

# Redes Neurais Convolucionais para detecção do Cancro Europeu das Pomáceas

Ântony de Aguiar Salvi<sup>1</sup>, Paulo C Camargo Jr<sup>1</sup>,  
Wilson Castello Branco Neto<sup>1</sup>, Carlos Andres Ferrero<sup>1</sup>,  
Andrey de Aguiar Salvi<sup>2</sup>

<sup>1</sup>Instituto Federal de Santa Catarina (IFSC)  
Rua Heitor Vila Lobos, 225 - São Francisco - 88506-400 - Lages - SC - Brasil

<sup>2</sup>Lojas Renner S.A  
Rua Joaquim Pôrto Vilanova, 401 Jardim do Salso - 91410-400 - Porto Alegre - RS - Brasil

{antony.s, paulo.ccj}@aluno.ifsc.edu.br

{wilson.castello, andres.ferrero}@ifsc.edu.br

andrey.salvi@lojasrenner.com.br

**Abstract.** *European Apple Canker affects orchards from several countries and emerged in Brazil in Rio Grande do Sul spreading to other states that represent a large portion of the production of apple. The disease is a threat due to its rapid spread and for the difficulty to detect it. In order to solve this problem, a CNN was proposed for disease classification. The result in validation set for F-Beta was 0.795, obtained by VGG19, with 0.878 in precision and 0.591 in recall. The results show the possibility of using CNN for european canker detection, although they still can be improved.*

**Resumo.** *O Cancro Europeu das Pomáceas afeta os pomares de diversos países e surgiu no Brasil no Rio Grande do Sul espalhando-se em outros estados que representam grande parcela da produção de maçã. A doença é uma ameaça por sua rápida propagação e pela dificuldade na sua detecção. Para solucionar este problema foi proposta uma CNN para classificação da doença. O resultado no conjunto de validação para o F-Beta foi de 0.795, obtido pelo VGG19, com precisão de 0.878 e recall de 0.591. Os resultados demonstram a possibilidade de utilizar CNN para a detecção do cancro europeu, embora ainda haja espaço para melhorias.*

## 1. Introdução

O cancro europeu das pomáceas é uma doença que afeta os pomares de diversas regiões do mundo (Weber, 2014). Causada pelo fungo *Neonectria ditissima*, a doença representa uma grave ameaça se não combatida devido a sua alta capacidade de propagação e ao fato dos cultivos de macieiras estarem em locais muito próximos uns dos outros, o que facilita a sua disseminação, causando grandes prejuízos aos proprietários (Alves et al., 2019).

Para identificar a possível presença do fungo é necessária uma inspeção por fitopatologistas experientes através de uma amostra da planta. Esta tarefa, normalmente, é feita pelos fitopatologistas da Epagri ou da Embrapa, o que acaba

dificultando o diagnóstico, pois nem todos os agricultores se encontram em regiões próximas a estas empresas e a etapa de deslocamento torna-se um empecilho.

Para conscientizar os produtores, auxiliar na detecção da doença e acelerar o diagnóstico, foi criada a plataforma Cancontrol que é composta por um aplicativo móvel e um sistema web. Uma de suas funcionalidades presente nos dois sistemas, é uma seção composta por textos, imagens e vídeos sobre o cancro europeu. Há também um módulo de monitoramento no aplicativo móvel, por meio do qual o produtor pode cadastrar um monitoramento com dados sobre uma planta específica, além de realizar o *upload* de imagens para posterior análise de um especialista da Epagri. A disponibilização da plataforma em 2021 facilitou o diagnóstico da doença porque os produtores não precisam mais levar as amostras da planta presencialmente na Epagri, mas apesar de as imagens serem enviadas por meio do aplicativo do Cancontrol, a análise para determinar se há indícios da doença pode levar algumas semanas, pois ainda depende de análise humana (Branco Neto et al., 2021).

