Training Data Filtering for Deep Learning Applied to Inspection of Welded Joints in Oil Pipelines

Rafael Rosa da Silva[1], **Fernando Moreira Suyama**[1]
Ricardo Dutra da Silva[1],**Myriam Regattieri Delgado**[1],

[1] Universidade Tecnológica Federal do Paraná (UTFPR) – Curitiba, PR – Brasil

rafaelsilva.2000@alunos.utfpr.edu.br, rdsilva@utfpr.edu.br,
fernando@suyama.com.br, myriamdelg@gmail.com

Abstract. *Inspection of weld radiographs is an essential task for preventing leaks in oil industry pipelines. Automatic inspection is important to assist specialists whose performance can be influenced by fatigue, visual acuity and experience. A challenging issue in automatic inspection is the availability of laudable images that can be used for training classifiers. Data augmentation is an alternative, however it can produce images with little or even negative impact on the training. This article presents a method for filtering images, aiming to favor the classifier relearning process, which is done by eliminating images with low performance from the training set and further retraining the segmentation model, with a positive impact for its generalization capabilities. Experimental results of the method show that data filtering accelerates the learning process and produces an improvement in the segmentation performance.*

1. Introdução

A inspeção radiográfica de juntas soldadas em tubulações de petróleo é essencial para prevenir possíveis vazamentos, impactando em diferentes setores: econômico, ambiental e da saúde [Rana et al. 2008, Atlas e Hazen 2011]. Tal atividade é comumente realizada por inspetores de radiografia industrial que analisam radiografias de juntas soldadas para identificar defeitos como porosidades e trincas [Davis 1989]. Estes inspetores avaliam e interpretam ensaios radiográficos de juntas soldadas na forma de laudos de acordo com os critérios de aceitação baseados em normas e especificações técnicas [NBR 2015].

No Brasil, tais critérios são estabelecidos pelo Comitê Brasileiro de Soldagem (ABNT/CB-042) da Associação Brasileira de Normas Técnicas (ABNT), cujas diretrizes normalizam o campo da soldagem e técnicas afins. Apesar das normas e especificações técnicas, a interpretação do inspetor laudista pode variar em decorrência da repetitividade e da reprodutibilidade do procedimento de inspeção radiográfica como a fadiga, a acuidade visual e a experiência [Fucsok et al. 2002].

Diante da relevância deste tema para os setores supracitados e pelas subjetividades envolvidas no processo de interpretação de juntas soldadas em radiografias por especialistas humanos, a criação de métodos auxiliares e de automatização dessa atividade tem sido foco da comunidade científica [Liao 2003, Nacereddine e Tridi 2005, Baniukiewicz 2014, Wang et al. 2019, Duan et al. 2019]. Assim, esta pesquisa trata de um subcampo de aprendizado de máquina em voga nos anos recentes: o aprendizado profundo, atuando na identificação de defeitos de soldagem de tubulações de petróleo.

Especificamente, o presente trabalho explora uma abordagem de detecção de defeitos com base na rede U-Net [Ronneberger et al. 2015] que, por se tratar de uma arquitetura profunda, não necessita do *design* manual da extração de características relevantes ao problema e permite realizar a segmentação semântica que localiza os defeitos pixel a pixel sobre a radiografia. Tal refinamento se faz necessário, tendo em vista que há uma grande quantidade de defeitos de tamanho reduzido nas bases de dados deste estudo.

Neste artigo, apresenta-se uma proposta de filtragem da base de dados utilizada no treinamento da U-Net, base esta composta por imagens radiográficas de juntas soldadas de tubulações de petróleo. O processo de filtragem remove, ao final do treinamento, imagens dessa base considerando-se uma métrica de desempenho usual para problemas de segmentação semântica, o IoU (*Intersection over Union*). O objetivo é eliminar imagens com baixo valor de IoU que podem contribuir negativamente (*outliers*) para a segmentação de defeitos de soldagem.

As motivações para desenvolver essa filtragem são as propriedades das radiografias presentes nas base de dados deste estudo que são obtidas a partir de ambientes reais de operação com variados níveis de ruído, luminosidade e contraste. Vale destacar que o intuito deste estudo não é comparar arquiteturas profundas nem propor um método estado-da-arte para a aplicação, mas sim atuar na preparação de uma base de dados, tendo em vista a complexidade das radiografias presentes nas bases de dados em análise e as subjetividades associadas aos laudos utilizados.

O artigo está estruturado da seguinte forma. Após esta seção introdutória, a Seção 2 detalha os trabalhos relacionados. A Seção 3 descreve os materiais e métodos usados, incluindo a elaboração da base de dados e o detalhamento do método de filtragem que está sendo proposto neste artigo. Além disso, a Seção 4 traz os resultados da filtragem e na Seção 5 apresentam-se as conclusões e trabalhos futuros.

