

Segmentação de madeira utilizando visão computacional clássica e CNN

Felipe Alfredo Nack

Programa de Pós-Graduação em Engenharia de Automação e Sistemas

felipe.a.nack@gmail.com

Abstract

O objetivo deste trabalho é aplicar técnicas de visão computacional clássica e redes neurais convolucionais na segmentação de madeira em imagens. A segmentação por visão clássica é feita através de uma limiarização no espaço de cor HSV, seguida de operações morfológicas de fechamento, abertura e a aplicação do algoritmo de envoltório convexo. A aplicação da CNN trata de uma segmentação semântica seguindo a estrutura da U-Net em conjunto com o modelo Resnet34. Ambas abordagens apresentaram bons resultados. Escolheu-se a segmentação semântica da CNN em conjunto com um pós processamento envolvendo os algoritmos clássicos de visão computacional. Os códigos e datasets podem ser encontrados em <https://códigos.ufsc.br/felipe.nack/vision>

1. Introdução

O setor de painéis de madeira no Brasil apresenta grande relevância, produzindo em 2005 mais de 7,7 milhões de metros cúbicos de painéis. Esse valor é bastante expressivo, representando um crescimento de 9,5% ao ano desde 1995. No setor de painéis de madeira é comum classificar as lâminas utilizadas para a confecção dos painéis. No geral, essa classificação depende das propriedades da tora de madeira laminada.

Um dos desafios dessa indústria é a extração dessas características de forma automática, permitindo que as toras de madeira possam ser classificadas de acordo com seu potencial de gerar boas lâminas. Um exemplo de tora de madeira que pode ter suas características extraídas está apresentado na Figura 1.



Figure 1: Tronco de madeira recém cortado.

Contudo, é comum que esse tipo de imagem contenha ruído, como os elementos do fundo da Figura 1, dificultando sua análise computacional. Nesse sentido, o presente trabalho propõe resolver o problema dos ruídos presentes nestas imagens, através de técnicas de segmentação. Um possível resultado da segmentação da Figura 1 é aquele apresentado na Figura 2.

A partir dessa nova imagem é possível extrair características com maior facilidade como, por exemplo, a dimensão da face da tora e a quantidade de anéis de crescimento presentes.



Figure 2: Madeira segmentada na imagem.

2. Metodologia

Uma primeira abordagem possível é a utilização de técnicas disponíveis na visão computacional clássica. O

método aqui empregado faz inicialmente uma limiarização no espaço de cor HSV, buscando filtrar da melhor forma possível os pixels que possuem tons parecidos com o da madeira. É natural que esse tipo de abordagem deixe ruídos na imagem, que aqui são tratados através de operações morfológicas de fechamento e abertura, respectivamente. As operações de fechamento e abertura, quando realizadas em sequência com um kernel de mesmo tamanho, tem a propriedade de conservar o tamanho dos objetos e eliminar ruídos que ligam objetos diferentes. Se as operações de fechamento e abertura forem realizadas com sucesso, a imagem resultante deve ser parecida com a da Figura 3.

Na sequência basta realizar a análise de componentes conexos, selecionar o maior elemento presente na imagem e, então, aplicar um algoritmo de geração de envoltório convexo para reconstruir o formato da madeira. Perceba que essa sequência de processamento funciona bem para toras de madeira cuja face tem formato próximo ao de uma elipse. O resultado do algoritmo de envoltório convexo deve ser parecido com aquele visto na Figura 4. Perceba que o elemento maior em tons de cinza é justamente a madeira segmentada.