Visando reduzir o número de análises manuais e diminuir a sobrecarga dos fitopatologistas da Epagri, foi desenvolvida uma Rede Neural Convolutacional (CNN - *Convolutional Neural Network*) para processar as imagens e obter resultados automatizados. As CNN são um tipo de Rede Neural Artificial (ANN) com foco em processamento de dados, como imagens, vídeos e faixas de áudio, que ganharam notoriedade após o trabalho de Lecun et al. (1998). CNNs aplicam uma operação matemática chamada convolução, que é um tipo de operação linear (Goodfellow et al., 2016). Uma convolução consiste em aplicar uma série de filtros em uma pequena região de *pixels* da imagem, ao longo de toda a sua extensão. O resultado de uma convolução é chamado de mapa de características.

Este trabalho apresenta o resultado de experimentos realizados com diferentes arquiteturas de CNN pré-treinadas aplicadas ao diagnóstico do Cancro Europeu das Pomáceas. Ele está dividido em três etapas, a primeira etapa corresponde ao estudo de diferentes modelos de CNN. Na segunda etapa foi realizada a implementação dos modelos, os escolhidos foram o *VGG*, em suas versões de 16 e 19 camadas, o *ResNet*, nas versões de 18 e 50 camadas e o *InceptionV3*. Na terceira etapa foram testados os modelos e comparados os resultados obtidos. Dos modelos treinados, o que obteve os melhores resultados foi o VGG19, com um F-Beta de 0.7955, enquanto o VGG16 atingiu um F-Beta de 0.7629, ambos com learning rate  $1e-4$ .

O artigo é constituído de cinco seções. Após esta introdução, na Seção 2 são apresentados os trabalhos similares. Na Seção 3 é apresentada a implementação dos modelos escolhidos. Na Seção 4 são avaliados os resultados obtidos de cada modelo. Na Seção 5 são apresentadas as conclusões.

## 2. Trabalhos similares

Os trabalhos apresentados foram obtidos através de uma pesquisa no Google Scholar, utilizando a seguinte string de pesquisa: ("*Convolutional Neural Networks*" OR "*CNN*") AND ("*plant disease detection*" OR "*plant disease diagnosis*") publicados entre 2018 e 2023. Não foram encontrados artigos que visem detectar o cancro europeu, por este motivo, os trabalhos apresentados identificam outras doenças em diferentes tipos de culturas. Foram selecionados os 30 primeiros resultados obtidos

pela pesquisa e, em seguida, os *abstracts* dos trabalhos foram lidos e 10 deles foram selecionados para serem lidos completamente. Dos trabalhos lidos os mais relevantes são apresentados a seguir.

Liang et al. (2019) propuseram uma nova rede neural baseada na arquitetura do *ResNet50* para o diagnóstico de doenças e classificação de severidade chamada PD<sup>2</sup>SE-Net. A base de dados utilizada nos experimentos realizados pelos autores continha nove espécies de plantas, cada uma com vários tipos de doenças e de duas categorias de severidade para cada doença, com dados gerados de forma artificial. O modelo mostrou bons resultados nas métricas *loss* e acurácia em todas as épocas do treinamento, alcançando 99%, 98% e 91% na classificação das espécies das plantas, diagnóstico da doença e severidade, respectivamente. A acurácia em todos os diagnósticos de plantas ficou superior a 90%, exceto em uma das classes.

Sagar e Jacob (2021) realizaram um estudo para comparar cinco diferentes arquiteturas pré-treinadas de redes neurais convolucionais para detecção de 38 tipos de doenças em plantas. As arquiteturas avaliadas foram *MobileNet*, *ResNet50*, *InceptionV3*, *InceptionResNetV2* e *DenseNet169*. As imagens utilizadas pelos autores foram coletadas da base do *PlantVillage* e foram escolhidas 13 espécies de culturas, incluindo maçã, mirtilo, cereja, uva, laranja, pêssego, pimenta, batata, framboesa, soja, abóbora, morango e tomate. Foram utilizadas técnicas para visualização das camadas da rede a fim de acompanhar a aprendizagem do modelo e realizar os ajustes de parâmetros mais rapidamente. Após finalizar o treinamento dos modelos, os autores apresentaram os resultados obtidos comparando acurácia, Precisão, *Recall* e *F1-score*. O modelo que obteve os melhores resultados foi o *ResNet50*, alcançando 98% de acurácia, 94% de Precisão, 94% de *Recall* e 94% de *F1-score*.