2. Trabalhos Relacionados

Na área de pesquisa relacionada ao suporte e à automação da inspeção radiográfica de juntas soldadas, a maior parte da literatura emprega técnicas rasas de aprendizado de máquina (como Sistemas *Fuzzy* [Baniukiewicz 2014, Liao 2003], redes *Multilayer Perceptron* [Zapata et al. 2012, Nacereddine e Tridi 2005] e *Support Vector Machine* [Duan et al. 2019]). Tais técnicas necessitam do *design* manual dos atributos relevantes ao problema, resultando em abordagens, geralmente, com alta dependência de parâmetros manuais. Além disso, são altamente dependentes de métodos específicos de processamento digital de imagens (definidos também de forma manual) para identificar regiões e depois classificá-las.

Em contraste, as abordagens profundas extraem os atributos por meio de filtros aprendidos em camadas de convolução. As primeiras abordagens profundas na detecção de defeitos de soldagem classificam janelas recortadas sobre a radiografia, verificando a presença ou não de defeitos nestas janelas [Mery e Arteta 2017, Liu et al. 2017]. Geralmente, estes métodos usam redes pré-treinadas (com filtros aprendidos para cenários distintos da aplicação atual, abordagem conhecida como *transfer-learning*) com diferentes *backbones* como: Alexnet [Krizhevsky et al. 2012], VGG [Simonyan e Zisserman 2014], GoogLeNet [Szegedy et al. 2015] e ResNet [He et al. 2016].

Outras técnicas profundas tratam também desse problema de detecção de defei-

tos com redes pré-treinadas, porém, por meio dos *feature maps* (resultados de camadas intermediárias da rede) para realizar localização de regiões de defeitos [Lin et al. 2017, Wang et al. 2019]. Esta técnica é conhecida como detecção de objetos, pois destaca as regiões aproximadas dos defeitos de soldagem sobre a radiografia com retângulos englobantes. Wang et al. (2019) aplicam a rede RetinaNet [Lin et al. 2017] para localizar os defeitos de soldagem com apenas um estágio (os objetos são localizados como uma tarefa de regressão), empregando como *backbone* a rede Resnet-50 [He et al. 2016].

Diferente dos trabalhos anteriores citados, este artigo apresenta uma abordagem baseada na rede profunda de segmentação semântica U-Net [Ronneberger et al. 2015], que permite obter uma localização com maior grau de refinamento, pois os defeitos de soldagem são detectados em nível de pixel nas imagens radiográficas.

3. Materiais e Métodos

Nesta seção, serão descritas a base de dados, a criação das máscaras de segmentação de defeitos, a técnica de *Data Augmentation* utilizada, o método de filtragem que está sendo proposto e sua aplicação no processo de treino e retreino.

3.1. Descrição da Base de Dados

Neste artigo são consideradas 85 imagens radiográficas obtidas em campo (imagens reais), cujas características variam de acordo com o fornecedor do equipamento: 72 foram obtidas por meio de um equipamento do fornecedor A e 13 obtidas pelo equipamento do fornecedor B. Conforme ilustra a Figura 1, a diferença da origem dos equipamentos impacta as radiografias produzidas gerando disparidades, como alterações no contraste e visibilidade do cordão de solda. As imagens possuem 8 bits/pixel (256 níveis de cinza), formato *.tif*, alturas que variam de 253 a 1076 e larguras entre 695 e 1805.

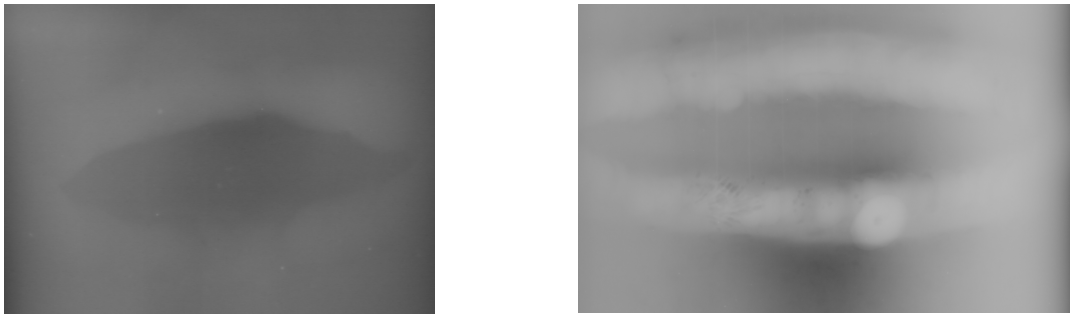


Imagem do Equipamento A (a)

Imagem do Equipamento B (b)

Figura 1. Exemplo das diferenças entre imagens radiográficas obtidas pelos equipamentos dos fornecedores A e B.

3.1.1. Criação das Máscaras de Segmentação

As máscaras de segmentação (gabaritos) foram criadas manualmente a partir de laudos de defeitos disponibilizados por especialistas. A Figura 2(a) mostra um exemplo de laudo

no qual as regiões destacadas pelo especialista estão em amarelo e contêm defeitos como porosidade (PO) e falta de fusão (FF). Observa-se na Figura 2(b) que, nas máscaras, as regiões de defeitos são caracterizadas por pixels com valor 1 (brancos) e as regiões sem defeito são caracterizadas por pixels com valor 0 (pretos).