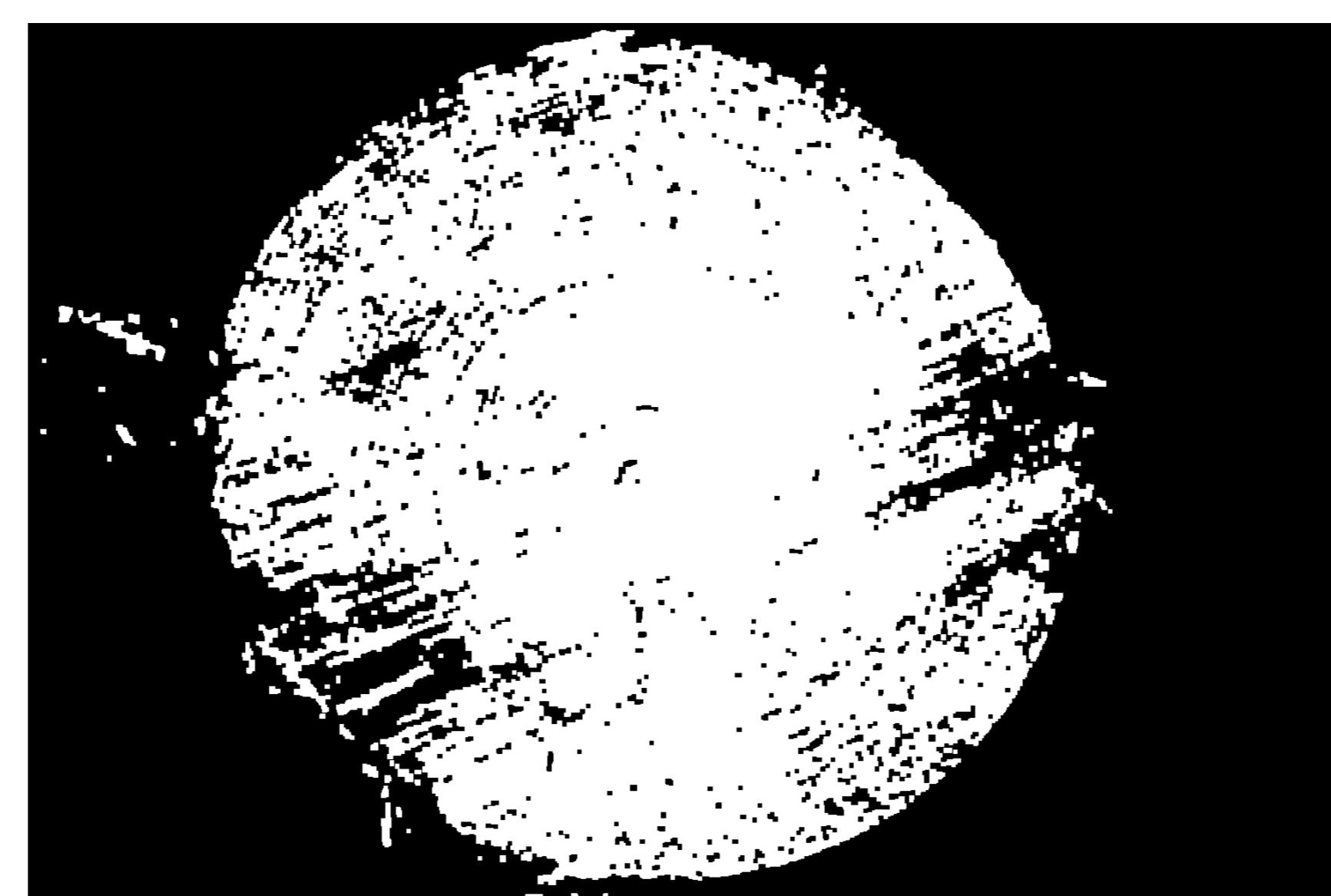


Figure 3: Operações morfológicas.

