

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE GEOCIÊNCIAS

**APLICAÇÃO DE *DEEP LEARNING* PARA ANÁLISE MORFOLÓGICA  
DE ZIRCÃO DETRÍTICO**

**Thiago da Cunha Rodrigues**

Orientadora: Profa. Dra. Maria Helena Bezerra Maia de Hollanda

Dissertação de Mestrado

Programa de Pós-Graduação em Geoquímica e Geotectônica.

SÃO PAULO

2023

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE GEOCIÊNCIAS

**Aplicação de *Deep Learning* para análise morfológica de zircão detrítico**

**THIAGO DA CUNHA RODRIGUES**

Dissertação apresentada ao Programa de Geociências – Geoquímica e Geotectônica para obtenção do título de Mestre em Ciências.

Área de concentração: Geotectônica.

Orientadora: Profa. Dra. Maria Helena Bezerra Maia de Hollanda

SÃO PAULO

2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Serviço de Biblioteca e Documentação do IGc/USP

Ficha catalográfica gerada automaticamente com dados fornecidos pelo(a) autor(a) via programa desenvolvido pela Seção Técnica de Informática do ICMC/USP

Bibliotecários responsáveis pela estrutura de catalogação da publicação:

Sonia Regina Yole Guerra - CRB-8/4208 | Anderson de Santana - CRB-8/6658

Rodrigues, Thiago da Cunha    Aplicação de Deep Learning para análise morfológica de zircão detrítico / Thiago da Cunha Rodrigues; orientadora Maria Helena Bezerra Maia de Hollanda. -- São Paulo, 2023.  
94 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Geoquímica e Geotectônica) -- Instituto de Geociências, Universidade de São Paulo, 2023.

1. Aprendizado de máquina. 2. Detecção de objetos. 3. Proveniência Sedimentar. I. Hollanda, Maria Helena Bezerra Maia de , orient. II. Título.

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE GEOCIÊNCIAS

**Aplicação de Deep Learning para análise morfológica  
de zircão detrítico**

**THIAGO DA CUNHA RODRIGUES**

Orientador: Profa. Dra. Maria Helena Bezerra Maia de Hollanda

Dissertação de Mestrado

**Nº 904**

COMISSÃO JULGADORA

Dra. Maria Helena Bezerra Maia de Hollanda

Dr. Cleyton de Carvalho Carneiro

Dr. Carlos Conforti Ferreira Guedes

SÃO PAULO  
2023

## AGRADECIMENTOS

Outra longa etapa de trabalho concluída, e dessa vez, ao longo de toda uma pandemia e em outra área do conhecimento além da geologia. Este trabalho só foi possível a sua concepção e conclusão graças ao suporte de profissionais, entidades, amigos, mentores e familiares. Portanto gostaria de deixar meus sinceros agradecimentos a todas as pessoas que fizeram parte dessa etapa e que contribuíram de forma positiva para o trabalho e na minha formação.

Em primeiro lugar, gostaria de agradecer a minha orientadora, Profa. Dra. Maria Helena B. M. de Hollanda, e, a minha co-orientadora Profa. Dra. Ana Maria Góes. Por todas as conversas edificantes, correções, sugestões e suporte pessoal e profissional, e acima de tudo, por acreditarem em mim. Agradeço a oportunidade de trabalhar em conjunto neste projeto com duas das mulheres mais incríveis que já conheci.

Gostaria de agradecer as entidades Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pela concessão de bolsa de pesquisa de mestrado. A todos os profissionais do Instituto de Geociências – IGc -USP que deram o suporte em uma época tão conturbada, que foi a pandemia de COVID-19. Em especial ao professor Dr. Paulo César Fonseca Giannini por suas aulas excepcionais e por ajudar na concepção deste projeto. Agradeço também aos pesquisadores Dr. André Zular e Msc. Ana Maria Sforcin por contribuir em uma etapa importante deste trabalho, mesmo sendo necessário uma demanda significativa de tempo em um curto prazo.

A todos os companheiros de pós-graduação no IGc-USP por todo suporte acadêmico e emocional, em especial a Juan Sebastian, Kauê Seoane, Paula Fraga, Carolina Bedoya (Caro), Dinarte Lucas (Dino), Caio de Freitas e em especial Ravi Sampaio, que foi o santo que aceitou a missão de entregar essa dissertação pessoalmente. A próxima ida ao Seven Kings é por minha conta Ravi.

A Marcos Jacinto (Coscos), um dos meus melhores amigos, o meu mais sincero agradecimento. Principalmente por ser a pessoa a me trazer à área de ciência de dados, pelas conversas edificantes, pela mentoria na área e na orientação quanto a dissertação.

Aos meus sogros, a Elenilda e Reginaldo, que deram todo o suporte nesta nova fase da minha vida, sempre me tratando como um filho. A minha cunhada Maria Luiza e ao meu irmão Eric, pela convivência em terras paulistas, sabendo que eu poderia contar com eles em qualquer momento. Aos meus pais, Lúcia e Jorge, por todo o apoio que me

permitiu realizar um mestrado em outra cidade, e por sempre demonstrarem se importar comigo, seja ligando todo domingo, ou conversando pelas redes sociais.

Por último e mais importante, a minha melhor amiga e namorada, Alana Régia Dantas o meu eterno amor e agradecimento. A pessoa por trás do design artístico deste trabalho, que confeccionou todas as imagens e participou até da classificação desenvolvida. Por estar sempre ao meu lado, principalmente nessa nova etapa de nossa vida, nos momentos de procrastinação ou me incentivando a ser produtivo, nos momentos mais difíceis e nos de felicidade. Obrigado por ser a pessoa que aguenta todos os meus inúmeros defeitos e egoísmos. De todas as mulheres excepcionais que estão na minha vida, você é a maior delas. *“Never opened myself this way, Life is ours, we live it our way, All this words i just don’t say, And nothing else matters”*.

*“Pass on what you have learned. Strength. Mastery. But weakness, folly, failure also. Yes, failure most of all. The greatest teacher, failure is. Luke, we are what they grow beyond. That is the true burden of all masters.”*

*— Yoda, Star Wars: The Last Jedi.*

## RESUMO

Parâmetros morfológicos de grãos de zircão podem ser aplicados ao estudo de proveniência de bacias em combinação com dados isotópicos e análise convencional de minerais pesados. Neste caso, os estudos morfológico e geocronológico do zircão detrítico estão incluídos na análise de variedades de minerais pesados. Tradicionalmente, os métodos de classificação morfológica foram desenvolvidos para quartzo detrítico e adaptada para grãos de zircão, minerais resistatos em arenitos, O processo demanda longo tempo para o usuário alcançar a representatividade desejada e, ainda assim, carrega significativa subjetividade. Advoga-se que a automação do procedimento através de redes neurais (*Deep Learning*) tornará esta pesquisa mais eficiente, minimizando a subjetividade humana. Esta hipótese foi testada em banco de imagens de catodoluminescência de zircão detrítico previamente datado. Para atender a natureza deste banco de dados, foi desenvolvida uma classificação morfológica que relaciona hábito e arredondamento do zircão detrítico, que resultou em três categorias: euédricos e subédricos (1), anédricos arredondados (2) e anédricos quebrados, que não dão informação de transporte (zero). Como as categorias zero e 1 ocorrem em menor quantidade em relação à 2, foi necessário equilibrar as quantidades de dados através da aplicação de técnicas de *data augmentation*. Foram elaborados dois conjuntos de dados diferentes: (i) composto por grãos de zircão imagens de catodoluminescência, (ii) composto por imagens de elétrons secundários. Esses conjuntos foram treinados em diversas arquiteturas de detecção de objetos, sendo a principal e com os melhores resultados a YOLOv5. Foi realizado um ajuste dos hiperparâmetros para otimizar as métricas dos modelos treinados. A validação dos dados adquiridos nos dois modelos foi feita com as métricas de *mean average precision* (mAP), Precisão, sensibilidade e *F1 score*. O primeiro conjunto de dados apresentou, para todas as categorias, mAP = 0,69 nos dados de teste, enquanto o segundo alcançou mAP = 0,66. Para comprovar se o modelo estava utilizando apenas as características morfológicas dos grãos, e não outros aspectos que poderiam imputar um viés à categorização planejada, efetuamos discretas mudanças artificiais em alguns grãos das categorias 1 e 2 com o software *Photoshop*. Essas mudanças demonstraram que as imagens de catodoluminescência exibem propriedades, além da morfologia, que afeta a classificação do modelo (e.g., padrão de zonação interna dos grãos). Além das métricas de avaliação, as predições do modelo foram comparadas a classificações de dois sedimentólogos (A e B). Quando comparados os valores do modelo e dos usuários, o modelo se saiu superior na classificação ao apresentar um valor *F1 score* = 75,61% contra 63,31% do usuário A e contra 69,41% do usuário B. O teste comparativo entre usuários mostrou uma discrepância alta nos resultados, demonstrando a necessidade de refinamento diante da alta subjetividade que o estudo de morfologia carrega, além de *feedback* maior do público alvo. O resultado geral foi positivo, pois o modelo não encontrou problemas em detectar todos os grãos e teve facilidade em classificar grãos da categoria 2, porém com dificuldades em discriminar os grãos das categorias zero e 1, devido à falta representatividade da categoria zero.

**Palavras-chave:** Aprendizado de máquina, Detecção de objetos, Proveniência Sedimentar.



## ABSTRACT

Morphological parameters of zircon grains can be applied to the study of basin provenance in combination with isotopic data and conventional analysis of heavy minerals. In this case, the morphological and geochronological studies of the detrital zircon are included in the analysis of heavy mineral varieties. Traditionally, morphological classification methods were developed for detrital quartz and adapted for zircon grains, resistate minerals in sandstones. The process demands a long time for the user to reach the desired representativeness and, even so, it carries significant subjectivity. It is argued that the automation of the procedure through neural networks (Deep Learning) will make this research more efficient, minimizing human subjectivity. This hypothesis was tested on a previously dated detrital zircon cathodoluminescence database. To meet the nature of this database, a morphological classification was developed that relates habit and roundness of detrital zircon, which resulted in three categories: euhedral and subhedral (1), rounded anhedral (2) and broken anhedral, which do not give transport information (zero). As categories zero and 1 occur in smaller amounts compared to 2, it was necessary to balance the amounts of data through the application of data augmentation techniques. Two different datasets were developed: (i) composed of zircon grains cathodoluminescence images, (ii) composed of secondary electron images. These sets were trained on several object detection architectures, the main one with the best results being YOLOv5. An adjustment of the hyperparameters was performed to optimize the metrics of the trained models. The validation of the data acquired in the two models was performed with the metrics of mean average precision (mAP), Precision, sensitivity and F1 score. The first set of data presented, for all categories, mAP = 0.69 in the test data, while the second reached mAP = 0.66. To check whether the model was using only the morphological characteristics of the grains, and no other aspects that could impute a bias to the planned categorization, we made discreet artificial changes in some grains of categories 1 and 2 with Photoshop software. These changes demonstrated that the cathodoluminescence images exhibit properties, in addition to morphology, that affect the model classification (e.g., grain internal zonation pattern). In addition to the evaluation metrics, the model predictions were compared to classifications by two sedimentologists (A and B). When comparing the values of the model and users, the model performed better in the classification by presenting an F1 score value = 75.61% against 63.31% of user A and against 69.41% of user B. The comparative test between users indicated a high discrepancy in the results, demonstrating the need for refinement in view of the high subjectivity that the morphology study carries, in addition to greater feedback from the target audience. The overall result was positive, since the model did not present any problems in detecting all grains. Besides that, it was easy for the model to classify grains in category 2, but with difficulties in discriminating grains in categories zero and 1, due to the lack of representativeness of category zero.

**Keywords:** Machine Learning, Object Detection, Basin Provenance.

# SUMÁRIO

<b>1. INTRODUÇÃO</b>	1
<b>1.1. Apresentação</b>	1
<b>1.2. Justificativa e objetivos</b>	3
<b>2. REVISÃO BIBLIOGRÁFICA</b>	5
<b>2.1. Morfologia de grãos detríticos</b>	5
<b>2.2. Relevância do zircão em estudos de proveniência</b>	8
<b>2.3. Proveniência a partir de minerais pesados</b>	8
<b>2.4. <i>Deep learning</i></b>	10
<b>2.5. Funcionamento de uma CNN</b>	11
<b>2.5.1. <i>Artificial Neural Network</i></b>	12
<b>2.5.2. A arquitetura de uma CNN</b>	13
<b>2.5.3. Tipos de camadas</b>	16
<b>2.5.3.1. <i>Convolutional Layer</i></b>	16
<b>2.5.3.2. <i>Pooling Layer</i></b>	17
<b>2.5.3.3. <i>Fully Connected Layer</i></b>	18
<b>2.6. Transferência de Aprendizado</b>	18
<b>2.7. Métricas de validação e desempenho</b>	19
<b>2.8. Detecção de objetos</b>	23
<b>2.8.1. Arquiteturas em dois estágios</b>	24
<b>2.8.2. Arquiteturas em um estágio</b>	25
<b>3. MÉTODOS</b>	30
<b>3.1. Pré-processamento</b>	31
<b>3.2. Treinamento</b>	36
<b>3.3. Pós-processamento</b>	38
<b>4. RESULTADOS</b>	39
<b>4.1. Conjunto de dados de catodoluminescência</b>	40
<b>4.1.1. Escolha do modelo</b>	40
<b>4.1.2. Ajuste dos hiperparâmetros</b>	43
<b>4.1.3. Inferência (detecção e classificação) do conjunto de dados de teste</b>	48
<b>4.1.4. Comportamento do modelo</b>	49
<b>4.1.5. Validação dos testes – análise de usuários</b>	54
<b>4.2. Conjunto de dados de Elétrons Secundários</b>	55
<b>5. DISCUSSÕES E CONCLUSÃO</b>	63
<b>5.1. Limitações do processo de classificação morfológica</b>	63
<b>5.2. Desempenho do modelo</b>	67
<b>5.3. Comparação do desempenho do modelo versus usuários</b>	67
<b>5.4. Perspectivas futuras</b>	70
<b>5.5. Comentários finais</b>	72
<b>REFERÊNCIAS</b>	74

## LISTA DE FIGURAS

<b>Figura 1</b>	Exemplo de quantificação de medidas de arredondamento de partículas (A; Wadell, 1932) e de circularidade (B; Riley, 1941). Os números internos representam os raios de cada círculo (Fonte: Blott e Pye, 2008).	6
<b>Figura 2</b>	Escala de arredondamento por comparação visual (Fonte: Powers, 1953).	7
<b>Figura 3</b>	Classes morfológicas desenvolvidas para minerais pesados resistentes, baseadas em Krumbein (1941), Powers (1953) e, possivelmente, Mange-Rajetzky (1995) e seguidores. O sentido da seta preta indica que quanto mais distante da forma eudral, maior é o grau de arredondamento dos grãos anédricos. A barra preta representa uma escala de 63 µm, em todos os grãos. (Fonte: Andò et al., 2012).	9
<b>Figura 4</b>	Representação esquemática de uma <i>artificial neural network</i> : (a) neurônio humano; (b) neurônio artificial; (c) sinapse biológica; e (d) sinapse de uma ANN (Fonte: Meng et al., 2020).	11
<b>Figura 5</b>	<i>Backpropagation</i> em uma ANN. Os valores dos pesos são atualizados a cada novo treino (sinal) que a rede executa, gerando o melhor resultado e permitindo o “aprendizado” da rede. (Adaptado de Meng et al., 2020).	13
<b>Figura 6</b>	Exemplo das camadas de uma CNN (Fonte: Wu, 2017).	13
<b>Figura 7</b>	<i>Cost function</i> convergindo para o menor valor possível (mínimo global), conforme a sua learning rate (Fonte: Bhattarai, 2018).	14
<b>Figura 8</b>	Ilustração esquemática de uma operação de convolução utilizada nas CNNs, tendo como exemplo um <i>kernel</i> 2x2 com um <i>stride</i> igual a 1 (Fonte: Wu, 2017).	17
<b>Figura 9</b>	Ilustração de um processo de aprendizado normal (a), e de um processo de transferência de aprendizado (b). (Fonte: Pan e Yang, 2009).	19
<b>Figura 10</b>	Matriz de confusão de uma classificação de múltiplas classes. Os campos apresentados utilizam como referencial a categoria da coluna do meio, onde as classificações corretas daquela categoria são chamadas TP e as classificações corretas de outras categorias são chamados TN.	20
<b>Figura 11</b>	Ilustração de grãos de zircão retratados por sua resposta à catodoluminescência. Notar que nesse tipo de imageamento eletrônico destaca-se a estrutura interna de crescimento deste mineral, o que pode ser um viés para a classificação morfológica automatizada (ver capítulos seguintes). Da esquerda para a direita temos uma sequência esquemática de: uma detecção correta (VP) onde se detectou a classe “Grão”; três detecções incorretas (FP) onde se detectou, na sequência, um erro de classe “Não grão”, pouca sobreposição (i.e., limiar abaixo do padrão) e nenhuma sobreposição, onde a caixa delimitadora circunscreve o grão no background; uma ‘não predição’ (FN).	21

<b>Figura 12</b>	Descrição e fórmula da IOU. O quadrado com a borda verde representa a caixa delimitadora original, e o de borda vermelha representa a caixa delimitadora detectada (Fonte: Padilla et al., 2019).	21
<b>Figura 13</b>	Exemplo de interpolação de todos os pontos da curva precisão-sensibilidade, ou P-R. O algoritmo contabiliza sucessivamente P em função de R. Em pontos de “ruídos”, utiliza-se P máximo em função de um R maior (à direita no gráfico), ou seja, maior ou igual a $R_{n+1}$ . (Fonte: Padilla et al., 2019).	23
<b>Figura 14</b>	Etapas de uma arquitetura de dois estágios <i>region-based</i> . (Adaptado de Girshick et al., 2014).	25
<b>Figura 15</b>	Exemplo de como funciona a detecção de objetos dentro de uma imagem na arquitetura YOLO. (Adaptado de Redmon et al., 2016).	26
<b>Figura 16</b>	Estrutura comum da arquitetura YOLO. A etapa de detecção e classificação ocorre na parte <i>head</i> onde a ‘YOLO loss’ (a <i>loss function</i> da arquitetura) é minimizada (Fonte: Long et al., 2020).	29
<b>Figura 17</b>	Fluxograma ilustrando como foi desenvolvido o projeto. O pré-processamento e o processamento ocorreram de forma cíclica. Toda vez que se obtinha um resultado abaixo do desejado, era refeito o pré-processamento do banco de dados afim de se obter valores melhores. O processo de pós-processamento ocorreu após os treinamentos e escolha do modelo final.	31
<b>Figura 18</b>	Fluxograma ilustrando como foi desenvolvida a classificação proposta neste projeto. A ‘classificação inicial’ foi uma tentativa de adotar as definições morfológicas da literatura, enquanto a ‘classificação final’ representa uma combinação de critérios morfológicos consensualmente utilizados na literatura, ajustada às características do banco de imagens utilizado neste trabalho.	33
<b>Figura 19</b>	Exemplos de morfologias de grãos de zircão comentadas no texto, e organizadas nas três categorias definidas para este trabalho.	33
<b>Figura 20</b>	Representação da “remoção” dos grãos sub-representados (neste exemplo para a categoria zero) e de limpeza da imagem, individualizando-se através de caixas delimitadoras os grãos das categorias super-representadas (à esquerda) e os da sub-representadas (à direita). As cores das caixas delimitadoras descrevem a que categoria cada grão pertence.	34
<b>Figura 21</b>	Imagem original e seus respectivos produtos resultantes das técnicas de <i>data augmentation</i> aplicada a imagens de catodoluminescência de grãos de zircão: espelhamento horizontal, espelhamento vertical, rotação 90° em sentido horário e anti-horário, e redimensionamento das imagens.	35
<b>Figura 22</b>	Comparação entre os modelos com configurações iniciais de treinamento diferentes, com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, i.e., com <i>confidence score</i> variando entre 0,5 e 0,95. (B) Valores de mAP a 0,5 de <i>confidence score</i> . (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.	41

<b>Figura 23</b>	Curvas de <i>loss functions</i> que compõem o YOLOv5 para os dados de treinamento (linha contínua, na legenda) e de validação (linha pontilhada). O efeito de <i>overfitting</i> pode ser visto na <i>classification loss</i> , nos gráficos C, G e J.	42
<b>Figura 24</b>	Comparação entre os modelos derivados do CL-Y5-1 (em laranja) com diferentes hiperparâmetros com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, i.e., com <i>confidence score</i> variando entre 0,5 e 0,95. (B) Valores de mAP a 0,5 de <i>confidence score</i> . (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.	45
<b>Figura 25</b>	Curvas de <i>loss functions</i> que compõem o YOLOv5 para os dados de treinamento e de validação (pontilhado). O efeito de <i>overfitting</i> é ilustrado pela função <i>classification loss</i> , especialmente destacado nos gráficos C, G, e incipiente em J.	46
<b>Figura 26</b>	Comparação da matriz de confusão dos modelos CL-Y5-1-C2 e -VOC vistas em forma de escala de cores, e diferenciadas com os dados de validação e de teste.	47
<b>Figura 27</b>	Curva F1 para variações do <i>confidence score</i> no momento da inferência.	48
<b>Figura 28</b>	Exemplos de detecção realizada pelo modelo escolhido (CL-Y5-1-C2). À esquerda são ilustradas três pranchas com as classificações originais ( <i>ground truth detection</i> ) feitas sem a função algorítmica. À direita tem-se as mesmas imagens com a detecção feita pelo modelo. Em (A) e (B) exemplos de grãos dos arenitos eólicos da Formação Sambaíba. Em (C) e (D) exemplos de grãos dos arenitos marinho-fluviais do Grupo Serra Grande. Em (E) e (F) exemplos de grãos dos arenitos fluviais da Bacia do Cocóci.	50
<b>Figura 29</b>	Limpeza artificial da zonação de crescimento interna dos grãos detríticos. Em (A), classificação original feita pelo autor. Em (B), inferência realizada pelo modelo. Em (C), inferência realizada pelo modelo após a limpeza das zonações dos grãos.	51
<b>Figura 30</b>	Pranchas exemplificando o efeito da orientação dos grãos na classificação do modelo. (A) e (D) são pranchas com grãos distintos onde ilustra-se a classificação original feita pelo autor. (B) e (E) são as mesmas pranchas, mas com classificação feita pelo modelo. (C) e (F) são as mesmas pranchas com classificação feita pelo modelo, mas com os grãos rotacionados 90° no sentido horário.	52
<b>Figura 31</b>	Pranchas editadas ilustrando modificações artificiais na morfologia de alguns grãos. O grau de alteração aumenta do topo para baixo, na prancha. (A) Alterações em grãos euédricos (categoria 1). (B) Alterações em grãos arredondados (categoria 2).	53
<b>Figura 32</b>	(A) Matriz de confusão do pesquisador A. (B) Matriz de confusão do pesquisador B.	54

<b>Figura 33</b>	Comparação entre os modelos com configurações iniciais de treinamento diferentes, com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, i.e., com <i>confidence score</i> variando entre 0,5 e 0,95. (B) Valores de mAP a 0,5 de <i>confidence score</i> . (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.	56
<b>Figura 34</b>	Curvas de <i>loss functions</i> que compõem o YOLOv5 para os dados de treinamento e de validação (pontilhado). O efeito de <i>overfitting</i> pode ser visto na <i>classification loss</i> , nos gráficos C, G e J.	57
<b>Figura 35</b>	Comparação entre os modelos derivados do ES-Y5-1 (em laranja) com diferentes hiperparâmetros com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, para valores de <i>confidence score</i> 0,5 a 0,95. (B) Valores de mAP a 0,5 de <i>confidence score</i> . (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.	58
<b>Figura 36</b>	Curvas de <i>loss functions</i> que compõem o YOLOv5 para os dados de treinamento e de validação (pontilhado). O efeito de <i>overfitting</i> é ilustrado pela função <i>classification loss</i> , especialmente destacado nos gráficos C, J, e menor em F.	59
<b>Figura 37</b>	Comparação da matriz de confusão de dois modelos, resultante da detecção realizada nos dados de validação e teste.	60
<b>Figura 38</b>	Curva F1 para variações do <i>confidence score</i> no momento da inferência.	61
<b>Figura 39</b>	Exemplo de detecção feita pelo modelo ES-Y5-1-B3. À esquerda vê-se as classificações originais. À direita, as detecções do modelo. Em (A) e (B), têm-se grãos da Formação Sambaíba. Em (C) e (D), grãos do Grupo Serra Grande com predominância da categoria 1. Em (E) e (F), têm-se grãos do Grupo Serra Grande (mistura das categorias 0, 1 e 2).	62
<b>Figura 40</b>	Pranchas exemplificando situações em que o modelo CL-Y5-1-C2 ‘errou’ na classificação dos grãos entre as categorias pré-definidas. À esquerda (pranchas A e C), são mostradas as classificações originais feitas pelo autor; à direita (pranchas B e D) estão as classificações entregues pelo modelo. No par A-B, os grãos identificados com 2-3-4 foram classificados pelo modelo na categoria 1 provavelmente devido à confusão de arestas artificiais com faces cristalográficas. Note que o grão 1, em (B), também foi classificado erroneamente na categoria 2 provavelmente devido ao padrão interno (zonação) irregular, dando uma ‘impressão’ de grão anédrico. No par C-D, os grãos 5-6-8-9 também são classificados erroneamente na categoria 1 dado o efeito de zonação interna, o inverso acontecendo com o grão 7.	64
<b>Figura 41</b>	Pranchas com imagens de elétrons secundários de grãos exemplificando situações em que o modelo ES-Y5-1-B3 não classificou corretamente. Em (A) tem-se a classificação original, e em (B) a detecção feita pelo modelo com o valor de <i>confidence score</i> . Em amarelo estão indicados os grãos classificados incorretamente: grãos originalmente da categoria zero classificados na categoria 1 (grãos 1, 2, 3 e 9) e vice-versa (grãos 5, 6, 7 e 8); o grão 4 passou da categoria 2 para a categoria zero.	70

## LISTA DE TABELAS

<b>Tabela 1</b>	Índices mais comuns de arredondamento (Fonte: Blot e Pye, 2008).	6
<b>Tabela 2</b>	Métodos de classificação por classificação visual baseados no método de medida de Wadell (1932) (Modificado de Blott e Pye, 2008).	7
<b>Tabela 3</b>	Proporção da quantidade de grãos antes e depois do processo de <i>data augmentation</i> .	36
<b>Tabela 4</b>	Modelos inicialmente desenvolvidos para este trabalho.	37
<b>Tabela 5</b>	Comparação entre as arquiteturas com melhores desempenhos na literatura. Métricas de desempenho dos modelos referentes ao conjunto de dados de validação. Destaque em negrito foi dado para o melhor desempenho.	40
<b>Tabela 6</b>	Comparação entre as configurações iniciais e métricas de desempenho dos modelos referentes ao conjunto de dados de teste. Destaque em negrito foi dado para o melhor desempenho.	41

<b>Tabela 7</b>	Configuração do otimizador SGD e ADAM para aprimorar o modelo de linha de base.	44
<b>Tabela 8</b>	Tabela comparando o desempenho entre os modelos com os melhores desempenhos, tendo o modelo CL-Y5-1 como linha de base. <i>Epoch</i> indica o momento em que esse desempenho foi alcançado no treinamento. Destaque em negrito foi dado para o melhor desempenho.	44
<b>Tabela 9</b>	Comparação das métricas: precisão (P), sensibilidade (R) e pontuação F1 entre o modelo e os usuários A e B.	55
<b>Tabela 10</b>	Comparação entre as configurações iniciais e métricas de desempenho dos modelos referentes ao conjunto de dados de teste. Destaque em negrito para o melhor desempenho entre os modelos.	56
<b>Tabela 11</b>	Configuração do otimizador SGD e ADAM para aprimorar o modelo de linha base.	57
<b>Tabela 12</b>	Tabela comparando o desempenho entre os modelos com os melhores desempenhos, tendo o modelo ES-Y5-1 como linha de base. <i>Epoch</i> indica o momento em que esse desempenho foi alcançado no treinamento. Destaque foi dado para o melhor desempenho.	60
<b>Tabela 13</b>	Valores de P e R derivados da matriz de confusão do modelo ES-Y5-1-B3	61

## LISTA DE SIGLAS

<b>ADAM</b>	Adaptative Moment Estimation.
<b>CL</b>	Catodoluminescencia.
<b>CNN</b>	Convolutional Neural Network.
<b>COLAB</b>	Colaboratory jupyter notebook environment
<b>DL</b>	Deep Learning
<b>ES</b>	Elétron Secundário
<b>FN</b>	Falso Negativo
<b>FP</b>	Falso Positivo
<b>HRHMA</b>	High-Resolution Heavy Mineral Analysis
<b>IoU</b>	Intersection over Union
<b>LA-ICPMS</b>	Laser Ablation Inductively Coupled Plasma Mass Spectrometry
<b>mAP</b>	Mean Average Precision
<b>ML</b>	Machine Learning



<b>NMS</b>	Non Max Suppression
<b>RCNN</b>	Region-based Convolutional Neural Network
<b>SGD</b>	Stochastic Gradient Descent
<b>SVM</b>	Suport Vector Machine
<b>VN</b>	Verdadeiro Negativo
<b>VP</b>	Verdadeiro Positivo
<b>YOLO</b>	You can Only Look Once

# 1. INTRODUÇÃO

## 1.1. Apresentação

A morfologia indica o estudo da forma de objetos pertencentes à mesma população e é uma importante propriedade textural em petrologia sedimentar. Ela é descrita de várias maneiras, como por exemplo o arredondamento e a esfericidade (Dewey e Mange, 1999). O arredondamento refere-se ao grau de agudeza das arestas das bordas do grão, enquanto a esfericidade se relaciona ao quanto o contorno do grão se aproxima de uma esfera perfeita, i.e., trata da equidimensionalidade do grão (ver Blott e Pry, 2008). Ambas as propriedades variam com o tipo e dimensões do grão mineral, suas propriedades físicas e químicas, e a ação efetiva de transporte da partícula no ciclo sedimentar (tipo, distância e meio). Quando voltados para areias e arenitos, os estudos morfológicos estiveram inicialmente centralizados no quartzo detrítico dada sua abundância nestas rochas e reconhecida resiliência física e química. Desde então, muitas classificações foram propostas, uma das quais a de Powers (1953) que classifica arredondamento e esfericidade utilizando uma escala de comparação visual. No presente trabalho utilizamos essa classificação devido sua simplicidade e consenso entre sedimentólogos.

Grãos guardam informações cruciais de todo o processo relativo à formação de rochas sedimentares, i.e., do intemperismo, erosão à sedimentação e, como tal, representam um registro do clima e da tectônica da área-fonte. Consequentemente, o estudo da forma de grãos constitui-se em uma das ferramentas para estudos de proveniência sedimentar, um tema da Geologia Sedimentar com reconhecida relevância em reconstruções paleogeográficas e na predição da qualidade de reservatórios (Haughton et al., 1991; Weltje e Eynatten, 2004; Garzanti et al., 2014; Bhattacharya et al., 2016).

As últimas décadas registraram aumento significativo no interesse do setor produtivo e da academia em pesquisas relacionadas à proveniência sedimentar de bacias-reservatórios. Uma das maiores bacias intracratônicas brasileiras paleozoicas é a Bacia do Parnaíba, com espessura da pilha sedimentar de cerca de 3.500 m, área total que recobre seis estados brasileiros, e hospedeira de reservas importantes de gás natural. Estas razões a fizeram alvo para diversos estudos de proveniência utilizando, principalmente, geocronologia de grãos de zircão combinada com análise de outros minerais pesados detríticos (Hollanda et al., 2014, 2018; Menzies et al., 2018; Oliveira e Moura, 2019; Oliveira e Truckenbrodt, 2019; Cardoso et al., 2019; Cerri et al., 2020; Del Rio et al., 2021).

Quando a temática é avaliar a contribuição de fontes internas no preenchimento da bacia, os objetos de investigação selecionados são minerais (pesados) acessórios ultra-estáveis, seja em datação e/ou classificações tipológicas ou morfológicas. Nesses contextos, destaca-se novamente o zircão. Esse mineral é comparável ao quartzo pelo seu comportamento como resistato e, portanto, por guardar particularidades que podem diretamente correlacioná-lo às rochas fontes. Dada a sua relevância, a década de 1990 presenciou desenvolvimento de um método analítico chamado *High-Resolution Heavy Mineral Analysis* (HRHMA) que categoriza os grãos de zircão através de parâmetros que remetem à cristalização dos minerais originais na rocha – i.e., tamanho, hábito, cor, e estrutura interna, associado a parâmetros que podem sinalizar os efeitos do transporte sedimentar como o arredondamento (p.ex., Mange-Rajetzky 1995; Dewey e Mange 1999; Mange e Otvos, 2005).

