscopic level fungi for the identification of diseases related to skin, nails, and hair. The computer vision method used was a training of a Resnet50 with weights trained with the Imagenet. The results were 99.5% for the skin, 95.2% for the nails, and 60% for the hair. The paper [23] addresses a very current problem, which is keratopathy (Non-inflammatory eye disease) caused by fungal infection, whose rise is due to the abuse of hormones, antibiotics, and the use of contact lenses. The training was executed with the Alexnet and VGG16 network models with an accuracy of 99.95 and 99.98% respectively.

Another topic addressed in this systematic review was fungi in infestations and plantations, discussed in four of the fourteen articles. In the article [24] the authors proposed the analysis of fungi that affect mango and apple plantations, an important activity for Indian agriculture. For this work, pre-trained models of AlexNet and SqueezeNet were used, achieving 93.8% and 87.5% respectively. Another example of work with fungi in agriculture is [25], now located in Brazil, where neural networks are used to identify coffee rust, a disease of the coffee plant caused by the fungus *Hemileia vastatrix*. In this context, they used a proprietary neural network architecture, containing only two convolutional layers, followed by ReLU and max pooling, obtaining an average Dice coefficient of 0.79. An example of a combination of computer vision and a very current topic is [26] where the authors propose a joint system of IoT and machine learning based on mushroom-producing farms, which often have difficulty identifying possible poisonous mushrooms. In relation to training, 6 different techniques were used: Decision Trees, Logistic Regression, K-nearest neighbor, SVM, Bayesian Networks, and Random Forest, achieving an accuracy of 100%, 92.78%, 100%, 93.37%, 90.93% and 100% respectively. The paper [27] uses neural networks for agriculture, more specifically in seeds for planting Rapeseed, a plant rich in vegetable oil. An MLP and a CNN were used, in this process it was

found that the MLP and the CNN had similar results, with an accuracy of 89%, however, the MLP won in the test cases with an accuracy of 90%.

In addition to these contexts, a group of the articles collected had an objective of differentiation between species. The article [28] uses computer vision due to the difficulty of differentiating species of the genus *Aspergillus*, especially by the naked eye, in this sense, they use neural networks in microscopic structures in an attempt to differentiate them, in the author's case 9 classes of *Aspergillus* species are used. As a classifier, a proprietary CNN was used where the model was trained for 20 epochs and generated an accuracy of 94.2%. Another relevant application to this subject is the work of [29] that uses neural networks to classify microscopic spores of different fungi, more specifically 5 species: *Penicillioides*, *Restrictus*, *Versicolor*, *Cladosporium* and *Eurotium*. For this work, a proprietary network was used that resulted in 94.8% accuracy. The work [30] uses a new technique called Two-dimensional correlation spectroscopy (2DCOS) for species discrimination linked to Deep learning, more specifically of boletes. This method was based on creating spectral images through an FT-MIR spectrometer and classifying them through a neural network, which in this case was applied to a Resnet with 20 layers, achieving 99.76% accuracy. The paper [31] is a work from Victoria University Sydney, which aims to classify a DeFungi Dataset, which is based on 2500 images in 5 classes. In this research, a pre-trained Resnet50 with ImageNet and K-fold cross-validation testing was used, generating a result of 73.2% accuracy without transfer learning and 85.04% with transfer learning.

Lastly, an article [32] was verified whose work on fungus classification is related to their FungiVision application, which according to them has increased the general public's interest in fungi. For the dataset, they used the Atlas of Danish Fungi, a database that has hundreds of thousands of data with 8300 different species. In terms of training, they used the new technique of computer vision, the Vision Transformers,

surpassing famous models like Inception and Resnets, reaching 91.68% top3 accuracy and 80.45% top1 accuracy.

### **3. Material and Methods**

In this section we will describe our dataset and the approach we followed to acquire the images, build the dataset and curate it. We will also describe the image classification CNN we choose to develop our macro fungi classifier and also describe the approach followed to develop the web application to host the image classifier and also the image dataset.

#### *3.1. Macrofungi Image Dataset*

The MIND.Funga dataset is composed of 17,467 curated images that represent 511 fungal species [33]. Briefly, the raw image dataset is composed of images received via email or asked for incorporation in the dataset, where sources included partner mycologists, volunteers, citizen scientists, online mycology groups, and past research projects focused on macrofungi diversity. All images were submitted to a quality control and treatment pipeline, which involved three critical steps: removal of poor-quality, unfocused images and images with foreign objects (e.g. scales, human parts, knives), specimen reframing to the image center, and 1:1 aspect ratio patronization. Initially, images had their original background replaced with green and white backgrounds (HEX codes) to improve CNN training and specimen recognition. As the background replacement didn't significantly improve the CNN performance, new additions were not modified. When multiple images of the same specimen were available, a subset that shows all specimen angles was selected. The dataset is organized in individual taxon directories names with current accepted scientific names for species or genera. Each taxon directory has associated metadata composed of taxonomy, morphology, phylogeny, ecology, and distribution information.