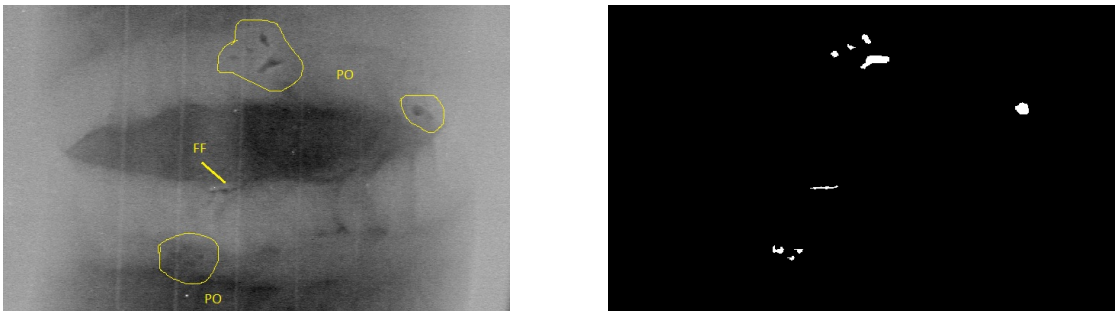


Imagem com Laudo

(a)

Máscara de Segmentação

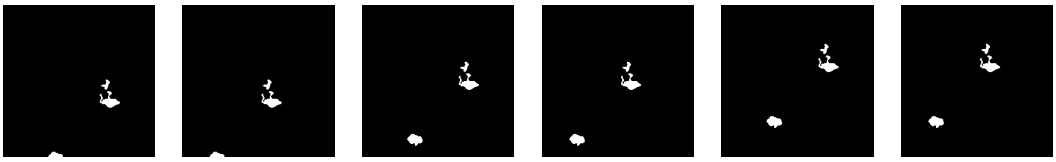
(b)

Figura 2. Imagem laudada e a respectiva máscara de segmentação criada.

É importante ressaltar que há possibilidade de imprecisões em relação às bordas nas máscaras definidas no gabarito, visto que, nos laudos, as bordas não são claramente definidas e os limites das máscaras estão sujeitos à subjetividade, podendo variar entre os especialistas em laudos de defeitos em tubulações de petróleo.

3.1.2. Geração da Base de Dados

Para a composição da base de dados, a partir do total de 85 imagens, foram escolhidas 68 para a base de treinamento (80% do total) e 17 para a base de teste (20% do total). Entretanto, essas 85 imagens possuem dimensões (altura e largura) diferentes entre si. Buscando padronizar o tamanho e disponibilizar mais exemplos na base de dados, para cada imagem foi feito o recorte de subimagens de 256 $\times$ 256 pixels por meio de uma janela deslizante com um *stride* (passo) de 30 pixels, conforme ilustrado na Figura 3. Após esses recortes, foram obtidas 7276 imagens, sendo 5317 pertencentes à base de treinamento (73% do total) e 1959 à base de teste (27% do total).



Janela 1 *Janela 2* *Janela 3* *Janela 4* *Janela 5* *Janela 6*

Figura 3. Representação do recorte de uma imagem em janelas de 256 $\times$ 256 pixels com *stride* de 30 pixels.

3.2. Técnica de Aumento de Dados

As técnicas de aumento de dados (*Data Augmentation*) permitem que se aumente significativamente a diversidade de dados disponíveis para modelos de treinamento em redes

de aprendizagem profunda, sem realmente coletar novos dados. Algumas técnicas usadas para treinar grandes redes neurais envolvem recortes de imagens para criar um conjunto maior de exemplos. Segundo [Shorten e Khoshgoftaar 2019], os algoritmos para aumento de dados mais comuns envolvem ainda transformações geométricas, aumento do espaço de cores e/ou características, aplicação de filtros, mistura de imagens, apagamento aleatório, treinamento de redes adversárias e generativas, e até meta-aprendizado.

3.3. Métrica para Filtragem

A métrica usada para avaliar o desempenho das imagens, com a finalidade de filtrar as piores imagens para o treinamento, é o IoU (*Intersection over Union*), que é, conforme a Figura 4, a área da intersecção dividida pela área da união de duas imagens, sendo elas, nesse caso, a imagem de saída da rede (s) e seu gabarito (g).

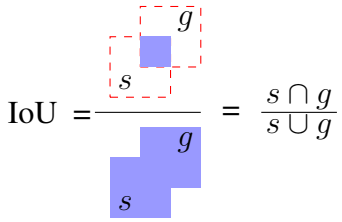


Figura 4. Métrica da filtragem: IoU.

3.4. Processo de Filtragem

O processo de filtragem proposto neste trabalho consiste em se treinar a rede com a base completa (BC) considerando-se um número E_{TC} de épocas e ao final escolher as imagens de pior desempenho para serem eliminadas. O objetivo é retirar de BC as imagens irrelevantes e/ou prejudiciais, para que um novo treinamento (retreinamento) da rede possa ser feito, agora utilizando-se a base filtrada (BF). Espera-se que este retreinamento produza melhores resultados sobre os dados de teste (BT) quando comparado com aqueles produzidos pelo treinamento usando a base completa.