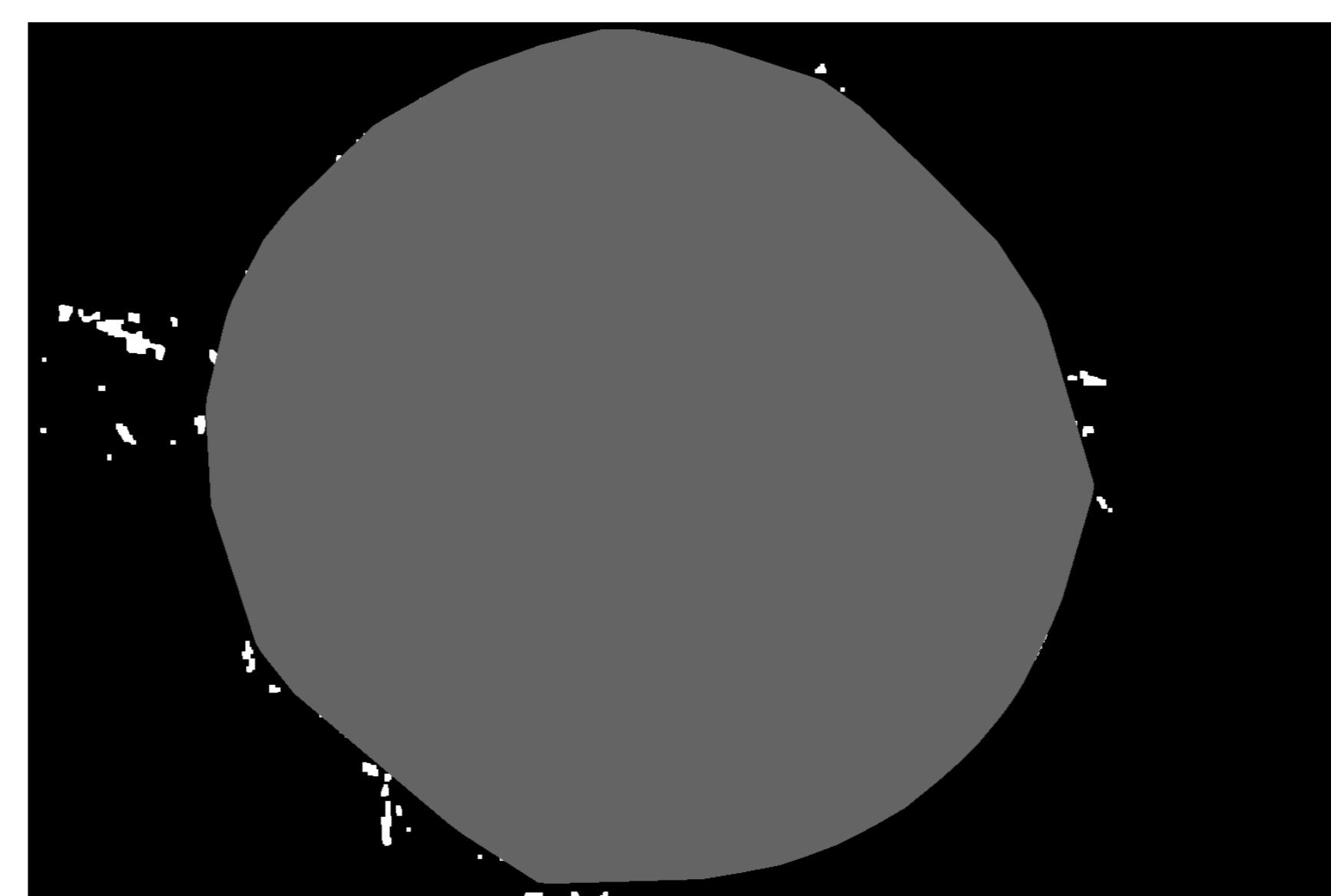


Figure 4: Algoritmo de envoltório convexo.

A segunda abordagem possível para o problema da segmentação é o uso de uma CNN. A arquitetura de CNN escolhida para esse trabalho é a da U-Net, utilizando-se do modelo Resnet34. O uso de ambos os elementos em conjunto permite realizar a segmentação de objetos em imagens. Conforme indica a Figura 5, a arquitetura da U-Net é um modelo de CNN ligado ao seu equivalente invertido, fazendo com que a saída da rede seja uma imagem de tamanho igual ao da imagem de entrada.

