

Automação do Teste de Compressão de Pelotas de Minério de Ferro Através de Visão Computacional

Janderson Antonio F. dos Santos, Vinícius da R. Motta, Carlos T. Valadão, Marco Antonio de S. L. Cuadros

Resumo Este artigo apresenta a aplicação do algoritmo Mask-RCNN para detectar pelotas de minério, enquanto estão em movimento em uma máquina para teste de compressão. Uma vez detectadas as pelotas dentro da imagem, seus diâmetros ajudarão a verificar se atendem aos requisitos da empresa. O sistema usa uma câmera da Logitech – Brio Ultra HD Pro Webcam – com framerate de 5 fps em ultra HD e as imagens passam por um pré-processamento que contempla a conversão para escala de cinza e, finalmente, a aplicação do algoritmo Mask R-CNN. O modelo foi treinado com 5 imagens, e isto foi o suficiente, pois foi usada a transferência de aprendizado de uma aplicação similar e o resultado atingiu uma acurácia superior a 96% nos testes.

Abstract This work shows the application of the Mask R-CNN algorithm to detect ore pellets while they are in movement in a disc pelletizer. Once the pellets are detected in the image, their diameters were verified to ensure they comply with the company standards. The system uses a Logitech –Brio Ultra HD Pro Webcam- camera with a framerate of 5 fps in ultra HD and the images acquired pass through a preprocessing that comprises converting into grayscale image and, finally, applying the Mask R-CNN algorithm. The model was trained with 5 images, it was sufficient because has been used transfer learning of the similar application and achieved an accuracy rate over 96% accuracy rate in the tests.

1 Introdução

O desafio mais básico da visão computacional é a detecção de objetos, sendo que esta tecnologia está sendo amplamente usada em diversas aplicações (LIU et al., 2020) Devido a isso,

Janderson Antonio F. dos Santos
Grupo da Automação Industrial, Instituto Federal do Espírito Santo (IFES), Serra, Espírito Santo, Brasil.
e-mail: janderson.eng.mec@gmail.com

Vinícius da R. Motta
Grupo da Automação Industrial, Instituto Federal do Espírito Santo (IFES), Serra, Espírito Santo, Brasil.

Carlos T. Valadão
Grupo da Automação Industrial, Instituto Federal do Espírito Santo (IFES), Serra, Espírito Santo, Brasil.
e-mail: carlostvaladao@gmail.com

Marco Antonio de S. L. Cuadros
Grupo da Automação Industrial, Instituto Federal do Espírito Santo (IFES), Serra, Espírito Santo, Brasil.
e-mail: marcoantonio@ifes.edu.br

Anais do XV Encontro Anual de Ciência da Computação (EnAComp 2020). ISSN: 2178-6992.

Catalão, Goiás, Brasil. 25 a 27 de Novembro de 2020.

Copyright © autores. Publicado pela Universidade Federal de Catalão.

Este é um artigo de acesso aberto sob a licença CC BY-NC (<http://creativecommons.org/licenses/by-nc/4.0/>).

surge grande interesse em desenvolver formas de localizar instâncias de objetos dentre muitas categorias predefinidas em imagens diversas (LIU et al., 2020). Neste artigo, é apresentado um sistema de reconhecimento, dimensionamento e localização de pelotas usando *machine learning* para encontrar o diâmetro de uma pelota e complementar a análise de qualidade feita por uma máquina de testes de compressão de pelotas de minério verde.

O teste de compressão é realizado para testar as propriedades de resistência mecânica das pelotas verdes e queimadas, quantificando a carga de ruptura necessária para sua fratura. Esse teste é realizado inicialmente nas pelotas verdes, logo após sua confecção no disco de pelotamento e passagem por esteiras separadoras dimensionais. O técnico realiza o teste através de uma amostragem dessas pelotas, nas quais se verifica as características dimensionais, que de acordo com o trabalho de Meyer (1980) deverá ser entre 8 e 16 mm de espessura. Além disso, é feita a verificação do limite mínimo de compressão que deve ser de 1 a 2kgf para garantir que essas recentes pelotas aguentem o processo de transporte por esteiras transportadoras até os fornos de queima. Atingindo as especificações necessárias, as pelotas seguem o processo industrial, caso contrário retornam para o beneficiamento. Finalizado o processo de fabricação, as pelotas verdes passam por um processo de cura chamado “queima”, realizado em fornos de atmosfera e temperatura controlada que garantirá que as pelotas atinjam as propriedades químicas e mecânicas necessárias para a maior produção do produto final, o aço, sendo a resistência mínima de compressão de 210kgf para aguentar as cargas nos altos-fornos (ZHU et al., 2015).