É importante destacar que não é possível realizar uma comparação direta entre o atual trabalho e os trabalhos apresentados, por conta de variações consideráveis tanto na detecção da doença quanto nas imagens utilizadas para o treinamento dos modelos. Enquanto os trabalhos citados utilizaram bases de dados com grandes quantidades de imagens para treinar seus modelos, este trabalho usou uma quantidade aproximadamente 18 vezes menor do que o menor *dataset* utilizado dentro os trabalhos apresentados. Além disso, as imagens que foram utilizadas no atual trabalho foram produzidas por agricultores e agrônomos em ambientes diferentes e utilizando diferentes dispositivos, enquanto que os demais trabalhos possuíam, além de uma maior quantidade de imagens, maior qualidade e padronização. Além disso, este trabalho se concentra na detecção de uma doença que se manifesta principalmente no tronco e nos ramos das macieiras, enquanto os demais trabalhos, em sua maior parte, focam em detectar doenças através das folhas, onde as variações de cores causadas pelas doenças são mais aparentes e, portanto, mais facilmente percebidas. Já neste trabalho, as diferenças de cor causadas no tronco e nos ramos são muito mais sutis, tornando mais difícil a detecção.

### 3. Materiais e Métodos

Quanto à natureza deste trabalho, ele se classifica como uma pesquisa aplicada quantitativa quanto a sua abordagem, explicativa quanto aos seus objetivos e bibliográfica e experimental quanto aos seus procedimentos técnicos.

O trabalho foi desenvolvido em três etapas. A partir das imagens obtidas pelos fitopatologistas da Epagri e das imagens enviadas à plataforma Cancontrol através do *upload* pelos agricultores, deu-se início à primeira etapa. Nesta etapa, um dos objetivos foi realizar uma inspeção manual nas imagens, a fim de eliminar àquelas que não eram adequadas. Também foi feita a preparação das imagens, de forma que sua utilização permita o melhor desempenho possível da rede.

A segunda etapa consistiu em treinar uma rede neural para que ela fosse capaz de classificar corretamente quais imagens estão contaminadas com a presença do fungo *Neonectria ditissima*. A rede foi desenvolvida na linguagem de programação Python, com o uso da biblioteca *Pytorch*. Os dados estatísticos foram criados com a biblioteca *scikit-learn*, as métricas salvas foram Acurácia, Precisão, Recall, FBeta com valor  $\beta$  de 0,5, F-Score, Área Abaixo da Curva (AUC) e Erro. Neste artigo, as análises foram realizadas com base nos valores de FBeta, Precisão e Recall e a Acurácia é apresentada por ser uma medida de fácil interpretação. O FBeta com valor  $\beta$  de 0,5 é a métrica escolhida para avaliar a qualidade do modelo, pois um valor  $\beta < 1$  dá maior peso para a Precisão do que para o Recall, nesse caso a Precisão tem o dobro da importância no cálculo da métrica.

Por fim, na terceira etapa os resultados dos diferentes modelos testados foram comparados.

### 3.1. Obtenção e Preparação das Imagens

Foram obtidas 135 imagens, sendo 58 delas obtidas junto aos fitopatologistas da Epagri e as 77 restantes retiradas da base de dados da plataforma Cancontrol. Três imagens foram descartadas devido à natureza do problema. Como a doença se manifesta primeiramente em troncos e galhos, a maioria dos diagnósticos não são feitos analisando o fruto da árvore. Por este motivo, apenas uma pequena quantidade das imagens têm o fruto presente nelas. Do total de imagens, somente três delas possuíam somente o fruto, sem a árvore. Para simplificar o modelo, essas três imagens não foram utilizadas.

A Tabela 1 mostra a distribuição das imagens. Elas foram divididas de forma aleatória em 60% para treino, 20% para validação e os 20% restantes para teste.