Seguindo a linha de pesquisa do método HRHMA, a ideia central deste trabalho foi desenvolver um método de visão computacional para a classificação morfológica de grãos de zircão considerando o hábito cristalino preservado (euédrico a subédrico) e os aspectos do arredondamento relacionados aos grãos anédricos. A base foi um banco de dados de imagens de 5.344 grãos de zircão datados pela técnica U-Pb *in situ* por ablação a laser (LA-ICPMS, Laser Ablation Inductively Coupled Plasma Mass Spectrometry) para fins de reconhecimento da proveniência sedimentar de arenitos paleozoicos da Bacia do Parnaíba. Dos mais de 5 mil grãos analisados à época de execução desse projeto – PRO-PARNAÍBA, foram validadas mais de quatro mil idades individuais que forneceram contribuições valiosas sobre as áreas-fontes daqueles arenitos e padrões de preenchimento sedimentar da bacia (Hollanda et al., 2014, 2018). Os grãos foram imageados de três formas: catodoluminescência, elétrons retroespalhados e luz transmitida, cujas imagens foram consideradas como objeto para a análise morfológica desenvolvida neste presente trabalho.

De forma resumida, adotamos o método *Convolutional Neural Networks* (CNNs) que utiliza técnicas de *Deep Learning* (DL) para detectar e classificar, de forma automatizada, objetos e suas propriedades a partir de imagens, encontrando e extraindo padrões, relações e estruturas do alvo de estudo. O processo é desenvolvido em linguagem computacional com a finalidade de, em um curto intervalo de tempo, se obter uma classificação satisfatória – i.e., ter uma quantidade de ‘acertos’ igual ou maior que a de um humano treinado (geólogo) e com expertise em sistemas de classificação convencionais. Como referência para treinar o algoritmo, desenvolvemos uma

classificação morfológica para grãos de zircão, inspirada na técnica de HRHMA para categorizar grãos euédricos e subédricos, aliada a propostas de Krumbein (1941) e Powers (1953) para classificar grãos anédricos. As adaptações foram feitas para adequar as classes criadas à natureza do banco de dados usados.

## 1.2. Justificativa e objetivos

A utilização do arredondamento é extensamente aplicada no estudo da análise de proveniência sedimentar quando a ênfase é discriminação entre a contribuição de fontes internas e externas à bacia, especialmente para sedimentos e rochas sedimentares que sofreram transporte de tração ou carga de fundo. No entanto, ainda hoje o método mais empregado para classificação morfológica de grãos é a observação direta via lupa binocular ou microscópio óptico, um processo reconhecidamente lento e muito susceptível à subjetividade da percepção de formas do observador humano. Buscando tornar esse processo mais dinâmico e minimizar o viés decorrente da influência humana, essa dissertação ilustra os desafios e avanços da implementação de técnicas de DL para tratar imagens de grãos (zircão) e avaliar o arredondamento durante o ciclo sedimentar. DL permite que arquiteturas computacionais, compostas de múltiplas camadas de processamento, “aprendam” representações de dados com múltiplos níveis de abstração (LeCun et al., 2015).

Adotamos o método CNN pelo sucesso em aplicações práticas para detecção de objetos em visão computacional, e seu nome “*Convolutional*” indica uma rede que emprega uma operação chamada de convolução – um tipo específico de operação linear (Goodfellow et al., 2016). Ao utilizar CNNs é desejável que o banco de dados (imagens) seja volumoso e representativo das classes que pertencem. Outro requisito desejável é que as categorias do conjunto de dados sejam igualmente distribuídas, i.e., que o banco de dados seja balanceado. Tomando-se como exemplo um conjunto de dados distribuídos em 4 classes, 25% desses dados deveriam se distribuir em cada uma das categorias discriminadas. Entretanto, faz-se necessário mencionar que bancos de dados ‘naturais’ normalmente são desbalanceados, de forma que o cientista deve considerar a utilização de artifícios para reverter ou minimizar esta condição.

Por isso, essa dissertação teve como objetivo central desenvolver sistemática para análise de proveniência sedimentar a partir da classificação morfológica automática de grãos em sucessões siliciclásticas, com ênfase no hábito mineral e as modificações

ocorridas devido ao transporte sedimentar que é refletido pelo grau de arredondamento dos grãos.

A análise de proveniência sedimentar pode ser tornar mais robusta através da adição de estudos do arredondamento em grãos de zircão datados. Esta ferramenta agrega valiosas informações na identificação de fontes externas cristalinas e a relação com reciclagem sedimentar ocorrida em bacias sedimentares de longa duração. A importância deste estudo não se restringe apenas à compreensão da história geológica da bacia sedimentar, mas também tem implicações importantes para a exploração de recursos naturais. Em especial, destaca-se a identificação de áreas promissoras para a exploração de petróleo e gás natural, bem como para a avaliação da qualidade dos reservatórios. A aplicação de técnicas de DL para análise de arredondamento de grãos pode, portanto, trazer benefícios para a indústria de exploração de petróleo e gás, ao permitir uma análise mais rápida da proveniência sedimentar e uma melhor avaliação do potencial dos reservatórios.

Os objetivos específicos foram:

1. Avaliar a consistência da base de dados buscando equilíbrio entre as instâncias das classes utilizadas para treinamento das CNNs;
2. Analisar modelos a partir de diferentes hiperparâmetros, buscando aqueles que apresentam maior acurácia na avaliação em dados de teste;
3. Comparar as classes obtidas por modelos computacionais e aquelas obtidas por interpretação de sedimentólogos;
4. Avaliar a sistemática desenvolvida quanto à adequação relacionada ao estudo de proveniência sedimentar.

Embora consideremos que este projeto ainda se encontra em construção, almejamos ampliar/ajustar as estratégias aqui desenvolvidas para classificar a morfologia de outros minerais detríticos, como quartzo e feldspatos, como forma de agregar informações ao estudo de proveniência sedimentar de sucessões siliciclásticas.

## 2. REVISÃO BIBLIOGRÁFICA

A partir da segunda metade do século 20, o crescente aumento no preço dos derivados de petróleo promoveu uma aceleração no investimento da indústria petrolífera em busca de novas reservas de combustíveis fósseis e gás natural. Como consequência, a indústria ampliou a distribuição de recursos para pesquisa científica em bacias-reservatórios. Nas bacias intracratônicas paleozoicas, parte significativa dos investimentos estiveram direcionados a estudos de proveniência sedimentar objetivando identificar as fontes responsáveis pelo preenchimento sedimentar, sejam externas à bacia ou mesmo internas (reciclagem ou ‘canibalismo’). Um dos métodos mais difundidos em estudos de proveniência é a análise morfológica de grãos em rochas siliciclásticas (Blatt 1967; Johnsson 1993). Historicamente, esses estudos têm sido desenvolvidos em quartzo detrítico por sua abundância em rochas sedimentares e reconhecida resiliência física e química. Mais recentemente, os estudos voltaram-se para os minerais pesados resistentes, cuja morfologia reflete a história do ciclo sedimentar (Mange-Rajetzky, 1995; Dewey e Mange, 1999; Mange e Otvos, 2005; Morton e Hallsworth, 1994, 1999; Morton 2012).

A seguir, apresentamos um breve estado da arte sobre morfologia de grãos detríticos para fins de estudos de proveniência sedimentar, desde quartzo até zircão. Na sequência, apresentamos alguns aspectos do *Machine Learning* mais relevantes para o desenvolvimento deste trabalho.

### 2.1. Morfologia de grãos detríticos

Para a Sedimentologia, a morfologia é uma das propriedades intrínsecas do grão detrítico que descreve a textura de sedimentos e rochas sedimentares terrígenas em termos de forma, arredondamento e esfericidade (Tucker, 2004). Arredondamento é um atributo usado para inferir o desgaste de partículas durante o transporte por carga de fundo em diferentes tipos de fluidos (Folk, 1978; Sagga, 1993; Kasper-Zubillaga et al., 2005), referindo-se diretamente ao grau de agudeza das arestas e vértices de grãos (e.g. Wadell, 1932; Pryor, 1971; Pettijohn et al., 1972; Blott e Pye, 2008). Necessário destacar que o arredondamento se difere da esfericidade, que diz respeito a quanto uma partícula se aproxima de uma esfera. A Tabela 1, abaixo, resume algumas propostas publicadas para quantificar o arredondamento de grãos, enquanto a Figura 1 ilustra um exemplo visual de quantificação.

Tabela 1 – Índices mais comuns de arredondamento (Fonte: Blot e Pye, 2008).

Fórmula	Range	Referência	Notas
$\frac{D_k}{L_w}$	0 a 1	Wentworth (1919)	Razão do diâmetro de curvatura do canto mais agudo até o eixo mais longo do grão
$\frac{D_k}{(L+1)/2}$	0 a 1	Wentworth (1922b)	Razão do diâmetro de curvatura do canto mais agudo até o diâmetro médio do seixo
$\sum_{i=1}^n \frac{D_i}{D_i} = \frac{(\sum D_i)}{D_i}$	0 a 1	Wadell (1932) (primeira fórmula)	Razão média do diâmetro da curvatura de todos os cantos até o diâmetro do maior círculo inscrito
$\sum_{i=1}^n \frac{D_i}{D_i}$	0 a 1	Wadell (1932) (segunda fórmula)	Recíproca da razão média do diâmetro da curvatura de todos os cantos do diâmetro do maior círculo inscrito
$\frac{D_k}{I}$	0 a 1	Kuening (1956)	Razão do diâmetro de curvatura do canto mais agudo até o eixo intermediário do grão
$\frac{D_k}{D_i}$	0 a 1	Dobkins e Folk (1970)	Razão do diâmetro de curvatura do canto mais agudo até o diâmetro do maior círculo inscrito

**As notações foram padronizadas para facilitar a comparação:**  
 $n$  é o número de cantos, incluindo aqueles que o diâmetro é zero  
 $D_r$  é o diâmetro da curvatura de qualquer saliência  
 $D_k$  é o diâmetro de curvatura da saliência com angulo mais agudo  
 $D_i$  é o diâmetro do maior círculo inscrito  
 $L_w$  é o diâmetro mais longo medido pelo canto de angulo mais agudo.  
 $L$  é o comprimento do grão, no plano projetado  
 $I$  é a largura do seixo, no plano projetado

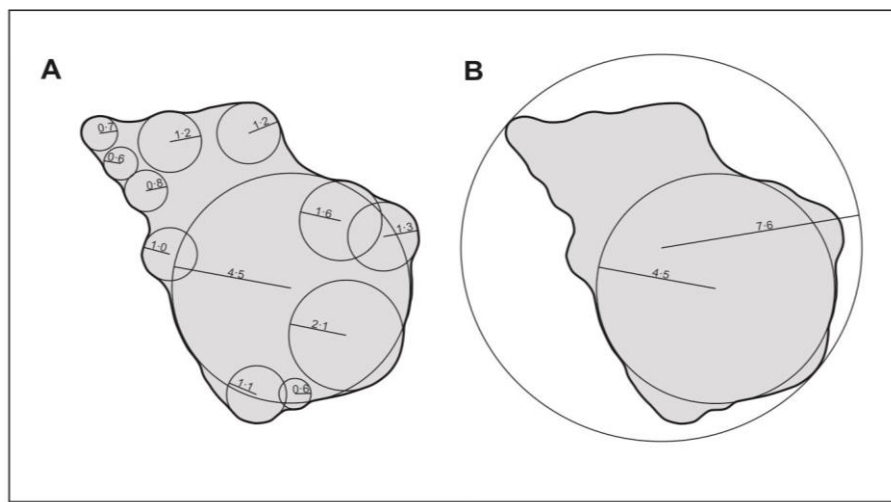


Figura 1 – Exemplo de quantificação de medidas de arredondamento de partículas (A; Wadell, 1932) e de circularidade (B; Riley, 1941). Os números internos representam os raios de cada círculo (Fonte: Blott e Pye, 2008).

Estas estratégias consomem muito tempo quando se considera o grande volume de grãos necessários para análise, além de gerar ambiguidades. Como resultado, essas tentativas foram gradualmente substituídas por escalas de classificação comparativa visual, sendo uma das primeiras a de Russel e Taylor (1937). Estes pesquisadores dividiram as amostras de estudo em cinco classes utilizando a fórmula para arredondamento definida no método de Wadell (1932). Pettijhon (1949) modificou a classificação de Russel e Taylor (1937) propondo classes em escala geométrica (log na base 0.6) em conjunto com o esboço de partículas representativas de cada classe. Powers (1953) modificou a classificação de Pettijhon (1949) utilizando uma escala de log na base 0.7, adicionando às classes já existentes a classe “muito angulosa” como forma de

melhorar a representatividade de grãos com baixos graus de arredondamento (Figura 2). A classificação de Powers ganhou popularidade desde então, atualmente sendo amplamente usada de forma adaptada (p.ex., Andò et al., 2012).

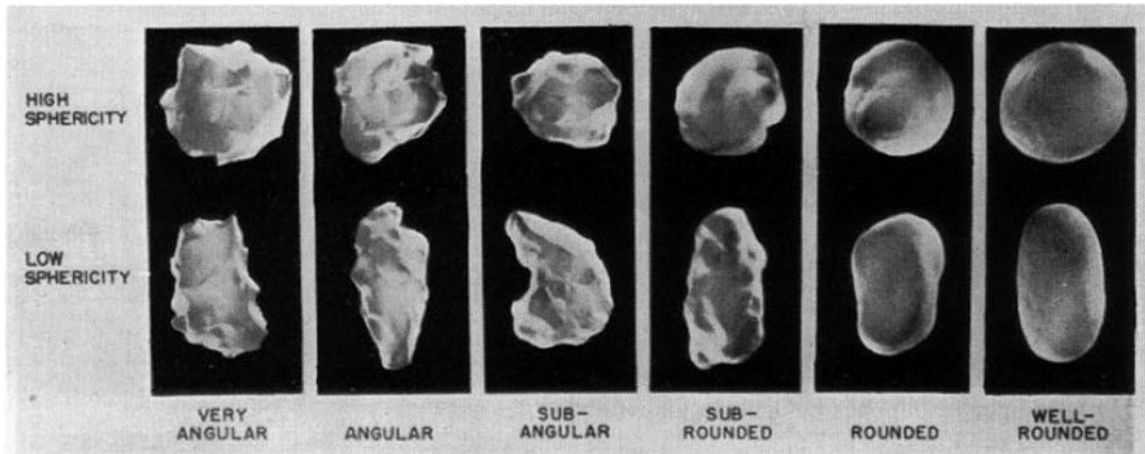


Figura 2 – Escala de arredondamento por comparação visual (Fonte: Powers, 1953).

Como avanço mais recente, Blott e Pye (2008) propõem uma forma otimizada de calcular as medidas geométricas (razão de alongação, razão de achatamento) de partículas a partir da análise computacional de imagens oferecendo, assim, um novo método para quantificar a forma de grãos detríticos, cujos valores são modificados daqueles propostos nos trabalhos anteriores (ver Tabela 2). Os autores também oferecem uma escala visual de imagens para guiar a discriminação das classes morfológicas. No entanto, a complexidade do esquema proposto, que inclui não só uma abordagem profunda sobre o arredondamento, mas também esfericidade (ou ‘circularidade’ cf. os autores) e irregularidade, impediu seu uso corriqueiro pela comunidade sedimentológica.

Tabela 2 – Métodos de classificação por classificação visual baseados no método de medida de Wadell (1932) (Modificado de Blott e Pye, 2008).

	Russel e Taylor (1937)		Pettijohn (1949)		Powers (1953)		Blott e Pye (2008)	
	Limite das classes	Ponto médio aritmético	Limite das classes	Ponto médio geométrico	Limite das classes	Ponto médio geométrico	Limite das classes	Ponto médio geométrico
Muito anguloso	–	–	–	–	0.12–0.17	0.14	–	–
Anguloso	0.00–0.15	0.075	0.00–0.15	0.125	0.17–0.25	0.21	0–0.13	0.09
Subanguloso	0.15–0.30	0.225	0.15–0.25	0.200	0.25–0.35	0.30	0.13–0.25	0.18
Subarredondado	0.30–0.50	0.400	0.25–0.40	0.315	0.35–0.49	0.41	0.25–0.50	0.35
Arredondado	0.50–0.70	0.600	0.40–0.60	0.500	0.49–0.70	0.59	0.50–1.00	0.71
Bem arredondado	0.70–1.00	0.850	0.60–1.00	0.800	0.70–1.00	0.84	–	–

## 2.2. Relevância do zircão em estudos de proveniência



O zircão é um silicato ( $ZrSiO_4$ , nesossilicato) que ocorre amplamente como mineral acessório ( $\ll 1\%$ ) em uma diversidade de rochas crustais, sejam de natureza ígnea ou metamórfica. Sua formação está relacionada aos estágios iniciais da cristalização magmática, mas também pode se originar da recristalização em estado sólido durante processos metamórficos de alto grau, ou ainda por hidrotermalismo. É, portanto, um dos minerais pesados mais comuns na assembléia mineralógica acessória das rochas siliciclásticas (Deer et al., 2013). A similaridade eletrônica do urânio tetravalente ( $U^{4+}$ ) com  $Zr^{4+}$  e outros cátions (p.ex., Th, Hf e elementos terras raras pesados) e sua resiliência física e química faz do zircão um mineral de excepcional aplicabilidade em estudos geocronológicos e de petrogênese crustal (Hoskin e Schaltegger, 2003), traduzindo-se em informações diretas sobre as rochas (ou áreas) fontes que alimentaram a bacia. Portanto, quando presente como fase detrítica, o zircão fornece pistas valiosas sobre a proveniência sedimentar, e ainda permite a inferência da idade máxima de deposição quando bioestratigrafia e vulcanismo não são aplicáveis (Dickinson e Gehrels, 2009). Nesses casos, a idade é estimada a partir da idade do zircão (ou população de) mais jovens obtida.

### **2.3. Proveniência a partir de minerais pesados**

A análise de alta resolução da assembleia de minerais pesados (*High-Resolution Heavy Mineral Analysis, HRHMA*) é fundamentada no reconhecimento de que certas propriedades dos minerais pesados (composição, estrutura interna, cor, características ópticas e hábito mineral) registram as condições de cristalização destes em sua rocha-fonte (Mange-Rajetzky, 1995; Dewey e Mange, 1999; Mange e Otvos, 2005). Esses trabalhos não necessariamente agruparam os minerais detríticos em classes pré-definidas quanto ao arredondamento, mas propuseram uma classificação mais simples – euédricos, subédricos, anédricos (angulosos ou arredondados) como forma de expressar o grau de retrabalhamento das arestas de minerais (ou grupo de) individuais durante o ciclo sedimentar. Turmalina, rutilo, zircão e outros minerais ultra-estáveis (resistatos) foram alguns dos minerais-alvo do método HRHMA. Por sua simplicidade, o método teve significativa aceitação pela comunidade científica (ver trabalhos de Morton e Hallsworth, 1994, 1999; Morton, 2012), hoje conhecido genericamente como ‘estudo de variedades’ de minerais pesados. Considerando o zircão detrítico como um dos minerais pesados mais abundantes em rochas siliciclásticas, o método HRHMA tornou-se complementar à abordagem morfológica já disseminada para quartzo, se mostrando mais simples e,

portanto, mais aplicável do que outras classificações tipológicas, como por exemplo a de Pupin (1980). Um exemplo interessante disso, é a publicação de Andò et al. (2012) que apresenta diversas classes morfológicas ao estudar as variedades de zircão, turmalina e rutilo detríticos presentes em sedimentos fluviais (Figura 3, ver abaixo).

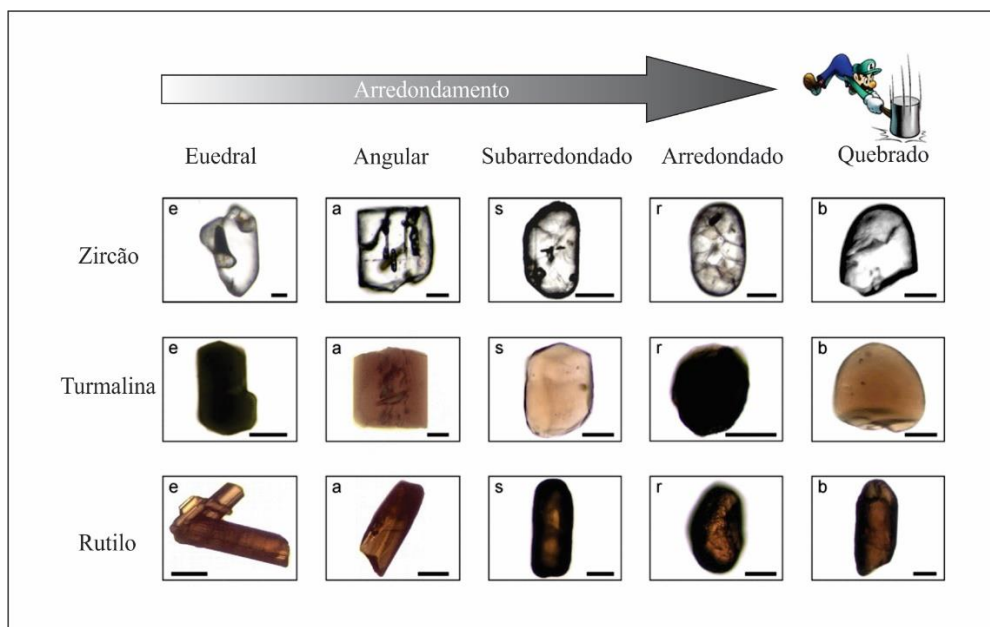


Figura 3 – Classes morfológicas desenvolvidas para minerais pesados resistentes, baseadas em Krumbein (1941), Powers (1953) e, possivelmente, Mange-Rajetzky (1995) e seguidores. O sentido da seta preta indica que quanto mais distante da forma eudral, maior é o grau de arredondamento dos grãos anédricos. A barra preta representa uma escala de 63 µm, em todos os grãos. (Fonte: Andò et al., 2012).

A classificação simplificada de Powers (1953) foi adotada como referência nesta pesquisa de dissertação de mestrado, também incluindo a forma eudral como definida no método HRHMA. Além disso, também adicionamos uma adaptação da classe de grãos quebrados, baseada nos critérios definidos por Krumbein (1941), para avaliar os efeitos do fraturamento em grãos arredondados através de processos naturais.

A breve revisão aqui apresentada sobre as classificações morfológicas, especialmente aquelas aplicadas ao zircão detrítico, mostra a necessidade de adaptações das mesmas dependendo do caso a ser estudado (vide p.ex., Andò et al. 2012). Independente dos ajustes a serem feitos, a morfologia de grãos de zircão é ferramenta importante a ser considerada em estudos de proveniência em bacias sedimentares onde o fator reciclagem precisa ser ponderado junto com as fontes externas.

Por outro lado, esta revisão mostra que ainda não existem pesquisas publicadas com dados de morfologia de zircão detrítico previamente datado, o que é estimulante para o desenvolvimento do mestrado em questão. Este adicional trará mais subsídios para o

entendimento das diversas gerações dos grãos de zircão. Ainda que as idades do zircão possam ser as mesmas obtidas em grãos euédricos e em grãos anédricos arredondados de níveis estratigráficos diferentes, a interpretação mudará significativamente. As fontes das primeiras sucessões sedimentares podem passar de rochas cristalinas a sedimentares nas sucessões mais jovens, com implicações nas mudanças das bordas originais ao longo do tempo. Estas mudanças refletem movimentações tectônicas de escala regional ainda não bem descritas nas bacias intracratônicas brasileiras.

## ***2.4. Deep learning***

*Machine Learning* (ML) é um campo da área de Inteligência Artificial que, por meio de algoritmos, dá ao computador a capacidade de identificar padrões de forma autônoma a partir de um grande conjunto de dados, e fazer análises preditivas. Os algoritmos ML dividem-se em duas categorias: aprendizagem supervisionada e não supervisionada. Algoritmos de aprendizagem supervisionada são algoritmos que aprendem a associar o dado de entrada com algum alvo ou rótulo que é decidido por um “supervisor”. Geralmente é utilizado para tarefas de regressão e de classificação. Já no aprendizado não supervisionado os algoritmos recebem os dados de entrada sem qualquer tipo de associação, aprendendo, ao longo do treinamento, propriedades úteis da estrutura dos dados de entrada, e dando ao final um tipo de resultado. Esses são normalmente utilizados em tarefas de clusterização, quando se divide o conjunto de dados em categorias semelhantes (Goodfellow et al., 2016).

Algumas tarefas apresentam problemas que possuem complexidade e quantidade de dados superiores ao que pode ser tratado com ML, implicando em uma dimensionalidade muito elevada dos atributos de análise. Para esses tipos de problemas criou-se uma ‘sub-área’ chamada de *Deep Learning* (DL), cujos algoritmos são inspirados na estrutura dos neurônios e em funções cerebrais chamadas de *Artificial Neural Network* (ANN; Figura 4). A principal diferença dos algoritmos de DL e ML está na quantidade de camadas ocultas da ANN. Para os problemas mais complexos, redes mais profundas geram modelos que conseguem produzir resultados melhores que os modelos antecessores.

Das técnicas desenvolvidas em DL, duas tem destaque por seu uso mais específico. A primeira é *Recurrent Neural Networks*, amplamente utilizada em tarefas que envolvem linguagem (i.e., reconhecimento de fala e textos, tradutor automático, etc.).

A segunda é *Convolutional Neural Networks* (CNNs) que realiza tarefas relacionadas a imagens ou vídeos.

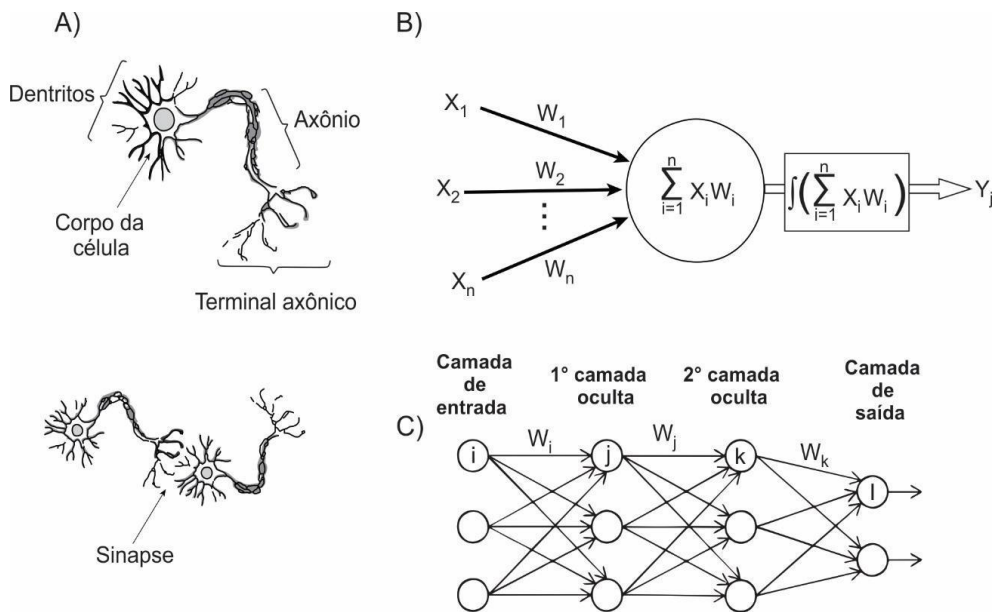


Figura 4 – Representação esquemática de uma *artificial neural network*: (a) neurônio humano; (b) neurônio artificial; (c) sinapse biológica; e (d) sinapse de uma ANN (Fonte: Meng et al., 2020).

## 2.5. Funcionamento de uma CNN

Uma CNN pode ser dividida em duas partes. A primeira parte é a de extração de feições ou características de um dado de entrada (comumente uma imagem), e a segunda é uma ANN tradicional que recebe essas feições e produz um resultado. Antes de abordar a arquitetura de uma CNN, se faz necessário entender o que é e como se dá o funcionamento de uma ANN.

### 2.5.1. Artificial Neural Network

ANN é um método inspirado na rede neural biológica que constitui o cérebro humano. Uma rede neural possui camadas, e quanto maior a quantidade de camadas maior é a complexidade estrutural do modelo resultante, portanto, maior o potencial de resolução de problemas pelo algoritmo. O método utiliza unidades interconectadas dentro

das camadas chamadas de “nós”, “nódulos” ou “*perceptrons*”, onde cada nó consegue emitir um sinal (i.e., número calculado por uma função) em um processo chamado de *forward propagation*. Cada nó apresenta a multiplicação de dois valores: um atributo de entrada  $x_i$  e um peso  $\omega_i$  que serve como um coeficiente, de forma similar ao funcionamento de uma função linear. A vantagem de uma rede neural vem da possibilidade de utilizar uma “função de ativação” que pode transformar uma função linear em uma função não-linear. Ao final da rede temos um resultado de saída contendo a soma de todos os nós (Goodfellow et al., 2016).

A maioria dos algoritmos de DL envolve algum tipo de otimização. Essa tarefa de aprimoramento envolve minimizar ou maximizar algum tipo de função  $f(x)$  ao se alterar  $x$ . Na ciência de dados, a maioria dessas tarefas se dá ao minimizar  $f(x)$ ; neste trabalho o algoritmo desenvolvido segue a mesma lógica. A função  $f(x)$  que se busca minimizar é chamada de *loss function* ou *cost function* (Goodfellow et al., 2016). As redes neurais minimizam a *loss function* e encontram os melhores pesos possíveis quando usada a derivada das funções de ativação, o que pode ser resumido através da técnica de *backpropagation error* (ou, simplesmente, *backpropagation* cf. Bishop, 1995), popularizada a partir do trabalho de Rumelhart et al. (1986). Quando a rede neural é treinada, a *loss function* é calculada em relação a cada um dos pesos da rede, possibilitando a retro-propagação do sinal de erro e a atualização de cada peso individualmente (Figura 5).

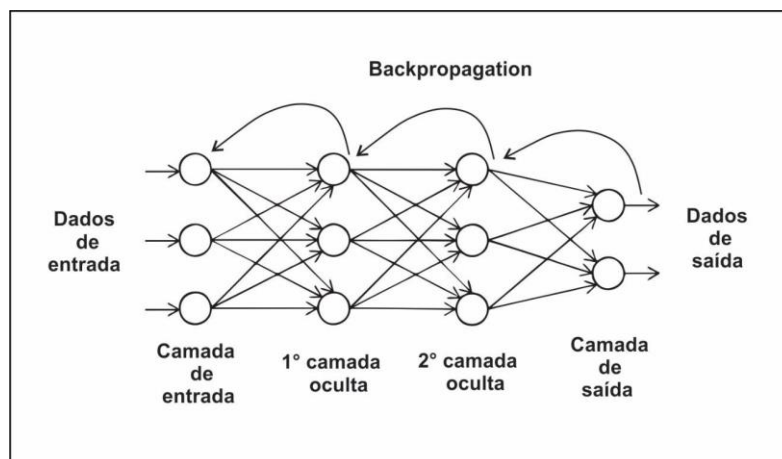


Figura 5 – *Backpropagation* em uma ANN. Os valores dos pesos são atualizados a cada novo treino (sinal) que a rede executa, gerando o melhor resultado e permitindo o “aprendizado” da rede. (Adaptado de Meng et al., 2020).

### 2.5.2. A arquitetura de uma CNN

Uma CNN funciona tal qual uma ANN, tendo como entrada uma imagem que recebe como sinal três tensores: uma imagem com H linhas, W colunas e L canais (em geral, os canais de cores). A imagem passa por diversos pontos chamados de ‘camadas’ que são discriminadas como *convolutional layer*, *pooling layer* e *fully connected layer*. Usando o exemplo abstrato da Figura 6, temos que cada caixa da imagem é uma camada e  $x^1$  é o sinal de entrada. Logo,  $w^1$  é o primeiro peso calculado na primeira camada e o resultado dela seria  $x^2$ , i.e., uma imagem resultante que serve como um “mapa” das feições e características principais da imagem de entrada. Esse “mapa” passa a ser o sinal de entrada para a segunda camada  $w^2$ , e assim sucessivamente. O produto resultante de uma camada de extração é chamado de *feature map*. O último resultado de saída é  $x^L$  que corresponde a um *feature map* da CNN que passa pela última camada  $w^L$  (Wu, 2017). É nessa camada que se encontra a rede neural tradicional, e onde se utiliza uma função de ativação para se obter o resultado desejado.

$$x^1 \longrightarrow \boxed{w^1} \longrightarrow x^2 \longrightarrow \dots \longrightarrow x^{L-1} \longrightarrow \boxed{w^{L-1}} \longrightarrow x^L \longrightarrow \boxed{w^L} \longrightarrow z$$

Figura 6 – Exemplo das camadas de uma CNN (Fonte: Wu, 2017).

O resultado é um número representado por  $z$ , onde se aplica a *loss function* que nada mais é do que uma medida para calcular o desempenho do modelo. A equação de erro quadrático médio para um problema de regressão, descrita abaixo, pode ser entendida como exemplo dessa função. A saber:

$$z = \frac{1}{2} \|k - x\|^2$$

onde  $t$  é a resposta esperada e  $x^L$  é a predição (Wu, 2017).