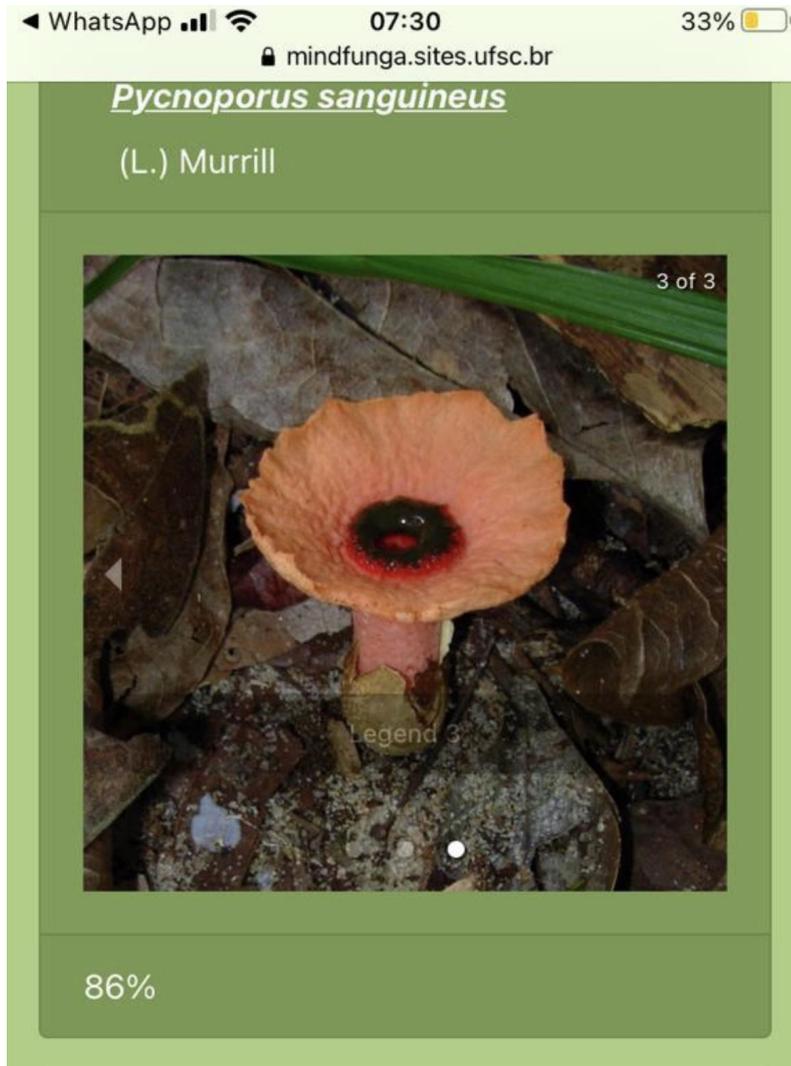


Figure 1 - Alpha Version of the App

### 3.2. Image Classification CNN

The identification of macro fungi specimens in the wild is a typical image classification (IC) task. In order to implement our approach, we choose a state of the art image classification CNN, the Residual Convolutional Neural Network (ResNet) [8]. ResNets are CNN models that present a very homogeneous, elegant and scalable architecture, which allows the development of extremely deep and powerful networks: a ResNet is typically composed of a sequence of *residual blocks* of varying length. A residual block is a structured sequence of two or more convolutional layers that possess an extra set of bypass connections that allow part of the retro-propagated error to jump over the block and thus be sent back to the beginning of the network with less signal loss during the training procedure. This allows ResNets to solve a common problem of deep neural networks called *vanishing gradients*. This homogeneous, building block-based architectural principle also allows one to easily design customized ResNets with different depths.

For the purpose of this work we selected a set of standard, well-proven ResNet architectures, because they are available as pre-trained networks, trained on the Imagenet data set and, thus, present internal connection and weight structures that reflect image characteristics found on most natural and common objects. CNNs pre-trained on the Imagenet dataset are much easier to train on a new data set and the procedure to use them as the starting point in a new project has become the state of the practice in modern computer vision projects.

In order to train the pre-trained ResNets on our macrofungi dataset, we followed the standard two-stage procedure for adapting pre-trained CNNs to a new image domain:

**Stage 1: Transfer Learning.** A new pair of fully-connected classification output layers, customized for the number of classes of our image domain, was generated for each network. The networks were then trained with only the connections from the network body to its input layer and the connections of the last two fully-connected classification layers allowed to learn. This first step is fast and maps the pre-trained network to the new image data.

**Stage 2: Fine Tuning.** All connections of the networks were unfrozen and a full, much longer, training cycle was performed. This second step adapts the whole network internally to better reflect the characteristics of new image data.