Conjunto	Positivas	Negativas	Total
Treino	35	42	77
Validação	12	15	27
Teste	13	15	28
Total	60	72	132

**Tabela 1. Distribuição das imagens considerando os conjuntos e as classificações das imagens.**

As imagens de treino são usadas pelo modelo para calcular as derivadas e atualizar os pesos da rede. As imagens de validação são utilizadas para avaliar o modelo, calcular o erro e outras métricas que permitem inferir se o modelo está melhorando ou piorando seu desempenho. Os resultados obtidos na validação servem

como base para escolher o melhor modelo dado as peculiaridades do problema. Após o modelo ser escolhido, os dados de teste foram utilizados para analisar o desempenho de teste do modelo e para simular qual seria a assertividade do modelo caso fosse posto em produção.

### 3.2. Indução dos modelos de classificação

O treinamento da rede neural foi feito utilizando cinco modelos pré-treinados diferentes, disponíveis no *framework torchvision* (TorchVision, 2016). A escolha desse *framework*, se deve a sua maior flexibilidade na construção e manutenção de arquiteturas, se comparado a outros *frameworks*. Ele também conta com uma grande comunidade ativa, o que confere mais facilidade na busca por informações ou solução de problemas.

Os modelos são: *Resnet*, nas versões com 18 e 50 camadas; *VGG*, nas versões com 16 e 19 camadas; e *Inception* na versão v3. Os modelos pré-treinados atingem bons resultados em geral e são treinados para classificar diversos tipos de imagens diferentes. A partir desses modelos pré-treinados foi dado início ao treinamento do modelo deste trabalho.

Na primeira rodada de experimentos, nove combinações de parâmetros foram utilizadas em cada um dos cinco modelos, totalizando 45 experimentos. Para obter um grau de confiança maior, cada experimento foi executado dez vezes. Os resultados foram utilizados para analisar se os diferentes fatores de *learning rate* (LR), *weight decay* (WD) e arquiteturas influenciam no resultado final e se essa influência ocorre de forma isolada, ou se há interação entre os fatores. Os diferentes modelos e valores de LR e WD utilizados estão disponíveis na lista abaixo, sendo que os experimentos foram realizados de forma fatorial completa.

- Modelos: ResNet18, ResNet50, VGG16, VGG19 e InceptionV3;
- LR:  $1e^{-4}$ ,  $1e^{-5}$  e  $5e^{-5}$ ;
- WD:  $1e^{-7}$ ,  $1e^{-8}$  e  $1e^{-9}$ .

### 3.3. Avaliação dos Modelos

A matriz de confusão é a forma mais comum de visualizar os resultados de um modelo em um problema de classificação. Ela mostra o desempenho do modelo em cada classe, mostrando quantas entradas positivas foram classificadas corretamente (VP), quantas entradas positivas foram classificadas incorretamente (FN), quantas entradas negativas foram classificadas corretamente (VN) e quantas foram classificadas incorretamente (FP).

As métricas de avaliação de um sistema de classificação são calculadas a partir da matriz de confusão. Para avaliar e selecionar os melhores modelos, diversas métricas são salvas durante o treinamento do modelo, tais como, acurácia, precisão *recall*, *fscore* e curva ROC.

A Acurácia mede a relação entre todos os acertos e a amostra inteira. Na maioria dos casos, esta métrica isoladamente não é suficiente para medir a qualidade de uma modelo. Os conjuntos de dados podem ser desbalanceados - o número de ocorrências é desproporcional entre as classes - ou o problema pode exigir uma maior importância para a classe negativa, por exemplo.