A escolha das imagens a serem filtradas é feita com base em um limiar L relativo à métrica definida para a segmentação semântica (neste caso o IoU): ao final do treino completo, as imagens cujo IoU nos dados de treinamento da base completa (BC) ficam abaixo desse limiar são eliminadas para se formar uma base menor, a filtrada (BF). Para avaliar as imagens com base em seu desempenho de IoU, foi calculada a média do IoU para cada imagem dentro de um intervalo de épocas, com início E_i (época inicial) e fim E_f (época final), sendo esses limites menores do que o total de épocas do treinamento completo realizado (E_{TC}). Subsequentemente, a fim de gerar BF_L, ou seja, a base de treinamento filtrada para o limiar L, foram removidas da base de treinamento completa (BC) as imagens com IoU médio inferior a L.

O processo de filtragem para cada imagem $I_j \in BC$, com base filtrada inicial vazia ($BF_L = \emptyset$), pode ser formalizado como a seguir

$$BF_L = \begin{cases} \{BF_L\} \cup \{I_j\} & \text{se } I_j \in BC \text{ e } \overline{IoU_j}(E_i, E_f) \geq L, \quad j = 1, \cdots, |BC| \\ \{BF_L\} & \text{caso contrário} \end{cases} \tag{1}$$

onde, assumindo-se $E_i < E_f \leq E_{TC}$, o IoU médio da imagem I_j no intervalo de épocas $[E_i, E_f]$ é dado por

$$\overline{IoU}_j(E_i, E_f) = \frac{\sum_{k=i}^{f} IoU_j(E_k)}{(E_f - E_i + 1)} \tag{2}$$

e $IoU_j(E_k)$ representa o valor de IoU para I_j na época E_k.

Para se analisar o efeito qualitativo da filtragem, as imagens retiradas foram categorizadas de acordo com a localização de suas regiões de defeitos em relação a uma borda de b pixels a partir dos limites da imagem de dimensão $d \times d$. A Figura 5 ilustra essa divisão, na qual as regiões em branco correspondem às regiões em que há defeitos e as regiões em preto, às regiões onde não há defeitos. As categorias de gabaritos e/ou imagens de saída da U-NET consideradas no trabalho são *borda*, quando os defeitos se situam exclusivamente dentro da borda de b pixels; *meio*, quando há defeitos somente fora da borda; e *misto*, quando existem defeitos na borda e no meio.

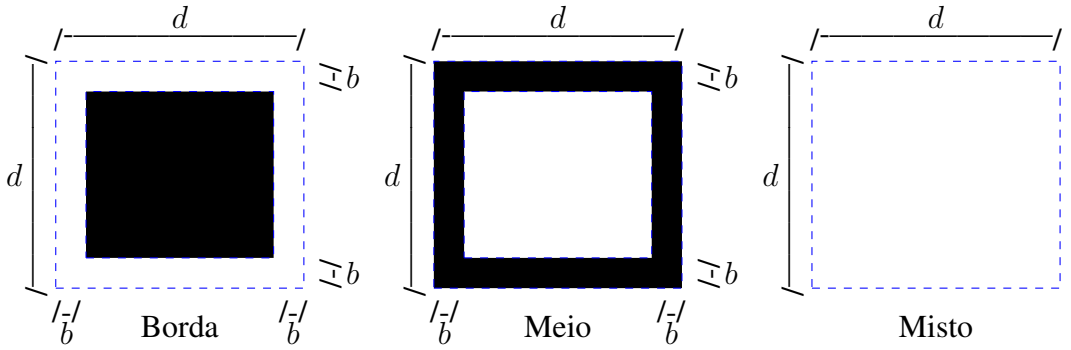


Figura 5. Ilustração das categorias de borda, meio e misto.

3.5. Aprendizado com BC e Reaprendizado com BF

Conforme discutido na seção anterior, para filtrar a base de treino, faz-se necessário identificar quais imagens devem ser removidas de BC após um treinamento completo e quais são mantidas em BF para que se possa realizar o retreinamento da rede. Assim sendo, realiza-se um treinamento na rede com E_{TC} épocas, utilizando-se como base completa de treinamento (BC), por exemplo, a base de dados especificada na Seção 3.1.2. Após o treinamento em E_{TC} épocas, as piores imagens de BC são filtradas e um conjunto de bases filtradas $BF = \{BF_{L_1}, BF_{L_2}, \ldots BF_{L_n}\}$ é obtido, onde cada base filtrada BF_{L_i} é definida de acordo com o limiar L_i. Após a filtragem, realiza-se novamente o treinamento da rede, utilizando-se uma determinada base de treino filtrada (BF_{L_i}), tendo em comum os pesos iniciais da rede usados no treinamento completo.