We additionally configured our ResNet models with an ADAM optimizer [9] and the cross entropy loss function. ADAM is an algorithm for first-order gradient-based optimization of stochastic objective functions that can be used instead of the classical stochastic gradient descent procedure to update neural network weights iteratively based on training data. It computes individual adaptive learning rates for different parameters from estimates of first and second moments of the gradients calculated at a given moment during the training process.

To accelerate training and further automate the optimization of the learning rates and network momentum during training, we applied the *One Cycle Policy* on both training stages: transfer-learning and fine-tuning [10][11]. One cycle policy is a training methodology introduced by Leslie Smith in a paper [12]. It suggests a way of training deep learning models with cyclic learning rates, where we start from a low learning rate, increase it for the first half of the training, and then decrease it for the second half. Furthermore this policy includes a brief period of further decreased learning rate at the very end of training. In addition to using Resnet, a traditional method and generally seen as a base solution, we used state of the art neural networks such as Efficientnets, HRnets and Mobilenets, all of them where trained in Fastai with the `fit_one_cycle` policy.

EfficientNet is a family of deep neural network architectures proposed in the paper [13], designed to maximize both computational efficiency and model accuracy. This is achieved through a scaling method composed of width, depth, resolution, and the use of efficient building blocks called MBConv. Therefore, the result is a highly efficient architecture that can be easily adapted to different sizes and tasks. There are 7 models of this architecture, ranging from Efficientnet B0 to Efficientnet B7 which basically differ in network depth. Furthermore, EfficientNets use a scaling coefficient to increase the depth, width, and resolution of the model in a balanced way. First, a base model (EfficientNet-B0) is trained using automatic neural architecture search to find the best network architecture in terms of efficiency and accuracy, and then the base model is scaled to different sizes (EfficientNet-B1 to B7) using compound model scaling.

Hrnet (High-Resolution Network) is a convolutional neural network architecture proposed in the paper [14] with the aim of improving the accuracy of understanding the human body and its poses. However, it is also used for other purposes due to its high efficiency. Hrnet, unlike most neural networks, maintains the highest resolution throughout the computing process due to multiple parallel processing paths. First the HRNet begins with an initial stage where input images are processed at high resolution, necessary because it helps to preserve detailed information present. Following this, the network divides into several branches, each processing the image at a different resolution. These parallel branches allow the network to maintain high resolution throughout the inference process. Additionally, there are cross-connections between each pair of branches at each stage. Therefore, these paths do not operate alone, they communicate and share information with each other. High-resolution features are downsampled and fused with the low-resolution features, and the reverse also occurs. In the end, the outputs from different branches at different resolutions are fused together to form a single representation. This representation then goes through a classification layer to produce the final output.

MobileNetV3, introduced by Andrew Howard and colleagues in a 2019 paper [15], is the third version of the MobileNet series. The MobileNet architecture is known for its efficiency in terms of speed and size, making it ideal for use on mobile devices and other applications with limited resources. The architecture uses a combination of expansive and point convolutions to maintain efficiency, but also uses h-swish activation and h-sigmoid activation, which are optimized for mobile devices. First, the input goes through a 3x3 convolution layer followed by batch normalization and h-swish activation. After the initial layers, MobileNetV3 uses a series of point and expansion convolution modules and Squeeze-and-Excitation (SE) modules, which are essential for the network to focus on important features by dynamically adjusting the channel weights. The final layers of the network include a point convolution and batch normalization, h-swish activation, and a global average pooling layer. The output of this layer is then fed into a fully connected layer to produce the final output of the network.

## 4. Results

In this section we will discuss the classification accuracy results we obtained with the CNN models we tested and also present the on-line macrofungi classification system we developed to host the image dataset and the CNN and also allow users to classify their photographs and submit new specimens to the database.

For this work, four libraries were used to train the neural networks, which are:

- Fastai -> [www.fast.ai](http://www.fast.ai)
- Torch Vision -> [pytorch.org/vision/stable/index.html](https://pytorch.org/vision/stable/index.html)
- Timm-> [github.com/huggingface/pytorch-image-models/tree/main/timm](https://github.com/huggingface/pytorch-image-models/tree/main/timm)
- Wandb -> [wandb.ai/home](https://wandb.ai/home)