A Precisão mede a porcentagem de acertos positivos, em relação a todos os resultados classificados como positivos. Esta medida é dada pela Equação

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (1)$$

O *Recall* mede a taxa de acertos positivos, em relação a todos os dados positivos, o que inclui os dados incorretamente classificados como negativos. Esta medida é definida pela Equação

$$\text{Recall} = \frac{VP}{VP + FN} \quad (2)$$

Normalmente as medidas de Precisão e *Recall* são antagônicas, dependendo do problema um modelo pode apresentar um baixo *Recall*, mas uma alta Precisão e vice-versa. A Precisão mostra o quanto um modelo acerta quando ele diz que uma entrada é positiva. O modelo pode ter uma alta taxa de acertos, mas fazer poucas predições. Em contrapartida, o *Recall* mostra quantas imagens ele deixou passar, isto é, quantos dados positivos foram classificados como negativos. Um modelo pode não deixar nenhum dado positivo para trás, simplesmente classificando todos como positivos, porém isso reduz a Precisão consideravelmente.

O *F-Score* é uma medida que leva em conta tanto a Precisão quanto o *Recall*.

$$F\text{-Score} = \frac{(1 + \beta^2) * \text{Precisão} * \text{Recall}}{\beta^2 * \text{Precisão} + \text{Recall}} \quad (3)$$

O *F-Score* é uma média harmônica entre a Precisão e o *Recall*. O valor  $\beta$  pode ser utilizado com valores diferentes de 1 para dar maior ênfase para o *Recall* ou para a Precisão, sendo  $\beta > 1$  mais ênfase para o *Recall* e  $\beta < 1$  mais importância para a Precisão.

Neste trabalho é utilizado o valor de  $\beta$  igual a 0.5 para o *F-Beta* a fim de dar mais importância para a Precisão já que dado o contexto do problema abordado ela é mais relevante do que o *Recall*, uma vez que os casos rotulados como saudáveis serão analisados por especialistas, com intuito de evitar os falsos negativos e permitir que uma árvore supostamente saudável contamine outras.

#### 4. Resultados

Conforme detalhado na Seção 3, os resultados apresentados nessa seção foram obtidos através do treinamento dos cinco modelos de CNN selecionados: *InceptionV3*, *VGG16*, *VGG19*, *ResNet* e *ResNet50*. As imagens utilizadas para treinamento foram separadas em três partes distintas, uma para treinamento com 60% das imagens, outra para a validação com 20%, e outra para teste com os 20% restantes. Além disso, foram realizados testes com combinações diferentes dos hiperparâmetros *learning rate* e *weight decay* a fim de encontrar a configuração que gerasse os melhores resultados.

Após a finalização de todos os experimentos foi realizada a análise de variância *ANOVA*, que pode ser visualizada na Tabela 2, com o objetivo de medir a influência de cada fator (Modelo de CNN, *learning rate* e *weight decay*) e suas interações nos resultados obtidos.

A coluna *df* representa os graus de liberdade de cada fator ou interação e a coluna *sum\_sq* a soma dos seus quadrados. Com estes valores é calculado o quadrado médio de cada fator ou interação (coluna *mean\_sq*) e a partir da divisão deste valor pelo quadrado médio dos resíduos e calculado o valor do teste F e o valor-p (coluna *PR(>F)*) que possibilitam identificar se o fator ou a interação em análise influenciam de forma significativa nos resultados.