Dependendo do problema, outros métodos de calcular o desempenho podem ser escolhidos, sendo que o mais comumente utilizado é o método *stochastic gradient descent* (SGD), que constitui um método iterativo para encontrar os melhores pesos. Um método iterativo é um procedimento matemático que usa um valor inicial para gerar uma sequência de soluções aproximadas de melhoria para uma classe de problemas (nesse caso, *loss function*), em que a enésima aproximação é derivada das anteriores. Nesse

método, o sinal de entrada  $x^l$  segue toda a CNN até encontrar um  $x^L$ , comparado ao resultado alvo  $t$ . Após isso, ocorre uma *backpropagation* do sinal de erro e atualização de cada peso individualmente até o mínimo global atingir o melhor (e menor) resultado possível (Figura 7). SGD é calculado pela fórmula:

$$\omega^i \leftarrow \omega^i - \eta \frac{\partial z}{\partial w^i}$$

onde,  $\eta$  é o *learning rate* e  $\frac{\partial z}{\partial w^i}$  é a derivada parcial da função que mede a variação de  $z$  em respeito às mudanças nas diferentes dimensões de  $w^i$ . A seta  $\leftarrow$  sinaliza a atualização do peso de um tempo  $t$  para  $t + 1$  (Wu J., 2017).

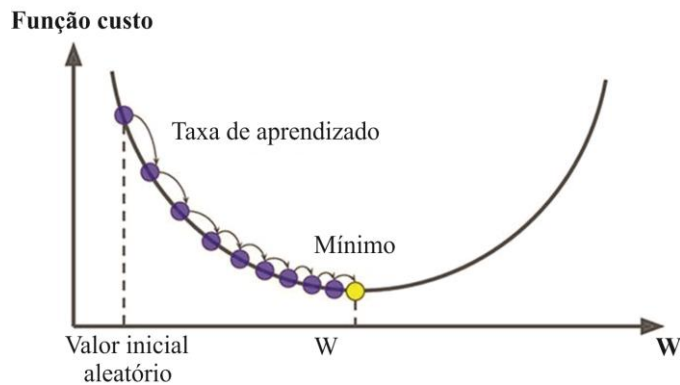


Figura 7 – *Cost function* convergindo para o menor valor possível (mínimo global), conforme a sua *learning rate* (Fonte: Bhattarai, 2018).

Os pesos calculados são chamados de parâmetros e se referem à parte matemática que o algoritmo aprende ao fazer o treinamento. Existem também parâmetros que não são aprendidos pelo algoritmo, mas precisam ser definidos pelo cientista uma vez que podem afetar o desempenho do modelo de várias maneiras e, portanto, são importantes para otimizar o modelo durante o processo de treinamento. Em ML, esses parâmetros são referidos hiperparâmetros. Alguns hiperparâmetros relevantes nesse projeto são: quantidade de *epochs*, tamanho do *batch*, *learning rate*, *momentum*, coeficiente de confiança do objeto, e o limiar de *Intersect Over Union* (IOU). O algoritmo de ML pode fazer a iteração dos dados ao longo do treino de duas formas: (1) por cada exemplo de entrada (e.g., uma imagem de um *mount* com vários grãos de zircão), ou (2) por *batches* (lotes) de imagens que são processadas em conjunto pelo algoritmo. Essa prática acontece para deixar o processamento mais eficiente e pode afetar também o desempenho do modelo resultante. Em geral um tamanho de *batch* maior pode levar a um processo de treinamento mais rápido, mas, caso muito grande, pode afetar negativamente a precisão

do modelo. Quando o algoritmo faz uma iteração completa por todos os *batches*, i.e., por todo o conjunto de dados de treinamento, o algoritmo realizou uma *epoch* (Goodfellow et al., 2016). *Momentum* é um valor designado para acelerar a convergência de técnicas de otimização baseada em gradiente (i.e., SGD). É um valor projetado para acelerar o aprendizado em direções de baixa curvatura, sem se tornar instável em direções de alta curvatura da *loss function* (Goodfellow et al., 2016). O coeficiente de confiança e o IOU serão explicados mais a frente ao longo da revisão bibliográfica.

Essas formas de atualizar os pesos por meio de um algoritmo durante o treinamento, é chamado de otimizadores. O otimizador mais utilizado, já citado, é o SGD, mas existem diversos outros. Um outro otimizador muito utilizado na ciência de dados é o Adaptive Moment Estimation (ADAM) que é baseado na ideia de um *learning rate* adaptativo de acordo com o parâmetro calculado em cada passo (ver Kingma e Ba, 2015).

Uma prática comum na área de ciência de dados é a divisão dos dados em três conjuntos, chamados de treinamento, validação e teste. Esta prática é importante para garantir que o conjunto de dados está sendo treinado pelo algoritmo e o modelo resultante está sendo avaliado de maneira correta. Um conjunto de dados de treinamento é um conjunto de dados de entrada usado pelo algoritmo para ajustar os parâmetros (os pesos) de um modelo de ML. Esse conjunto pode ser rodado pelo algoritmo, que usa os exemplos para aprender a identificar padrões nos dados. Um conjunto de dados de validação é um conjunto de dados usado para avaliar o desempenho do modelo que está sendo gerado durante o processo de treinamento, além disso esse conjunto é usado para ajustar hiperparâmetros. Esse processo é importante para evitar que o modelo resultante tenha um *overfitting* no conjunto de treinamento, dando a ilusão que o resultado tem um bom desempenho, mas na verdade apresentando um desempenho ruim em dados nunca vistos. Esse *overfitting* pode ocorrer quando os valores das *loss functions* dos dados de treinamento e validação apresentam uma diferença de valor muito considerável (Goodfellow et al., 2016). Existem estratégias de atenuar o efeito de *overfitting* ao longo do treinamento do modelo, uma dessas formas é chamada de *early stopping*, uma técnica de regularização onde se aplica uma penalização no treinamento. Mais detalhes sobre como o *early stopping* foi aplicado é descrito no capítulo de metodologia. Um conjunto de dados de teste é um conjunto de dados usado para avaliar o desempenho do modelo final após o treinamento. O conjunto de teste é usado para estimar a capacidade do modelo de generalizar para dados novos e é, geralmente, o conjunto de dados mais importante para avaliar o desempenho do modelo.



### 2.5.3. Tipos de camadas

#### 2.5.3.1. Convolutional Layer

Em uma *convolutional layer* existem *kernels* que realizam a operação de convolução. *Kernels* são filtros usados para extrair os aspectos ou feições dos dados de entrada, i.e., constituem uma matriz que se move ao longo dos dados realizando o produto escalar ao longo da região escaneada. O movimento do *kernel* é chamado de *stride*; se o filtro tem um *stride* igual a 2, ele se move a cada 2 números ao longo da matriz que compõe a imagem (Figura 8). Na figura, o tamanho espacial da imagem resultante do *kernel* é menor que o de entrada, e isso é válido sempre que o tamanho do *kernel* for maior que 1x1. A principal ideia por trás dos *kernels* em *convolutional layer* é que estando os dados (imagens) em forma de matriz, os grupos de valores são altamente correlacionados formando padrões distintos que são facilmente detectáveis. Outra característica é que as estatísticas das imagens e outros sinais são invariáveis para a localização, ou seja, se um padrão aparecer em uma parte da imagem, pode também aparecer em qualquer lugar, resultando em unidades em posições diferentes compartilhando o mesmo peso e detectando os mesmos padrões em diferentes partes da matriz. Matematicamente, a operação de filtragem realizada por *kernel* dando origem a um *feature map* é uma convolução discreta, batizando, portanto, o nome da rede neural (LeCun et al., 2015).

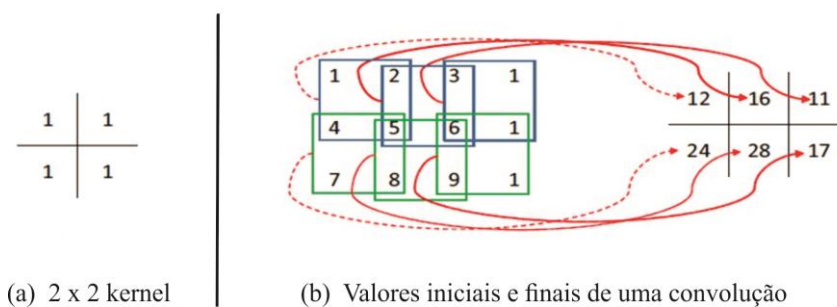


Figura 8 – Ilustração esquemática de uma operação de convolução utilizada nas CNNs, tendo como exemplo um *kernel* 2x2 com um *stride* igual a 1 (Fonte: Wu, 2017).

Depois de cada *convolutional layer* convencionam-se aplicar uma camada não-linear (ou camada de ativação) com o propósito de introduzir uma não-linearidade no sistema que, até então, só tem computado operações lineares (multiplicações e somas

elementares). A função de ativação mais utilizada é a função *ReLU*, capaz de treinar a rede bem mais rápido sem mudar a acurácia do resultado. A saber,

$$f(x) = (0, x),$$

para todos os valores de entrada, produzindo basicamente resultados nos intervalos  $[0, \infty]$ , e transformando os valores negativos em 0 (Hinton e Nair, 2010).

### 2.5.3.2. Pooling Layer

Em uma *pooling layer* ocorre a redução de dimensionalidade dos *features maps*. A principal função dessa camada é efetuar um *downsampling* e diminuir sua variância, além de reduzir a quantidade de parâmetros treinados pela rede e, portanto, minimizar problemas de *overfitting*. É prática comum inserir periodicamente uma *pooling layer* entre as camadas de convolução.

Existem três tipos de *pooling layer*: *MaxPooling*, *SumPooling* e *AveragePooling*. Todos os tipos seguem o mesmo princípio de redução de dimensionalidade, mas se diferem na forma como se calcula o valor final. O *MaxPooling* retira o maior elemento de determinada região da matriz e submete este objeto a um *stride* para, em seguida, aplicar uma nova operação, da mesma forma que os *kernels* nas camadas de convolução. Os outros tipos de *pooling layers* fazem exatamente o que o nome indica, ou seja, *SumPooling* retira a soma dos elementos e *AveragePooling* aplica a média aritmética aos elementos (Lecun et al., 1998).

### 2.5.3.3. Fully Connected Layer

A última camada de uma CNN é normalmente representada por uma ANN comum, adicionada para facilitar o aprendizado do algoritmo em combinações não lineares das feições de alto nível, quando é possível aplicar uma função de ativação (função *softmax*) para classificações múltiplas. No processo de transição de uma *pooling layer* para uma *fully connected layer* ocorre o processo de *flattening* que, como expressa o nome, implica no achatamento de um *feature map* em matriz, transformando-a de números em um vetor contínuo. Esse tipo de rede recebe cada número desse vetor contínuo em um nódulo diferente, e todos os nódulos são conectados, exatamente como

em uma ANN. O dado resultante é classificado após passar pela função de ativação (Lecun et al., 1998).

## 2.6. Transferência de Aprendizado

Para que o treinamento de um algoritmo utilizando CNNs seja correto a literatura dita que é necessário um conjunto de dados de treinamento, de validação e de teste bem representativos e volumoso (10.000 objetos para cada classe) de uma mesma população e possuindo iguais distribuições. Quando há uma mudança nesse conjunto de dados, seja na frequência ou na “representatividade” dos objetos, é necessário que se proceda um novo treinamento. Quanto maior o número de objetos (banco de dados) e quanto maior a CNN em que é treinada, maior é o tempo necessário para se treinar, a partir do zero, um novo modelo de ML. Obviamente, deve-se considerar também o desempenho do *hardware* em que é feito o treinamento. Em um cenário de *big data* pode ser inviável treinar sempre do zero, toda vez que um conjunto de dados atualizado/adaptado/melhorado, já que cada treinamento pode levar dias ou até semanas para ser completado. A transferência de aprendizado é um processo no qual um modelo de aprendizado de máquina treinado para resolver um problema, é reutilizado para resolver outro problema. Isso pode ser útil quando os conjuntos de dados e distribuições utilizados nos dois problemas são diferentes (Pan e Yang, 2009). O objetivo é permitir que o conhecimento adquirido pelo modelo durante o treinamento de uma tarefa seja aplicado para a solução de uma nova tarefa, mesmo que essas tarefas não estejam relacionadas diretamente (Figura 9). Pode-se tomar como exemplo, as primeiras camadas de uma CNN que usualmente detectam as características genéricas de uma rede (como retas, pontas, textura etc.), ao passo que as últimas camadas detectam propriedades mais específicas do problema em questão (Pan e Yang, 2009). A transferência de aprendizado é, portanto, uma ferramenta importante para se ganhar tempo no treinamento de novos problemas através de problemas já resolvidos, como a detecção de objetos comuns.

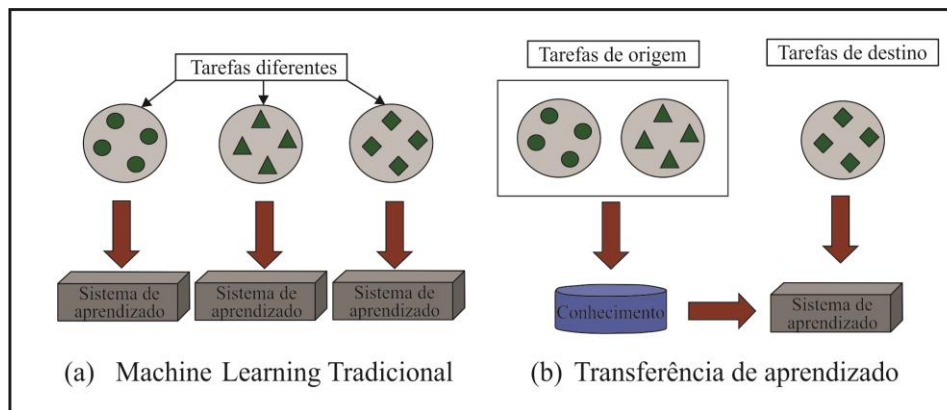


Figura 9 – Ilustração de um processo de aprendizado normal (a), e de um processo de transferência de aprendizado (b). (Fonte: Pan e Yang, 2009)

## 2.7. Métricas de validação e desempenho

Para tarefas de classificação, é comum utilizar ‘matrizes de confusão’ para fazer uma análise da classificação realizada. Matrizes de confusão são tabelas usadas para resumir o desempenho de um modelo de classificação (Kohavi e Provost, 1998). Ela mostra o número de vezes em que o modelo fez previsões corretas ou incorretas para cada classe, sendo útil para entender onde o modelo está tendo dificuldades em fazer previsões precisas. Também é possível calcular os valores de verdadeiro positivo (VP), falso positivo (FP), falso negativo (FN) e verdadeiro negativo (VN) através de uma matriz de confusão, e conseqüentemente, os valores de acurácia, precisão (P), sensibilidade (*recall*/R) e pontuação F1. Na Figura 10 vemos um exemplo de uma matriz de confusão para uma classificação de múltiplas classes e seus respectivos campos.

Quando um modelo de ML ou DL é criado, é indispensável estabelecer uma métrica de validação estatística para comprovar sua eficiência, seja para compará-la entre modelos, seja para melhorar seu desempenho durante o desenvolvimento do projeto (Zhu, 2004). Para casos de classificação de objetos normalmente se utilizam as métricas de acurácia, P, R e a pontuação F1, enquanto que para casos de detecção de objetos é comum a aplicação do *mean average precision* (mAP). A mAP é uma métrica que melhor caracteriza a eficiência de modelos que utilizam caixas delimitadoras, como é o caso deste trabalho.

		Predito			
		C0...Ck-1	Ck	Ck+1...Cn	
Classificação original	Ck+1...Cn	VN	FP	TN	VN true negative
	Ck	FN	VP	FN	VP true positive
	C0...Ck-1	TN	FP	TN	FN false negative
					FP false positive

Figura 10 – Matriz de confusão de uma classificação de múltiplas classes. Os campos apresentados utilizam como referencial a categoria da coluna do meio, onde as classificações corretas daquela categoria são chamadas TP e as classificações corretas de outras categorias são chamados TN.

Para entender a métrica mAP é preciso entender os conceitos de:

- Verdadeiro Positivo (VP): quando a detecção correta é efetuada por uma caixa delimitadora
- Falso Positivo (FP): quando a classificação é incorreta, ou ocorre em uma região totalmente fora da caixa delimitadora original, ou que tenha uma pequena sobreposição com a caixa delimitadora original.
- Falso Negativo (FN): quando um objeto dentro de uma caixa delimitadora original não é detectado.

O Verdadeiro Negativo (VN) não pode ser aplicado no contexto de detecção de objetos. De forma simplificada, o VN pode ser entendido como ‘o algoritmo afirmando corretamente que a área escaneada não contém um objeto’, ou seja, seria o mesmo que prever o *background*, tornando a quantidade de VN muito vasta para ser possível contabilizar (Padilla et al., 2019). A definição de VP, FP, FN na detecção de objetos é exemplificada na figura 11.

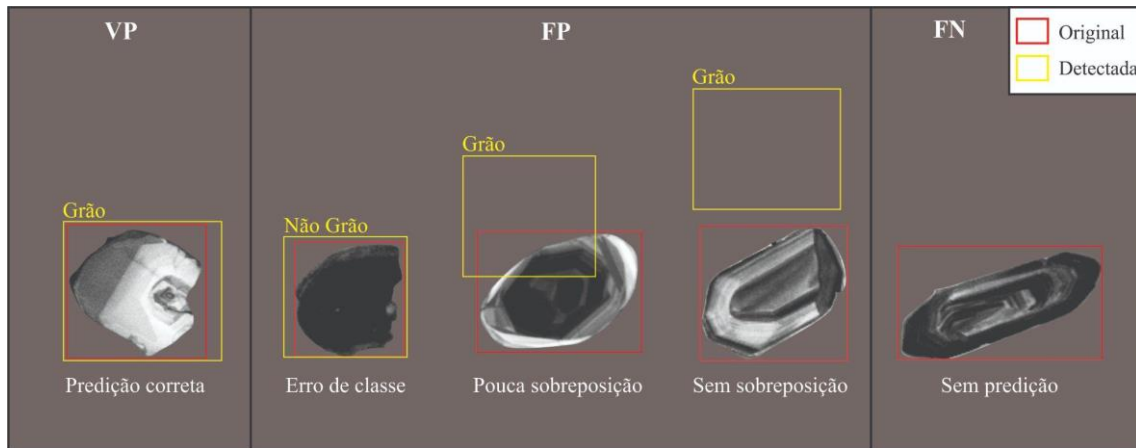


Figura 11 – Ilustração de imagens de grãos de zircão retratados por sua resposta à catodoluminescência. Notar que nesse tipo de imageamento eletrônico destaca-se a estrutura interna de crescimento deste mineral, o que pode ser um viés para a classificação morfológica automatizada (ver capítulos seguintes). Da esquerda para a direita temos uma sequência esquemática de: uma detecção correta (VP) onde se detectou a classe “Grão”; três detecções incorretas (FP) onde se detectou, na sequência, um erro de classe “Não grão”, pouca sobreposição (i.e., limiar abaixo do padrão) e nenhuma sobreposição, onde a caixa delimitadora circunscreve o grão no *background*; uma ‘não predição’ (FN).

Em modelos de detecção de objetos que utilizam caixas delimitadoras é necessário estabelecer o que é uma detecção ‘correta’ e ‘incorreta’ com base na interseção das caixas delimitadoras (Figura 12). Essa interseção é abreviada por IOU (em inglês, *Intersection Over Union*) que consiste no limiar definido pela área de sobreposição entre a caixa prevista e a caixa original. Esse limiar  $t$  é decidido pelo usuário, sendo que comumente adota-se 0,5 como valor padrão. Se o  $IOU \geq t$ , então a detecção é considerada correta; se  $IOU < t$ , então a detecção é considerada incorreta. A interseção é uma medição baseada no índice de similaridade de Jaccard, entre dois conjuntos (Padilla et al., 2019).

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{área de sobreposição}}{\text{área da união}}$$

Figura 12 – Descrição e fórmula da IOU. O quadrado com a borda verde representa a caixa delimitadora original, e o de borda vermelha representa a caixa delimitadora detectada (Fonte: Padilla et al., 2019).

Com as métricas descritas, é possível avaliar e validar os métodos de detecção de objetos visto que eles têm por base a relação entre sensibilidade  $R$  e a precisão  $P$ , definida como F1. A saber:

$$R = \frac{VP}{VP + FN}$$

$$P = \frac{VP}{VP + FP}$$

$$F1 = \frac{2 * P * R}{P + R},$$

onde R descreve a habilidade que o modelo tem de encontrar as classificações originais (todas as caixas delimitadoras originais), e P é a porcentagem de predições positivas corretas.

Ao final da etapa de detecção obtêm-se um gráfico que ilustra a relação entre as métricas R e P (Figura 13), de onde é extraído o parâmetro AP (ver equação abaixo). AP representa a média da precisão entre todos os valores de sensibilidade, os quais variam de 0 a 1. O valor de AP corresponde à área da curva P-R que, por vezes, pode apresentar “ruídos” que atrapalham o cálculo de AP. Nesses casos utiliza-se uma aproximação alternativa para AP em que se interpola os valores de precisão (P) para cada ponto de sensibilidade (R), utilizando o valor de precisão máximo cujo valor de R deve ser maior ou igual a  $R_{n+1}$  (Padilla et al., 2019). Em outras palavras, em vez de usar a precisão que foi realmente observada em um ponto de corte específico  $k$ , utiliza-se a precisão máxima observada à direita do ponto de maior  $R$ . Esse tipo de interpolação é descrito por:

$$AP_{all} = \sum_n (R_{n+1} - R_n) P_{interp}(R_{n+1}),$$

Onde  $AP_{all}$  é o AP da área total da curva e  $P_{interp}(R_{n+1})$  pode ser calculado através de:

$$P_{interp}(R_{n+1}) = \max_{\tilde{R}: \tilde{R} \geq R_{n+1}} P(\tilde{R}).$$

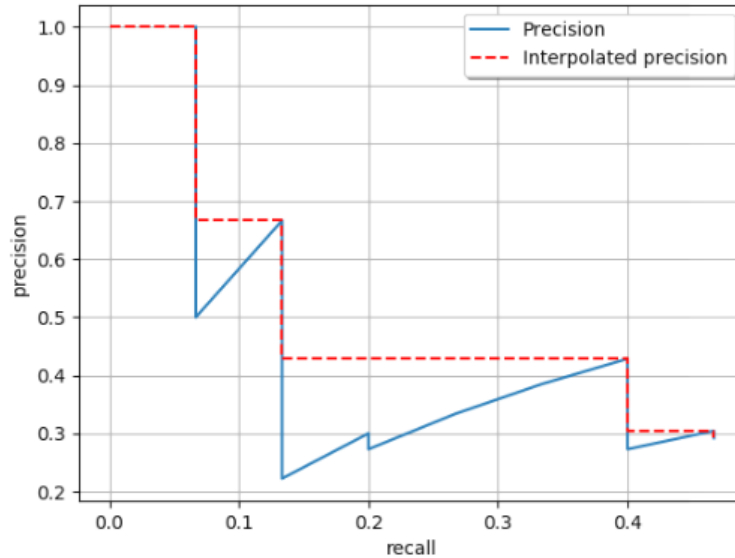


Figura 13 – Exemplo de interpolação de todos os pontos da curva precisão-sensibilidade, ou P-R. O algoritmo contabiliza sucessivamente P em função de R. Em pontos de “ruídos”, utiliza-se P máximo em função de um R maior (à direita no gráfico), ou seja, maior ou igual a  $R_{n+1}$ . (Fonte: Padilla et al., 2019).

Após o cálculo de AP individual para cada classe, tira-se a média de AP (mAP) sobre todas as classes computadas, o que representa o desempenho do modelo (Padilla et al. 2019) segundo a equação:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

onde  $AP_i$  é o AP na  $i$ th classe, e  $N$  o número total de classes avaliadas.

Geralmente escolhe-se um valor de IoU (normalmente o de 0,5) para fazer a interpolação da curva P-R ao longo de 11 pontos (0,0 até 1,0 de sensibilidade em um intervalo de 0,1) gerando então o mAP@0.5. Para deixar esse valor mais robusto, passou-se a utilizar uma média desses valores de mAP no intervalo de 0,5 – 0,95 a cada 0,05 de IoU, totalizando 111 pontos. O nome dado para essa média e o que consideramos neste trabalho como a métrica definitiva é mAP@0.5:.95.

## 2.8. Detecção de objetos

A detecção de objetos é uma função importante da visão computacional. Nada mais é do que a identificação de objetos de uma classe pré-definida a partir de imagens ou quadros/seg de vídeos. Com o avanço de conhecimento nas técnicas de DL, os métodos



de detecção de objetos usando CNNs passaram a ganhar maior destaque por conseguir executar tarefas mais complexas. Hoje, uma CNN de detecção de objetos é composta por três partes: *backbone*, *neck* e *head*, onde cada parte é composta por um algoritmo com uma função diferente dentro da rede (ver Bochkovsky et al., 2020). A parte do *backbone* se encarrega de extrair as principais características e feições da imagem ao longo das camadas, dando origem ao *feature map* que será utilizado na sequência. A função do *neck* é fundir todas as características extraídas da parte anterior (*backbone*) e passar adiante. Por fim, a função do *head* é prever, além de localizar corretamente e classificar o objeto estudado e onde as arquiteturas classificadoras atuam. As arquiteturas de classificadores na predição podem ser divididas em dois tipos. O primeiro tipo funciona em dois estágios distintos, um de localização e outro de classificação propriamente dito. O outro tipo funciona em apenas um estágio, executando duas funções simultaneamente, o que se traduz com ganho de tempo em detrimento de precisão da métrica de validação.

### 2.8.1. Arquiteturas em dois estágios

O *Region-based Convolutional Neural Network* (R-CNN) é um algoritmo de detecção de objetos que se divide em dois estágios (ver Girshick et al. 2014). No primeiro estágio, o R-CNN utiliza a técnica de Busca Seletiva para propor regiões de interesse na imagem de entrada. Essas regiões são então passadas para o segundo estágio, onde são extraídas características utilizando redes neurais convolucionais (CNNs) e, em seguida, classificadas usando máquinas de vetores de suporte (SVMs, ver Cortes e Vapnik, 1995). A figura 14 ilustra um exemplo do processo do R-CNN, no qual (1) uma imagem de entrada é recebida pelo algoritmo, (2) as regiões de interesse são extraídas, (3) características são extraídas usando CNNs e, por fim, (4) as regiões são classificadas usando SVMs. A arquitetura com um dos melhores desempenhos em duas etapas é a *Faster region-based CNN* (ver Ren et al, 2016).

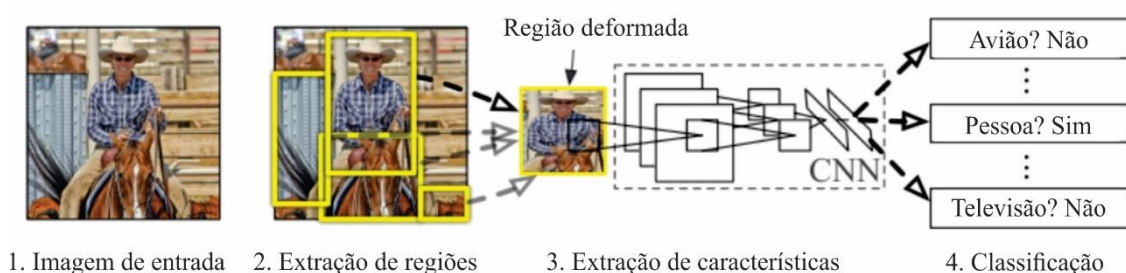


Figura 14 – Etapas de uma arquitetura de dois estágios *region-based*. (Adaptado de Girshick et al., 2014).

## 2.8.2. Arquiteturas em um estágio

Na categoria dos algoritmos em um estágio, tem-se o *You Only Look Once* (YOLO). O algoritmo de detecção de objetos YOLO é um método rápido e preciso para detectar objetos em imagens e vídeos, projetado para processar a imagem inteira em uma única passagem ao invés de processar várias regiões da imagem separadamente e tentar combinar os resultados (Redmon et al. 2016). O YOLO utiliza "*anchor boxes*" ou caixas delimitadoras pré-definidas que nada mais são que retângulos com diferentes tamanhos e orientações usados para ajustar objetos de diferentes formas e tamanhos. Durante o treinamento, o modelo YOLO é apresentado com um conjunto de *anchor boxes* e um conjunto de caixas delimitadoras verdadeiras para os objetos nas imagens de treinamento. O modelo então tenta prever caixas delimitadoras que sejam similares às caixas delimitadoras verdadeiras e bem-ajustadas aos *anchor boxes*, dividindo uma imagem de entrada em uma grade de células  $S \times S$  e, para cada célula, prevendo um conjunto de caixas delimitadoras e probabilidades de classe (Figura 15). As caixas delimitadoras representam a localização e o tamanho dos objetos na imagem, enquanto as probabilidades de classe representam a probabilidade de um objeto pertencer a uma classe específica caso exista um objeto naquela caixa representado por  $\Pr(\text{object})$  (por exemplo, pessoa, carro, cachorro). Para cada caixa delimitadora, o YOLO prevê um nível de confiança da existência de um objeto na célula correspondente para cada classe  $C$ ,  $\Pr(\text{class}_i|\text{object})$ , estima o centro da caixa delimitadora  $(x,y)$  e sua largura e altura  $(w,h)$ .

As informações de todas as caixas delimitadoras da imagem são então reunidas em um tensor de saída unificado com dimensões  $S \times S \times (B * 5 + C)$ , em que  $S \times S$  é o tamanho da grade,  $B$  é o número de caixas delimitadoras por célula e  $C$  é o número de classes possíveis. A constante 5 representa as cinco variáveis  $(\Pr, x, y, w, h)$  para cada *bounding box*. Ao final é aplicado um algoritmo chamado *Non Max Suppression* (NMS) que, por meio das IOU das inúmeras caixas delimitadoras detectadas, seleciona apenas aquelas caixas com IOU acima de um limiar escolhido, ou seja, caixas que se sobrepõem a um certo nível. Por outro lado, suprime as caixas delimitadoras que possuem os menores valores de  $\Pr$ , deixando apenas as caixas com os melhores valores de detecção para cada objeto. Esse 'Pr' reflete a confiabilidade do modelo em relação à caixa delimitadora que contém um objeto, bem como a acurácia do modelo sobre a categorização do objeto.

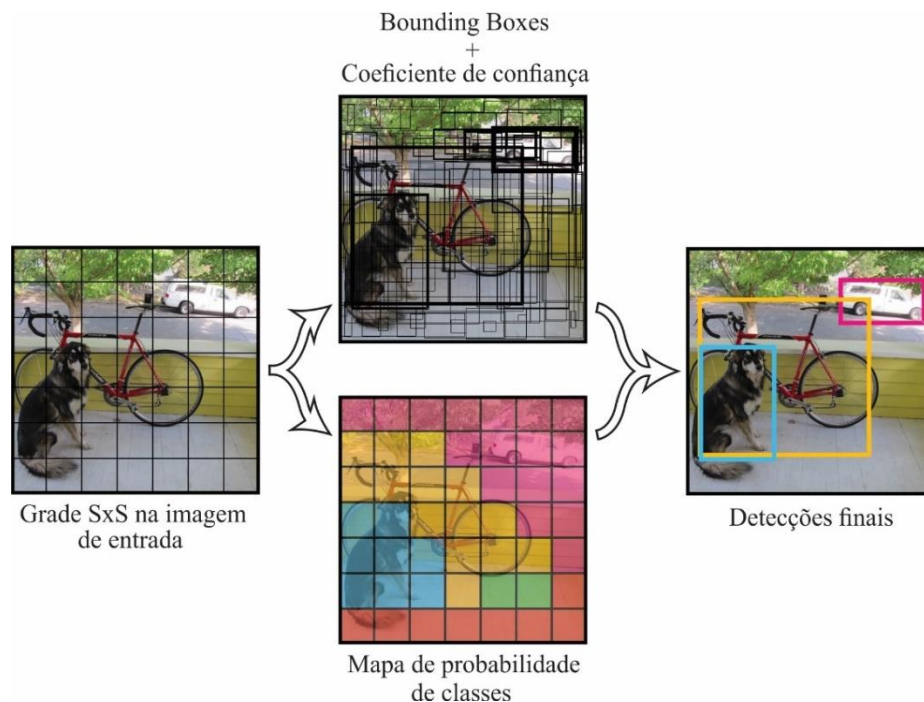


Figura 15 – Exemplo de como funciona a detecção de objetos dentro de uma imagem na arquitetura YOLO. (Adaptado de Redmon et al., 2016).

O algoritmo NMS pode ser utilizado na detecção entre as classes (chamado de NMS padrão), no qual não suprime sobreposições de detecções de classes diferentes. Ou então, pode ser utilizado em todas as detecções (chamado de *agnostic NMS*), onde suprime todas as caixas sobrepostas, deixando apenas os maiores valores de *confidence score*. A definição matemática de *confidence score* é:

$$\begin{aligned}
 \text{box confidence score} &= \Pr(\text{object}) \cdot \text{IoU} \\
 \text{conditional class probability} &= \Pr(\text{class}_i | \text{object}) \\
 \text{class confidence score} &= \Pr(\text{class}_i) \cdot \text{IoU} \\
 &= \text{box confidence score} \times \text{conditional class probability}
 \end{aligned}$$

Onde

$\Pr(\text{object})$  é a probabilidade de uma caixa conter um objeto

IoU é a interseção entre as caixas preditas e a caixa delimitadora

$\Pr(\text{class}_i | \text{object})$  é a probabilidade de um objeto pertencer a classe  $i$  dado que um objeto está presente.