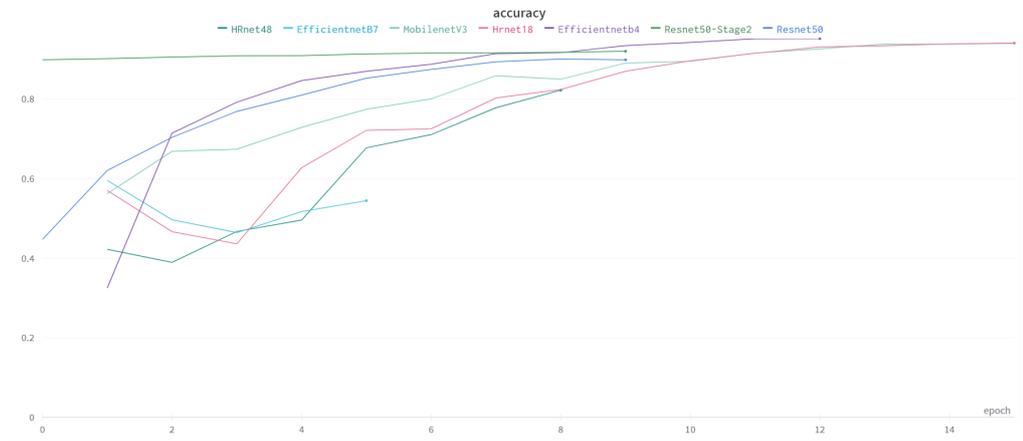
In this context, Fastai was the base platform for the project, as it defines the basic functions for training such as: setting the dataset, batch-size, applying transformations, and fit-one-cycle. However, the Torch Vision and Timm libraries were used to obtain neural network models not originally implemented on the platform such as HRNet, Efficientnet, and Mobilenet. In addition, a fourth library called Wandb was used, whose purpose is to plot and monitor metrics of different neural networks. On this platform, it is possible to see the comparison of Accuracy, Train-Loss, Val-loss, among others, of different models.

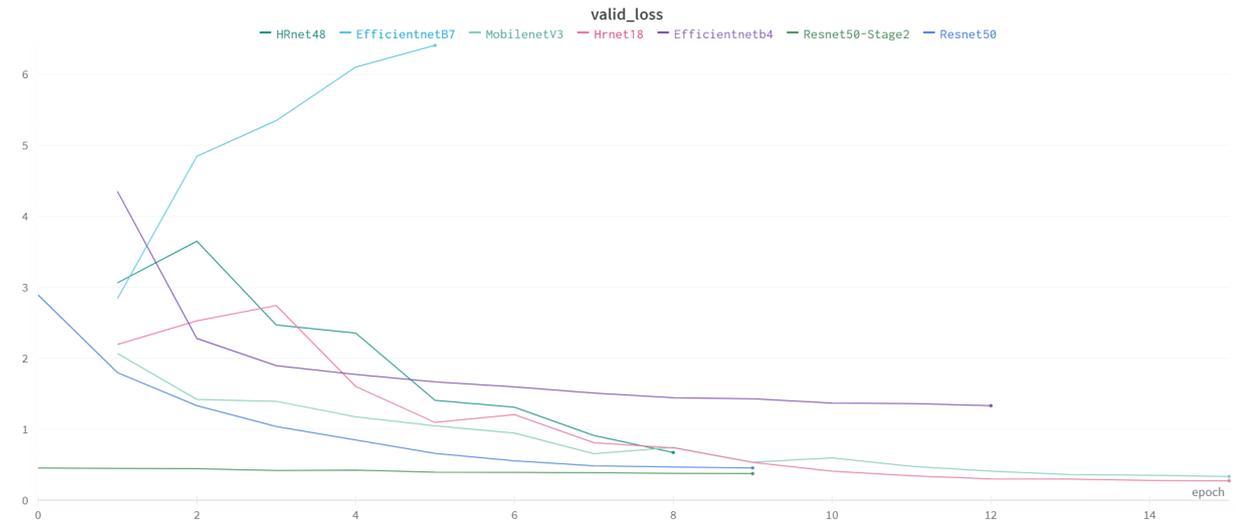
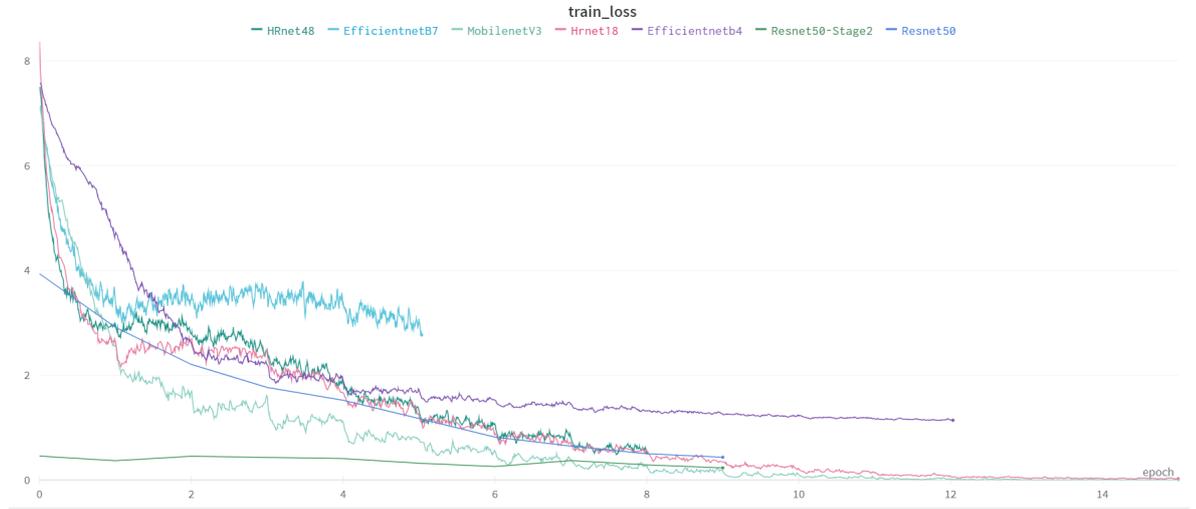
#### *4.1. Neural Network Performance*