Salienta-se que, apenas para visualizar os desempenhos das redes ao longo da evolução, podem ser executadas avaliações pontuais com a base B a cada T_t épocas, e o IoU médio de todas as imagens é calculado como:

$$IoU_{rede}(E_t) = \frac{1}{|B|} \sum_{j=1}^{|B|} IoU_j(E_t) \tag{3}$$

onde $IoU_j(E_t)$ é o IoU da imagem I_j calculado na época E_t.

Com os dados obtidos, constroem-se gráficos da performance da rede ao longo das épocas e, a partir destes, gráficos suavizados. Assim, após o treinamento completo ou filtrado por E_T épocas, e avaliações pontuais ao longo da evolução a cada T_t épocas, tem-se uma matriz M formada por E_T/T_t vetores, cada um com dimensão igual à cardinalidade ($|B|$) da base usada na avaliação e contendo as medidas de IoU para todas as imagens dessa base considerada. Portanto, $M_B = \{\mathbf{m}_1, \mathbf{m}_2, ..., \mathbf{m}_{E_T/T_t}\}$, onde $\mathbf{m}_t$ é dado por, $\mathbf{m}_t = (m_{t1}, \cdots, m_{|B|})^T$. Para os gráficos não-suaves, tem-se $m_{tj} = IoU_j(E_t)$, ou seja, IoU calculado para cada imagem I_j no instante da visualização (E_t). A suavização implementada consiste em uma média móvel sobre as últimas K medidas de IoU_j. Portanto, $m_{tj} = \overline{IoU}_j(E_t - K, E_t)$, com $\overline{IoU}_j(E_t - K, E_t)$ definido pela Eq. 2. Em outras palavras, a partir da época atual E_t calcula-se a média dos valores de IoU para a imagem I_j nas últimas K épocas e esse valor define o elemento m_{tj} do vetor $\mathbf{m}_t$ na época atual E_t. Ao final avalia-se o desempenho da rede ao longo da evolução a partir da matriz M_B, e usando-se a Eq. 3 com IoU_j dado pelo IoU instantâneo (gráficos não-suaves) ou IoU médio das últimas K épocas (gráficos suavizados).

4. Experimentos e Resultados

Para realizar os experimentos, foram usados Keras e TensorFlow no Python 3.6 e uma GPU GTX 1060. As bases de dados utilizadas foram as descritas na Seção 3. Neste trabalho os parâmetros do Keras para aumento de dados foram: rotação ($rotation_range = 0.2$), translação ($width_shift_range = 0.05$ e $height_shift_range = 0.05$), cisalhamento ($shear_range = 0.05$), escala ($zoom_range = 0.05$) e espelhamento horizontal, com modo de preenchimento do valor mais próximo.

4.1. Resultados de Treinamento e Retreinamento

Realizou-se um treinamento com $E_{TC} = 1800$, $|BC| = 5317$ (ou seja 73% do total dos dados são usados no treinamento). Calculou-se o $\overline{IoU}_j(E_i, E_f)$ de cada imagem I_j da base de treinamento BC, conforme Eq. 2. Foram testados diferentes intervalos $[E_i, E_f]$ para o cálculo do IoU médio de cada imagem variando-se o valor de E_i ($E_i = \{100, 800, 1000, 1200, 1400, 1600\}$), enquanto E_f permaneceu fixo em 1800. Subsequentemente, para cada intervalo testado, as imagens foram ordenadas (ordem crescente de IoU). A Figura 6 ilustra o resultado da ordenação de todas as imagens de BC para dois intervalos $[E_i, E_f]$ de avaliação de IoU médio($[100, 1800]$ e $[1600, 1800]$) e mostra: (i) que valores maiores de E_i resultam num IoU maior e (ii) que o comportamento das curvas é semelhante.

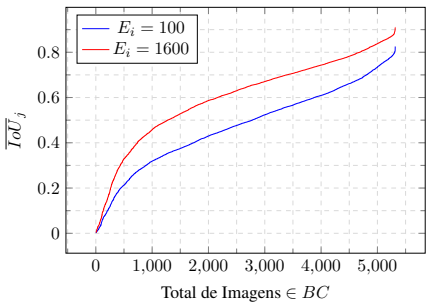


Figura 6. Gráfico das médias para os intervalos com $\mathbf{E_i}$ igual a 100 e 1600.