$\Pr(\text{class}_i)$  é a probabilidade de um objeto pertencer a classe  $i$ .

A arquitetura YOLO considera a detecção de objetos como um problema de regressão usando a soma dos quadrados dos erros entre as previsões e os dados originais como sua *loss function*. Essa função é uma soma de três componentes principais e pode ser separada da seguinte forma:

- 1) *Localization loss*: Mede a diferença entre as coordenadas das caixas delimitadoras detectadas e as originais ( $x, y$  para o centro e  $w, h$  para a largura e altura da caixa delimitadora).

$$\lambda_{coord} = \sum_{i=0}^{s^2} \sum_{j=0}^B [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2]$$

$$+ \lambda_{coord} = \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{1}_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2]$$

Onde

$\mathbb{1}_{ij}^{obj} = 1$  se um a  $j$  caixa delimitadora na célula  $i$  é responsável por detectar o objeto, caso contrário 0.

$\lambda_{coord}$  aumenta o peso para a função *loss* para as coordenadas das caixas delimitadoras.

- 2) *Objectness loss*: Mede a diferença entre o Pr e os valores das caixas delimitadoras originais (que são 0 ou 1 dependendo se a caixa contém um objeto ou não).

$$+ \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{1}_{ij}^{obj} (C_i - \hat{C}_i)^2$$

$$+ \lambda_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{1}_{ij}^{noobj} (C_i - \hat{C}_i)^2$$

Onde

$\hat{C}_i$  é a *box confidence score* da caixa  $j$  na célula  $i$

$\mathbb{1}_{ij}^{obj} = 1$  se um a  $j$  caixa delimitadora na célula  $i$  é responsável por detectar o objeto, caso contrário 0.

$\mathbb{1}_{ij}^{noobj}$  indica o complemento de  $\mathbb{1}_{ij}^{obj}$

$\lambda_{noobj}$  aumenta o peso para a função *loss* quando é detectado o *background*

3) *Classification loss*: Mede a diferença entre a probabilidade predita de uma classe (indicando a probabilidade de uma *anchor box* pertencer a uma certa classe) e a caixa delimitadora original (que são vetores indicando a verdadeira classe do objeto). Quando se tem valores de probabilidades em uma classificação, é calculado a *cross-entropy loss* comumente utilizada como *loss function* para tarefas de classificações, na qual é a medida de diferença entre duas distribuições probabilísticas.

$$+ \sum_{i=0}^{S^2} \mathbb{1}_i^{obj} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2$$

Onde

$\mathbb{1}_{ij}^{obj} = 1$  se um a  $j$  caixa delimitadora na célula  $i$  é responsável por detectar o objeto, caso contrário 0.

$\hat{p}_i(c)$  denota o *conditional class probability* para a classe  $c$  na célula  $i$

Cada uma dessas funções somadas forma a equação completa e é possível aplicar pesos a cada uma delas, de acordo com a finalidade da detecção. Se a classificação do objeto for mais importante, coloca-se um peso maior nesta parte (*classification loss*). Se for a detecção de objetos na caixa delimitadora (*Objectness loss*), coloca-se um peso maior nesta parte da equação. A parte de localização do objeto (*localization loss*) nas coordenadas da imagem sempre terá um peso alto, porque essa parte é a mais básica e será a mais importante. A equação final fica desta forma:

$$\begin{aligned} \lambda_{coord} &= \sum_{i=0}^{S^2} \sum_{j=0}^B [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\ + \lambda_{coord} &= \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\ &+ \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{obj} (C_i - \hat{C}_i)^2 \end{aligned}$$

$$\begin{aligned}
& + \lambda_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbb{1}_{ij}^{noobj} (C_i - \hat{C}_i)^2 \\
& + \sum_{i=0}^{s^2} \mathbb{1}_i^{obj} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2
\end{aligned}$$

Todo o processo da arquitetura YOLO, como foi descrito acima, acontece na parte do *head* de uma CNN e pode ser simplificada e ilustrada na Figura 16, abaixo.

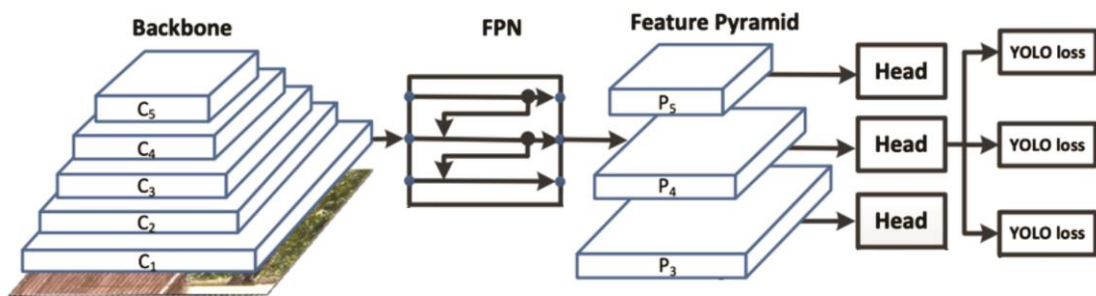


Figura 16 - Estrutura comum da arquitetura YOLO. A etapa de detecção e classificação ocorre na parte *head* onde a ‘YOLO loss’ (a *loss function* da arquitetura) é minimizada (Fonte: Long et al., 2020).

### 3. MÉTODOS

Antes da descrição dos métodos utilizados neste projeto, se faz necessário comentar que para construir um esquema de classificação baseado em ML é necessário que o banco de dados represente bem a população da categoria definida, e que seja de grande proporção - quanto maior a quantidade de dados melhor será o ajuste do modelo aos dados de teste, para servir como base para o modelo que será treinado pelo algoritmo. A definição de **algoritmo** é descrita como ‘uma sequência de instruções bem definidas, normalmente usadas para resolver problemas de matemática específicos, executar tarefas, ou para realizar cálculos e equações’ (p.ex., Medina e Fertig, 2005). Neste trabalho, o termo algoritmo é empregado para os comandos que realizam os cálculos para o treinamento dos dados de treinamento, faz o seu ajuste no conjunto de dados de validação que resulta em um modelo final, onde cada treinamento resulta em um modelo diferente.

Apesar de não ter uma definição específica na literatura, o termo **modelo** remete a uma descrição geral do conceito e de como ele é utilizado (ver <https://learn.microsoft.com/en-us/windows/ai/windows-ml/>). No contexto de ML, modelo quer dizer um conjunto de hipóteses, regras ou equações que são utilizadas para prever ou classificar os dados, i.e., é uma representação matemática de um fenômeno ou processo que pode ser utilizada para fazer previsões, classificar objetos, tomar decisões etc. Aqui, adotamos o termo ‘modelo’ para realizar a classificação dos grãos em categorias tendo, portanto, seu desempenho testado quando comparamos sua resposta com a classificação do mesmo conjunto de dados realizada pelo humano (este autor ou um colaborador, na fase de teste).

A metodologia deste trabalho ocorreu em três diferentes estágios de trabalho, cada um composto de diferentes etapas (Figura 17). O primeiro e segundo estágios ocorreram de forma cíclica, já que, dependendo dos resultados do processamento, era identificado ajustes necessários no banco de dados, resultando em um novo estágio de pré-processamento. Após se obter resultados adequados, seguiu-se para o terceiro estágio, o pós-processamento.

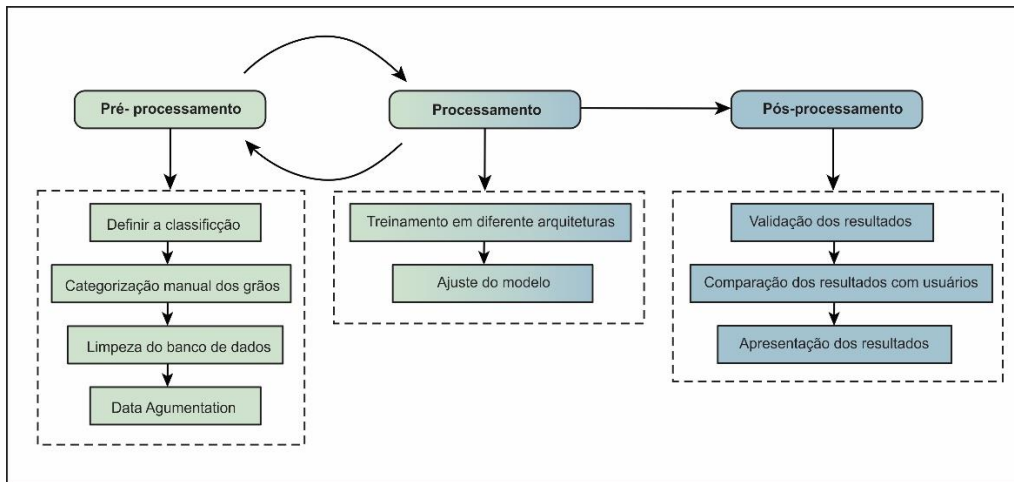


Figura 17 - Fluxograma ilustrando como foi desenvolvido o projeto. O pré-processamento e o processamento ocorreram de forma cíclica. Toda vez que se obtinha um resultado abaixo do desejado, era refeito o pré-processamento do banco de dados afim de se obter valores melhores. O processo de pós-processamento ocorreu após os treinamentos e escolha do modelo final.

### 3.1. Pré-processamento

Com o objetivo definido e metas físicas traçadas, a primeira etapa de desenvolvimento foi analisar visualmente todos os grãos de zircão estudados no projeto PRO-PARNAÍBA e selecionar quais seriam usados para treinar o algoritmo. Considerando que o argumento científico que motivou o interesse no estudo morfológico do zircão detrítico é o de atestar a história de reciclagem na bacia, selecionamos 5.355 grãos de zircão dos arenitos fluviais do Grupo Serra Grande (indiviso) e eólicos da Formação Sambaíba, os quais correspondem respectivamente às unidades paleozoica-triássica basal e de topo da Bacia do Parnaíba. De forma complementar, também usamos um banco de dados menor de imagens de zircões de uma amostra de arenito imaturo do *Graben* de Cococí (Brito Neves, 1998), um dos vários *grabens* cambro-ordovicianos supostamente precursores da sinéclise paleozoica, instalado ao longo do atual limite erosivo sudeste da bacia. A utilização dessa amostra resolveu o problema de baixa representatividade volumétrica de grãos euédricos nos arenitos Serra Grande e Sambaíba. A segunda etapa consistiu na classificação das imagens propriamente dita. Tendo em vista a adequação ao banco de dados usados nesta pesquisa, as categorias selecionadas seguiram os preceitos propostos por Krumbein (1941, grãos quebrados) e Powers (1953, grãos angulosos a arredondados) combinados com a abordagem feita por Mange-Rajetzky (1995, grãos euédricos quando o hábito mineral foi preservado) para minerais pesados ultraestáveis utilizando a técnica *high-resolution heavy mineral analysis* (HRHMA). A classificação grão a grão foi feita ‘manualmente’ e de forma



visual, comparando imagens de luz transmitida com catodoluminescência de forma a garantir uma interpretação robusta do grau de arredondamento.

Na prática, para adequarmos as categorias analisadas com as características do banco de dados usados, inicialmente optamos: 1) diminuição das categorias (mix entre as três classificações anteriormente citadas) para minimizar a subjetividade (humana) na interpretação das formas (Figura 17); 2) diminuição do desbalanceamento entre as categorias, o que poderia complexar o treinamento do algoritmo. Considerando estas premissas, no decorrer do trabalho, notamos que a população de grãos quebrados era volumosa, cerca de 20-30% em relação ao total de grãos estudados. Por se tratar de banco de dados originário dos procedimentos para datação de zircão, onde alguns procedimentos podem quebrar artificialmente os grãos estudados, fica difícil identificar as formas quebradas através do processo natural de transporte sedimentar, como reporta Krumbein (1941). Este é o caso do nosso banco de dados onde é difícil diferenciar grãos quebrados naturalmente daqueles artificialmente, o que tornou a categoria de grãos quebrados ambígua à interpretação do transporte sedimentar.

Todas estas considerações resultaram em três categorias, as quais estão ilustradas nas Figuras 18 e 19. A **categoria 1** inclui grãos com hábito prismático, grãos euédricos a subédricos, incluindo formas que preservam o hábito do zircão com fácies e alongações preservadas ou parcialmente preservadas. A **categoria 2** inclui grãos anédricos arredondados a subarredondados, que podem guardar alongação (semi) original, e grãos arredondados com fraturas discretas sugestivas de serem anteriores à história completa de transporte no ciclo sedimentar. Os grãos que não pertencem essas duas categorias, e que mostram quebrados (angulosidade e/ou fraturamento pronunciados) foram alocados na “**categoria zero**”, i.e., grãos quebrados cujas formas difíceis de serem relacionadas a algum processo geológico, como foi explicado anteriormente.

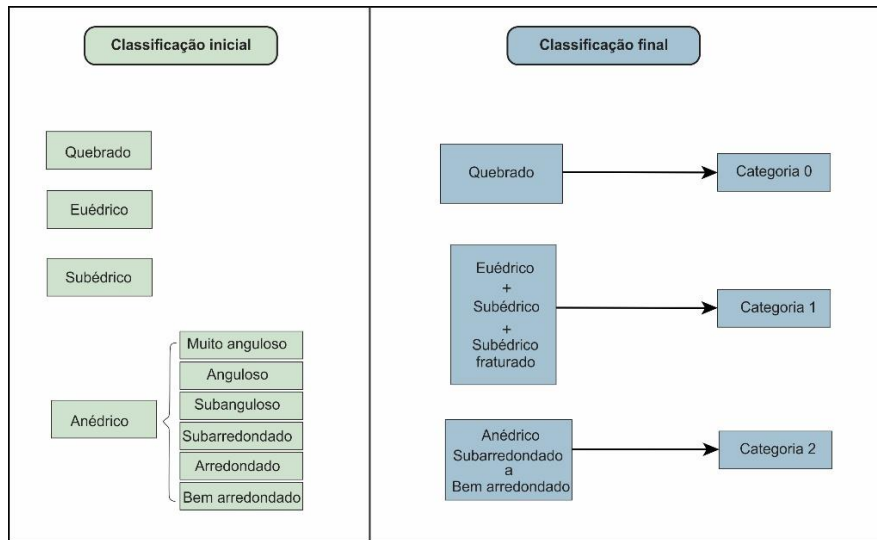


Figura 18 - Fluxograma ilustrando como foi desenvolvida a classificação proposta neste projeto. A ‘classificação inicial’ foi uma tentativa de adotar as definições morfológicas da literatura, enquanto a ‘classificação final’ representa uma combinação de critérios morfológicos consensualmente utilizados na literatura, ajustada às características do banco de imagens utilizado neste trabalho.

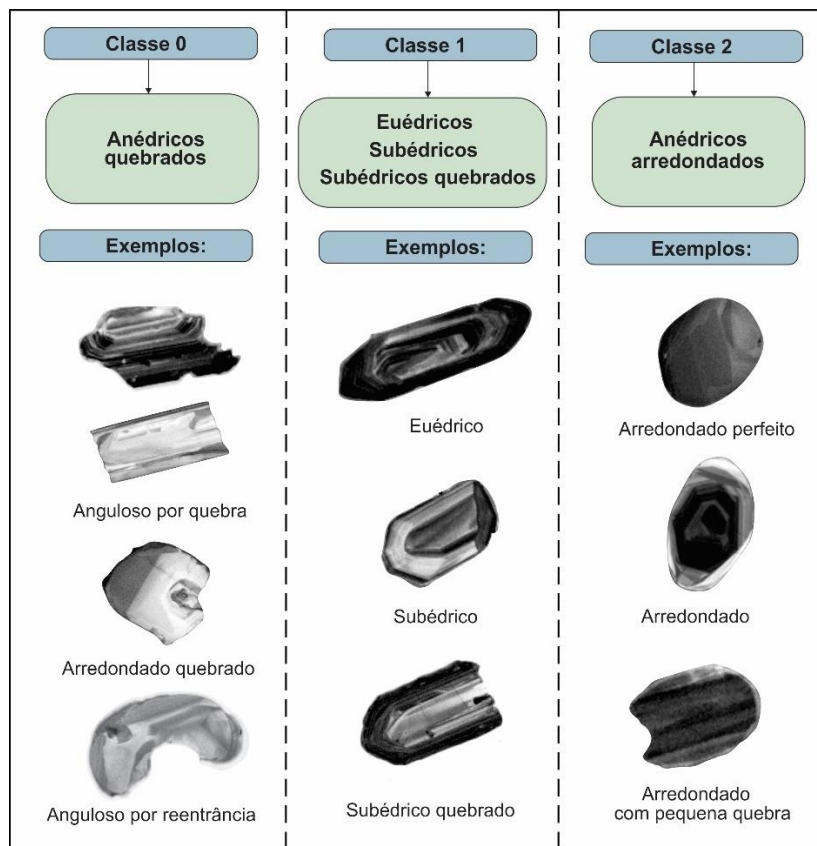


Figura 19 - Exemplos de morfologias de grãos de zircão comentadas no texto, e organizadas nas três categorias definidas para este trabalho.

A terceira etapa resumiu-se à seleção das imagens que iriam compor o banco de dados “limpo” a ser usado no modelo, separando imagens que seriam usadas para treinamento e para validação (80-90%) e para teste (10-20%). Nessa etapa, tivemos que rotular cada grão em uma das três categorias pré-definidas, individualizando-os por caixas delimitadoras. Para balancear a quantidade de indivíduos (grãos) entre as categorias, foram geradas imagens "artificiais" através de um processo chamado *data augmentation*. Esse processo consiste em adicionar cópias discretamente modificadas de imagens já existentes ao conjunto de dados de treino das categorias mais sub-representadas. A categorias zero e 1 se encontravam nesse padrão em relação a categoria 2, o que poderia tornar o treinamento do algoritmo deficiente. O procedimento de “remoção” de grãos para se confundir com o *background* foi feito com um editor de imagens (ver Figura 20).

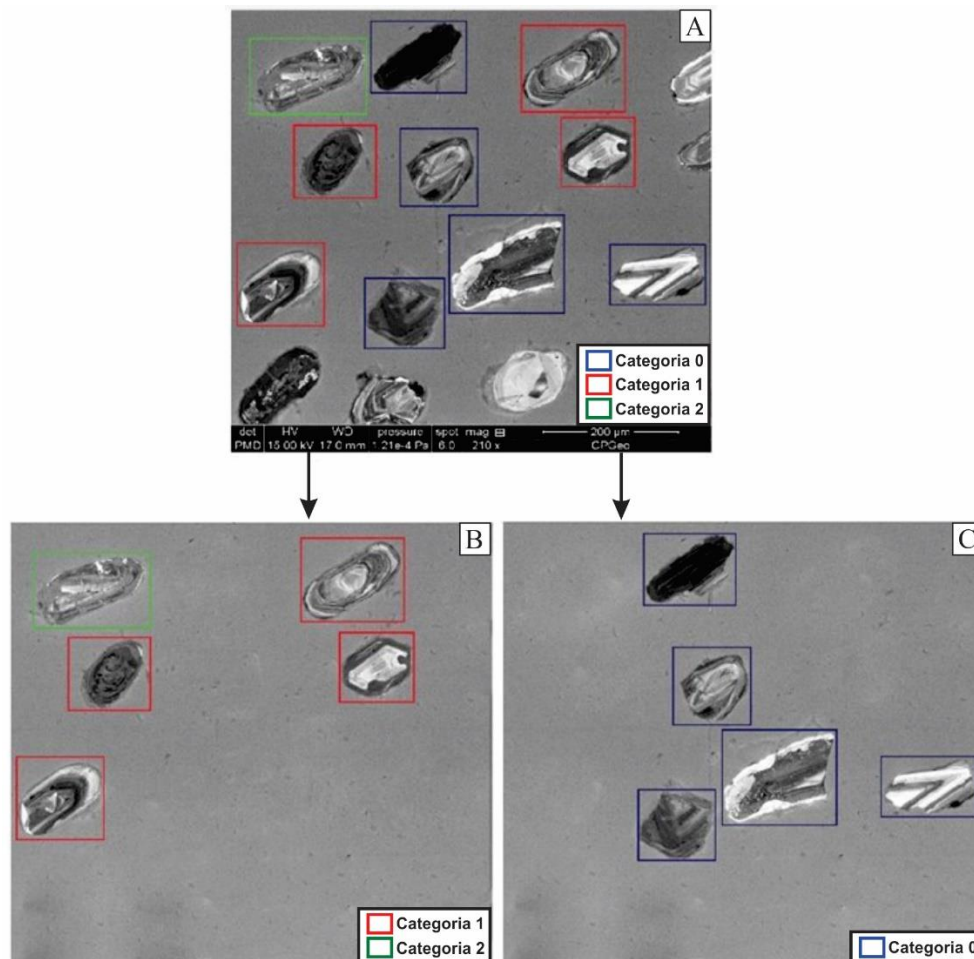


Figura 20 - Representação da “remoção” dos grãos sub-representados (neste exemplo para a categoria zero) e de limpeza da imagem, individualizando-se através de caixas delimitadoras os grãos das categorias super-representadas (à esquerda) e os da sub-representadas (à direita). As cores das caixas delimitadoras descrevem a que categoria cada grão pertence.

Com o processo de *data augmentation* foram “repetidos” 2.291 novos grãos das as três categorias juntas. A maior quantidade de reproduções foi de grãos da categoria zero que estava originalmente desbalanceada. Após a “remoção” dos grãos, as novas imagens foram modificadas através da combinação aleatória de efeitos de (i) rotação 90° (da imagem) no sentido horário ou anti-horário e (ii) espelhamento (da imagem) vertical e/ou horizontal (Figura 21).

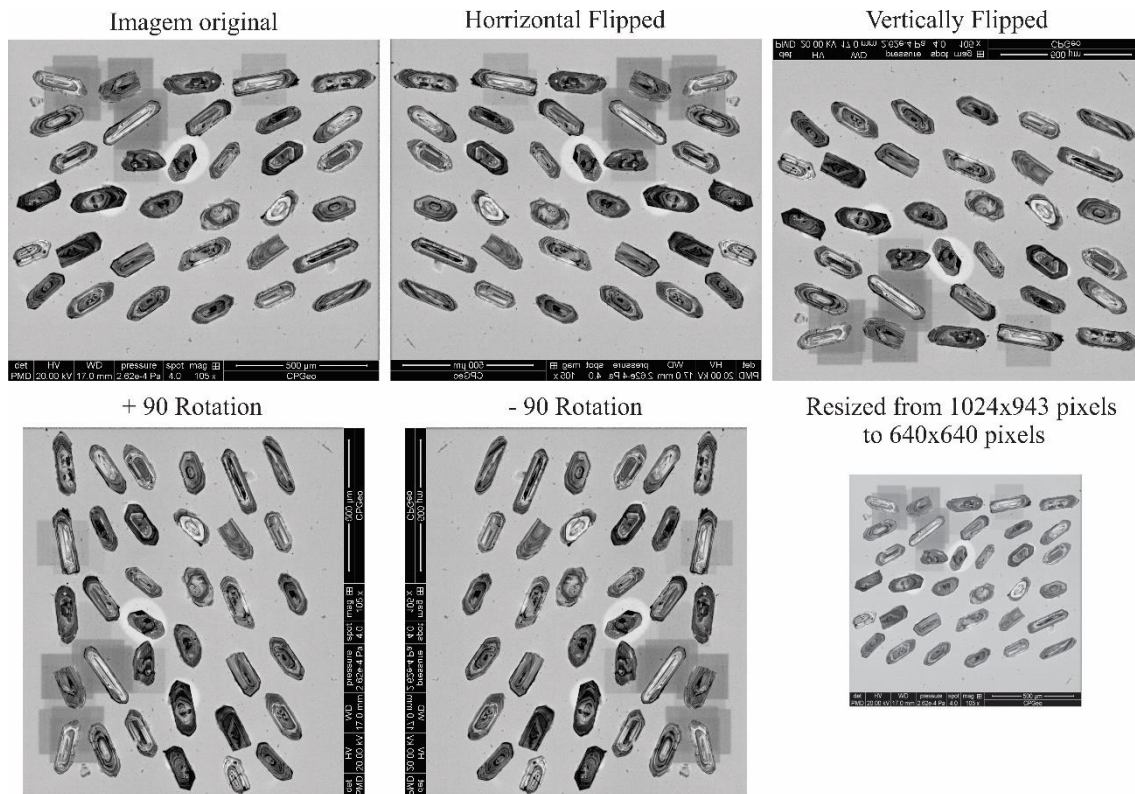


Figura 21 – Imagem original e seus respectivos produtos resultantes das técnicas de *data augmentation* aplicada a imagens de catodoluminescência de grãos de zircão: espelhamento horizontal, espelhamento vertical, rotação 90° em sentido horário e anti-horário, e redimensionamento das imagens.

A rotulagem dos grãos de acordo com as categorias e a criação das imagens artificiais usados na terceira etapa foi feita na plataforma *online* gratuita *Roboflow.com*, enquanto que a limpeza das imagens foi feita com *Adobe Photoshop*. Por limpeza entende-se a remoção de grãos pontuais que ora estavam localizados nas bordas das imagens, ora foram “cortados” ao captar as imagens. Feito isso, o banco de dados, i.e., a quantidade de grãos total somou 7.646 grãos (5.355 originais + 2.291 artificiais), aproximadamente 78% dos quais foram usados para o treino, 14% para validação e 8% para teste (ver Tabela 3).

Tabela 3 - Proporção da quantidade de grãos antes e depois do processo de *data augmentation*.

Classe	Sem <i>data augmentation</i>		Com <i>data augmentation</i>	
	nº grãos	nº grãos (%)	nº grãos	nº grãos (%)
0	1.191	22,24%	2.256	29,51%
1	1.475	27,54%	2.426	31,73%
2	2.689	50,21%	2.964	38,77%
todas as classes	5.355	100,00%	7.646	100,00%

### 3.2. Treinamento

Com o banco de dados pronto, a quarta etapa consistiu em testar e escolher uma das diversas arquiteturas disponíveis na literatura, dentre os quais as que mostraram os melhores resultados foram a YOLOv5 e a *FasterRCNN*.

Existem dois parâmetros que são relevantes quando se escolhe uma arquitetura para detecção de objetos, a saber: (i) a melhor métrica de validação (maior mAP) do modelo e (ii) a sua rapidez ao realizar a detecção. Para esse trabalho, a rapidez não é um parâmetro relevante visto que a classificação em tempo real ou a utilização de vídeos não se aplicam. Optamos por utilizar o modelo YOLOv5 (*You Only Look Once version 5*; Jocher et al., 2022) que é uma das últimas arquitetura de rede neural de detecção de objetos desenvolvida pelo grupo de pesquisa de aprendizado de máquina da Universidade de Washington e que apresentou bons resultados na literatura em detecções de objetos comuns ou *Common Objects in Context* (COCO; cocodataset.org). Essa escolha baseou-se na simplicidade de uso da arquitetura, na qualidade de seu desenvolvimento (com correção de *bugs*, inclusão de novas ferramentas, tutoriais, etc., compatibilidade com outros bancos de dados e com outras arquiteturas), e na possibilidade de implementação de *agnostic* NMS. Os conjuntos de dados foram treinados pelo algoritmo em *hardwares* do *Google colab pro* em uma GPU Tesla P100-PCIE-16GB.

Todos os treinamentos dos modelos foram realizados utilizando o *framework* de ML *PyTorch*, baseada na biblioteca *Torch*, que é um *framework open-source* da linguagem de programação Python. O treinamento padrão dos modelos foi realizado utilizando 700 *epochs* e com um *early stopping* de 50-100 *epochs* (100 ao se treinar os modelos iniciais e 50 para o ajuste dos hiperparâmetros). Isso indica que o treinamento foi interrompido quando não se observou uma melhora na métrica mAP no conjunto de dados de validação depois de 100 *epochs*. A ideia por trás do *early stopping* é monitorar o desempenho do modelo no conjunto de dados de validação para impedir o modelo de

ter um *overfitting*. Para a arquitetura *faster* R-CNN foram utilizadas 1500 iterações (equivalente a 60 *epochs*). Todos os modelos treinados pelo YOLOv5 utilizaram *agnostic* NMS na detecção, e o treinamento realizado pela arquitetura *faster* R-CNN usou o NMS padrão.

Neste trabalho foram produzidos 8 modelos diferentes (Tabela 4). O primeiro e o segundo utilizam a inicialização de pesos em dois conjuntos de dados diferentes, o de catodoluminescência e elétrons secundários utilizando a arquitetura yolov5. O terceiro e quarto utilizam o mesmo conjunto de dados e a mesma arquitetura, mas não utilizam a transferência de aprendizado e nem a inicialização de pesos. O quinto e o sexto modelos utilizam a transferência de aprendizado ao se “congelar” todas as camadas do *backbone* da rede YOLOv5, tendo um treinamento apenas nas camadas do *head*. E o sétimo e oitavo modelo foram treinados na rede *Faster RCNN* com apenas a inicialização de pesos. O primeiro e o segundo foram os que apresentaram o melhor desempenho e, portanto, foram aqueles cujos resultados de classificação são aqui discutidos. O restante dos modelos foi utilizado apenas para comparar desempenhos discutidos em relação ao primeiro e segundo.

Tabela 4 – Modelos inicialmente desenvolvidos para este trabalho.

Modelo	Tipo de Imagem	Transferência de aprendizado	Inicialização de pesos	Arquitetura
1	catodoluminescência.	não	sim	YOLOv5
2	elétrons secundários	não	sim	YOLOv5
3	catodoluminescência.	não	não	YOLOv5
4	elétrons secundários	não	não	YOLOv5
5	catodoluminescência.	sim	sim	YOLOv5
6	elétrons secundários	sim	sim	YOLOv5
7	catodoluminescência	não	sim	<i>Faster RCNN</i>
8	elétrons secundários	não	sim	<i>Faster RCNN</i>

É comum as arquiteturas disponibilizarem os pesos para inicialização e os hiperparâmetros padrões utilizados na obtenção dos resultados publicados, para efeitos de replicabilidade. Logo, os modelos tiveram seu treinamento inicializado com esses pesos e com os hiperparâmetros padrões. Entretanto, com o desenvolvimento de novas técnicas automáticas de ajuste de hiperparâmetros é comum que as arquiteturas consigam demonstrar melhorias de rendimento como é o caso do YOLOv5 que disponibilizou hiperparâmetros melhores que os anteriores por meio de um algoritmo genético (ver

Holland, 1975). Por essa razão, após a fase de treinamento realizamos um ajuste dos hiperparâmetros do modelo, testados manualmente modificando apenas o *learning rate*, o *momentum* e o otimizador (utilizando também a combinação de hiperparâmetros disponibilizados pelos autores), fazendo uma comparação entre os modelos resultantes.

### 3.3 Pós-processamento

Com a escolha do modelo foi necessário ajustar os produtos para melhor visualizar e compreender os resultados. Para comprovar se o modelo estava, de fato, utilizando as características morfológicas dos grãos, e não outros aspectos que poderiam imputar um viés à categorização planejada, efetuamos discretas mudanças em alguns grãos das categorias 1 e 2 com o *software Photoshop*. Por exemplo, para a categoria 1 excluímos uma de suas terminações (bipirâmides), em seguida cortamos a metade do grão, e depois deixamos apenas uma das terminações bipiramidais, o que provocou a diminuição do grão original em aproximadamente 75%. Ainda ‘desenhamos’ fraturas, pontas e ‘arredondamos’ a forma final a fim de ter certeza que se essas mudanças seriam suficientes para alterar a classificação morfológica realizada pelo modelo. Já para os grãos da categoria 2 não procedemos com modificações graduais, mas ‘desenhamos’ simultaneamente uma fratura reta no meio do grão e pontas em uma de suas terminações criando um aspecto que lembra reentrâncias. Para as duas categorias, mascaramos a zonação (de crescimento) interna destacada nas imagens de catodoluminescência com uma tonalidade homogênea similar ao plano de fundo (*background*) da imagem, além de mudar a orientação da alongação máxima do grão em 90°, em sentido horário.

A penúltima etapa do trabalho, foi a de comparar a classificação realizada pelo modelo escolhido em relação à classificação realizada pelos dois sedimentólogos (A e B). Esses pesquisadores receberam e classificaram todos os 685 grãos de zircão do conjunto de dados de teste das imagens de catodoluminescência e um texto explicativo (se encontra no repositório github). A comparação foi realizada pelas matrizes de confusão resultantes de cada classificação e a métrica de pontuação F1 foi utilizada como métrica de comparação. Por fim, este projeto foi apresentado em forma de uma dissertação de mestrado.