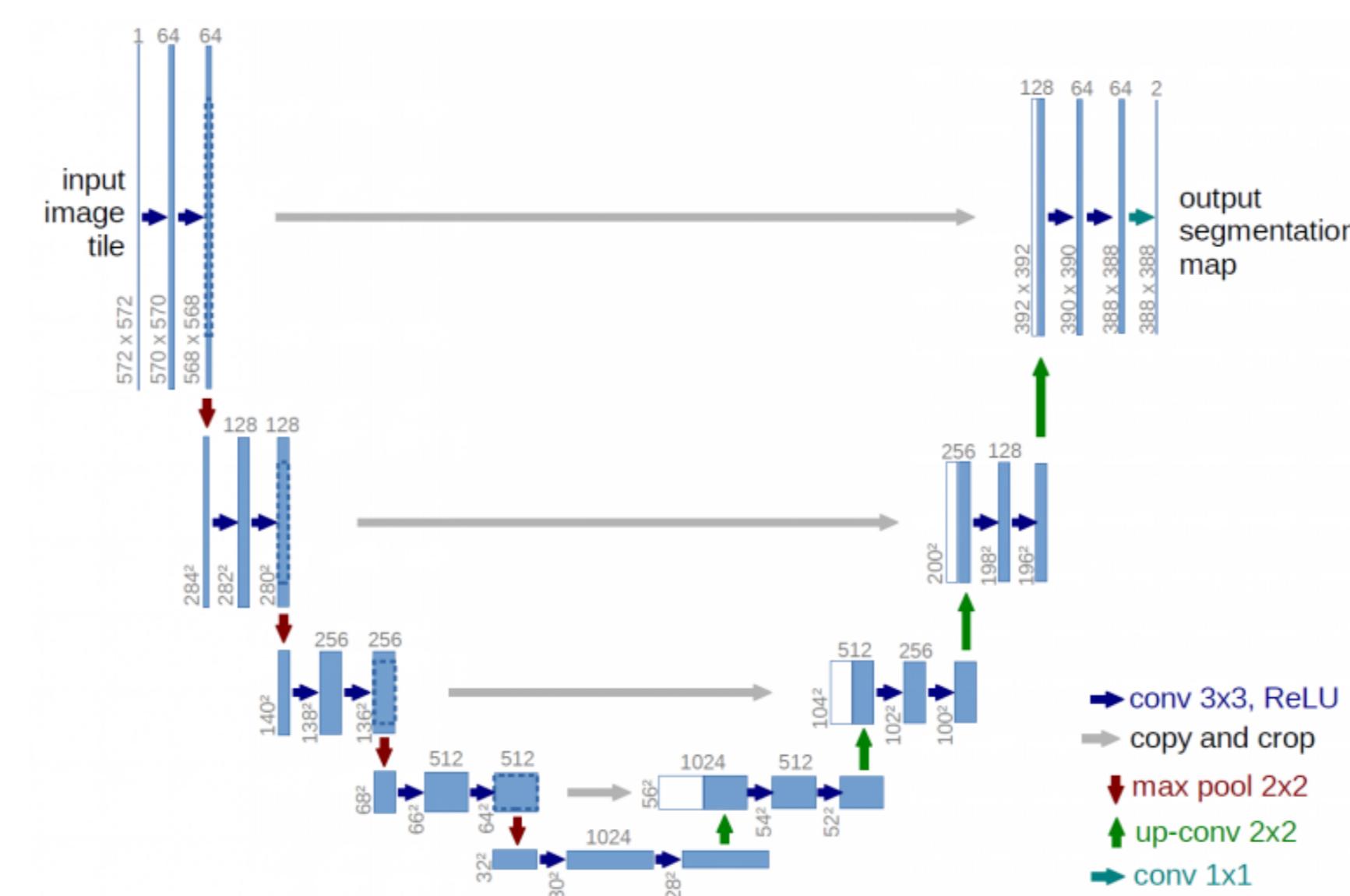


Figure 5: Arquitetura de uma U-Net.

A forma de treinamento escolhida é dividida em três partes:

- Treinar a rede com uma resolução 4 vezes menor;
- Treinar a rede com uma resolução 2 vezes menor;
- Treinar a rede com a resolução original.

Esse treinamento permite alcançar bons resultados com menor tempo de processamento. A métrica utilizada no treinamento é conhecida como IoU (*Intersection over Union*).

3. Resultados e Discussão

O algoritmo de visão computacional clássica se mostrou suficiente para o processo de segmentação, contudo, apresenta limitações. Um dos resultados obtidos pelo algoritmo clássico é a própria Figura 2. Apesar do bom resultado, o algoritmo é altamente dependente da cor da madeira, necessitando que em condições de luminosidade adversa ajustes finos precisem ser feitos manualmente. Essa solução é específica e também não é capaz de realizar a segmentação para espécies de madeira muito diferentes. A CNN, por outro lado, obteve resultados semelhantes. O resultado da inferência da CNN quando a entrada é a Figura 1 pode ser visto na Figura 6. Esse resultado pode ser facilmente melhorado aplicando o mesmo processamento feito no algoritmo clássico.