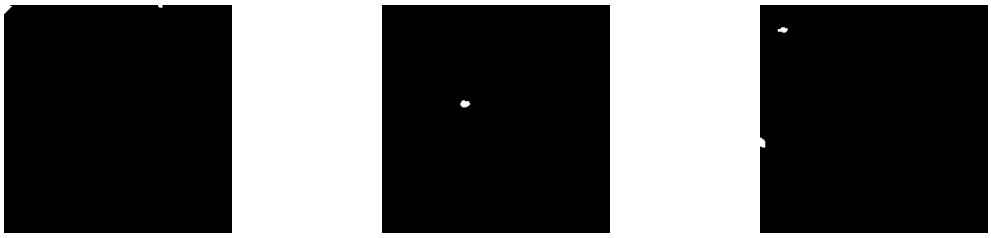
O primeiro resultado já era esperado uma vez que os estágios finais da evolução (ou seja E_i maior definindo intervalos mais estreitos) tendem a produzir IoU mais altos. O segundo resultado mostra que ao se fixar um determinado limiar L para a definição da base filtrada, mais imagens serão eliminadas utilizando-se o intervalo mais amplo ([100,1800]) do que o intervalo mais estreito ([1600,1800]). Para se verificar a influência do intervalo $[E_i, E_f]$ sobre os tipos de imagens escolhidas para serem retiradas da BC, foi testado o percentual de semelhança no conjunto de piores imagens para cada par subsequente no conjunto de valores E_i, ou seja, $\{100 \text{ e } 800, 800 \text{ e } 1000, \ldots, 1400 \text{ e } 1600\}$. A porcentagem de imagens iguais foi maior que 95% em todas as comparações, concluindo-se que pequenas variações de E_i têm pouco efeito sobre quais imagens são eliminadas. Ao fim destes testes, escolheu-se o intervalo mais estreito ($[1600, 1800]$) por este ser mais representativo dos estágios finais do treinamento.

Para se analisar a influência de L sobre a base filtrada BF_L foram realizados diferentes testes considerando-se o intervalo fixado anteriormente ($[E_i, E_f] = [1600, 1800]$) e $L = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$. As quantidades de imagens totais filtradas de BC, assim como as quantidades classificadas nas categorias (*borda*, *meio* e *misto*, com $b = 10$ pixels, conforme discutido na Seção 2.4), são mostradas na Tabela 1.

Tabela 1. Quantidade de imagens retiradas no total e por categoria (*borda*, *meio* e *misto*, com $\mathrm{b} = 10$ pixels).

Base	|Base|	Total de imagens filtradas	Borda	Meio	Misto
$BF_{0.1}$	5185	132	132	0	0
$BF_{0.2}$	5056	261	250	6	5
$BF_{0.3}$	4884	433	351	35	47
$BF_{0.4}$	4578	739	445	139	135
$BF_{0.5}$	4040	1277	501	331	445
$BF_{0.6}$	3157	2160	529	709	922
$BF_{0.7}$	1926	3391	539	1175	1677

Observa-se que as imagens com pior desempenho, removidas pelo limiar $L = 0.1$, são todas da categoria de *borda*. Esse tipo de imagem, assim como exemplos de *meio* e *misto*, retirados do limiar $L = 0.2$, estão ilustrados na Figura 7.



Borda Meio Misto

Figura 7. Máscaras de segmentação de imagens das categorias de *borda*, *meio* e *misto*.

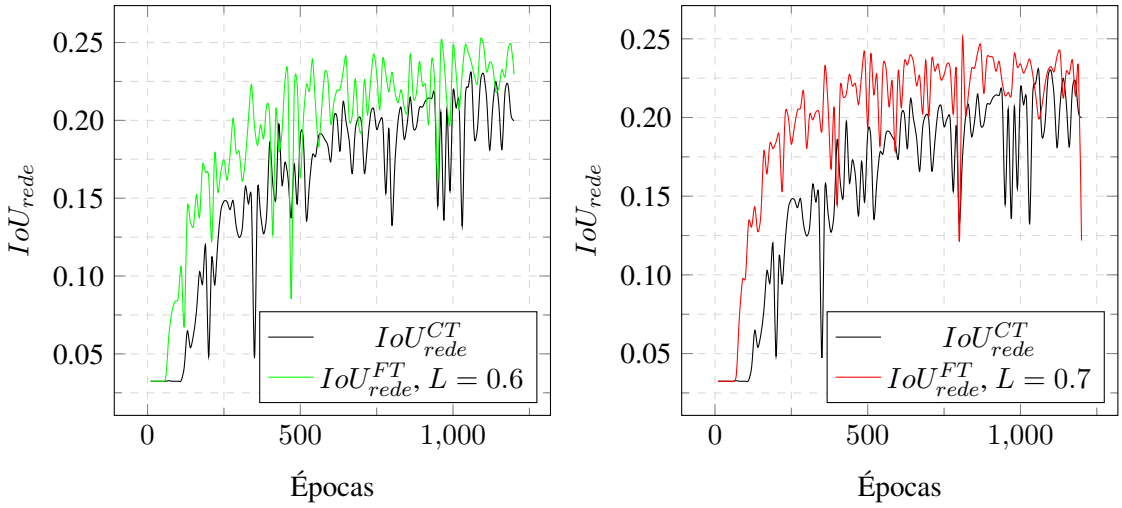
4.2. Reaprendizagem com as bases filtradas

Nesta seção, será feita referência aos resultados dos testes obtidos com $|BT| = 1959$, ou seja, 27% dos dados, divididos em duas categorias conforme a combinação treinamento-

teste:

- IoU_{rede}^{CT}: resultado da rede treinada com a base completa (BC) e testada com a base de teste (BT).
- IoU_{rede}^{FT}: resultado da rede treinada com a base filtrada com o limiar L (BF_L) e testada com a base de teste (BT).