## 4. RESULTADOS

Os resultados deste projeto foram obtidos do treinamento de modelos para a detecção dos objetos (i.e., imagens digitais dos grãos de zircão detríticos) partindo-se de (1) imagens dos grãos obtidas por catodoluminescência e (2) imagens dos grãos geradas por elétrons secundários.

As arquiteturas escolhidas para o treinamento foram aquelas com o maior desempenho e maior facilidade de aplicação dentre as disponíveis na literatura. Desempenho porque influencia diretamente na precisão da detecção do objeto, e a facilidade de aplicação pensando em um público alvo não familiarizado com a área de ciência de dados. Com isso, as arquiteturas YOLOv5, *Faster* R-CNN foram escolhidas para treinar o modelo numa tentativa de encontrar aquela que melhor se adequasse ao tipo de dados e oferecesse as melhores ferramentas para o alcançar os objetivos deste projeto. Apesar de ser um classificador apenas em uma etapa, a arquitetura YOLOv5 apresentou resultados equiparados com a *Faster* R-CNN para os dois conjuntos de dados (ver Tabela 5).

A nomenclatura adotada para discriminar cada modelo considera que o primeiro termo representa o conjunto de dados que esse modelo foi treinado (i.e., **CL** para catodoluminescência, e **ES** para elétrons secundários), e o segundo termo representa a arquitetura em que ela foi treinada (i.e., **Y5** para YOLOV5, **FR** para *faster* R-CNN). Quanto à sua inicialização, a discriminação é feita por meio de números, sendo **1** para modelos com inicialização de pesos, **2** para modelos sem inicialização de pesos e sem transferência de aprendizado, e **3** para modelos com inicialização de pesos e com transferência de aprendizado. Por fim, a partir do modelo escolhido de acordo e sua inicialização, os hiperparâmetros foram ajustados com uma combinação de *learning rate* (Lr.) com *momentum*, onde Lr. se mantém fixo e o *momentum* é alterado (i.e., A1, A2 e A3 correspondem a modelos com Lr. de 0,01 com momentum alternando respectivamente em 0,937, 0,9 e 0,99; ver Tabela 6). A sigla VOC ao final da nomenclatura, indica o uso da melhor combinação de hiperparâmetros encontrada por algoritmo, no treinamento do conjunto de dados PASCAL *Visual Object Classes Challenge* (PASCAL VOC), um predecessor e com tamanho reduzido quando comparados com o banco de dados COCO.

Dadas as diferenças nas respostas obtidas para as imagens de catodoluminescência e as imagens de elétrons secundários, as seções deste capítulo estão separadas com resultados de cada um dos conjuntos. Uma diferença é que para o conjunto de dados gerados a partir do imageamento por elétrons secundários não foi feita a etapa de comparações entre análises feitas por usuários (humanos) e o modelo gerado utilizando a



arquitetura YOLOv5. Para constar novamente, entenda-se por tais análises as classificações morfológicas feitas por dois pesquisadores, na fase de testes.

Tabela 5 – Comparação entre as arquiteturas com melhores desempenhos na literatura. Métricas de desempenho dos modelos referentes ao conjunto de dados de validação. Destaque em negrito foi dado para o melhor desempenho.

Modelo	Arquit.	Inic. de pesos	mAP@.5	mAP@.5:.95
CL-Y5-1	YOLOv5	Sim	<b>0,69</b>	<b>0,64</b>
ES-Y5-1	YOLOv5	Sim	0,66	0,55
CL-Y5-2	YOLOv5	Não	0,62	0,55
ES-Y5-1	YOLOv5	Não	0,59	0,52
CL-Y5-3	YOLOv5	Não	0,58	0,51
ES-Y5-1	YOLOv5	Não	0,59	0,53
CL-FR-1	Faster-RCNN	Sim	0,72	0,63
ES-FR-1	Faster-RCNN	Sim	0,66	0,59

## 4.1. Conjunto de dados de catodoluminescência

### 4.1.1. Escolha do modelo

Foram treinados três modelos diferentes: CL-Y5-1, CL-Y5-2 e CL-Y5-3, conforme detalhamento ilustrado na Tabela 5. O modelo CL-Y5-1 foi inicializado com os melhores pesos pré-treinados do conjunto de dados COCO e os hiperparâmetros comumente utilizados. O modelo CL-Y5-2 teve seus pesos inicializados aleatoriamente, sem utilizar pesos pré-treinados. Finalmente o modelo CL-Y5-3 foi treinado com transferência de aprendizado, onde os pesos do *backbone* foram "congelados" e apenas as camadas do *head* foram treinadas, ou seja, os pesos das camadas do *backbone* não são atualizados ao longo do treinamento. A ideia do "congelamento" é utilizar os mesmos valores de pesos das camadas do *backbone* do treinamento original. Os valores de P, R, mAP@.5, mAP e da *loss function* para os dados de treinamento e validação de cada modelo ao longo das *epochs* podem ser vistos na Tabela 6, e nas figuras 22 e 23.

Tabela 6 – Comparação entre as configurações iniciais e métricas de desempenho dos modelos referentes ao conjunto de dados de teste. Destaque em negrito foi dado para o melhor desempenho.

Modelo	Inic. de pesos	Transf. aprendizado	P	R	mAP@.5	mAP@.5:.95
--------	----------------	---------------------	---	---	--------	------------

CL-Y5-1	sim	não	0,70	0,69	0,69	<b>0,64</b>
CL-Y5-2	não	não	0,67	0,65	0,62	0,55
CL-Y5-3	não	sim	0,64	0,63	0,58	0,51

Nota: os valores dos parâmetros foram arredondados para duas casas decimais.

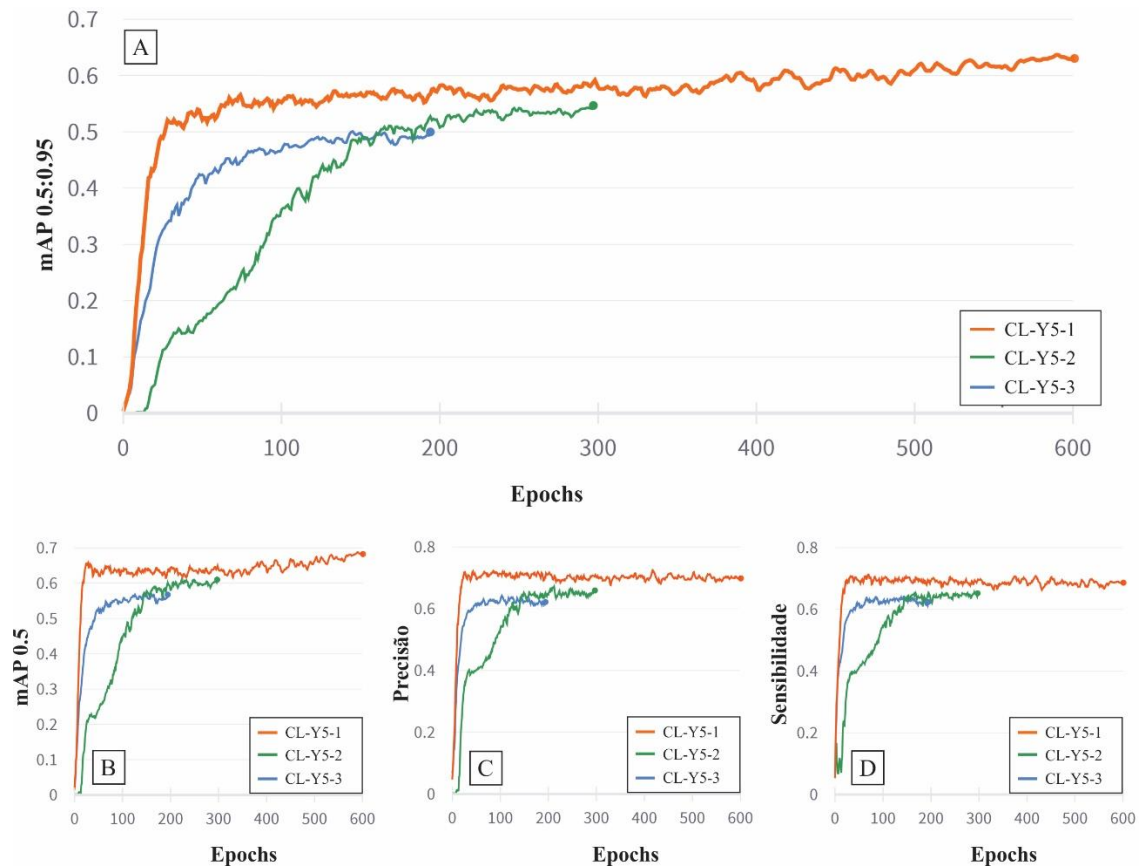


Figura 22 – Comparação entre os modelos com configurações iniciais de treinamento diferentes, com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, i.e., com *confidence score* variando entre 0,5 e 0,95. (B) Valores de mAP a 0,5 de *confidence score*. (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.

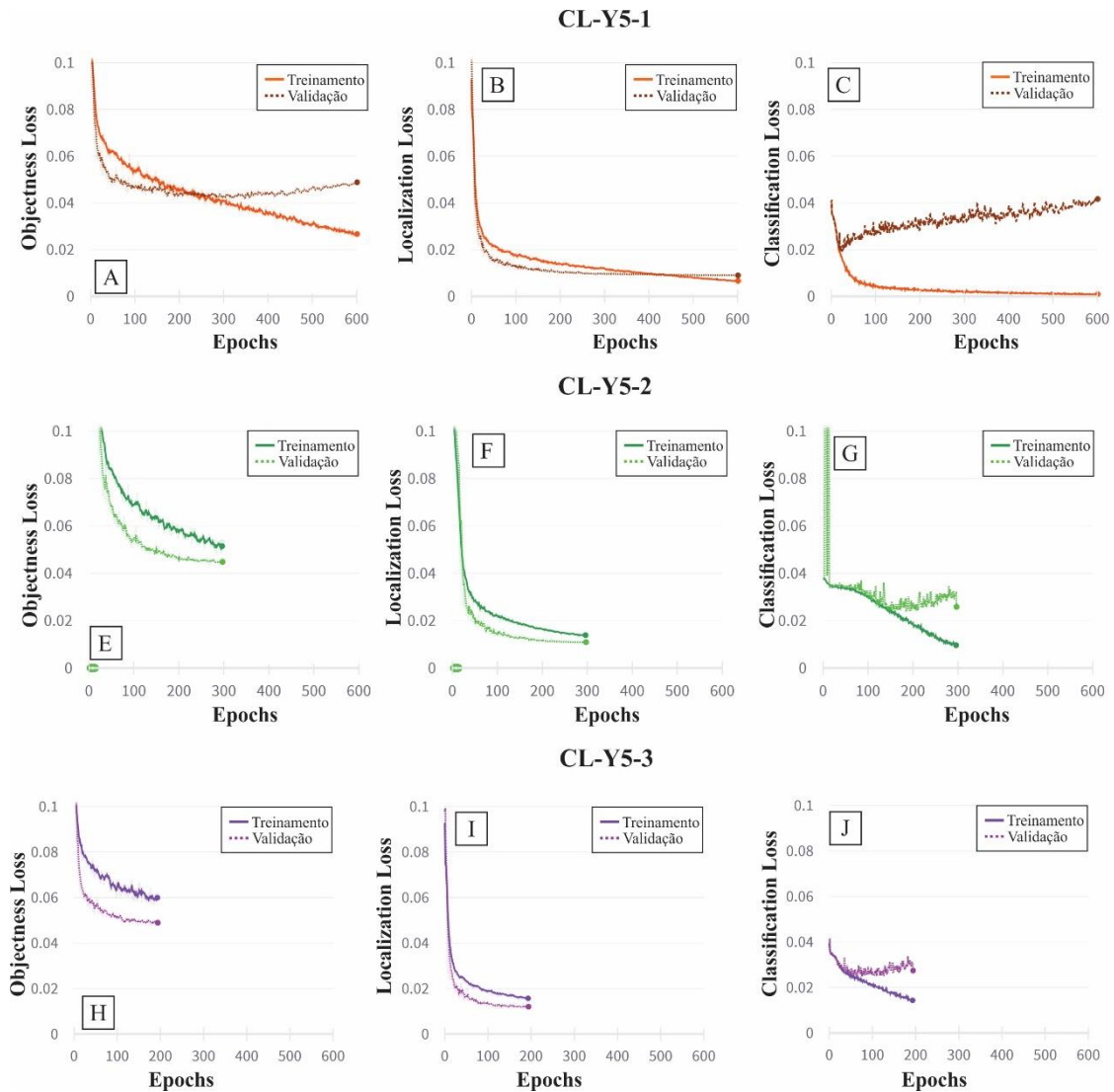


Figura 23 – Curvas de *loss functions* que compõem o YOLOv5 para os dados de treinamento (linha contínua, na legenda) e de validação (linha pontilhada). O efeito de *overfitting* pode ser visto na *classification loss*, nos gráficos C, G e J.

Analisando os gráficos, nota-se que o primeiro modelo apresentou o melhor desempenho, com um  $mAP@.5:.95$  ao longo do treinamento de 0,6417 em comparação aos outros dois modelos, os quais tiveram valores de 0,55 e 0,51. Na avaliação da *loss function* dos 3 modelos, principalmente para a função *classification loss*, nota-se que a partir de certo ponto (*epoch* 20 para o modelo CL-Y5-1, 150 para o CL-Y5-2 e 36 para o CL-Y5-3) os valores de validação e treino se distanciam muito influenciado pelo aumento dos valores de validação, quando deveriam diminuir. Esse comportamento reflete um problema de *overfitting* dos dados onde, ao longo do treinamento, o modelo vai apresentando valores cada vez melhores para os dados de treino e cada vez piores para os dados de validação, indicando uma dificuldade em classificar dados novos. É possível atenuar esse efeito com várias iniciativas, seja aumentando o número de dados para

treinamento, utilizando formas de penalização no treinamento com, por exemplo, *early stopping*. Aqui, foi aplicada a *early stopping* de 50-100 para os modelos CL-Y5-2 e CL-Y5-3, os quais tiveram treinamento interrompido na faixa de 200 e 300 *epochs*, respectivamente. Ainda com a penalização, esses dois modelos apresentaram desempenhos fracos e foram descartados.

Para o modelo CL-Y5-1 esse tipo de penalização não surtiu efeito, apresentando valores progressivamente mais altos com a continuação do treinamento. Então, foram feitos novos treinamentos neste modelo, mas adotando hiperparâmetros diferentes para encontrar versões sem *overfitting* e, ao mesmo tempo, com melhoria do desempenho.

#### 4.1.2. Ajuste dos hiperparâmetros

Para o modelo CL-Y5-1, pudemos encontrar os hiperparâmetros mais importantes tanto manual quanto automaticamente (via algoritmo) para aumentar o desempenho na detecção dos objetos. O YOLOv5 disponibiliza uma lista de hiperparâmetros considerados mais ajustados (ideais), após passarem por um processo de automatização que testou centenas de combinações diferentes (consumindo um poder computacional e de tempo muito alto), até encontrar a combinação que apresentou o melhor desempenho, medido pelo mAP<sub>0.5-0.95</sub> com base PASCAL VOC <sup>1</sup>. No entanto, é de escolha do usuário utilizar os tais hiperparâmetros “ideais”, o “padrão” ou então escolher a combinação que lhe convier. Os resultados dos treinamentos utilizando várias combinações de hiperparâmetros para o modelo CL-Y5-1, adotadas neste trabalho, estão discriminadas na Tabela 7. Nota-se que os modelos com menores Lr. e maiores valores de *momentum* foram os que apresentaram os melhores valores de mAP<sub>0.5-0.95</sub>, com exceção do modelo CL-Y5-1-VOC. Assumimos, portanto, que os modelos CL-Y5-1-B3, CL-Y5-1-C2 e CL-Y5-1-VOC foram as melhores combinações. Os parâmetros desses modelos estão listados na Tabela 8 e os gráficos de desempenho ilustrados na Figura 24. As demais combinações com valores de mAP<sub>0.5-0.95</sub> inferiores ou iguais ao modelo CL-Y5-1 ( $\leq 0,64$ ) não serão comentadas.

---

<sup>1</sup> <https://github.com/ultralytics/yolov5/blob/master/data/hyps/hyp.VOC.yaml>

Tabela 7 – Configuração do otimizador SGD e ADAM para aprimorar o modelo de linha de base.

Modelo	Otimizador	Learning rate	Momentum	P	R	mAP@.5	mAP@.5:.95
CL-Y5-1	SGD	0,01*	0,937*	0,70	0,69	0,69	0,64
CL-Y5-1-A1	SGD	0,01	0,9	0,73	0,72	0,66	0,58
CL-Y5-1-A2	SGD	0,01	0,99	0,73	0,72	0,70	0,64
CL-Y5-1-B1	SGD	0,001	0,937	0,75	0,75	0,72	0,66
CL-Y5-1-B2	SGD	0,001	0,9	0,76	0,73	0,72	0,64
CL-Y5-1-B3	SGD	0,001	0,99	0,74	0,73	0,74	<b>0,68</b>
CL-Y5-1-C1	SGD	0,0001	0,9	0,62	0,63	0,54	0,46
CL-Y5-1-C2	SGD	0,0001	0,99	0,75	0,74	0,74	<b>0,69</b>
CL-Y5-1-D1	ADAM	0,001**	0,937**	0,72	0,70	0,70	0,63
CL-Y5-1-D2	ADAM	0,001	0,9	0,73	0,70	0,70	0,61
CL-Y5-1-D3	ADAM	0,001	0,99	0,71	0,72	0,71	0,63
CL-Y5-1-VOC	SGD	0,00334	0,7483	0,74	0,72	0,75	<b>0,69</b>

NOTA: \*Valores padrão para o otimizador SGD; \*\*Valores padrão para o otimizador ADAM. As siglas A, B, C, D (seguidas de 1, 2, 3) se referem a modelos com Lr. de 0,01 com momentum alternando respectivamente em 0,937, 0,9 e 0,99; VOC significa a combinação de hiperparâmetros no treinamento do banco de dados PASCAL VOC

Tabela 8 – Tabela comparando o desempenho entre os modelos com os melhores desempenhos, tendo o modelo CL-Y5-1 como linha de base. *Epoch* indica o momento em que esse desempenho foi alcançado no treinamento. Destaque em negrito foi dado para o melhor desempenho.

Dados de Validação					
Modelo	<i>epoch</i>	P	R	mAP@.5	mAP@.5:.95
CL-Y5-1	589	0,70	0,69	0,69	0,64
CL-Y5-1-B3	112	0,74	0,73	0,74	0,68
CL-Y5-1-C2	208	0,74	0,74	0,74	<b>0,685</b>
CL-Y5-1-VOC	307	0,73	0,72	0,75	<b>0,685</b>
Dados de Teste					
Modelo	<i>epoch</i>	P	R	mAP@.5	mAP@.5:.95
CL-Y5-1	589	0,693	0,696	0,663	0,623
CL-Y5-1-B3	112	0,63	0,80	0,74	0,65
CL-Y5-1-C2	208	0,77	0,74	0,75	<b>0,69</b>
CL-Y5-1-VOC	307	0,75	0,69	0,72	0,66

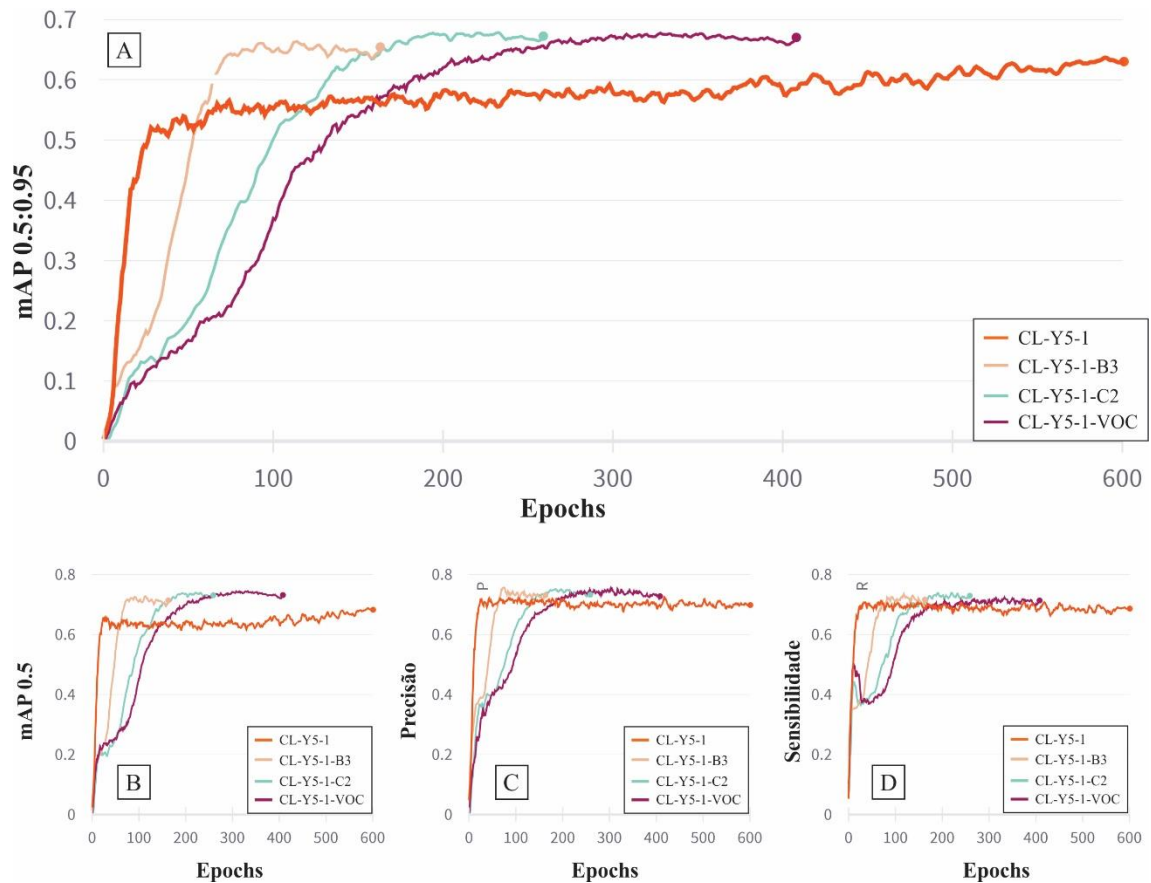


Figura 24 – Comparação entre os modelos derivados do CL-Y5-1 (em laranja) com diferentes hiperparâmetros com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, i.e., com *confidence score* variando entre 0,5 e 0,95. (B) Valores de mAP a 0,5 de *confidence score*. (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.

Assim como para o modelo CL-Y5-1 (vide Figura 22), os modelos CL-Y5-1-B3 e CL-Y5-1-C2 (ver Tabela 8) apresentaram *overfitting* com treinamento interrompido entre 200-300 *epochs*, mas com uma diferença (Figura 25). Os dados de treino e validação do modelo CL-Y5-1-C2 no seu ponto de maior mAP (*epoch* 208) mostrou uma diferença entre *classification loss* de validação e teste menor em comparação com o modelo CL-Y5-1-B3 (*epoch* 112). Ou seja, o *overfitting* do modelo -B3 é graficamente maior que o do modelo -C2. Dentre todos, o modelo CL-Y5-1-VOC apresentou o mAP@.5:.95 de 0,69 e os menores valores de *loss functions* (*object*, *localization* e *classification loss*). Mantendo-se apenas os modelos CL-Y5-1-C2 e CL-Y5-1-VOC, a etapa seguinte consistiu em melhor avaliar o efeito de *overfitting* comparando-se os resultados e suas matrizes de confusão nos dados de validação e teste (Figura 26).

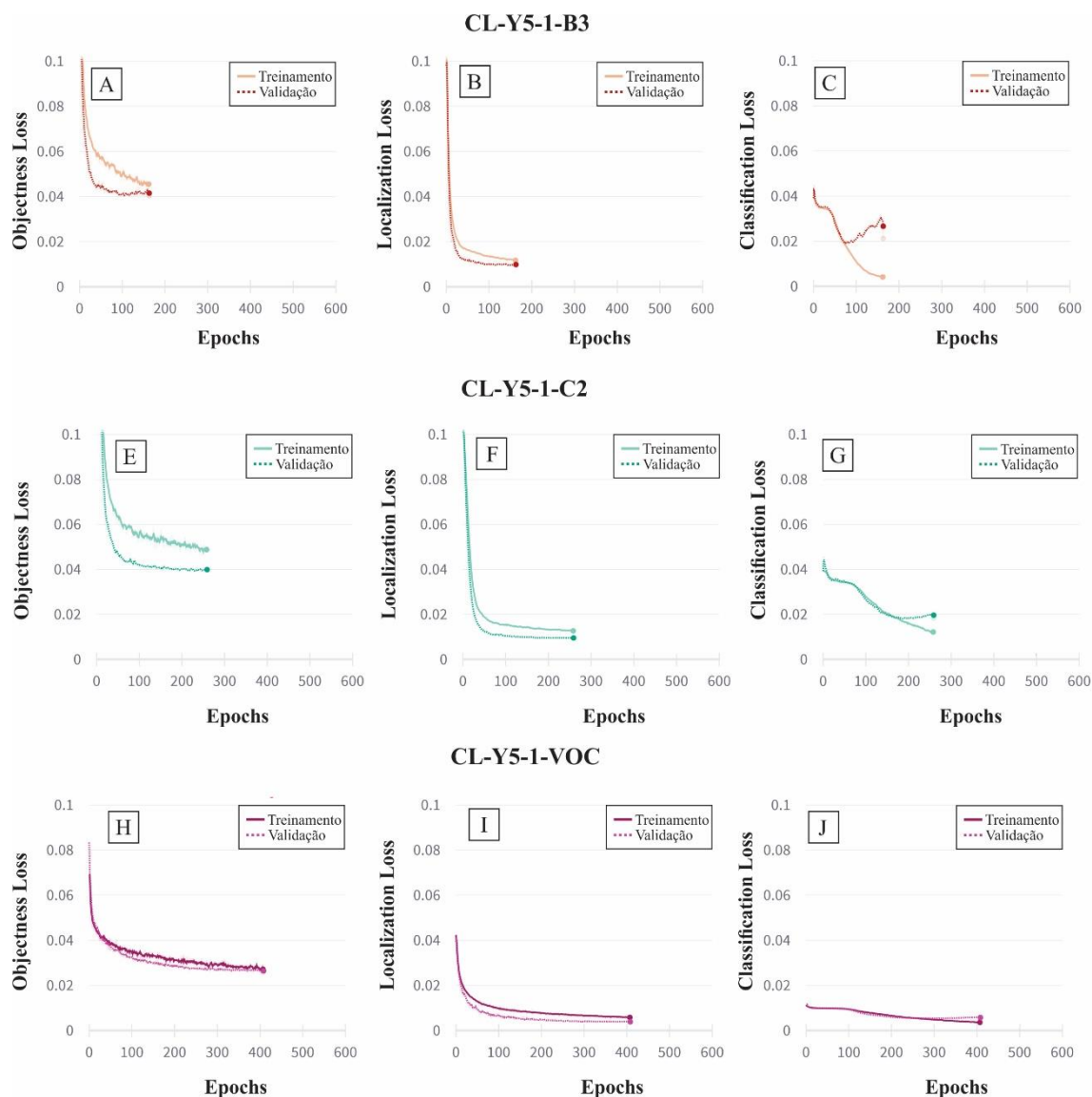


Figura 25 – Curvas de *loss functions* que compõem o YOLOv5 para os dados de treinamento e de validação (pontilhado). O efeito de *overfitting* é ilustrado pela função *classification loss*, especialmente destacado nos gráficos C, G, e incipiente em J.

A matriz de confusão dos dois modelos mostrou que a análise inicial de *overfitting* do modelo CL-Y5-1-C2 estava equivocada, pois este modelo apresentou resultados de verdadeiros positivos superiores ao modelo CL-Y5-1-VOC, tanto para os dados de validação como para os dados de teste – ver colunas 0-1-2 de classificações originais (eixo X). O modelo CL-Y5-1-C2 foi escolhido por apresentar um incremento total em relação ao CL-Y5-1 (linha base) de 0,045 nos dados de validação e de 0,07 nos dados de teste.

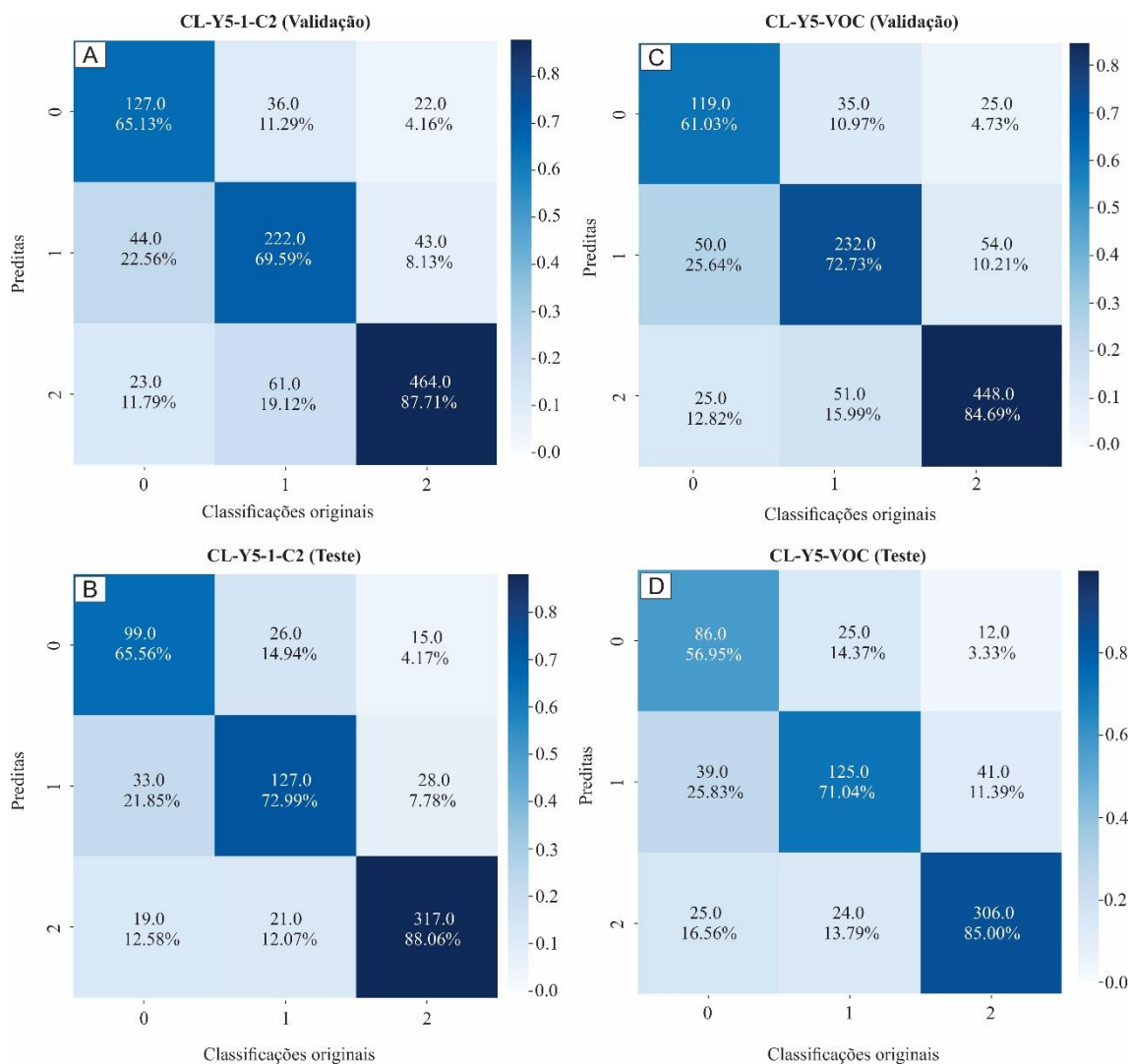


Figura 26 – Comparação da matriz de confusão dos modelos CL-Y5-1-C2 e -VOC vistas em forma de escala de cores, e diferenciadas com os dados de validação e de teste.

Por fim, o limiar do *confidence score* utilizado para fazer as inferências (i.e., delimitação e classificação dos objetos) foi ajustado com base no gráfico F1 (Figura 27). Esse valor se refere à multiplicação entre a probabilidade de uma caixa delimitadora conter um objeto e a probabilidade deste objeto ser classificado em uma das categorias<sup>2</sup>. O gráfico mostra a variação dos valores da média harmônica de P e R com respeito aos valores de *confidence score*, ao se realizar a detecção nos dados de teste.