Six different networks from four different architectures were trained:

- Resnet50
- Efficientnet B4
- Efficientnet B7
- Hrnet18
- Hrnet48
- MobilenetV3

The ResNet was utilized due to its popularity and frequent use as a base network. Specifically, we chose the ResNet50 due to the complexity of the dataset, which features approximately 500 classes and 17,000 images. The EfficientNets and HRNets were chosen in consideration of their current relevance and their proven superior performance in similar projects conducted at the Lapix laboratory. On the other hand, the MobileNetV3 was selected due to the need for a more efficient and lighter network, suitable for use in mobile applications, with a view to future public release of the app.





The networks were trained over a span of 9 to 15 epochs, with termination based on the analysis of the Train and Valid Loss graphs, aiming to reach a plateau. However, the EfficientnetB7 was prematurely halted as the Train Loss and accuracy did not generate significant results, while the Valid Loss merely increased. The best performance was achieved by the EfficientnetB4, with an accuracy of 95.23%, followed by the Hrnet18 at 94.11%, and the MobilenetV3 at 94.05%. The performance of the MobilenetV3 was particularly impressive compared to conventional networks, considering it is a network designed to be lightweight and efficient for the mobile market.

## 5. Discussion

In this work we tackled a problem for which, at least in various related fields, there already exist *ad hoc* solutions freely distributed as smartphone apps, such as plant and animal classification apps. This necessarily raises the need for comparison: are we as good as those apps? Unfortunately the documentation on the scientific and technical background of those solutions is as good as inexistent. Even if it seems extremely likely that those apps employ CNNs, which models they use and what their precision rates remain unknown. On the other side, solid scientific work on artificial intelligence solutions for fungal classification is also extremely sparse and most papers we analyzed are either extremely naive or present very poor results. These situations make it very difficult to discuss our work in the light of similar or previous works.

### 5.1. Limitations

As in many other AI systems, our CNN suffers from the Closed-World Assumption (CWA) problem: in a formal system of logic used for knowledge representation, CWA is the presumption that a statement that is computed as true is also known to be true and that there cannot be data outside the universe known to the system [1]. In an image classification CNN this manifests as the fact that the CNN will always classify an image, answering with its most similar class, even if what the image depicts lies completely outside the scope of the dataset used to train it. As an example, the picture below depicts the results of trying to classify an image of a computer keyboard and that of a Christmas ornament with our CNN: the system will always provide the list of the five fungal responses with the most likelihood. Since the network employs the *softmax* probability function at its output layer, the sum of all outputs will always be 100% [3]. In some cases, as it is in the keyboard example, the probability associated with the highest-ranked answer can be high, contrary to the common sense assumption that an image depicting a non-trained object will generate a diffuse and non-specific answer from the CNN. This indicates that a mere contrast analysis between the CNN outputs for each class, where a high contrast between the highest ranking value and the others would indicate certainty, whereas a diffuse answer would indicate an unknown object, as performed in approaches substituting the Softmax function with the OpenMax function, cannot offer guarantee that the image pertains *de facto* to a known class [4]. Another limitation of the project is the imbalance among the classes, where some possess significantly fewer images than others, potentially leading to a bias towards the more represented species. Specifically, 46 of the 511 species have only 1 to 3 images, while some have over 100.