Realizou-se o retreinamento utilizando-se as bases filtradas, com $E_T = 1200$, utilizando os mesmos pesos iniciais da base completa. A Figura 8 ilustra a comparação do treinamento realizado com $BC \times BF$, para os dois maiores valores de L, considerando-se os gráficos não suavizados. As curvas foram obtidas sobre base de teste BT em avaliações realizadas a cada $T_t = 10$ épocas ao longo do treinamento. Verifica-se que as bases filtradas $BF_{0.6}$ e $BF_{0.7}$ produzem resultados melhores do que aqueles produzidos pela base completa (BC).



(a) $\mathbf{L = 0.6}$ **(b)** $\mathbf{L = 0.7}$

Figura 8. Gráfico do desempenho da base completa e das bases filtradas.

A Figura 9 ilustra a suavização dos gráficos mostrados na Figura 8. Estes foram obtidos com janela de média móvel definida por $K = 100$. São mostrados os resultados obtidos tanto pelas bases filtradas IoU_{rede}^{FT} (com $L = \{0.1, 0.6, 0.7\}$) quanto pela base completa IoU_{rede}^{CT}. Observa-se que os melhores desempenhos foram obtidos pelas bases filtradas $BF_{0.6}$ e $BF_{0.7}$.

Conforme ilustrado nos gráficos da Figura 9, é possível observar que as filtragens realizadas com as bases $BF_{0.6}$ e $BF_{0.7}$ antecipam o desempenho da rede em várias gerações, em especial a base $BF_{0.7}$ que apresentou, aproximadamente na época 400, desempenho semelhante ao da base completa nas épocas finais. Portanto houve um adiantamento de cerca de 800 épocas no treino da $BF_{0.7}$ com relação à BC.

A Tabela 2 mostra a aproximação da área sob a curva para três estágios evolutivos dos gráficos mostrados na Figura 9: (i) início (intervalo de épocas [1-400]); (ii) intermediário (intervalo de épocas [401-800]) e (i) final (intervalo de épocas [801-1200]). É possível se observar que as duas bases filtradas superam a base completa em todos os

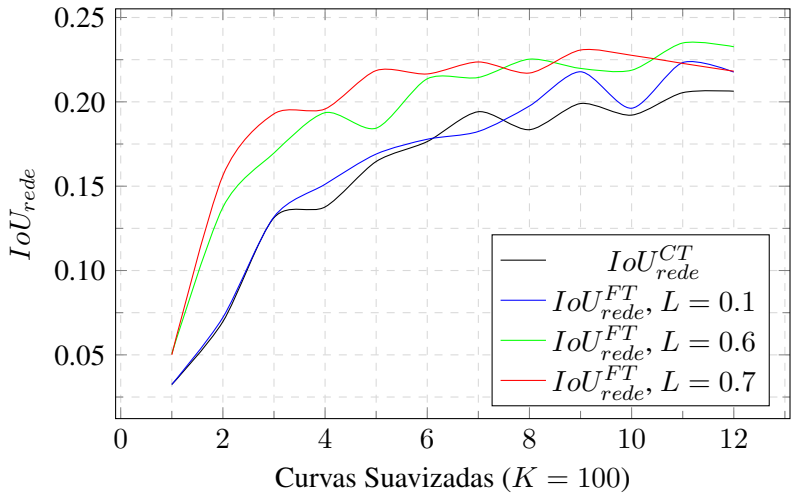


Figura 9. Gráfico das curvas suavizadas, com $\mathrm{K} = 100$, das bases filtradas com limiar 0.1, 0.6 e 0.7 e da base completa.

estágios e o ganho da base $BF_{0.6}$ é menor do que o ganho da base $BF_{0.7}$ nos estágios inicial e intermediário mas é melhor ao final do processo evolutivo. Portanto, a escolha da melhor base filtrada deve ser feita considerando-se a quantidade de recursos disponíveis para o retreinamento da rede: se forem poucos recursos ($\leq$ 800 épocas) a $BF_{0.7}$ pode ser mais adequada; e se houver mais recurso ($>$ 800 épocas) a base $BF_{0.6}$ pode fornecer melhores resultados.

Tabela 2. Área sob a curva de três intervalos evolutivos.