<sup>2</sup> Função *sigmoid* da formula:  $confidence\ score = P(object) * P(class\_c | object)$



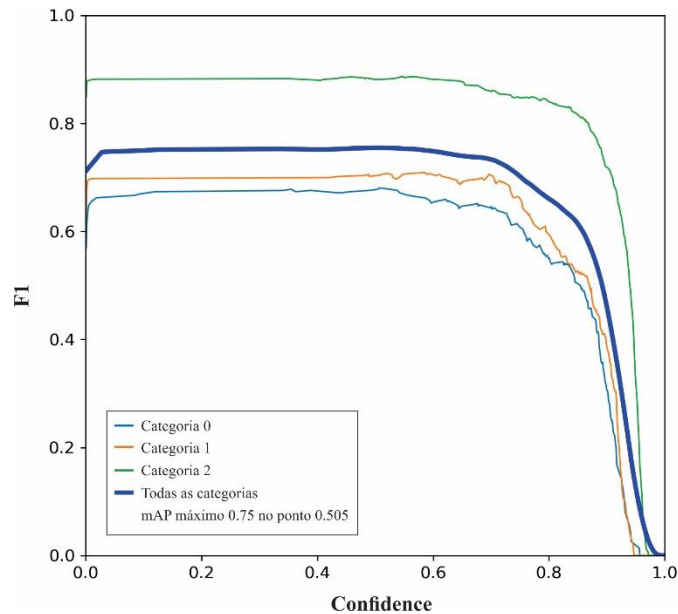


Figura 27 – Curva F1 para variações do *confidence score* no momento da inferência.

Todos os resultados tiveram *confidence score* de 0,01 indicando que qualquer detecção com valor  $\geq 0,01$  foi considerada na contagem de  $mAP@.5:.95$ . É importante definir um limiar para o *confidence score* para retirar inferências com valores muito baixos. Entretanto, pode acontecer do modelo inferir um FP (Falso Positivo) indicando a existência de um objeto em uma região sem objetos (i.e., apenas com *background*). Por essa razão, foi necessário analisar criticamente cada inferência para evitar aumento no número de FP e diminuição do valor de  $mAP@.5:.95$ . A Figura 26 mostra que o modelo possui maior  $mAP@.5$  quando o seu coeficiente é 0,5, muito embora o valor de F1 máximo seja constante desde valores próximos a 0 até próximos a 0,6. Pensando em um valor baixo o suficiente para não aumentar o número de FP e alto o suficiente para retirar classificações no *background* com valores muito baixos, foi escolhido o valor de 0,2 para todas as detecções realizadas por esse modelo.

#### 4.1.3. Inferência (detecção e classificação) do conjunto de dados de teste

Um pequeno exemplo das inferências do modelo CL-Y5-1-C2 para o conjunto de dados de teste está ilustrado na Figura 28, com grãos representativos das três unidades litoestratigráficas usadas neste projeto – duas unidades da Bacia do Parnaíba (Grupo Serra Grande indiviso e Formação Sambaíba) e a unidade basal da Bacia do Cococí. Cada prancha representa uma das três categorias morfológicas discriminadas: grãos predominantemente da categoria 1 (Bacia do Cococí), grãos predominantes da categoria

2 (Formação Sambaíba) e um cenário mais heterogêneo com mistura de grãos das três categorias, inclusive a de grãos quebrados representado pelo Grupo Serra Grande. Voltando para a Tabela 8 e para a matriz de confusão (Figura 26), nota-se que o modelo teve facilidade em classificar grãos da categoria 2, fornecendo altos valores de P e R. A dificuldade do modelo recaiu sobre as outras duas categorias 0 e 1, e principalmente na categoria 1 ao incluir nesta 33 grãos (de todo o conjunto testado) da categoria zero e 28 grãos da categoria 2.

#### **4.1.4. Comportamento do modelo**

Nesta etapa, foram feitas algumas mudanças (artificialmente) no tamanho e limites dos grãos com ferramentas de edição de imagens do Photoshop e CorelDraw para testar a funcionalidade do modelo e o peso de certas características da imagem/grão (estrutura interna, luminosidade etc.) na classificação das categorias (Figura 29). Notamos que a estrutura interna (zonação de crescimento oscilatória) possui grande peso para a categoria escolhida pelo modelo, sobretudo para a categoria 1. Ao ‘limpar’ a zonação interna, 19 dos 33 grãos da prancha tiveram uma mudança na classificação. Todos esses 19 grãos foram inicialmente classificados pelo modelo na categoria 1 (alguns grãos de forma correta, outros não), mudaram para as outras categorias ao se ‘limpar’ essa zonação.

A orientação do eixo maior do grão (máxima alongação) também tem efeito na categorização (Figura 30). Foi aplicada uma rotação de 90° em sentido horário em todos os grãos do conjunto de dados de teste. Dos 685 grãos, 128 grãos mudaram sua classificação inicial realizada pelo modelo. Isso indica que, ao se mudar a orientação do eixo maior do grão, 1 a cada 4-5 grãos terá sua classificação modificada pelo modelo sem seguir critérios morfológicos. A melhor resposta veio da alteração artificial da morfologia (Figura 31), ao ‘cortar’ o grão em frações de 25%, 50% e 75% de sua área, com o modelo ‘entendendo’ bem como deveria reclassificar as novas formas. A Figura 31A demonstra que a maioria dos grãos subédricos quebrados é reclassificada na categoria 1 corretamente, e que grãos mais deformados, ou quando perdem muito de sua “área”, acabam na categoria zero. Já a imagem 30B demonstra que a maioria dos grãos que era arredondada e foram fraturados artificialmente, é corretamente reclassificada na classe zero.

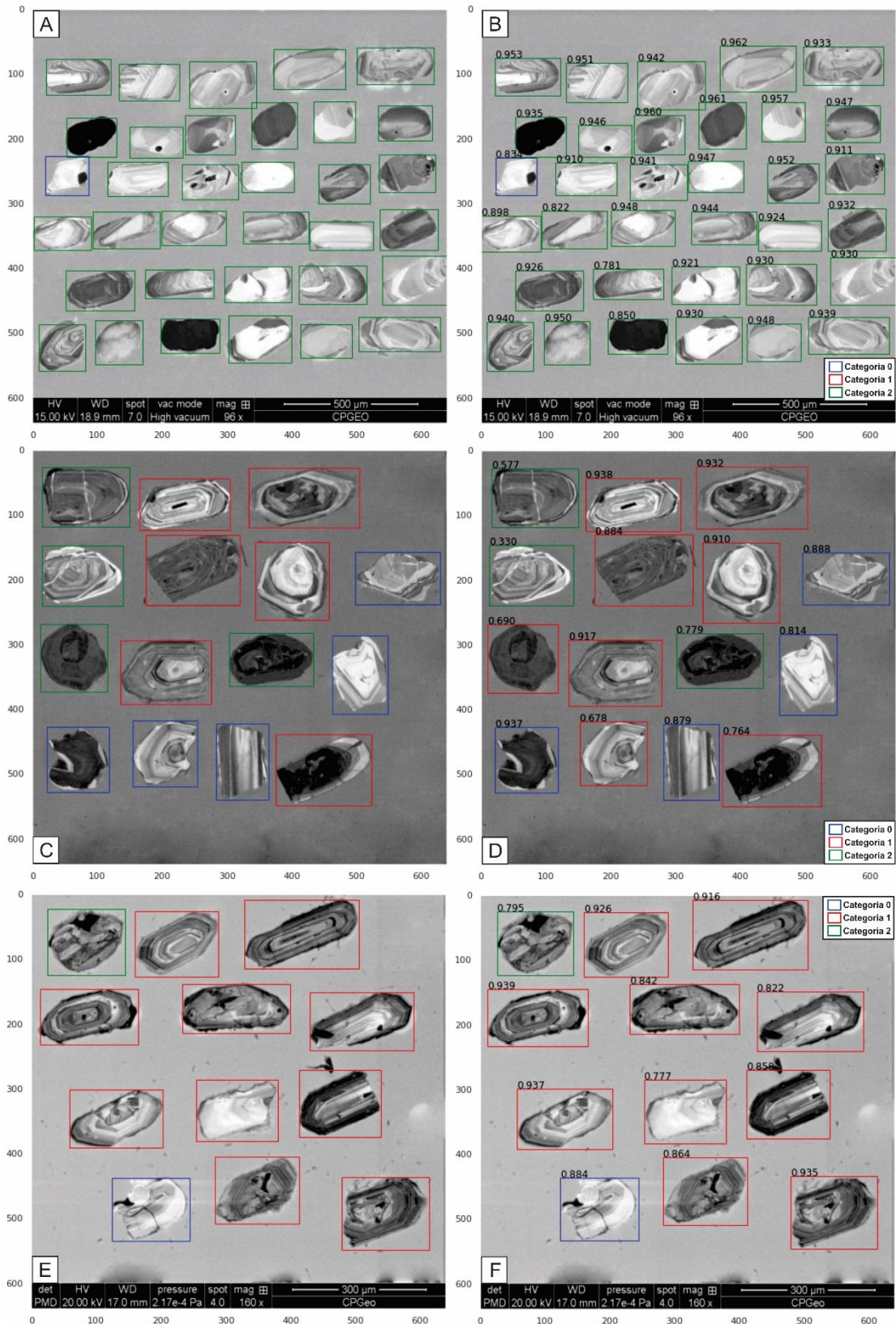


Figura 28 – Exemplos de detecção realizada pelo modelo escolhido (CL-Y5-1-C2). À esquerda são ilustradas três pranchas com as classificações originais (*ground truth detection*) feitas sem a função algorítmica. À direita tem-se as mesmas imagens com a detecção feita pelo modelo. Em (A) e (B) exemplos de grãos dos arenitos eólicos da Formação Sambaíba. Em (C) e (D) exemplos de grãos dos arenitos marinho-fluviais do Grupo Serra Grande. Em (E) e (F) exemplos de grãos dos arenitos fluviais da Bacia do Cocó.

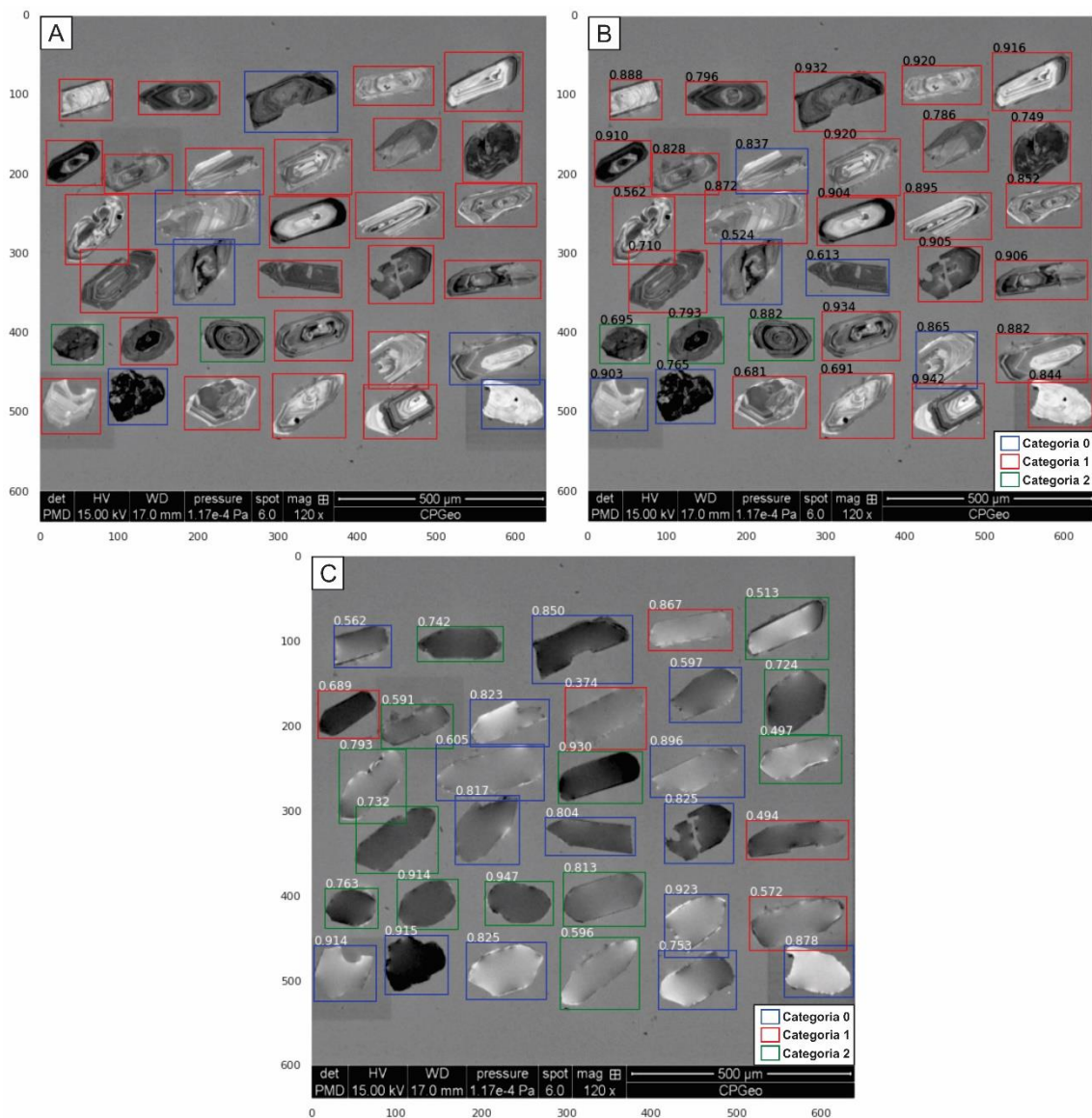


Figura 29 – Limpeza artificial da zonação de crescimento interna dos grãos detríticos. Em (A), classificação original feita pelo autor. Em (B), inferência realizada pelo modelo. Em (C), inferência realizada pelo modelo após a limpeza das zonações dos grãos.

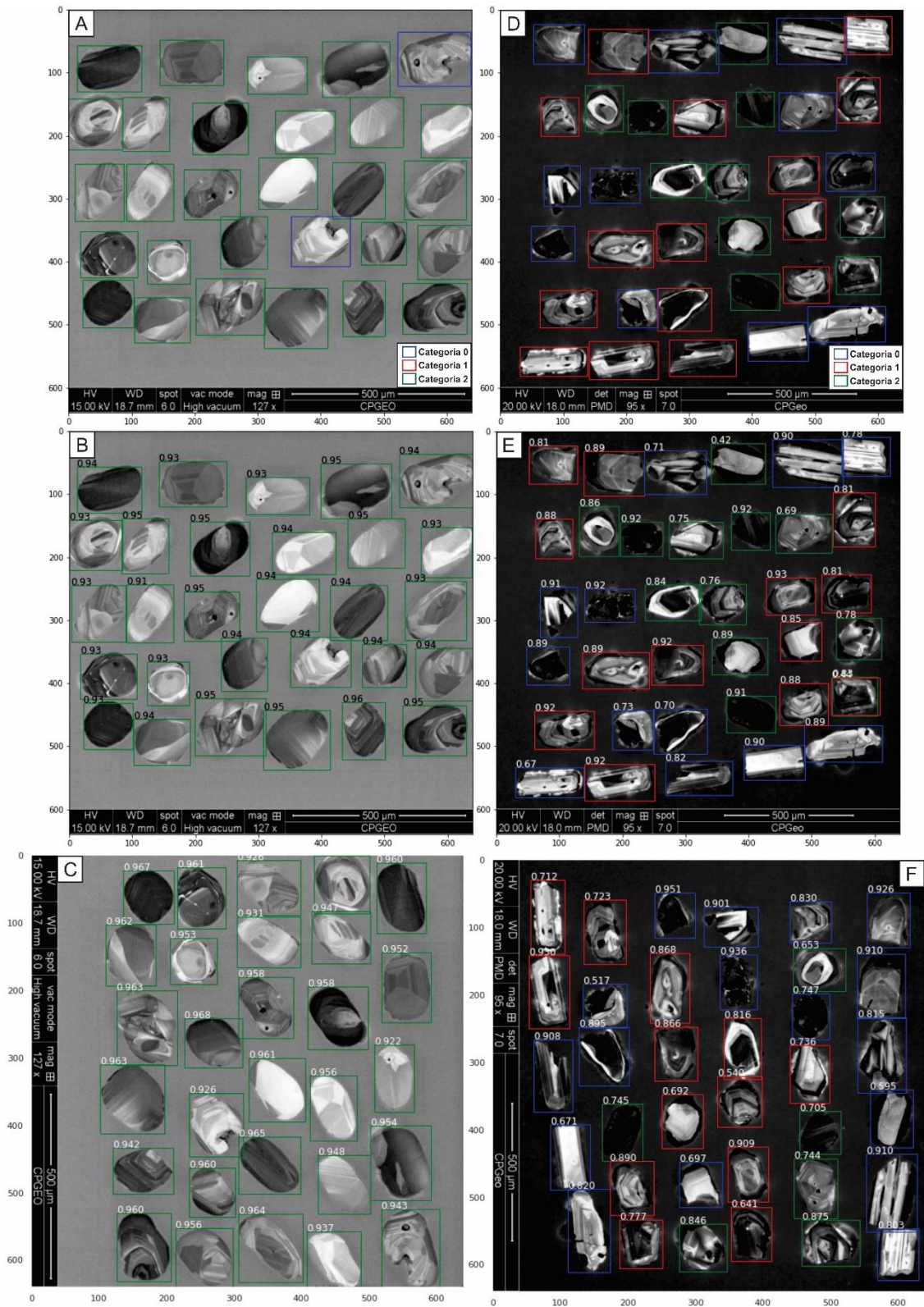


Figura 30 – Pranchas exemplificando o efeito da orientação dos grãos na classificação do modelo. (A) e (D) são pranchas com grãos distintos onde ilustra-se a classificação original feita pelo autor. (B) e (E) são as mesmas pranchas, mas com classificação feita pelo modelo. (C) e (F) são as mesmas pranchas com classificação feita pelo modelo, mas com os grãos rotacionados 90° no sentido horário.

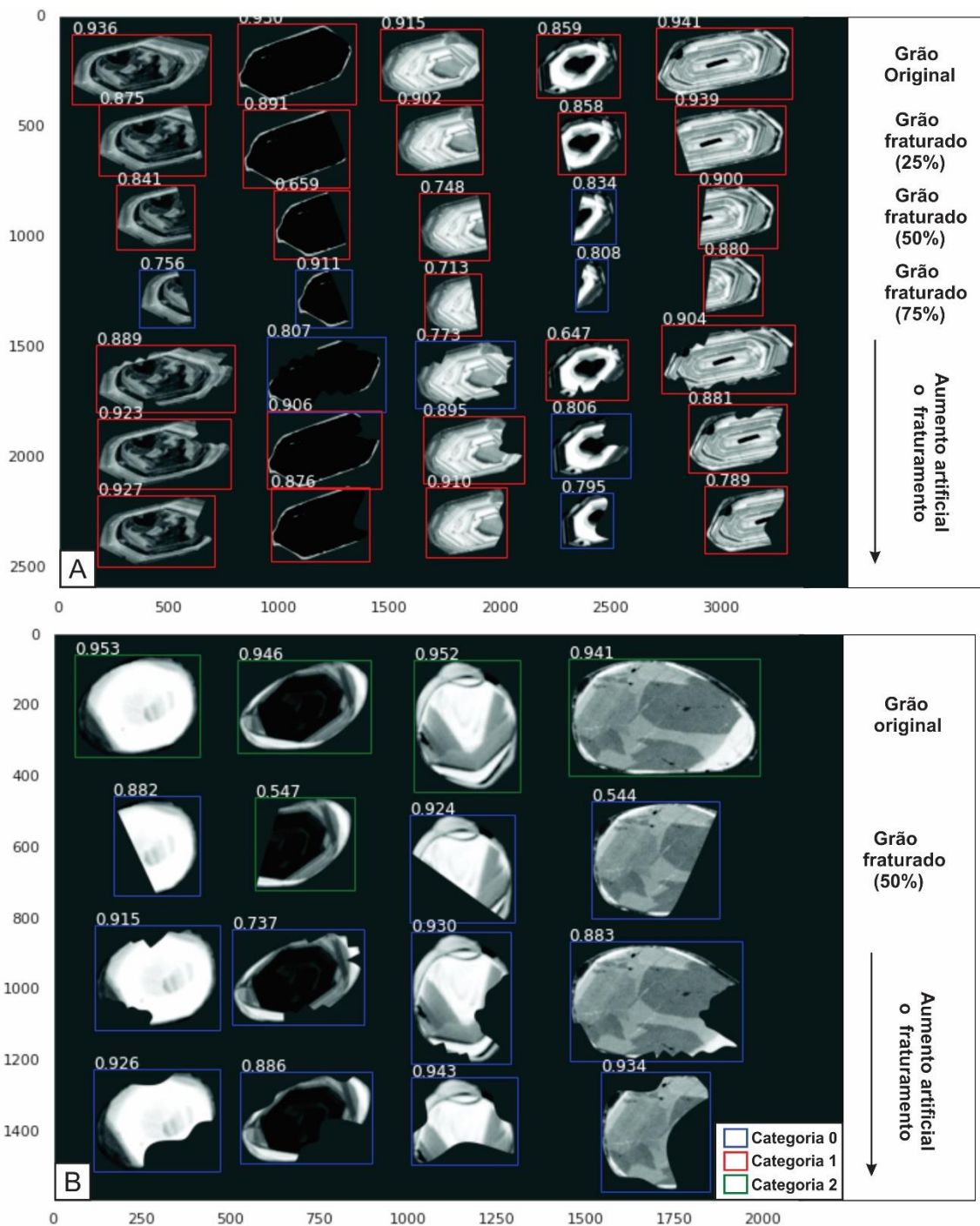


Figura 31 – Pranchas editadas ilustrando modificações artificiais na morfologia de alguns grãos. O grau de alteração aumenta do topo para baixo, na prancha. (A) Alterações em grãos euédricos (categoria 1). (B) Alterações em grãos arredondados (categoria 2).

#### 4.1.5. Validação dos testes – análise de usuários

A subjetividade na classificação morfológica de grãos é um fator de peso muito elevado, sendo um dos objetivos deste trabalho minimizar essa particularidade e, ao

mesmo tempo, diminuir significativamente o tempo para realizar a classificação em banco de dados envolvendo centenas de grãos. Como um ensaio preliminar de validação, solicitamos a contribuição de dois pesquisadores (nomeados A e B) com experiência em sedimentologia de rochas siliciclásticas na classificação manual dos 685 grãos detríticos usados na fase de teste do modelo. Os resultados dessa classificação ‘humana’ foram confrontados com a classificação fornecida pelo modelo, utilizando os critérios definidos na Figura 18. Os 685 grãos estavam dispostos em 27 pranchas, e os resultados ilustrados na Figura 32 e listados na Tabela 9. A discussão será apresentada no capítulo seguinte.

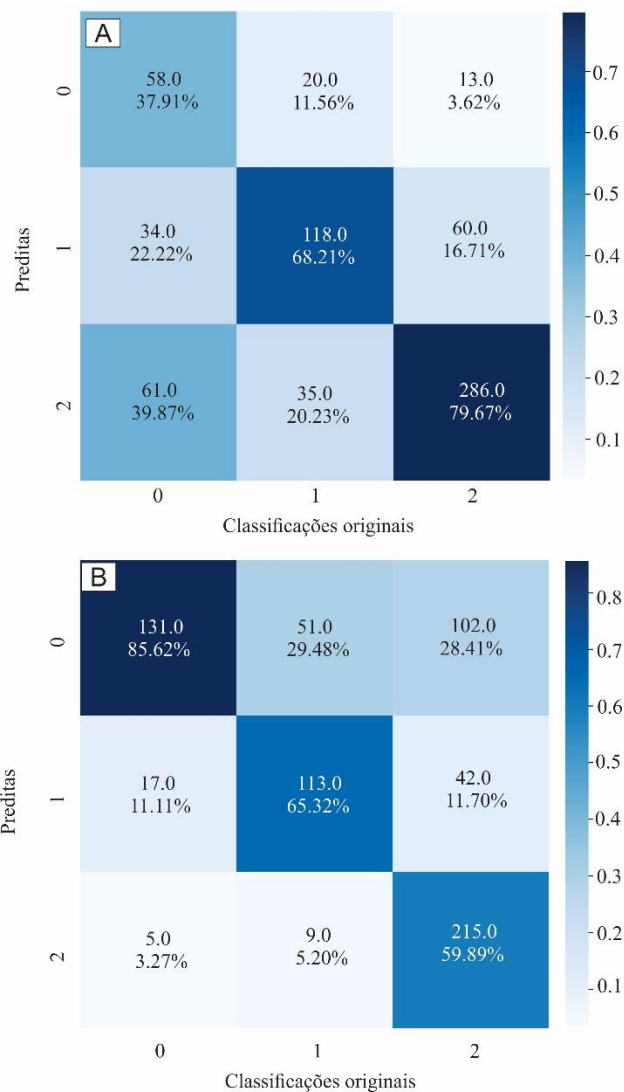


Figura 32 – (A) Matriz de confusão do pesquisador A. (B) Matriz de confusão do pesquisador B.

Tabela 9 – Comparação das métricas: precisão (P), sensibilidade (R) e pontuação F1 entre o modelo e os usuários A e B.

<b>Modelo CL-Y5-1-C2</b>			
Classe	Precisão (P)	Sensibilidade (R)	Pontuação F1
0	0,71	0,66	75,61%
1	0,68	0,73	
2	0,89	0,88	
média	0,76	0,76	
<b>Pesquisador A</b>			
Classe	Precisão (P)	Sensibilidade (R)	Pontuação F1
0	0,64	0,38	63,31%
1	0,56	0,68	
2	0,75	0,80	
média	0,65	0,62	
<b>Pesquisador B</b>			
Classe	Precisão (P)	Sensibilidade (R)	Pontuação F1
0	0,46	0,86	69,41%
1	0,66	0,65	
2	0,94	0,60	
média	0,69	0,70	

## 4.2. Conjunto de dados de Elétrons Secundários

As mesmas etapas descritas para o conjunto de dados de catodoluminescência foram aplicadas ao conjunto de dados de elétrons secundários, que será nomeado pela sigla ES-Y5. Foram treinados três modelos seguindo a mesma lógica: primeiro com pesos inicializados e sem transferência de aprendizado, segundo sem pesos e sem transferência, e terceiro com a presença de pesos e transferência (ver Tabela 10, e Figuras 33 e 34).

Da mesma forma que os modelos para o banco de dados de catodoluminescência, o modelo ES-Y5-1 demonstrou o melhor resultado e com uma diferença crescente entre os valores de validação e teste na função *classification loss*, um cenário característico de *overfitting* quando se categoriza os grãos detríticos (Figura 34C, G e J). Nota-se que os valores de mAP@.5:.95 dos modelos ES-Y5-2 e ES-Y5-3 são mais baixos, embora que progressivamente ascendentes, talvez requerendo maior período de treinamento para apresentar resultados similares (ou até melhores) que o modelo ES-Y5-1. Conhecido o melhor modelo, foram testadas várias combinações de hiperparâmetros na configuração inicial de ES-Y5-1 (Tabela 11; Figura 35).



Tabela 10 - Comparação entre as configurações iniciais e métricas de desempenho dos modelos referentes ao conjunto de dados de teste. Destaque em negrito para o melhor desempenho entre os modelos.

Modelo	Inic. de pesos	Transf. aprendido	P	R	mAP@.5	mAP@.5:.95
ES-Y5-1	sim	não	0,72	0,72	0,66	<b>0,55</b>
ES-Y5-2	não	não	0,64	0,65	0,59	0,52
ES-Y5-3	não	sim	0,64	0,64	0,59	0,53

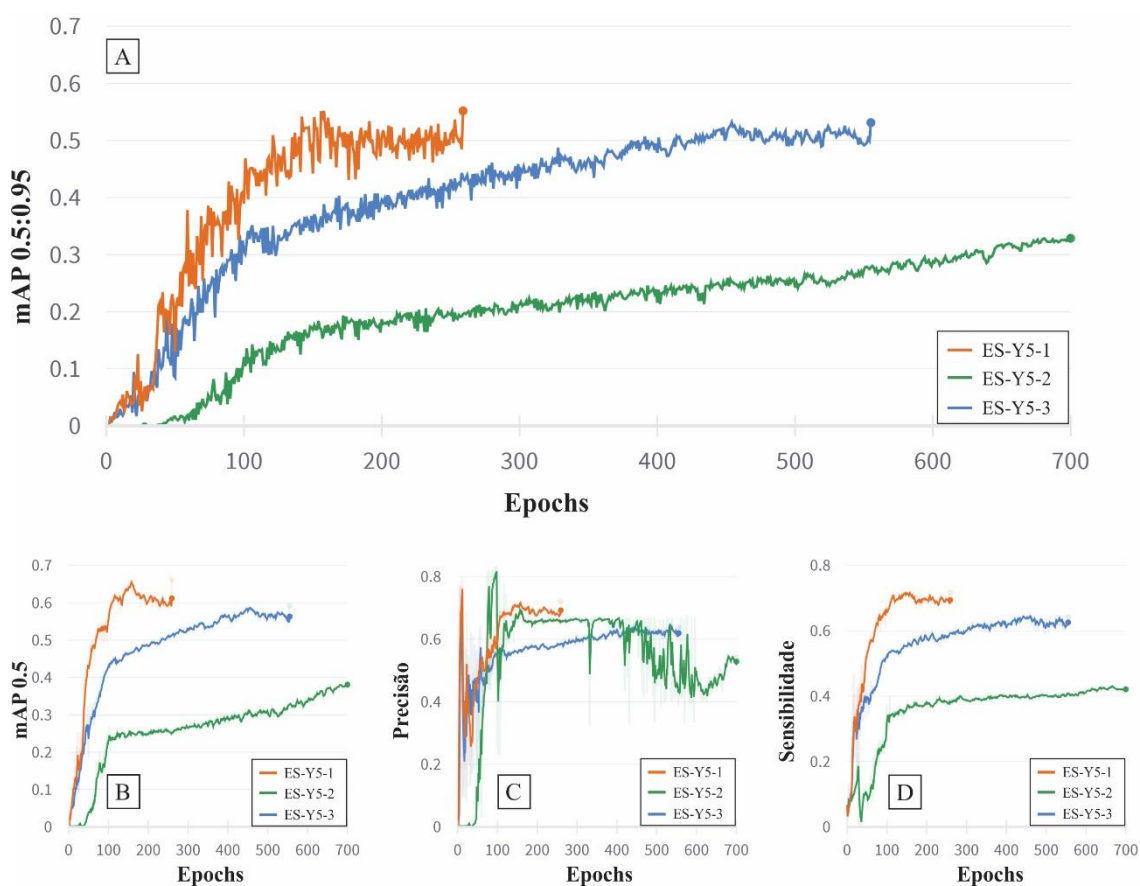


Figura 33 – Comparação entre os modelos com configurações iniciais de treinamento diferentes, com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, i.e., com *confidence score* variando entre 0,5 e 0,95. (B) Valores de mAP a 0,5 de *confidence score*. (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.

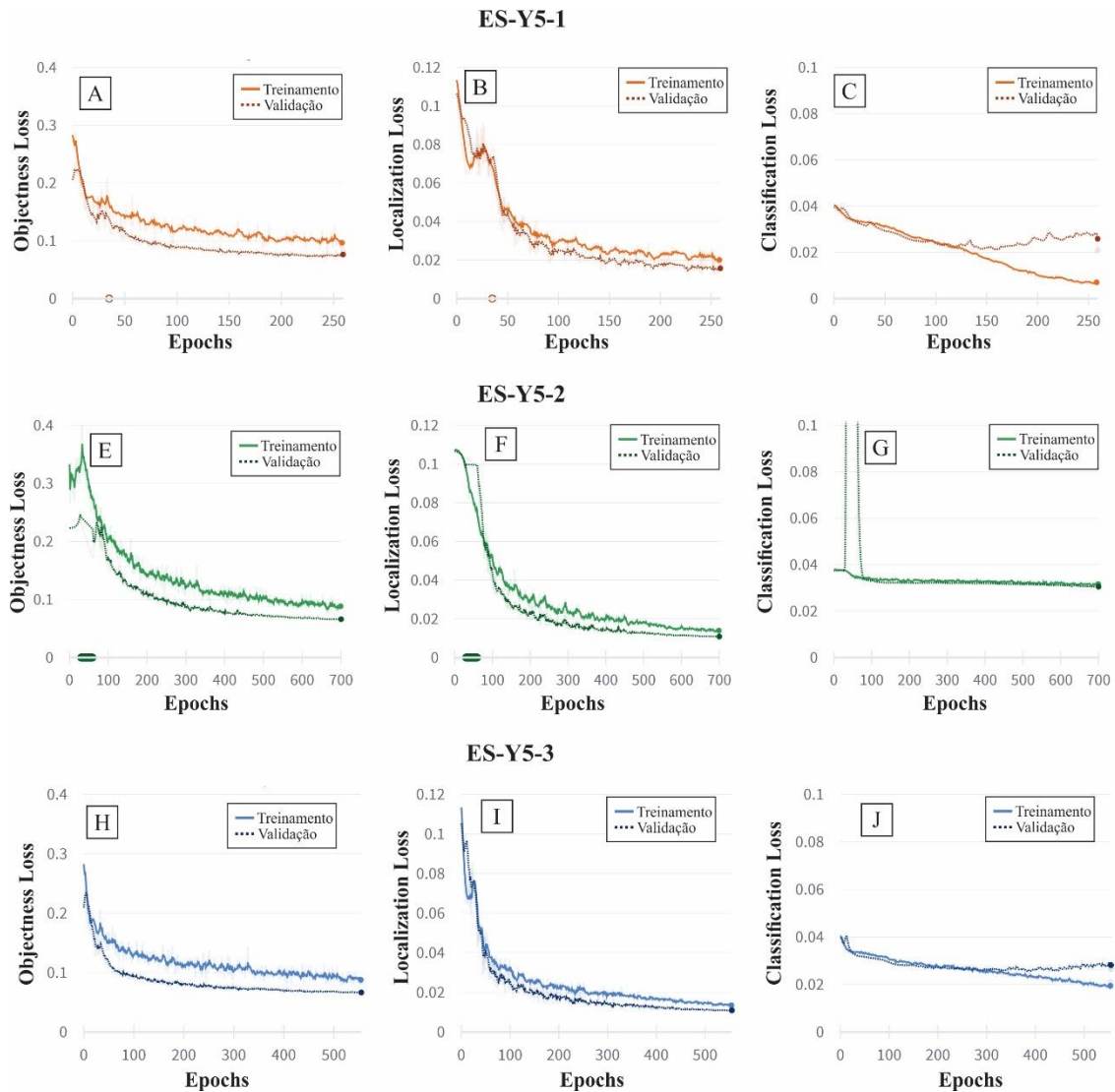


Figura 34 – Curvas de *loss functions* que compõem o YOLOv5 para os dados de treinamento e de validação (pontilhado). O efeito de *overfitting* pode ser visto na *classification loss*, nos gráficos C, G e J.