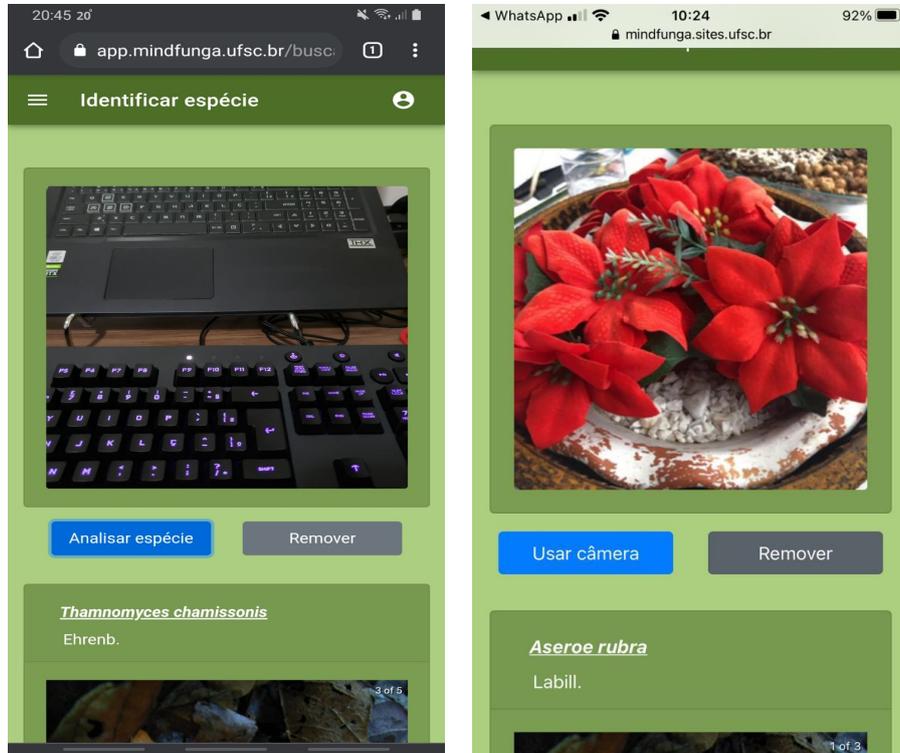


Figure 2,3 - Computer keyboard and Christmas ornament classified respectively as *Thammomyces chamissonis* (96%!) and *Aseroe rubra*.

Since the app is intended to be used in schools, the submission of pornographic and other NSFW pictures is another problem that has to be considered and images containing inadequate content should be discarded immediately. To resolve these current limitations, there are two potential approaches that could be followed. The first option involves expanding the classes in our dataset to include common objects and Not Safe for Work (NSFW) images, and then retraining our Convolutional Neural Network (CNN) with these additional images. Alternatively, the second approach would be to establish a pipeline with filters designed to pre-classify images into categories such as common objects, non-fungal living beings, and NSFW images. These images could subsequently be eliminated, or alternatively, the user could be given a crucial response before they submit an image to our macrofungi CNN.

The first approach has the disadvantage that, in order for it to work, it would need a huge database of non-fungal objects. Besides the complexities of building such a database, this would also introduce a bias into our neural network, since most of our fungal classes are composed of less than 30 specimens and these extra classes would necessarily contain thousands of objects. The second approach is much more promising and we will discuss it briefly in our Future Work section below.

## **6. Conclusions**

Upon reviewing the results, we observed a highly satisfactory performance from many neural networks, particularly the Efficientnet B4, which achieved 95.23% accuracy. This study concluded that for the problem at hand, the most effective networks weren't necessarily the most complex, such as the EfficientnetB7 and Hnet48. These models yielded results significantly below average, with accuracies of 55% and 82.26%, respectively. Furthermore, the MobilenetV3's training outcome was a pleasantly surprising revelation. Its performance mirrored that of the traditional, larger networks

and markedly surpassed its predecessor, MobilenetV2, which had been trained in the lab during previous experiments. Utilizing these trained neural networks, the lab aspires to develop an app for classifying fungi, specifically targeting Brazilian species, which sets it apart from numerous applications readily available in mobile app stores. This initiative could raise awareness of macrofungi among the general populace and may contribute to diminishing the extent of mycophobia that a portion of the population exhibits.

## 7. Future Work

In order to address the main CWA limitation we identified during the tests of the system, we are now working on a nonsense filtering pipeline. Following the second approach presented in the *Limitations* subsection, we are developing a cascaded filter, which will first filter out common objects and then, depending on this first classification, try to identify if there are NSFW images that were submitted to the system or pass the image on to our macrofungi CNN. The first task of this pipeline will be to identify common non-fungal objects. For this purpose we are investigating the possibility of using a standard image classification CNN, pre-trained on the 1,000 base-classes of common objects of the Imagenet<sup>3</sup> data set, as an image filter [2]. We understand that, since this dataset is composed of millions of images, a neural network pre-trained on it would provide us with much better capabilities of identifying common objects than we could achieve in training a few extra classes in our macrofungi neural network. For the second task, identifying NSFW images, we are partnering with the *Computing at School Initiative* at our University<sup>4</sup>, which is presently developing online CNN-based computer vision courses for high school students and also needs to develop an NSFW filter. Another necessary future work is the balancing of the dataset. Despite the fact that some species may be harder to come by than others, it's essential to ensure a roughly equal number of images across species to avoid any biases. In addition to that we want to test Vision Transformers in this context, which is a technology that has proven a good alternative in many projects.