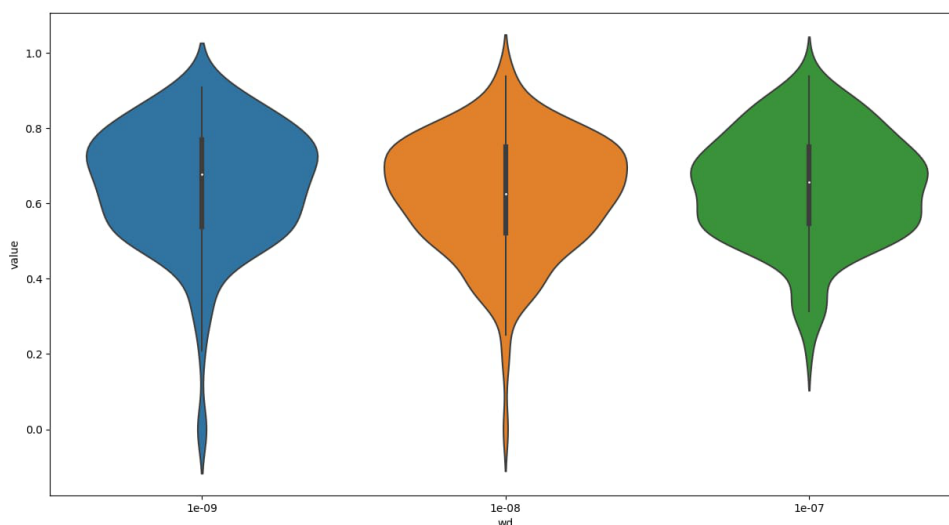
Ao realizar a inspeção dos resultados da análise de variância no conjunto de dados de validação, pode-se perceber que os diferentes modelos, diferentes valores de *learning rate* e as interações entre estes geraram uma influência nos resultados, como é possível observar na coluna *PR(> F)*, os quais atingiram  $9.0112318223e^{-24}$ ,  $4.3166784722e^{-18}$  e  $5.0261641571e^{-07}$  respectivamente, o que indica que tanto os fatores individuais quanto as interações entre eles foram significativas já que seus valores foram inferiores a 0.05. Por outro lado, tanto o fator individual do *weight decay*, quanto suas interações ficaram com valores acima de 0.05, o que indica que o *weight decay* não tem influência significativa nos resultados.

Factors	df	sum_sq	mean_sq	F	PR(>F)
C(model)	4.0	2.3793	0.5948	32.9081	9.0112e-24
C(lr)	2.0	1.5980	0.7990	44.2046	4.3166e-18
C(wd)	2.0	0.0540	0.0270	1.4964	2.2516e-01
C(model):C(lr)	8.0	0.8394	0.1049	5.8047	5.0261e-07
C(model):C(wd)	8.0	0.2736	0.0342	1.8921	5.9719e-02
C(lr):C(wd)	4.0	0.1345	0.0336	1.8611	1.1641e-01
C(model):C(lr):C(wd)	16.0	0.2470	0.0154	0.8543	6.2296e-01
Residual	405.0	7.3207	0.0180		

**Tabela 2. Análise de variância Anova para modelos, *learning rate* e *weight decay* no conjunto de validação.**

Por fim, foi gerado um gráfico de violino para observar a variação do *weight decay* para os valores de *F-Beta*, que pode ser visualizado na Figura 1, com intuito de selecionar apenas um dos valores testados para o *weight decay*, já que seus valores não apresentaram alterações significativas nos resultados. Ao analisar o gráfico pode-se notar que o valor  $1e^{-7}$  de *weight decay* conferiu uma maior estabilidade aos resultados, além disso este valor também gerou menos resultados mais baixos quando comparado com os demais valores de *weight decay*. Por isso o valor de  $1e^{-7}$  foi escolhido como parâmetro fixo para o *weight decay*.

A Tabela 3 apresenta a média dos resultados do teste com diferentes métricas dos experimentos realizados. É possível observar qual valor de *learning rate* levou cada modelo ao melhor resultado *F-Beta*, que foi considerada a métrica mais importante no contexto deste trabalho.



**Figura 1.** Gráfico de violino para *weight decay* com valores de *F-Beta* do conjunto de validação

Modelo	LR	Acurácia	Precisão	Recall	F-Beta
InceptionV3	$1e^{-4}$	0.6185	0.7498	0.4583	0.5718
InceptionV3	$1e^{-5}$	0.6037	0.6169	0.5333	0.5427
InceptionV3	$5e^{-5}$	0.5666	0.6265	0.45	0.4932
ResNet50	$1e^{-4}$	0.7259	0.7820	0.6	0.7224
ResNet50	$1e^{-5}$	0.6370	0.6223	0.5	0.5635
ResNet50	$5e^{-5}$	0.6259	0.6984	0.4166	0.5859
VGG16	$1e^{-4}$	0.7592	0.8356	0.5833	0.7629
VGG16	$1e^{-5}$	0.6518	0.6606	0.725	0.6299
VGG16	$5e^{-5}$	0.7296	0.8718	0.5583	0.7569
ResNet	$1e^{-4}$	0.6259	0.6019	0.6000	0.5917
ResNet	$1e^{-5}$	0.6333	0.6732	0.6416	0.6205
ResNet	$5e^{-5}$	0.6814	0.7003	0.575	0.6526
VGG19	$1e^{-4}$	0.7777	0.8788	0.5916	0.7955
VGG19	$1e^{-5}$	0.6518	0.7020	0.5	0.5924
VGG19	$5e^{-5}$	0.7407	0.8873	0.5083	0.7606

