

Segmentação e classificação de defeitos em chapas de aço

Felipe Sassi Del Sant¹

¹Mestrado em engenharia elétrica

Introdução

Segmentação semântica de imagens é um tema muito discutido atualmente dentro da área de visão computacional. Diversos exemplos podem ser mencionados, como carros autônomos que usam a segmentação para diferenciar carros de pedestres, sistemas de identificação de tumores e sistemas de detecção de plantas daninhas em lavouras. Com esses exemplos pode-se verificar a grande gama de aplicações dessa área. Nesse trabalho foram aplicados métodos de segmentação semântica para detecção de defeitos em chapas de aço. Quatro diferentes tipos de falha podem aparecer na mesma imagem. Para ilustrar, um exemplo é apresentado na figura 1. Duas técnicas serão exploradas, uma baseada em visão computacional clássica e outra baseada em redes neurais convolucionais, para que uma comparação entre os métodos possa ser realizada.



Figura 1: Chapa de aço com seus defeitos destacados.

Métodos

O primeiro método testado para a segmentação semântica foi baseado em visão computacional clássica. Isso significa que todas as características extraídas das imagens foram desenvolvidas manualmente e após extraídas essas características o conjunto de dados resultante foi aplicado em um método clássico de aprendizagem de máquina. A figura abaixo apresenta a *pipeline* desenvolvida para essa tarefa [1]. Para treinamento foram utilizadas 8 imagens e para teste foram utilizadas 16 imagens. Essa quantidade de imagens foi utilizada pois a classificação é realizada pixel a pixel, isso faz com que o conjunto de treinamento tenha uma grande quantidade de dados.

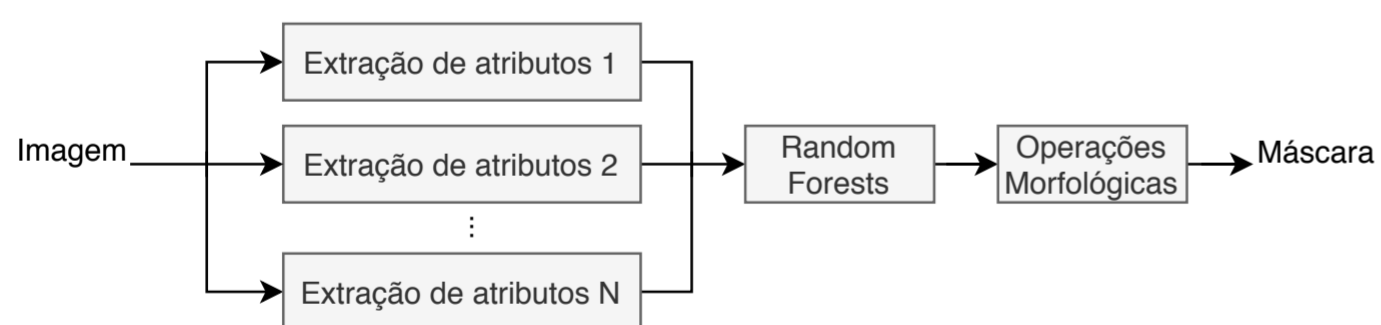


Figura 2: *Pipeline* desenvolvida utilizando visão computacional clássica.

O segundo método abordado foi baseado em redes neurais convolucionais. Foi utilizada uma arquitetura U-Net a qual é amplamente utilizada em problemas desse tipo [2]. No caso das redes neurais convolucionais, nenhuma engenharia de atributos precisa ser realizada, visto que a rede neural aprende todas as informações relevantes para cumprir a tarefa adequadamente. Foram utilizadas 13120 imagens para treinamento e 4373 imagens para validação/teste. A figura 3 apresenta a arquitetura utilizada para a rede neural.

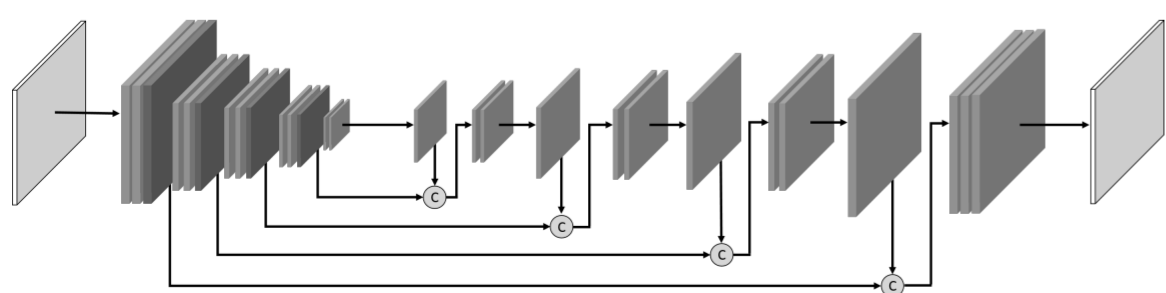


Figura 3: Arquitetura U-Net utilizada para segmentação semântica.