---

<sup>3</sup> <https://image-net.org/>

<sup>4</sup> <https://computacaonaescola.ufsc.br/>

## Supplementary Information

Selected Papers	SVM	Bag of Words	Alexnet	Own CNN	Lenet5	Vision Transformers	Resnet	Decision Trees	MLP
(ZIELINSKI et al., 2020)	X	X							
(GAIKWAD; RUMMA; HANGARGE, 2022)			X						
(MARCOS; RODOVALHO;BACKES, 2019)				X					
(PICEK et al., 2022)						X			
(HAO et al., 2019)					X				
(PROMMAKHOT; SRINONCHAT, 2020)				X					
(BILLONES et al., 2020)				X					
(TAHIR et al., 2018)				X					
(RAHMAN et al., 2022)	X							X	X
(DONG et al., 2021)							X		
(GAO et al.,2021)							X		
(SOPO; HAJATI; GHEISARI, 2021)							X		
(LIU et al., 2020)			X						
(PRZYBYŁ et al., 2020)				X					X

Table S1 - Papers organized by Method

Selected Papers	Human Infections	Plantation	App	Differentiate species
(ZIELINSKI et al., 2020)	X			
(GAIKWAD; RUMMA; HANGARGE, 2022)		X		
(MARCOS; RODOVALHO;BACKES, 2019)		X		
(PICEK et al., 2022)			X	
(HAO et al., 2019)	X			
(PROMMAKHOT; SRINONCHAT, 2020)	X			
(BILLONES et al., 2020)				X
(TAHIR et al., 2018)				X
(RAHMAN et al., 2022)		X		
(DONG et al., 2021)				X
(GAO et al.,2021)	X			
(SOPO; HAJATI; GHEISARI, 2021)				X
(LIU et al., 2020)	X			
(PRZYBYŁ et al., 2020)		X		

Table S2 - Papers organized by Matter

#	Base	Initial Search String
1	IEEE	((Deep Learning OR Microscopic) AND (Fungi OR Protozoa OR Micology)) OR ((Deep convolutional neural network OR MICROSCOPIC) AND (fungi OR protozoa OR micology))
2	SD	((("Deep Learning") OR ("Machine Learning") OR ("CNN")) AND (("Microscopic) OR ("Mushroom") OR ("Micology"))
3	SP	(mushroom OR micology OR fungi OR funga) AND (Deep Learning OR Machine Learning OR CNN)
4	GS	(mushroom OR micology OR fungi OR funga OR microscopic) AND (Deep Learning OR Machine Learning OR CNN)

Figure 1 - Search Strings

First Version of the Dataset:

<https://data.mendeley.com/datasets/sfrbdjvxcc/1>

## Acknowledgements

Acknowledgements and Reference heading should be left justified, bold, with the first letter capitalized but have no numbers. Text below continues as normal.

## References

- [1] Reiter, Raymond (1978). "On Closed World Data Bases". In Gallaire, Hervé; Minker, Jack. *Logic and Data Bases*. Plenum Press. pp. 119–140. ISBN 9780306400605.
- [2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, **ImageNet: A Large-Scale Hierarchical Image Database**. *IEEE Computer Vision and Pattern Recognition (CVPR), 2009*
- [3] Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron (2016). "6.2.2.3 Softmax Units for Multinoulli Output Distributions". *Deep Learning*. MIT Press. pp. 180–184. ISBN 978-0-26203561-3.
- [4] Abhijit Bendale and Terrance Boult. Towards Open Set Deep Networks. arXiv, e-print 1511.06233, 2015. Available at: <https://arxiv.org/abs/1511.06233>
- [5] D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Henry Holt and Co., Inc., USA, 1982. doi:10.5555/1095712.

- [6] Y. LeCun, Y. Bengio, G. Hinton. Deep learning, Nature 521 (2015) 436–444. DOI: 10.1038/nature14539.
- [7] Matias, A. V. ; Macarini, L. A. B. ; Amorim, J. G. A. ; Cerentini, A. ; Onofre, F. B. M. ; Onofre, A. S. C. ; Daltoe, F. P. ; Stemmer, M. R. ; von Wangenheim, Aldo . What is the State of the Art of Computer Vision-Assisted Cytology? A Systematic Literature Review. COMPUTERIZED MEDICAL IMAGING AND GRAPHICS, p. 101934, 2021. DOI: 10.1016/j.compmedimag.2021.101934
- [8] He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778. DOI: 10.1109/CVPR.2016.90
- [9] Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. URL <http://arxiv.org/abs/1412.6980>. Comment: Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015
- [10] Smith LN (2015) No more pesky learning rate guessing games. arXiv. CoRR abs/1506.01186, URL <http://arxiv.org/abs/1506.01186>, 1506.0118635.
- [11] Smith LN (2018) A disciplined approach to neural network hyper-parameters: Part 1 - learning rate, batch size, momentum, and weight decay. arXiv. CoRR abs/1803.09820, URL <http://arxiv.org/abs/1803.09820>, 1803.09820
- [12] A Disciplined Approach to Neural Network Hyper-parameters: Part 1 -- Learning Rate, Batch Size, Momentum, and Weight Decay