**Tabela 3.** Resultados dos modelos no conjunto de validação.

Os dois modelos com melhores resultados em seu F-Beta médio, foram o VGG16 e o VGG19. Em ambos os casos, a configuração com *learning rate*  $5e^{-5}$  trouxe melhores resultados na Precisão média, porém o *learning rate*  $1e^{-4}$  aumentou o Recall médio, o que fez com que o F-Beta médio dos modelos aumentasse nessa configuração. O melhor modelo, VGG19, atingiu um F-Beta de 0.7955, enquanto que o VGG16 atingiu um F-Beta de 0.7629, ambos com *learning rate*  $1e^{-4}$ .

De todas as combinações testadas, os experimentos que obtiveram o melhor F-Beta médio para cada modelo podem ser visualizados na Tabela 4, na qual são mostrados os menores e o maiores valores de F-Beta atingidos em cada modelo, bem



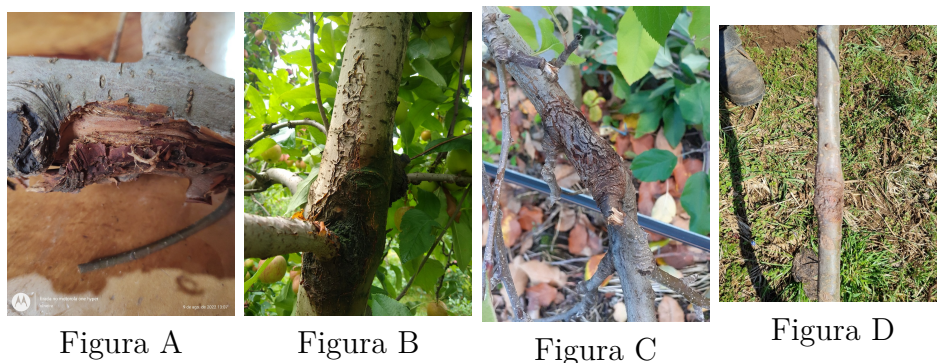
como a média do F-Beta nas 10 replicações.

Modelo	LR	F-Beta(Min)	F-Beta(Média)	F-Beta(Max)
InceptionV3	$1e^{-4}$	0.3125	0.5718	0.7291
ResNet50	$1e^{-4}$	0.5208	0.7224	0.875
VGG16	$1e^{-4}$	0.6730	0.7629	0.9090
ResNet	$5e^{-5}$	0.5357	0.6526	0.8653
VGG19	$1e^{-4}$	0.6730	0.7955	0.875

**Tabela 4. Resultados dos modelos no conjunto de validação.**

A partir dos resultados obtidos é possível observar que CNNs podem ser usadas para o diagnóstico do Cancro Europeu, com baixa taxa de falsos positivos, pois modelos como o VGG 19, obtiveram quase 90% de precisão. Como mencionado na Seção 3.3, neste trabalho priorizou-se a precisão em detrimento do Recall na escolha do modelo, por meio do valor de  $\beta$  igual a 0.5, pois ao ser integrado na plataforma Cancontrol, o sistema só apresentará automaticamente ao usuário o resultado gerado pela CNN caso ele seja positivo. Quando a CNN identificar que o resultado é negativo, as imagens serão enviadas aos fitopatologistas de Epagri para análise, seguindo o fluxo atual, e o produtor precisará aguardar o resultado. Portanto, os resultados falso negativos gerados pela CNN, neste momento, não representam um problema para a plataforma como um todo.