Tabela 11 – Configuração do otimizador SGD e ADAM para aprimorar o modelo de linha base.

Modelo	Otimizador	Learning rate	Momentum	P	R	mAP@.5	mAP@.5:.95
ES-Y5-1	SGD	0,01*	0,937*	0,7197	0,7217	0,6627	0,5519
ES-Y5-1-A1	SGD	0,01	0,9	0,7055	0,7168	0,6375	0,5324
ES-Y5-1-A2	SGD	0,01	0,99	0,6932	0,6915	0,6415	<b>0,5635</b>
ES-Y5-1-B1	SGD	0,001	0,937	0,655	0,659	0,571	0,511
ES-Y5-1-B2	SGD	0,001	0,9	0,628	0,6213	0,5261	0,46
ES-Y5-1-B3	SGD	0,001	0,99	0,6793	0,6946	0,6636	<b>0,6115</b>
ES-Y5-1-C1	SGD	0,0001	0,9	0,73	0,351	0,294	0,212
ES-Y5-1-C2	SGD	0,0001	0,99	0,62	0,582	0,535	0,48
ES-Y5-1-D1	ADAM	0,001**	0,937**	0,6763	0,6828	0,634	<b>0,5853</b>
ES-Y5-1-D2	ADAM	0,001	0,9	0,7	0,681	0,639	0,547
ES-Y5-1-D3	ADAM	0,001	0,99	0,663	0,389	0,338	0,279
ES-Y5-1-VOC	SGD	0,00334	0,74832	0,6426	0,6377	0,5957	0,5479

NOTA: \*Valores padrão para o otimizador SGD; \*\*Valores padrão para o otimizador ADAM. As siglas A, B, C, D (seguidas de 1, 2, 3) se referem a modelos com Lr. de 0,01 com *momentum* alternando respectivamente em 0,937, 0,9 e 0,99; VOC significa a combinação de hiperparâmetros no treinamento do banco de dados PASCAL VOC.

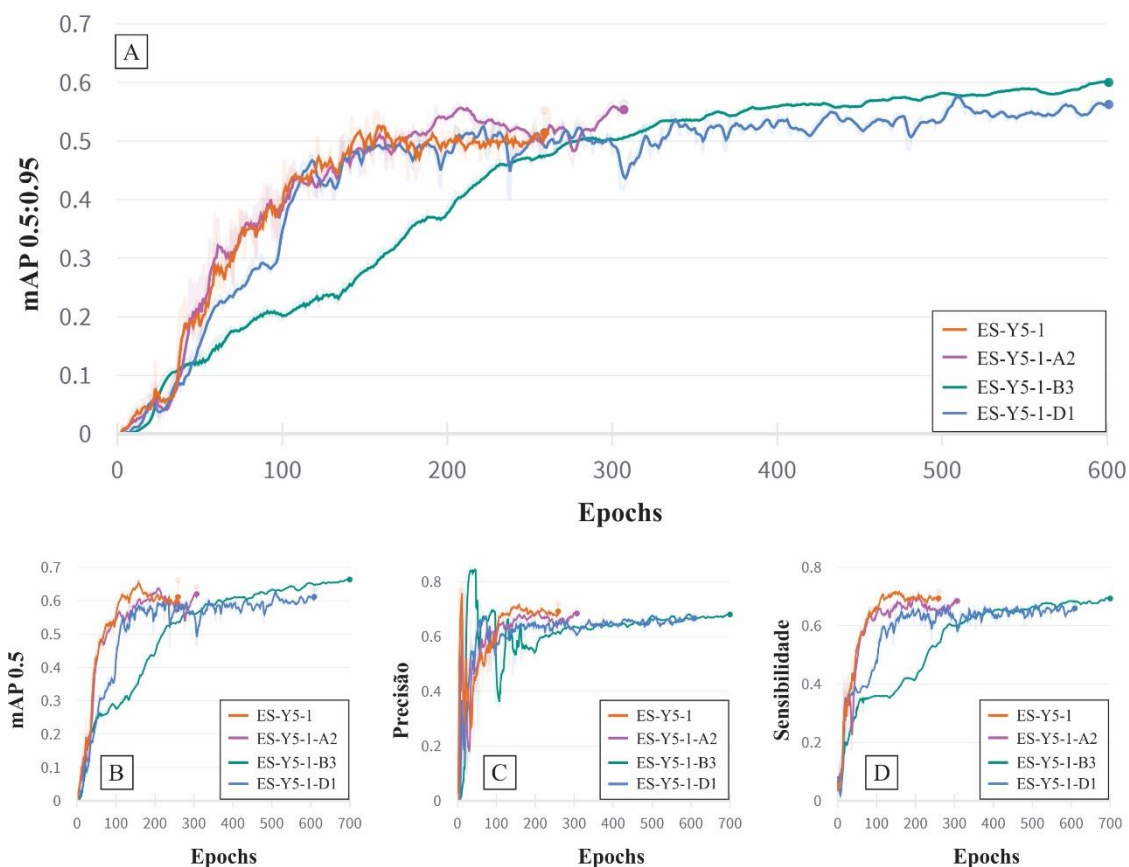


Figura 35 – Comparação entre os modelos derivados do ES-Y5-1 (em laranja) com diferentes hiperparâmetros com destaque para a métrica principal: (A) Valores da mAP ao longo do treinamento, para valores de *confidence score* 0,5 a 0,95. (B) Valores de mAP a 0,5 de *confidence score*. (C) Valores de P ao longo do treinamento. (D) Valores de R ao longo do treinamento.

Ao contrário do conjunto de dados de catodoluminescência, o modelo ES-Y5-1-VOC obteve resultados piores que o ES-Y5-1, com pelo menos 1500 *epochs* de treinamento. Os melhores resultados foram os modelos ES-Y5-1-A2, ES-Y5-1-B3 e ES-Y5-1-D1, todos apresentando *overfitting* ao longo do treinamento (Figura 36). Como opções para minimizar esse cenário idealizamos: (i) aplicar *early stopping* no modelo que apresentou o melhor resultado e no ponto com pouco *overfitting* nos dados de treino, (ii) melhorar o conjunto de dados aumentando a quantidade de objetos (via aplicação de *data augmentation*, p.ex.) ou (iii) utilizar o modelo que convergiu sem apresentar *overfitting*.

Apesar do *early stopping*, todos os modelos apresentaram uma diferença entre o *classification loss* (Figura 36F) entre os dados de validação e teste. Principalmente o modelo ES-Y5-1-B3 que teve seu treinamento sem interrupções, mostrando o melhor mAP@.5:.95 em 696 *epochs*. Portanto, apesar de ter uma pequena diferença de *classification loss*, o modelo estava conseguindo obter valores de mAP@.5:.95 cada vez melhores no conjunto de dados de validação (Figura 35A). Comparando-se os modelos -

A2 e -D1, estes apresentaram valores máximos de  $mAP@.5:.95$  menores e bem mais cedo no treinamento (*epoch* 206 e 508 respectivamente), ativando o *early stopping* e impedindo estes modelos de treinarem por mais tempo.

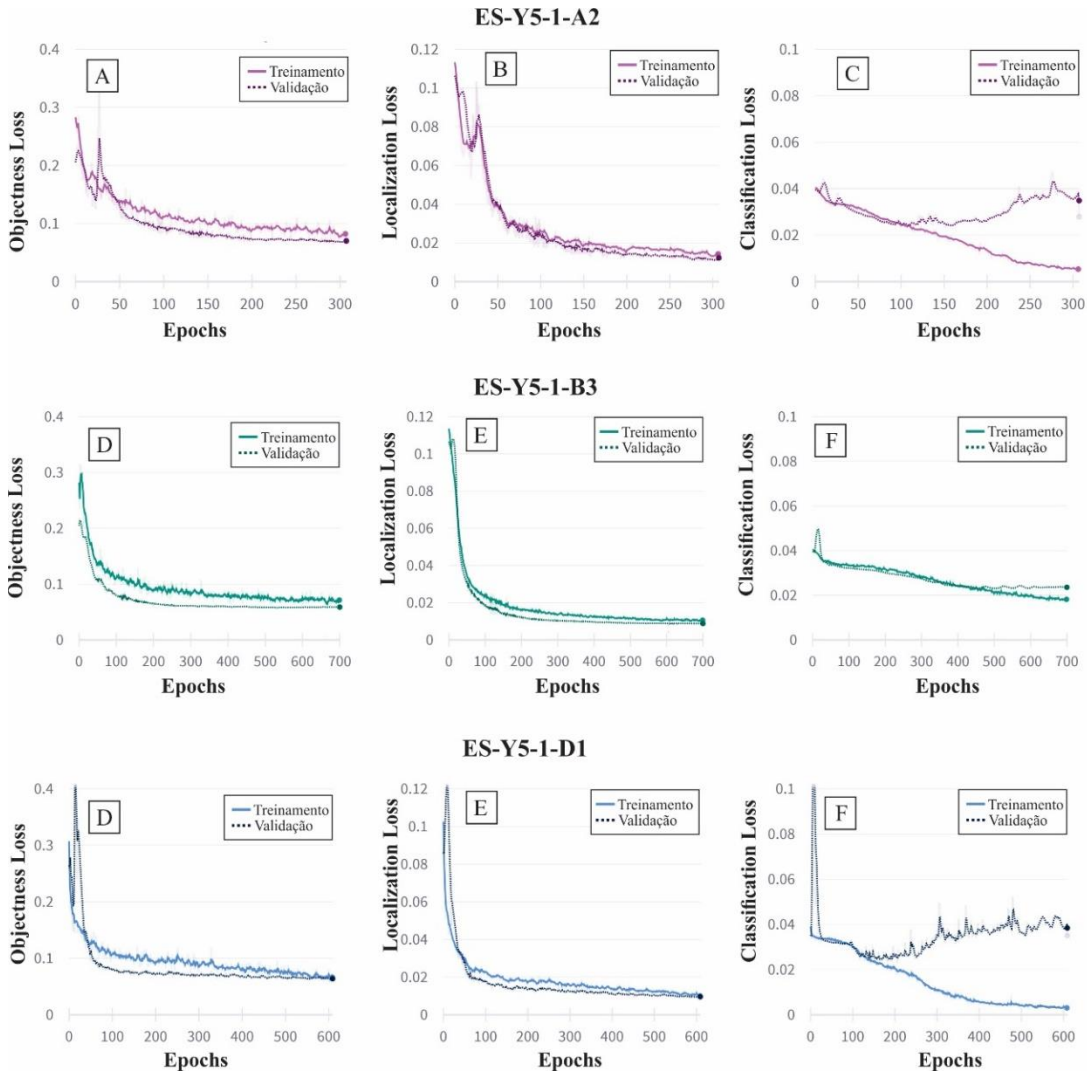


Figura 36 – Curvas de *loss functions* que compõem o YOLOv5 para os dados de treinamento e de validação (pontilhado). O efeito de *overfitting* é ilustrado pela função *classification loss*, especialmente destacado nos gráficos C, J, e menor em F.

A comparação entre os dois modelos que apresentaram o melhor desempenho, i.e. ES-Y5-1-B3 e ES-Y5-1-D1, foi feita entre os dados de validação e os dados de teste (ver Tabela 12 e Figura 37). Pela análise dos dados, o melhor modelo foi o ES-Y5-1-B3 com um ganho de 0,04 pontos em relação ao modelo ES-Y5-1, e com uma eficiência na classificação bem superior, principalmente na categoria 2 (Figura 37). Após essa escolha, e da mesma forma como feito para o conjunto de dados de catodoluminescência, escolheu-se o limiar de coeficiente pela curva F1 cujo melhor valor está no ponto 0,274

(Figura 38), e o valor de 0,2 para todas as detecções realizadas por este modelo como forma de evitar detecções FP.

Tabela 12 - Tabela comparando o desempenho entre os modelos com os melhores desempenhos, tendo o modelo ES-Y5-1 como linha de base. *Epoch* indica o momento em que esse desempenho foi alcançado no treinamento. Destaque foi dado para o melhor desempenho.

Dados de Validação					
Modelo	<i>epoch</i>	P	R	mAP@.5	mAP@.5:.95
ES-Y5-1	158	0,72	0,72	0,66	0,55
ES-Y5-1-B3	694	0,68	0,69	0,66	<b>0,611</b>
ES-Y5-1-D1	508	0,68	0,68	0,64	0,59
Dados de Teste					
Modelo	<i>epoch</i>	P	R	mAP@.5	mAP@.5:.95
ES-Y5-1	158	0,693	0,696	0,663	0,623
ES-Y5-1-B3	694	0,73	0,68	0,7	<b>0,66</b>
ES-Y5-1-D1	508	0,69	0,7	0,68	0,63

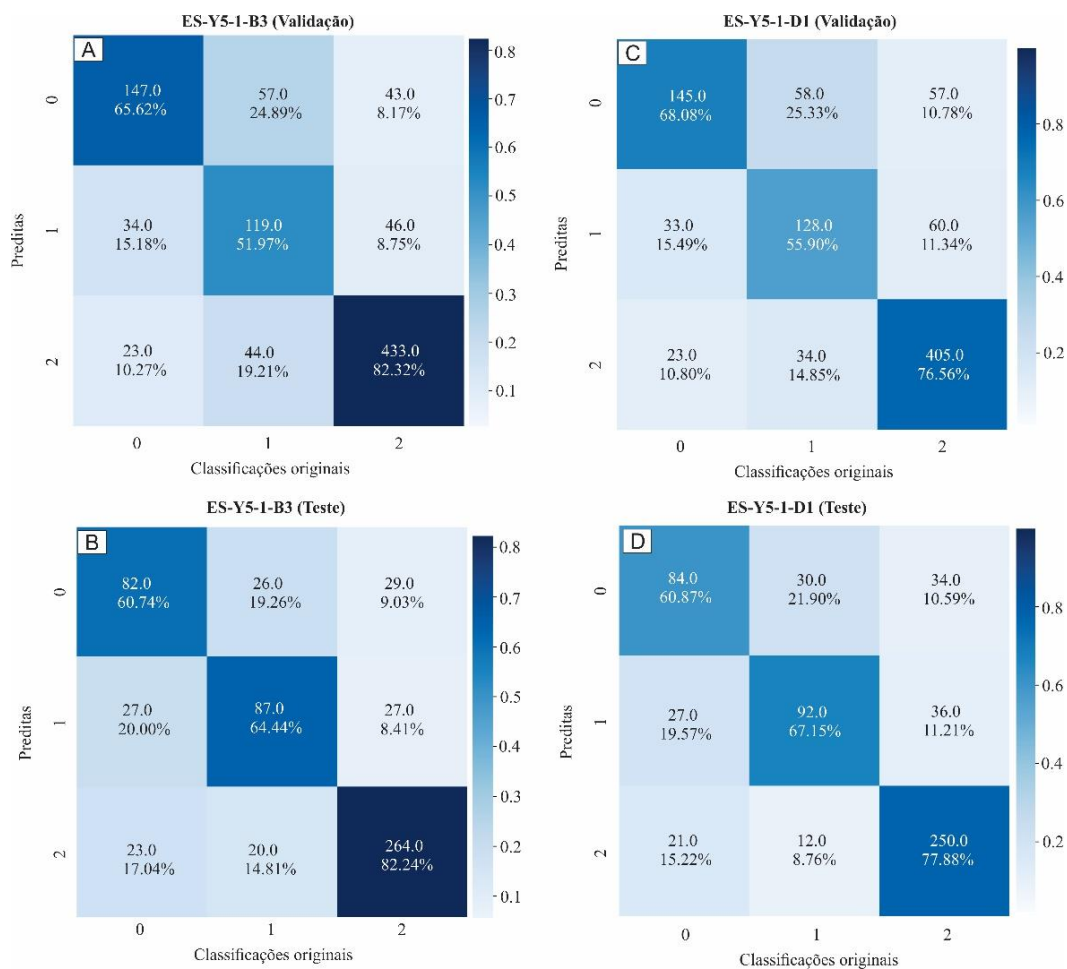


Figura 37 – Comparação da matriz de confusão de dois modelos, resultante da detecção realizada nos dados de validação e teste.

Tabela 13 – Valores de P e R derivados da matriz de confusão do modelo ES-Y5-1-B3

Dados de validação		
Classe	Precisão (P)	Sensibilidade (R)
0	0,60	0,72
1	0,60	0,54
2	0,87	0,83
Todas as classes	0,69	0,70
Dados de teste		
Classe	Precisão (P)	Sensibilidade (R)
0	0,60	0,62
1	0,61	0,65
2	0,86	0,82
Todas as classes	0,69	0,70

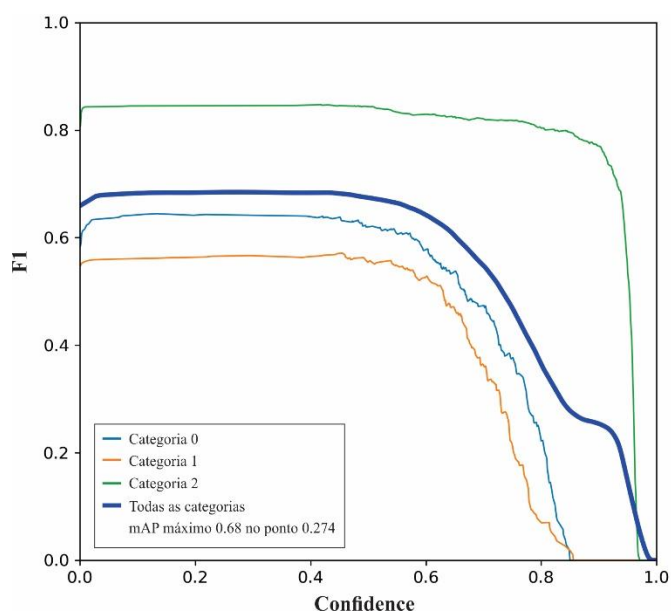


Figura 38 – Curva F1 para variações do *confidence score* no momento da inferência.

Um exemplo dos resultados das inferências feitas pelo modelo ES-Y5-B3 estão ilustrados na Figura 39. Necessário lembrar que o banco de dados de elétrons secundários foi reduzido (em aproximadamente metade) em comparação ao de catodoluminescência, não tendo sido incluídas as imagens de grãos dos arenitos da Bacia do Cocóci, caracteristicamente da categoria 1. Assim, o modelo ficou relativamente defasado em treinar essa categoria.

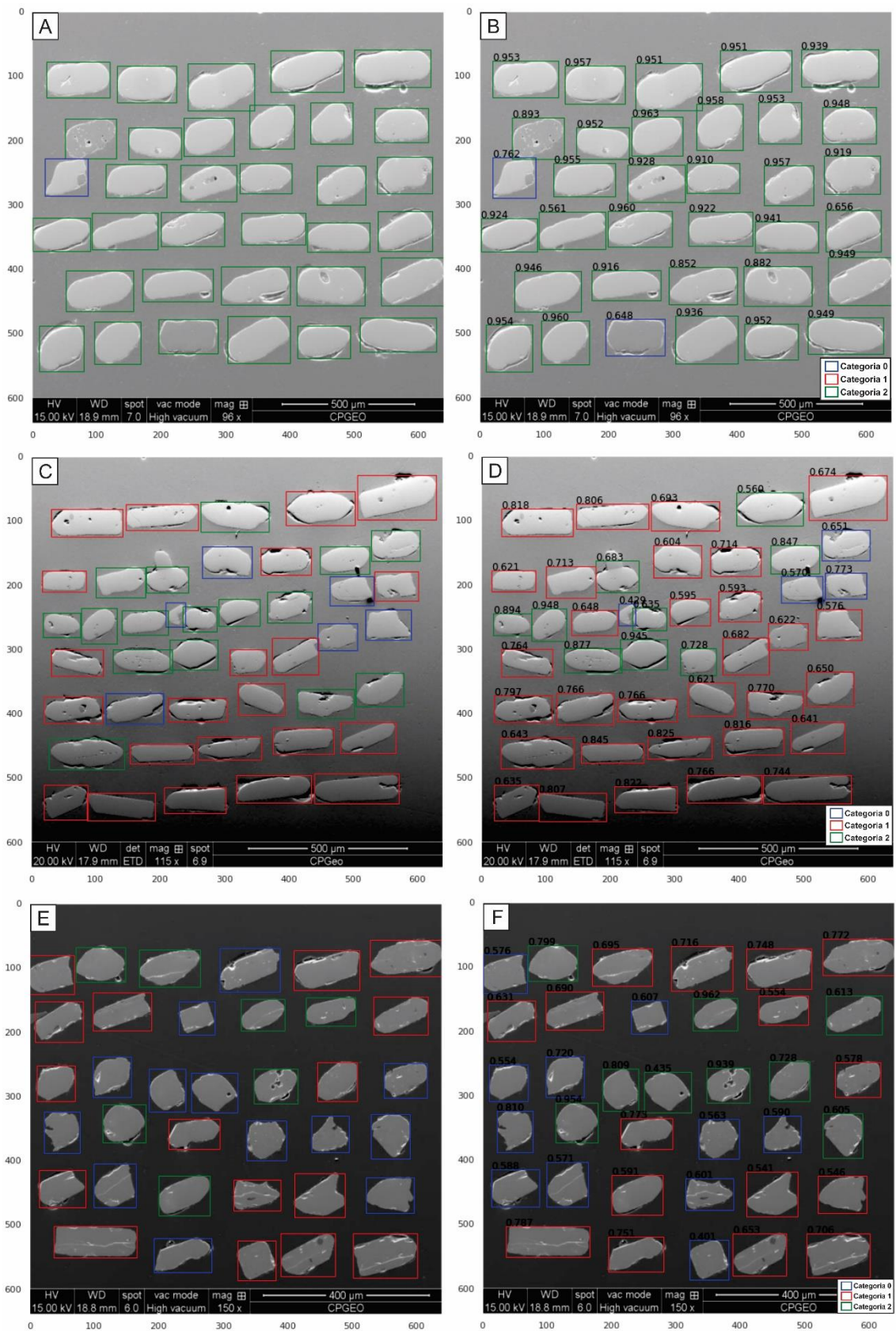


Figura 39 – Exemplo de detecção feita pelo modelo ES-Y5-1-B3. À esquerda vê-se as classificações originais. À direita, as detecções do modelo. Em (A) e (B), têm-se grãos da Formação Sambaíba. Em (C) e (D), grãos do Grupo Serra Grande com predominância da categoria 1. Em (E) e (F), têm-se grãos do Grupo Serra Grande (mistura das categorias 0, 1 e 2).

## 5. DISCUSSÕES E CONCLUSÃO

Neste capítulo são comentados os resultados de treinamento, validação e testes do modelo obtido para o conjunto de dados de catodoluminescência, que representou a primeira escolha de objetos a serem analisados neste projeto. De forma complementar, são comentados os resultados preliminares do treinamento com os dados de elétrons secundários, os quais só foram considerados para análise numa fase mais tardia do projeto e, por esta razão, ainda devem ser melhor processados futuramente.

### 5.1. Limitações do processo de classificação morfológica

Os parâmetros de desempenho listados na Tabela 8 e a matriz de confusão Figura 26, mostram que os conjuntos de validação e de teste dos modelos CL-Y5-1-C2 e CL-Y5-1-VOC têm comportamentos semelhantes, uma vez que para ambos os conjuntos houve uma classificação ‘errada’ de grãos das categorias 0 (quebrados) e 2 (anédricos arredondados a subarredondados) identificados como categoria 1 (euédricos a subédricos, i.e., aqueles mais preservados a despeito da ação do transporte sedimentar). Por outro lado, e especialmente para o conjunto de dados de validação, o modelo tende a classificar (muitos) grãos da categoria 1 dentro da categoria 2, um erro que atingiu 19% no conjunto de validação (i.e, 19% dos grãos da categoria 1 foram classificados erroneamente) contra 12% no conjunto de teste. Esse comportamento traduz-se numa pontuação discretamente maior de mAP nos dados de teste em relação aos dados de validação.

De forma geral, nota-se que os dois modelos tiveram um certo grau de dificuldade em reconhecer corretamente cada categoria devido à ‘confusão’ com atributos físicos não específicos (Figura 40), especialmente feições de quebraimento e zonação interna. Ademais, o modelo apresentou um elevado erro de classificação da categoria 1, presente nos dados de validação, mas não presente nos dados de teste, mostrando que apesar dos dois conjuntos de dados terem a mesma fonte (população), possuem diferenças entre si. Uma hipótese sobre o porquê desse erro seja a natureza do banco de dados usados nesta pesquisa de mestrado. O quebraimento (ou fraturamento) artificial é, por vezes, uma característica ‘herdada’ do processamento mecânico (britagem, moagem) das amostras de rocha para a separação do zircão, levando ao desenvolvimento de arestas que podem dar ao modelo a ‘impressão’ de serem faces cristalográficas naturais. Essa questão teve grande impacto na classificação dos grãos dentro da categoria zero (quebrados).



Observando a matriz de confusão (ver Figura 26), nota-se que o desbalanceamento entre as categorias permaneceu significativo mesmo após a aplicação do método de *data augmentation* para esta categoria.

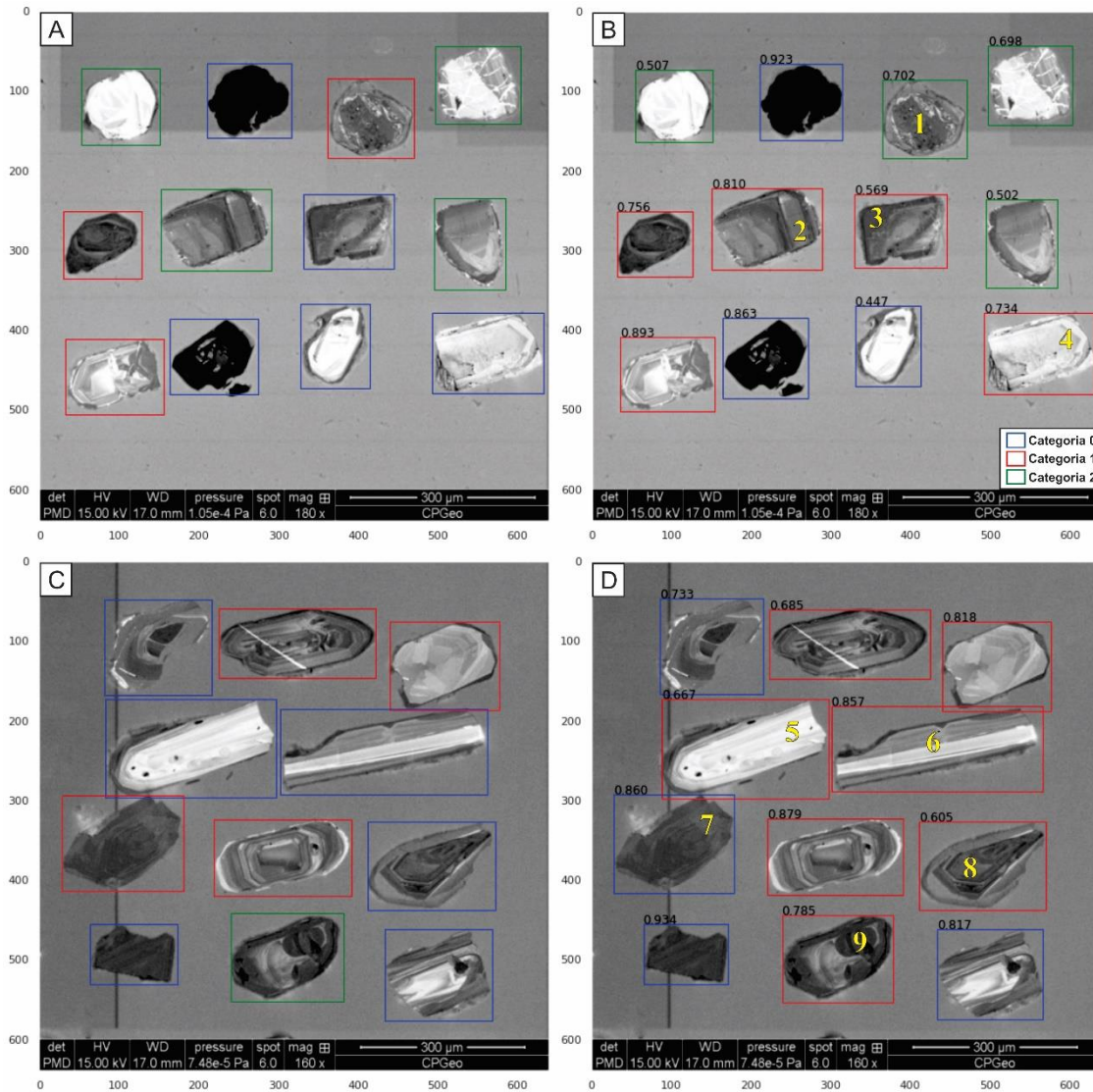


Figura 40 – Pranchas exemplificando situações em que o modelo CL-Y5-1-C2 ‘errou’ na classificação dos grãos entre as categorias pré-definidas. À esquerda (pranchas A e C), são mostradas as classificações originais feitas pelo autor; à direita (pranchas B e D) estão as classificações entregues pelo modelo. No par A-B, os grãos identificados com 2-3-4 foram classificados pelo modelo na categoria 1 provavelmente devido à confusão de arestas artificiais com faces cristalográficas. Note que o grão 1, em (B), também foi classificado erroneamente na categoria 2 provavelmente devido ao padrão interno (zonação) irregular, dando uma ‘impressão’ de grão anédrico. No par C-D, os grãos 5-6-8-9 também são classificados erroneamente na categoria 1 dado o efeito de zonação interna, o inverso acontecendo com o grão 7.

Um dos fatores que pode influenciar na dificuldade de rotulação de grãos quebrados é a presença de alguns elementos característicos deste tipo nas outras categorias, como por exemplo pontas/vértices, reentrâncias e fraturas. Os grãos euédricos/subédricos de

zircão da categoria 1 (e em menor proporção, da categoria 2) não raro também possuem vértices desenvolvidos quando de sua cristalização a partir do resfriamento do magma, ou gerados em decorrência do quebramento durante o transporte (por tração, p.ex.) sedimentar, desde sua área-fonte até o sítio de deposição. Fraturamento e quebramento parciais são mais comuns em minerais detríticos que têm originalmente hábito prismático, como é o caso do zircão. Esse mineral tem, em geral, alongação superior a 2:1, podendo atingir razões 5:1 ou 6:1 quando cristalizado em rochas vulcânicas (ver grão 7 da Figura 40; prancha D) (Corfu et al., 2003). O hábito prismático original do zircão é um aspecto que dificulta (se não, complexa) à análise de um outro aspecto morfológico usado em estudos de proveniência sedimentar já que é a esfericidade – tendência da morfologia do grão de se aproximar de uma esfera por retrabalhamento durante o transporte. Arredondamento e esfericidade são mais aplicáveis na análise de grãos originalmente anédricos como é o quartzo, cujo crescimento tardio (intersticial) em rochas ígneas dificulta o desenvolvimento do seu hábito e o espaço confinado favorece a aumentar o grau de esfericidade.

Apesar da limitação, a inclusão da categoria zero neste trabalho justifica-se para acomodar um conjunto razoável de grãos quebrados que, a princípio, poderiam tanto ser resultado do processamento das amostras em laboratório quanto do transporte sedimentar. Infelizmente, esse atributo físico não é próprio para bem treinar o modelo e bem distinguir a categoria zero das demais. Como consequência, entendemos que essa categoria pode ser sub- ou sobrestimada num banco de dados baseado em grãos de zircão detrítico proveniente de datação radiométrica. A Figura 40 ilustra os problemas de classificação equivocada do modelo especialmente em decorrência de quebramento/fraturamento e hábito dos grãos.