Base	[1-400]	[401-800]	[801-1200]
BC	3.712	7.187	8.029
$BF_{0.1}$	3.876	7.268	8.549
$BF_{0.5}$	4.965	8.079	8.723
$BF_{0.6}$	5.514	8.378	9.061
$BF_{0.7}$	5.951	8.758	8.993

5. Conclusão

Este trabalho apresentou uma proposta de filtragem de imagens na base de treinamento de uma rede U-Net buscando melhorar o desempenho de generalização quando a rede é retreinada com a base filtrada. Considerando o contexto de imagens de tubulações de dutos petrolíferos, foram criadas máscaras de segmentação a partir dos laudos das radiografias disponibilizados por especialistas. Buscando padronizar as dimensões das imagens de entrada da rede, as imagens originais foram recortadas em imagens menores de 256 $\times$ 256 pixels através de uma janela deslizante. Essas imagens foram divididas entre base de treinamento completa e de teste. A base completa foi utilizada para treinar a rede por um número elevado de épocas. A partir do desempenho de cada imagem da base de treino completa, medido pelo IoU, escolheram-se aquelas com pior desempenho para serem removidas. As imagens remanescentes (aquelas com desempenho acima de um limiar de IoU) formaram a base filtrada para aquele limiar específico, base esta usada

no retreinamento da rede. Foram realizados testes comparando-se os resultados da rede treinada com base completa com aqueles produzidos pelas redes treinadas com as bases filtradas, cada uma obtida com limiar de corte diferente.

Conforme mostram os resultados obtidos, apesar de remover uma grande parcela das imagens da base de treino, em especial para valores altos no limiar de IoU, utilizar a filtragem proposta gerou um adiantamento significativo no tempo de retreinamento da rede, além de ter melhorado seu desempenho no início, meio e fim do treino em comparação com o desempenho da rede treinada com a base completa. Os resultados mostraram que à medida que o limiar de IoU usado na filtragem aumentou, a capacidade de generalização acompanhou esse aumento. Isso permitiu se concluir que, mesmo para a complexa base de dados considerada neste trabalho (os valores de IoU não superaram o limite de 0.3), ao se implementar a técnica de filtragem de dados proposta, a qual se baseia no desempenho da rede na própria base de treino, há a possibilidade de se melhorar o desempenho da rede nos dados de teste.

No entanto, muitas questões permanecem para uma investigação futura: é possível se definir de forma automática o melhor limiar sem necessidade de dados de validação (ou seja vários retreinamentos com diferentes limiares)? O método é geral e pode ser aplicado a outros modelos e contextos de aplicação (incluindo redes CNN e problemas de classificação)? Estas e outras questões deverão ser avaliadas futuramente para que a proposta possa ser automatizada e sua utilização possa ser ampliada para outros modelos e aplicações.

6. Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento parcial obtido através dos projetos 309935/2017-2 e 439226/2018-0, bem como, ao Centro de Pesquisa e Desenvolvimento Leopoldo Américo Miguez de Mello (Cenpes) - Petróleo Brasileiro (Petrobras) pelas radiografias fornecidas.

Referências

(2015). *NBR 14842*. Associação Brasileira de Normas Técnicas, Rio de Janeiro.

Atlas, R. M. e Hazen, T. C. (2011). Oil biodegradation and bioremediation: a tale of the two worst spills in us history.

Baniukiewicz, P. (2014). Automated defect recognition and identification in digital radiography. *Journal of Nondestructive Evaluation*, 33(3):327–334.

Davis, J. R. (1989). *ASM Handbook: Nondestructive evaluation and quality control*, volume 17. ASM International.

Duan, F., Yin, S., Song, P., Zhang, W., Zhu, C., e Yokoi, H. (2019). Automatic welding defect detection of x-ray images by using cascade adaboost with penalty term. *IEEE Access*, 7:125929–125938.

Fucsok, F., Muller, C., e Scharmach, M. (2002). Reliability of routine radiographic film evaluation–an extended roc study of the human factor. In *8th European Conference on Non Destructive Testing, Barcelona, June*, pages 17–21.

He, K., Zhang, X., Ren, S., e Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Krizhevsky, A., Sutskever, I., e Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.

Liao, T. W. (2003). Classification of welding flaw types with fuzzy expert systems. *Expert Systems with Applications*, 25(1):101–111.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., e Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988.

Liu, B., Zhang, X., Gao, Z., e Chen, L. (2017). Weld defect images classification with vgg16-based neural network. In *International Forum on Digital TV and Wireless Multimedia Communications*, pages 215–223. Springer.

Mery, D. e Arteta, C. (2017). Automatic defect recognition in x-ray testing using computer vision. In *2017 IEEE winter conference on applications of computer vision (WACV)*, pages 1026–1035. IEEE.

Nacereddine, N. e Tridi, M. (2005). Computer-aided shape analysis and classification of weld defects in industrial radiography based invariant attributes and neural networks. In *ISPA 2005. Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, 2005.*, pages 88–93. IEEE.

Rana, S. et al. (2008). Facts and data on environmental risks-oil and gas drilling operations. In *SPE Asia Pacific Oil and Gas Conference and Exhibition*. Society of Petroleum Engineers.

Ronneberger, O., Fischer, P., e Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.

Shorten, C. e Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):60.

Simonyan, K. e Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., e Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.

Wang, Y., Shi, F., e Tong, X. (2019). A welding defect identification approach in x-ray images based on deep convolutional neural networks. In *International Conference on Intelligent Computing*, pages 53–64. Springer.

Zapata, J., Vilar, R., e Ruiz, R. (2012). Automatic inspection system of welding radiographic images based on ann under a regularisation process. *Journal of Nondestructive Evaluation*, 31(1):34–45.