Ambos os métodos foram avaliados pela métrica conhecida como *Dice coefficient*, a qual é representada pela equação 1:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

Resultados

Nas imagens abaixo as regiões em vermelho correspondem as falhas reais, as em azul as falhas previstas e em rosa a intersecção entre elas. Os resultados obtidos utilizando visão computacional clássica foram os seguintes:

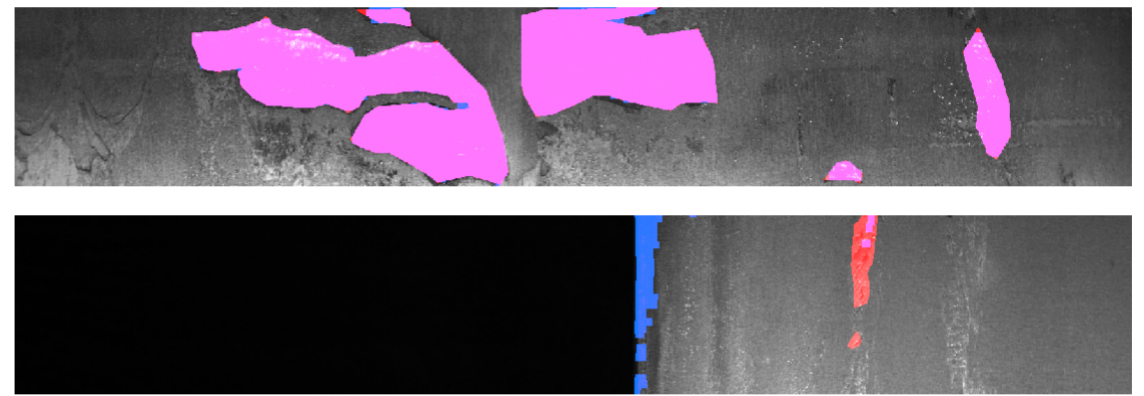


Figura 4: Resultados obtidos por meio de técnicas de visão computacional clássica (primeira imagem corresponde ao conjunto de treinamento e a segunda imagem ao conjunto de teste).

Além disso pode-se verificar o resultado para a métrica utilizada:

Treinamento	Validação
0,975	0,150

Nota-se claramente a ocorrência de *overfitting* com a utilização de técnicas clássicas de visão computacional.

Já com redes neurais convolucionais se obteve os seguintes resultados:

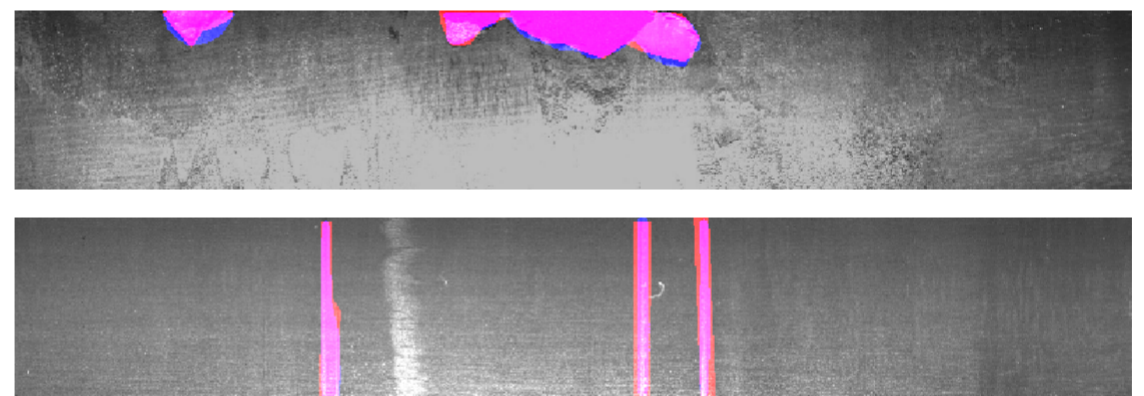


Figura 5: Resultados obtidos por meio de técnicas de visão computacional moderna (primeira imagem corresponde ao conjunto de treinamento e a segunda imagem ao conjunto de teste).

A tabela abaixo apresenta os resultados da métrica de avaliação. Nota-se a consistência entre os resultados no conjunto de treinamento e no conjunto de validação.

Treinamento	Validação
0,825	0,780

Conclusões

Com a análise dos resultados fica claro as vantagens proporcionadas pela utilização de redes neurais convolucionais para a resolução desse problema. Duas justificativas a isso devem ser frisadas:

- Redes neurais convolucionais não precisam de engenharia de atributos para extração de características relevantes nas imagens;
- Redes neurais convolucionais, nesse caso, tem menos chances de sofrerem de *overfitting*.

Referências

- SHOTTON, Jamie; JOHNSON, Matthew; CIPOLLA, Roberto. Semantic texton forests for image categorization and segmentation. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2008. p. 1-8.
- RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015. p. 234-241.