Dentre as três categorias, o modelo escolhido CL-Y5-1-C2 melhor distinguiu a categoria 2 quando comparados os valores de mAP listados na Tabela 9 e pela matriz de confusão B da Figura 26 muito embora o modelo também teve bom desempenho em discriminar a categoria 1. Os grãos da categoria 1 têm por atributo característico a presença de terminações bipiramidais bem desenvolvidas, facilitando o treinamento e testes (Figura 31A).

Dois outros pontos negativos na classificação merecem ser destacados. O primeiro refere-se ao uso, por parte do modelo, de atributos internos (zonação) dos grãos para rotular em uma ou outra categoria. Necessário dizer que a estrutura interna de grãos de zircão (ígneos e metamórficos) é retratada em imagens de catodoluminescência pela

variação de bandas de composições distintas formadas durante o crescimento do cristal no magma resfriando. Em geral, a variação composicional reflete alternância de elementos químicos (p.ex., U, Th) de alto peso atômico. Em geocronologia, o imageamento por catodoluminescência é uma etapa indispensável que precede a análise *in situ* propriamente dita, requerida para guiar o usuário a reconhecer o melhor sítio para datação dentro de cada grão. Portanto, indiferente à sua natureza, todo grão de zircão mostrará um determinado padrão de zonação. Tomando a Figura 39 para ilustrar, os padrões mais comuns são o oscilatório (grãos 2-5-7-8) e o paralelo (grão 6), mas um grão pode também apresentar padrão interno homogêneo (sem zonação) ou ainda adquirir padrões após sua completa formação. Nesses casos incluem-se grãos metamitizados (grão 1) ou com padrões mais complexos (grão 9) (Corfu et al., 2003).

A zonação interna teve peso significativo na classificação. Para grãos com uma zonação interna tipo oscilatório ou paralelo, o modelo demonstrou uma tendência à inclusão na categoria 1 (euédricos/subédricos). Num teste feito para ‘limpar’ a zonação, notou-se que vários desses grãos foram transferidos de categoria, geralmente para a categoria 2. Mesmo aplicando esse método alguns poucos casos ainda foram mantidos.

O segundo ponto negativo refere-se à influência que a orientação do grão (i.e., do eixo de máxima alongação) exerce na classificação. Como visto na Figura 30, a rotação artificial da imagem induziu o modelo a reclassificar os grãos com uma tendência ao incremento da categoria zero, indicando uma relativa predileção quando o eixo maior do grão é ‘verticalizado’. Ao passo que os grãos da categoria 2 são mais ‘resistentes’ a essa alteração por rotação, pois os mesmos apresentam eixo de maior alongação, geralmente, menor que os grãos da categoria zero e 1, em decorrência do processo de retrabalhamento ao longo do transporte.

Dos 685 grãos analisados na etapa de teste, cerca de 128 grãos foram reclassificados ao rotacionar as pranchas em 90° no sentido horário, resultando em um percentual de ~19% de modificação. Infelizmente não foi possível verificar um padrão consistente de ‘reclassificação’ de forma a tentar minimizar ou mesmo evitar essa interferência, algo que provavelmente será abordado com maior rigor em etapas futuras do desenvolvimento do modelo.

Outros aspectos do tipo de objeto também mostraram algum grau (em geral, pequeno) de interferência na classificação, como contraste de do sinal de luminescência (tons de cinza) entre grão e *background*, e a proximidade entre grãos.

Todas as limitações citadas acima foram gradualmente sendo percebidas durante o desenvolvimento deste projeto, se traduzindo como vieses do próprio banco de dados. Desde a seleção criteriosa dos grãos para treinamento-validação-teste até a aplicação de métodos e filtros que garantissem melhor desempenho do modelo e, lógico, maior eficiência na classificação, foram todas etapas que envolveram intenso trabalho. Ainda assim, talvez não tenham sido suficientes para contornar a complexidade inerente ao banco de dados. Essa percepção levou a iniciar uma segunda fase do trabalho que está aqui apresentada apenas como ensaio preliminar, e diz respeito a testes com as imagens de elétrons secundários. Essas imagens não retratam alguns atributos, como a zonação interna, sendo uma alternativa potencial às imagens de catodoluminescência (ver item 5.2, adiante).

## **5.2. Desempenho do modelo**

Como já mencionado, o modelo teve um bom desempenho na classificação dos grãos da categoria 2, apresentando altos valores de P e R. De todas as predições que o modelo realizou da categoria 2, 89% estavam corretas, e para o conjunto de dados de teste, o modelo classificou corretamente 88% dos grãos desta categoria. O pior desempenho deu-se na categoria zero para a qual R teve valor abaixo do esperado, de 66%. Já para a categoria 1 o modelo apresentou o menor valor de P devido à tendência de interpretar qualquer padrão de zonação (mesmo que incipiente) como as bordas bem formada de grãos euédricos; também influenciou para o pior desempenho para esta categoria a forte similaridade entre os grãos subédricos fraturados com os grãos quebrados da categoria zero. Infelizmente esse pode ser um viés incorporado durante a fase de treinamento proveniente da própria dificuldade em caracterizar corretamente a categoria zero.

## **5.3. Comparação do desempenho do modelo versus usuários**

O mesmo conjunto de dados (n=685 grãos) que usado para teste do modelo CL-Y5-1-C2 foi também apresentado a dois pesquisadores (ou usuários) que tiveram o primeiro contato com este projeto exclusivamente para contribuir nesta fase. O objetivo foi comparar a resposta dada pelo modelo com a percepção dos dois usuários independentes, e avaliar eventuais discrepâncias na classificação. Os critérios de

classificação foram os mesmos em que o modelo foi treinado, ou seja, as morfologias como estabelecidas para as categorias 0, 1 e 2. Algo a ser considerado, foi que devido a circunstâncias excepcionais da época em que ocorreu este estudo, não foi possível efetuar um treinamento prévio com os usuários para acostumá-los a classificação desenvolvida ou ao tipo de imagem utilizada. Mesmo assim, apesar de todas essas limitações, achou-se importante realizar uma comparação inicial, a fim de perceber o quão subjetivo poderia ser a classificação final (resumida em 3 categorias) para pessoas fora do estudo.

De maneira geral, os dois usuários tiveram forte inclinação a classificar o conjunto de grãos na categoria 2 (genericamente os grãos anédricos), seguindo a mesma tendência que o modelo, mas com diferenças fundamentais. Dos vários grãos rotulados como arredondados, muitos não eram condizentes com a classificação original estabelecido no treino do modelo. Por consequência, os valores de P e R foram inferiores, respectivamente 75% e 80%, quando comparados aos valores do modelo. Valores igualmente mais baixos de P e/ou R foram determinados para as categorias zero e 1 classificadas pelos usuários. Por exemplo, grãos quebrados foram classificados como arredondados (identificado pelo baixo valor de R da categoria zero para o usuário A), assim como grãos arredondados foram rotulados como euédricos (i.e., baixo valor de P da categoria 1, para o usuário B). As pranchas de grãos foram entregues aos usuários para classificação em um curto período após a definição final das categorias. Nesse meio tempo, a classificação definida neste trabalho não tinha tido o refino necessário para ficar claro os limites da morfologia que cada categoria abrange. Inclusive, quando a solicitação foi feita aos usuários foi enviado um texto com explicações que tinha apenas uma ilustração das categorias zero, 1 e 2 como da Figura 18, porém, de forma resumida. Neste sentido, é possível que a discrepância entre a classificação realizada pelos usuários tenha sido decorrente de uma descrição ainda imatura sobre estas categorias. De forma que levanta questionamentos se a classificação aqui definida não gera equívocos, já que em todas as classificações, tanto dos usuários, quanto do modelo, existe sempre uma certa “confusão” entre as categorias zero e 1, ou entre as categorias zero e 2, demonstrando que o problema maior provavelmente se encontra na categoria zero. Evidentemente, a possibilidade destes problemas decorrerem apenas da subjetividade inerente ao estudo morfológico também não foi excluída. Portanto, é importante um número maior de amostragem de usuários para esclarecer essas dúvidas levantadas. Devido ao grande trabalho decorrido da preparação do banco de dados, não foi possível conseguir mais usuários com experiência em sedimentologia para testar essa classificação.

Quando comparados os resultados dos dois usuários, nota-se que particularmente o usuário B deu peso maior para a categoria zero (quebrados), apesar do conjunto de dados ser constituído maioritariamente por grãos da categoria 2 (58% de grãos arredondados). Pela Tabela 9 vê-se que a classificação fornecida por este usuário obteve um valor alto de P (94%), enquanto um valor baixo de R (60%). Isso significa que, dentre todos os grãos da categoria 2 poucos foram de fato classificados como tal; todavia, os que foram classificados, o foram corretamente. Como consequência da tendência em classificar ‘erroneamente’ grãos na categoria zero (p.ex., 102 grãos de 2), os resultados forneceram baixo valor de P (46%) e alto valor de R (86%) para a categoria. Assim, dentre todos os grãos da categoria zero presentes no conjunto de dados, muitos foram corretamente classificados nesta categoria.

Quanto ao valor de F1, o qual indica qual das classificações obteve o melhor equilíbrio de P-R, o algoritmo sem dúvida teve mais acertos na classificação (ver Tabela 9). É necessário comentar que a capacidade visual humana em identificar variações sutis em formas naturais é em muito afetada pela subjetividade, um usuário não exatamente respondendo igual ao outro. A princípio, não se deve rotular ‘erros e acertos’ nesse tipo de classificação. Mesmo o modelo foi treinado em cima de informações inseridas por um humano (neste caso, o autor), mas com uma maior robustez pois foi feito cruzando as características de cada grão sob catodoluminescência e elétrons secundários, também combinando a luz transmitida. Por essa razão, é esperado que o modelo tenha melhor desempenho, mas não que possa ser considerado como um “classificador absoluto”, imune a inconsistências ou ‘erros’. O mais importante na contribuição dos usuários foi avaliar o desempenho do modelo e decidir (ver item 5.5) se ele pode ser oferecido à comunidade científica como uma alternativa automatizada e confiável que venha substituir a análise humana.

#### **5.4. Perspectivas futuras**

Três pontos podem ser mencionados como possíveis melhorias do algoritmo/modelo para análise morfológica de grãos detríticos (a princípio, zircão). O primeiro deles refere-se ao uso de imagens de elétrons secundários em substituição de

imagens de catodoluminescência. Seguindo a mesma linha de raciocínio dos itens 5.1 e 5.2, o modelo ES-Y5-1-B3 forneceu resultados de validação e de teste semelhantes ao modelo CL-Y5-1-C2, com os mesmos pontos positivos e limitações (ver parâmetros na Tabela 13; Figura 41). Um viés do modelo ES-Y5-1-B3 é colocar os grãos com alongação maior (i.e., razão comprimento:largura > 2:1) preferencialmente na categoria 1 e, subordinadamente, na categoria 2, e quase nada classificado na categoria zero, muito embora inúmeros representantes desta categoria tenham razão de alongação nesta ordem de grandeza. Mesmo tendo sido treinado com um número reduzido de objetos se comparado com o modelo CL-Y5-1-C2, o modelo ES-Y5-1-B3 teve valor de  $mAP@.05:.95$  superior a 0,6 tanto para os dados de validação como de teste. É, portanto, um resultado promissor que pode ser melhorado com a inclusão de um número maior de objetos ao banco de dados, e aplicação do método de *data augmentation* às categorias ou a alguma categoria em específico.

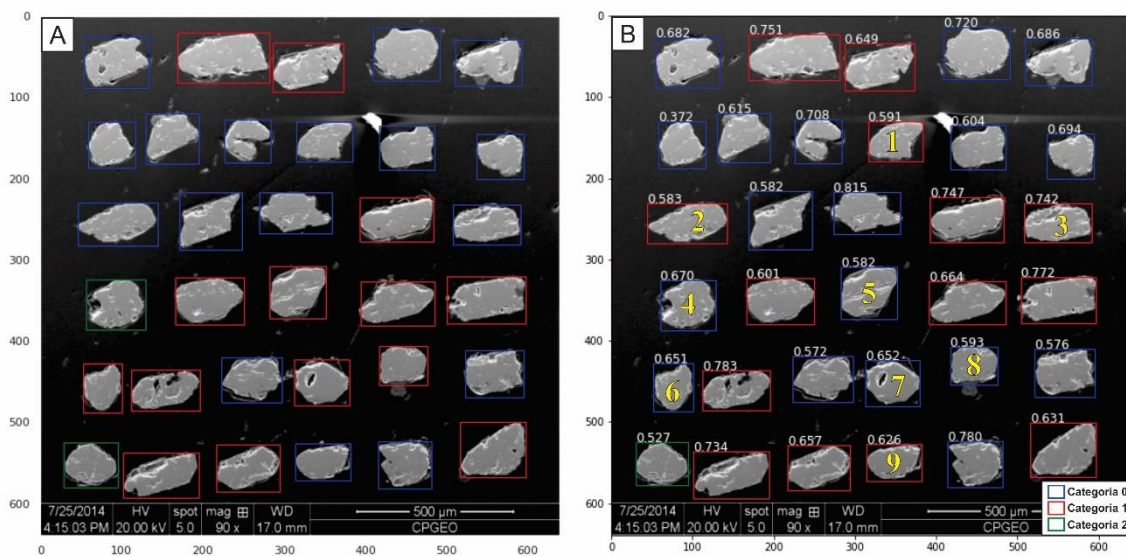


Figura 40 – Pranchas com imagens de elétrons secundários de grãos exemplificando situações em que o modelo ES-Y5-1-B3 não classificou corretamente. Em (A) tem-se a classificação original, e em (B) a detecção feita pelo modelo com o valor de *confidence score*. Em amarelo estão indicados os grãos classificados incorretamente: grãos originalmente da categoria zero classificados na categoria 1 (grãos 1, 2, 3 e 9) e vice-versa (grãos 5, 6, 7 e 8); o grão 4 passou da categoria 2 para a categoria zero.

Independente de se adotar, no futuro, imagens de elétrons secundários como forma de contornar as inconsistências na classificação influenciadas pela presença de zonação interna, é necessário considerar uma “limpeza” pré-treinamento de objetos (grãos) que não atendem os critérios definidos para as categorias. É real a existência de grãos que não se encaixam na classificação por estarem “na fronteira” entre as categorias,

algo comum quando se lida com banco de dados naturais. Obviamente, a limpeza deve ser considerada até um limite que não implique em redução significativa na quantidade de objetos, o que traria dificuldades para detecção – algumas categorias poderiam ser sub-representadas. Para este caso em que os objetos são imagens 2D de zircões detríticos (catodo ou elétrons secundários), o banco de dados pode ser expandido para um mínimo de 10 mil objetos por categoria. Obviamente, aumentando-se o número de objetos aumenta-se muito o trabalho de pré-processamento do banco de dados, o que torna o desenvolvimento do produto final ainda mais complexo e laborioso. De toda forma, é uma questão a ser considerada para um futuro próximo.

Não menos importante é o teste comparativo feito com usuários que mostrou diferenças entre eles, seja porque os critérios pré-definidos para as categorias são sutis e/ou dúbios para uma análise visual precisa, seja porque o grau de arredondamento é uma propriedade naturalmente subjetiva e de difícil consenso. Como dito anteriormente, a classificação para o treinamento do modelo foi realizada pelo autor com informações extraídas de três tipos de imagens, uma situação que não foi oferecida aos usuários que utilizaram apenas a imagem de catodoluminescência. Logo, espera-se organizar num futuro próximo nova rodada de testes com usuários em que tais discrepâncias sejam minimizadas ou completamente contornadas. Essa é uma etapa à parte do desenvolvimento do algoritmo que requer tempo e logística para garantir resultados mais robustos, e que deve essencialmente preceder a disponibilização do produto final para uso na comunidade científica.

Uma evolução natural do trabalho seria a de criar outros classificadores considerando mais de uma propriedade ao se definir as categorias. Portanto, com o banco de dados montado, tem-se como objetivo utilizar cada tipo de imagem para uma detecção diferente. Para as imagens de elétrons secundários, a detecção será da morfologia dos grãos de zircão. Com as imagens de catodoluminescência, existe a possibilidade de detectar, além da morfologia, a estrutura interna do grão (padrão de zonação) e a sua ‘luminosidade’, dessa forma, pode-se usá-las para discriminar grãos arredondados de origem metamórfica (com pouco transporte, ou retrabalhamento) dos grãos arredondados de origem ígnea que sofreu retrabalhamento.

Por fim, a continuação deste projeto deverá avaliar a utilização de duas novas gerações da arquitetura YOLO – YOLOv7, ou versão 7 da YOLO, que vem apresentando um melhor desempenho na literatura, na inferência do conjunto de dados COCO. Por



outro lado, a versão possui algumas críticas da comunidade, ao apresentar alguns erros na *loss function* ao longo do treinamento, ainda não reparados.

## 5.5. Comentários finais

O sucesso no desenvolvimento deste projeto dependeria de um conjunto de fatores que não puderam ser dimensionados à época de sua idealização, razão pela qual ainda não é possível entregar um produto que atenda plenamente às expectativas e objetivos traçados originalmente.

Comparando-se a matriz de confusão (Figura 32) dos dois usuários com o resultado do modelo, percebe-se que o arredondamento é uma propriedade física que carrega consigo uma dose razoavelmente alta de subjetividade quando avaliada através de imagens 2D. Adicionalmente, percebe-se que a classificação desenvolvida neste trabalho possui espaço para refinamento, e mesmo ela sendo bastante simplificada, ainda carrega subjetividade significativa, provocando imprecisão tanto para o modelo quanto para humanos. Esses dois fatores combinados, aliado com o fato de modelos da CNN considerarem todos os elementos que constituem o objeto (o que se encontra dentro da caixa delimitadora) naquela imagem, torna este tipo de técnica um pouco limitada para a classificar apenas o arredondamento de grãos de zircão. Para se tornar aplicável, existe a necessidade de o modelo ter, pelo menos, uma evolução na métrica da categoria 1. Se por ventura, o modelo distinguir bem essas duas categorias (categoria 1 e 2), que são os extremos do transporte sedimentar, acredita-se na sua aplicabilidade. Entretanto, é possível observar que o algoritmo, apesar de possuir dificuldades, apresentou resultados positivos dentro das métricas avaliadas, com resultados elevados para o tamanho dos dados de treino e um pouco superior as métricas dos usuários, algo esperado devido ao viés inerente ao tema e as dificuldades relatadas anteriormente.

O que se recomenda é a utilização desta técnica para identificação e classificação de objetos com características que, quando combinadas, os tornem objetos mais singulares. E neste sentido, as sugestões para o aprimoramento são elaborar um modelo para cada tipo de imagem detectando objetos com propriedades evidenciadas em cada estilo. A seguir, testá-los com especialistas da área e desta forma aumentar a amostragem estatística e com *feedbacks*, permitindo o refinamento da classificação. A melhoria também seria alcançada através da expansão e limpeza dos conjuntos de dados para aperfeiçoar os resultados presentes aqui. Finalmente, a meta seria criar um conjunto de

dados definitivo e bem representativo, para que esses modelos possam classificar qualquer grão de zircão de qualquer população.

Por fim, ressalta-se a importância das técnicas apresentadas aqui, mais especificamente a detecção de objetos. Onde a modelagem por meio de CNNs permite agilizar trabalhos manuais extensos, trazendo possibilidades de classificação de diferentes propriedades e minerais, onde quando correlacionadas, possibilita *insights* sobre os dados e o problema geológicos complexos/ambíguos.

## **REFERÊNCIAS**

Andò, S., Garzanti, E., Padoan, M., Limonta, M.M.P., 2012, Corrosion of heavy minerals during weathering and diagenesis: A catalog for optical analysis: *Sedimentary Geology*, v. 280, p. 165–178, doi:10.1016/j.sedgeo.2012.03.023.

Bhattacharya, J.P., Copeland, P., Lawton, T.F., Holbrook, J., 2016, Estimation of source area, river paleo-discharge, paleoslope, and sediment budgets of linked deep-time depositional systems and implications for hydrocarbon potential: *Earth-Science Reviews*, v. 153, p. 77–110, doi:10.1016/j.earscirev.2015.10.013.

Bhattacharai, S., 2018, What is gradient descent in machine learning?: <https://saugatbhattacharai.com/np/what-is-gradient-descent-in-machine-learning/> (acessado em setembro de 2020).

Bishop, C., 1995, *Neural networks for pattern recognition*, Oxford University Press, USA, 482p.

Blatt, H., 1967, Original characteristics of elastic quartz grains: *Journal Sedimentary Petrology*, v. 37, p. 401- 424.

Blott, S.J., Pye, K., 2008, Particle shape: a review and new methods of characterization and classification: *Sedimentology*, v. 55, p. 31-63, doi: 10.1111/j.1365-3091.2007.00892.x.

Bochkovskiy, A., Wang C.-Y., Liao, H.-Y.M., 2020, YOLOv4: Optimal Speed and Accuracy of Object Detection.

Brito Neves, B.B., 1998, O Cambro-Ordoviciano da Província Borborema: *Boletim IG-USP, Série Científica*, v. 29, p. 175-193, doi:10.11606/issn.2316-8986.v29i0p175-193.

Cardoso, A. R., Nogueira, A. C. R., Rabelo, C. E. N., Soares, J. L., Góes, A. M., 2019, Multi-approach provenance in stratigraphy: Implications for the Upper Mesozoic evolution of the Parnaíba Basin, NE Brazil: *Journal of South American Earth Sciences*, v. 96, 102386, doi:10.1016/j.jsames.2019.102386.

Cerri, R. I., Warren, L. V., Varejão, F. G., Marconato, A., Luvizotto, G. L., Assine, M. L., 2020, Unraveling the origin of the Parnaíba Basin: Testing the rift to sag hypothesis using a multi-proxy provenance analysis: *Journal of South American Earth Sciences*, 102625, doi:10.1016/j.jsames.2020.102625.

Corfu, F., Hanchar, J. M., Hoskin, P. W. O., Kinny, P., 2003, *Atlas of Zircon Textures: Zircon*, v. 16, p. 469–502, doi:10.1515/9781501509322-019.

Cortes, C., Vapnik, V., 1995, Support Vector Network: *Machine Learning*, v. 20(3), p. 273–297, doi:10.1023/a:1022627411411.

Deer, W.A., Howie, R.A., Zussman, J., 2013, *Introduction to the rock forming minerals Mineralogical Society*, 3<sup>o</sup> eds, 498p.

Del Rio, I., Sawakuchi, A.O., Góes, A. M., Hollanda, M.H.B.M., Furukawa, L.Y., Porat, N., Jain, M., Mineli, T.D., Negri, F.D.A., 2021, Luminescence signals of quartz and feldspar as new methods for stratigraphic discrimination and provenance analysis of siliciclastic successions: The case of the Parnaíba Basin (Brazil) of West Gondwana: *Basin Research*, p. 1–22, doi:10.1111/bre.12590.

Dewey, J., Mange, M., 1999, Petrography of Ordovician and Silurian sediments in the western Irish Caledonides: tracers of a short-lived Ordovician continent-arc collision orogeny and the evolution of the Laurentian Appalachian-Caledonian margin: *Geological*

Society, London, Special Publications, v. 164(1), p. 55–107, doi:10.1144/gsl.sp.1999.164.01.05.

Dickinson, W.R., Gehrels, G.E., 2009, U–Pb ages of detrital zircons in Jurassic eolian and associated sandstones of the Colorado Plateau: evidence for transcontinental dispersal and intraregional recycling of sediment: *Geological Society of America Bulletin*, v. 121, p. 408–433, doi: 10.1130/B26406.1.

Folk, R.L., 1978, Angularity and silica coatings of Simpson Desert and grains, Northern Territory Australia: *Journal of Sedimentary Petrology*, v. 48, p. 611–624.

Garzanti, E., Vermeesch, P., Andò, S., Lustrino, M., Padoan, M., Vezzoli, G., 2014, Ultra-long distance littoral transport of Orange sand and provenance of the Skeleton Coast Erg (Namibia): *Marine Geology*, v. 357, p. 25–36, doi:10.1016/j.margeo.2014.07.005.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014, Rich feature hierarchies for accurate object detection and semantic segmentation: 2014 IEEE Conference on Computer Vision and Pattern Recognition, doi:10.1109/cvpr.2014.81.

Goodfellow, I., Bengio, Y., Courville, A., 2016, *Deep Learning*: MIT Press, 775p.

Haughton, P.D.W., Todd, S. P., Morton, A.C., 1991, Sedimentary provenance studies: Geological Society, London, Special Publications, v. 57(1), p. 1–11, doi:10.1144/gsl.sp.1991.057.01.01.

Holland, H.J., 1975, *Adaptation in Natural and Artificial Systems*: University of Michigan Press, Ann Arbor, 2° eds.

Hollanda, M.H.B.M., Góes, A.M., Negri, F.A., 2018, Provenance of sandstones in the Parnaíba Basin through detrital zircon geochronology: Geological Society, London, Special Publications, SP472.16, doi:10.1144/sp472.16.

Hollanda, M.H.B.M., Góes, A.M., Silva, D.B., Negri, F.A., 2014, Proveniência sedimentar dos arenitos da Bacia do Parnaíba (NE do Brasil): *Boletim de Geociências da Petrobras*, v. 22, p. 191–211.

Hoskin, P.W.O., Schaltegger, U., 2003, The Composition of Zircon and Igneous and Metamorphic Petrogenesis: *Reviews in Mineralogy and Geochemistry*, v. 53, p. 27–62, doi:10.2113/0530027.

Jocher, G. et al., 2022, ultralytics/yolov5: v7.0 – YOLOv5 SOTA Realtime Instance Segmentation, doi: 10.5281/zenodo.3908559.

Johnsson, M.J., 1993, The system controlling the composition of clastic sediments, In: Johnsson, M.J., Basu, A., Processes controlling the composition of clastic sediments: Special Paper, Geological Society of America, v. 284, p. 1–19.

Kasper-Zubillaga, J.J., Dickinson, W.W., Carranza-Edwards, A., Hornelas-Orozco, Y., 2005, Petrography of quartz grains in beach and dune sands of Northland, North Island, New Zealand: *New Zealand Journal of Geology and Geophysics*, v. 48(4), p. 649–660, doi:10.1080/00288306.2005.9515139.

- Kingma, P.D., Ba, J., 2015, Adam: A Method for Stochastic Optimization: 2015 – 3<sup>o</sup> eds International Conference for Learning Representations, doi:10.48550/arXiv.1412.6980.
- Kohavi, R., Provost, F., 1998, Glossary of terms. Machine Learning: Special Issue on Applications of Machine Learning and the Knowledge Discovery Process, Machine Learning, v. 30, p. 271-274, doi: 10.1023/A:1017181826899.
- Krumbein, W.C., 1941, Measurement and geological significance of shape and roundness of sedimentary particles: Journal of Sedimentary Petrology, v. 11, p.64-72.
- LeCun, Y., Bengio, Y., Hinton, G., 2015, Deep learning: Nature, v. 521(7553), p. 436–444, doi:10.1038/nature14539.
- Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998, Gradient-based learning applied to document recognition: Proceedings of the IEEE, v. 86(11), p. 2278–2324, doi:10.1109/5.726791.
- Long, X., Deng, K., Wang, G., Zhang, Y., Dang, Q., Gao, Y., Shen, H., Ren, J., Han, S., Ding, E., Wen, S., 2020, PP-YOLO: An Effective and Efficient Implementation of Object Detector.
- Mange, M.A., Otvos, E.G., 2005, Gulf coastal plain evolution in West Louisiana: Heavy mineral provenance and Pleistocene alluvial chronology: Sedimentary Geology, v. 182 (1), p. 29-57, doi: 10.1016/j.sedgeo.2005.07.015.
- Mange-Rajetzky, M.A., 1995, Subdivision and correlation of monotonous sandstone sequences using high-resolution heavy mineral analysis, a case study: the Triassic of the Central Graben: Geological Society, London, Special Publications, v. 89(1), p. 23–30, doi:10.1144/gsl.sp.1995.089.01.03.
- Medina, M, Fertig, C., 2005, Algoritmos e Programação - Teoria e Prática, 2<sup>a</sup> eds. Editora Novatec, 384 p.
- Meng Z., Hu Y., Ancey C., 2020, Using a Data Driven Approach to Predict Waves Generated by Gravity Driven Mass Flows. Water, v. 12 (2), 600, doi: 10.3390/w12020600.
- Menzies, L.A., Carter, A., MacDonald, D.I.M., 2018, Evolution of a cratonic basin: insights from the stratal architecture and provenance history of the Parnaíba Basin: Geological Society, London, Special Publications, SP472.18, doi:10.1144/sp472.18.
- Morton, A.C., 2012. Value of heavy minerals in sediments and sedimentary rocks for provenance, transport history and stratigraphic correlation, In: Sylvester, P. (Ed.), Quantitative mineralogy and microanalysis of sediments and sedimentary rocks: Mineralogical Association of Canada Short Course Series, v. 42, p. 133–165.
- Morton, A.C., Hallsworth, C.R., 1994, Identifying provenance-specific features of detrital heavy mineral assemblages in sandstones: Sedimentary Geology, v. 90, p. 241-256.
- Morton, A.C., Hallsworth, C.R., 1999, Processes controlling the composition of heavy mineral assemblages in sandstones: Sedimentary Geology, v. 124, p. 3-29.

- Nair, V., Hinton, G., 2010, Rectified Linear Units Improve Restricted Boltzmann Machines Vinod Nair: Proceedings of ICML, v. 27, p. 807-814.
- Oliveira, C.V., Moura, C.V., 2019, Provenance of detrital zircons of the canindé group (Parnaíba basin), northeastern Brazil: Journal of South American Earth Sciences, v. 90, p. 162-180, doi:10.1016/j.jsames.2018.12.009.
- Oliveira, R. S. de, Truckenbrodt, W., 2019, Provenance and diagenesis of Guamá Sandstone, northeastern Pará, Brazil: A Silurian link between the Amazonas and Parnaíba basins: Journal of South American Earth Sciences, v. 90, p. 265–281. doi:10.1016/j.jsames.2018.12.014.
- Padilla, R., Netto, S., Silva, E., 2020, A Survey on Performance Metrics for Object-Detection Algorithms, doi: 10.1109/IWSSIP48289.2020.
- Pan, S.J., Yang, Q., 2010, A Survey on Transfer Learning: IEEE Transactions on Knowledge and Data Engineering, v. 22 (10), p. 1345-1359, doi: 10.1109/TKDE.2009.191.
- Pettijohn, F.J., 1949, Sedimentary Rocks: Harper & Bros, New York, 425p.
- Pettijohn, F.J., Potter, P.E., Siever, R., 1972, Sand and Sandstone: Springer, New York, N.Y., 618 p.
- Powers, M.C, 1953, A new roundness scale for sedimentary particles: Journal of Sedimentary Petrology, v. 23, p. 117–119.
- Pryor, W.A., 1971, Grain shape, In: Carver, R. E., Procedures in Sedimentary Petrology, Wiley, p. 131-150.
- Pupin, J.P., 1980, Zircon and granite petrology: Contributions to Mineralogy and Petrology, v. 73, p. 207–220, doi:10.1007/BF00381441.
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016, You Only Look Once: Unified, Real-Time Object Detection: 2016 IEEE Conference on Computer Vision and Pattern Recognition - CVPR, doi:10.1109/cvpr.2016.91.
- Ren, S., He, K., Girshick, R., Sun, J., 2016, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks: IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 39(6), p. 1137–1149, doi:10.1109/tpami.2016.2577031.
- Rumelhart, D. E., Hinton, G. E., Williams, R. J., 1986, Learning representations by back-propagating errors. Nature, v. 323(6088), p. 533–536, doi:10.1038/323533a0.
- Russell, R.D., Taylor, R.E., 1937, Roundness and shape of Mississippi River sands: Jour. Geology, v. 45, p. 225-267.
- Sagga, A., 1993, Roundness of Sand Grains of Longitudinal Dunes in Saudi Arabia: Sedimentary Geology, v. 87, p. 63-68, doi:10.1016/0037-0738(93)90036-5.
- Tucker, M.E., 2004, Sedimentary Rocks in the Field (Geological Field Guide), Wiley; 3<sup>o</sup> eds, 244p.

Wadell, H., 1932, Volume, shape and roundness of rock particles: *Brazilian Journal of Geology*, v. 40, p. 443-451.

Weltje, G.J., Eynatten, H.V., 2004, Quantitative Provenance Analysis of Sediments: Review and Outlook, *Sedimentary Geology*, v.171, p.1-11.

Wu, J., 2017, Introduction to Convolutional Neural Networks: LAMDA Group National Key Lab for Novel Software Technology Nanjing University, China, 31p.

Zhu, H.-T., Zhang, H., 2004, Hypothesis testing in mixture regression models: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 66(1), p. 3–16, doi:10.1046/j.1369-7412.2003.05379.x.