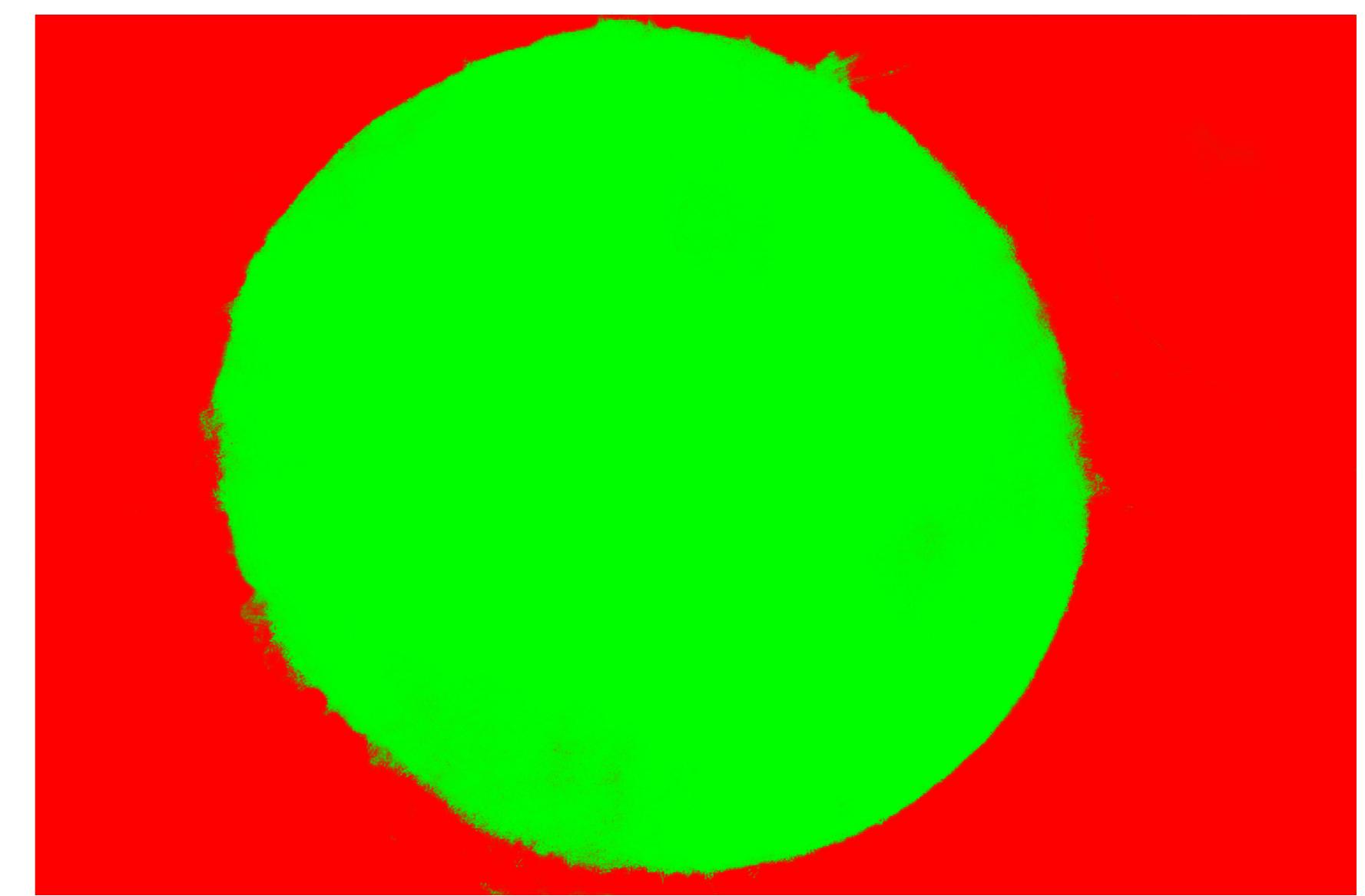


Figure 6: Resultado da CNN.

É importante ressaltar que os resultados da CNN foram obtidos com poucas épocas de treinamento. Para o último treinamento, realizado com a resolução original das imagens, obtiveram-se as seguintes métricas:

Table 1: Métricas de treinamento da CNN.

epoch	train loss	IoU metric	time (min)
0	0.169930	0.988759	1:39
1	0.141255	0.992406	1:37
...
8	0.042674	0.983961	1:37
9	0.036680	0.984380	1:37

O conjunto de treinamento utilizado é composto por 30 imagens distintas, sem o uso de *data augmentation*. O treinamento de mostrou rápido e apresentou boa métrica IoU.

4. Conclusão

Em geral ambas as abordagens apresentaram bons resultados. Contudo, a CNN se mostrou mais robusta para trabalhar com condições adversas e com madeiras de diferentes espécies. Essa robustez surge justamente porque o banco de imagens utilizado possui diversos exemplos de madeiras. Dentro do mesmo banco de imagens o algoritmo clássico não foi capaz de segmentar corretamente os diversos exemplos, necessitando de ajustes finos para cada espécie de madeira.

Os resultados apontam que o algoritmo clássico tem dificuldades em segmentar diferentes espécies de madeira sem ajustes finos, já a CNN é capaz de segmentar facilmente dentre as espécies, porém, deixa alguns ruídos na imagem. A solução ideal então é realizar a segmentação através da CNN e, dada essa saída, aplicar as operações clássicas de morfologia e envoltório convexo para obter o resultado final.