Uma hipótese que pode justificar a ausência de melhores resultados, como os obtidos nos trabalhos similares apresentados na Seção 2.5, é a natureza do problema. É comum utilizar um conjunto de dados onde a doença se encontra na folha ou no fruto da planta, o que gera uma maior variação de cores, por exemplo, facilitando a detecção. No caso do cancro europeu os sintomas estão no galho e no tronco e se manifestam de forma diferente na árvore, dependendo de quanto tempo a planta está infectada com a doença. Na Figura 2 é possível observar exemplos de falso positivo ao lado de um verdadeiro negativo e um verdadeiro positivo com falso negativo, respectivamente. Ao analisar estas imagens nota-se que as diferenças entre árvores saudáveis e árvores infectadas podem ser muito sutis, e portanto, dificilmente podem ser diferenciadas a olho nu.



**Figura 2. Galho sem a doença classificado como positivo (A). Galho sem a doença classificado corretamente (B). Galho com a doença classificado corretamente (C). Galho com a doença classificado como negativo (D).**

Fonte : Imagens retiradas da plataforma Cancontrol.

## 5. Conclusão

O cancro europeu é uma doença que vem se espalhando de forma acelerada nos pomares do sul do Brasil. No estado de Santa Catarina os produtores recorrem aos especialistas da Epagri para fazer o diagnóstico de uma amostra da planta. Devido ao alto índice de proliferação da doença, são muitos os casos a serem examinados pelos membros da Epagri, o que vem gerando sobrecarga nos funcionários.

Neste trabalho foram avaliadas experimentalmente diferentes arquiteturas de modelos de Redes Neurais Convolucionais pré-treinadas para o diagnóstico do Cancro Europeu das Pomáceas. Embora um número maior de imagens para análise seja importante para consolidar os achados, os resultados alcançados se mostram promissores e indicam que os modelos podem ser usados para auxiliar no diagnóstico da doença, podendo desempenhar um papel crucial para ajudar a combater e erradicar a doença.

Para consolidar esses resultados, trabalhos futuros incluem: a obtenção de um número maior de imagens para treinar e avaliar modelos apresentados; a utilização de novos modelos, como *Large Vision Models*; e o uso de técnicas de *Data Augmentation*. Também deve ser considerada a hipótese de treinar um modelo do zero, considerando que os modelos utilizados são pré-treinados no conjunto de dados ImageNet, composto de 1000 classes diferentes, envolvendo desde animais até a veículos. Por exemplo, o conjunto PlantVillage (Hughes e Salathe, 2016) contém apenas imagens de folhas de diferentes espécies de plantas e pode fornecer um conjunto de pesos mais adequado para o problema de detecção do cancro europeu.

## Referências

- Alves, S. A. M., Czermainski, A. B. C., Czermainski, A. B. C., et al. (2019). O cancro europeu no brasil. *Embrapa*.
- Branco Neto, W. C., Araújo, L., Pinto, F., Machado, R., Ribeiro, Y., Junior, W. C., e Mattos, K. (2021). Cancontrol: plataforma para diagnóstico do cancro europeu da macieira. In *Anais do XIII Congresso Brasileiro de Agroinformática*, pages 44–52, Porto Alegre, RS, Brasil. SBC.
- Goodfellow, I., Bengio, Y., e Courville, A. (2016). *Deep Learning*. MIT Press.
- Hughes, D. P. e Salathe, M. (2016). An open access repository of images on plant health to enable the development of mobile disease diagnostics.
- Lecun, Y., Bottou, L., Bengio, Y., e Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Liang, Q., Xiang, S., Hu, Y., Coppola, G., Zhang, D., e Sun, W. (2019). Pd2se-net: Computer-assisted plant disease diagnosis and severity estimation network. *Computers and Electronics in Agriculture*, 157:518–529.
- Sagar, A. e Jacob, D. (2021). On using transfer learning for plant disease detection. *bioRxiv*.
- TorchVision (2016). TorchVision: PyTorch’s Computer Vision library.
- Weber, R. (2014). Biology and control of the apple canker fungus *neonectria ditissima* (syn. n. *galligena*) from a northwestern european perspective. *Erwerbs-Obstbau*, 56:95–107.