- [13] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proceedings of the 36th International Conference on Machine Learning, 2019, vol. 97, pp. 6105–6114.
- [14] J. Wang et al., "Deep high-resolution representation learning for visual recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 1, pp. 334-347, 2020.
- [15] A. Howard et al., "Searching for MobileNetV3," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 1314-1324.
- [16] BRASIL. Secretaria de Educação Fundamental. Parâmetros Curriculares Nacionais: Ciências. Naturais /Secretaria de Educação Fundamental. Brasília: MEC /SEF, 1998
- [17] SILVA, C. H., RODRIGUES, C. W. M., OLIVEIRA, G. F., ARAÚJO M. L. F., Estudando fungos a partir de uma prática problematizadora e dialógica: Relato de uma experiência no ensino médio em uma escola pública. Universidade Federal Rural de Pernambuco: 2009.
- [18] MARTÍNEZ-AZNAR, M.M. Un estudio comparativo sobre El pensamiento profesional y La 'acción docente' de los profesores de ciencias de educación secundaria. Parte II. Enseñanza de las Ciencias, Barcelona, v. 20, n. 2, p. 243 -260, 2002

- [19] ZIELINSKI, B. et al. Deep learning approach to describing and classifying fungi microscopic images. 2020. Available at: <https://openreview.net/forum?id=qbNMvyepH>.
- [20] HAO, R. et al. Automatic detection of fungi in microscopic leucorrhea images based on convolutional neural network and morphological method. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). [S.l.:s.n.], 2019. p. 2491–2494.
- [21] PROMMAKHOT, A.; SRINONCHAT, J. Exploiting convolutional neural network for automatic fungus detection in microscope images. In: 2020 8th International Electrical Engineering Congress (iEECON). [S.l.: s.n.], 2020. p. 1–4
- [22] GAO, W. et al. The design and application of an automated microscope developed based on deep learning for fungal detection in dermatology. *Mycoses*, v. 64, n. 3, p. 245–251, 2021. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/myc.13209>
- [23] LIU, Z. et al. Automatic diagnosis of fungal keratitis using data augmentation and image fusion with deep convolutional neural network. *Computer Methods and Programs in Biomedicine*, v. 187, p. 105019, 2020. ISSN 0169-2607. Available at: <https://www.sciencedirect.com/science/article/pii/S0169260719308673>

- [24] GAIKWAD, S.; RUMMA, S.; HANGARGE, M. Classification of microscopic images of fungi using deep learning models. SAMRIDDHI A Journal of Physical Sciences Engineering and Technology, Vol 13, p. 105–110, 04 2022.
- [25] MARCOS, A. P.; RODOVALHO, N. L. S.; BACKES, A. R. Coffee leaf rust detection using convolutional neural network. In: 2019 XV Workshop de Visão Computacional (WVC) [S.l.: s.n.], 2019. p. 38–42
- [26] RAHMAN, H. et al. Iot enabled mushroom farm automation with machine learning to classify toxic mushrooms in bangladesh. Journal of Agriculture and Food Research, v. 7, p. 100267, 2022. ISSN 2666-1543. Available at: <https://www.sciencedirect.com/science/article/pii/S2666154321001691>
- [27] PRZYBYŁ, K. et al. Application of deep and machine learning using image analysis to detect fungal contamination of rapeseed. Sensors, v. 20, n. 24, 2020. ISSN 1424-8220. Available at: <https://www.mdpi.com/1424-8220/20/24/7305>.
- [28] BILLONES, R. K. C. et al. Aspergillus species fungi identification using microscopic scale images. In: 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). [S.l.: s.n.], 2020. p. 1–5.

- [29]TAHIR, M. W. et al. A fungus spores dataset and a convolutional neural network based approach for fungus detection. *IEEE Transactions on NanoBioscience*, v. 17, n. 3, p.281–290, 2018
- [30]DONG, J.-E. et al. Deep learning for species identification of bolete mushrooms with two-dimensional correlation spectral (2dcos) images. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, v. 249, p. 119211, 2021. ISSN 1386-1425. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1386142520311902>.
- [31]SOPO, C. J. P.; HAJATI, F.; GHEISARI, S. DeFungi: Direct Mycological Examination of Microscopic Fungi Images. arXiv, 2021. Disponível em: <https://arxiv.org/abs/2109.07322>
- [32]PICEK, L. et al. Automatic fungi recognition: Deep learning meets mycology. *Sensors*, v. 22, n. 2, 2022. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/22/2/633>
- [33]Drechsler-Santos, Elisandro Ricardo; Karstedt, Fernanda; Zimmermann Loureiro Chaves, Thiago; Titton, Mahatmã; Martins da Cunha, Kelmer; Leopoldo, Eloísa; Alves-Silva, Genivaldo; Rezende, Diogo; Kossmann, Thiago; Gumboski, Emerson Luiz; von Wangenheim, Aldo (2022), “MIND.Funga App: images dataset from Neotropical macrofungi used to train an artificial neural network to recognize fungal species”, Mendeley Data, V1, doi: 10.17632/sfrbdjvxcc.1