Automatizar esse tipo de tarefa permite a empresa maior confiabilidade nas análises, hoje realizada por um técnico que coloca uma pelota por vez em uma máquina de ensaio de compressão universal, a qual, após o final do teste manual, gera um relatório das propriedades mecânicas. Este teste é monótono para seu realizador, portanto, acaba sendo mais demorado e propício a falhas humanas, o que traz desvantagens em termos financeiros para a empresa. No momento, não existe um fornecedor de máquinas de ensaios de compressão para pelotas de minério verde no mercado que atenda de forma personalizada as necessidades da área.

Com tal situação em vista, está sendo desenvolvido um equipamento de teste que contém um disco giratório em posição horizontal, perpendicular a uma torre de compressão. Desta maneira, esta realizará a pressão em cada pelota por vez e, assim ao terminar, o disco realizará um leve giro posicionando uma nova amostra de material. Sobre um disco são dispostas as pelotas que passarão pelo teste de compressão a ser realizado na torre de compressão. Após analisadas, as amostras são deslocadas para zona de descarte. Finalizando o teste, é gerado um relatório com os dados pertinentes de resistência mecânica do lote, além de registros gráficos visuais vantajosos para aprimoramentos do processo. O equipamento utilizado pode ser visualizado na Figura 1.

Por meio do equipamento mostrado na Figura 1, este estudo propõe completar a sua operação utilizando visão computacional para localização das pelotas, dimensionamento e o seu deslocamento até a zona de compressão. Neste trabalho, são apresentados resultados do algoritmo de localização e obtenção da dimensão das pelotas como resultados preliminares do desenvolvimento do protótipo.

Existem estudos correlacionados, os quais apresentam um método experimental para medir as propriedades do volume e as relações de carregamento de fratura para pelotas de minério de ferro, como pode ser visto em Gustafsson et al. (2013). Já no artigo de Castellanos et al.

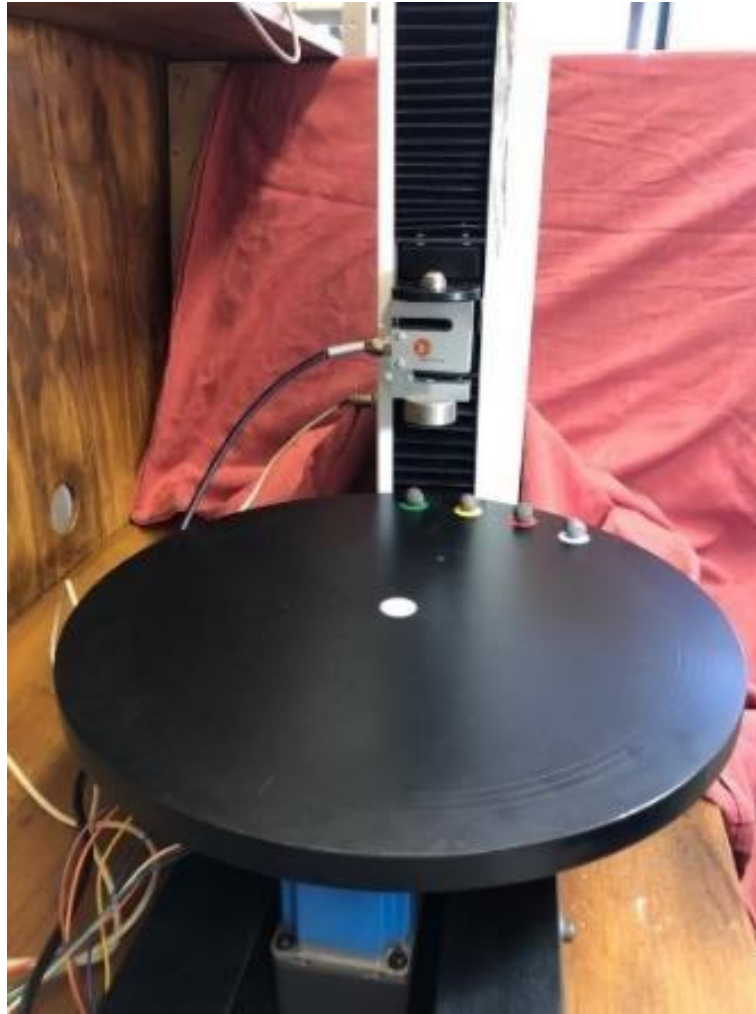


Figura 1: Protótipo em desenvolvimento.

(2018), é mencionada uma abordagem correlativa empregando microscopia óptica e microscopia eletrônica de varredura para a quantificação de fases e poros em pelotas de minério de ferro.

Este artigo está organizado da seguinte forma: esta primeira seção é a introdução, onde o projeto é explicado de forma geral e a organização do texto é mostrada; na segunda seção é apresentada a metodologia utilizada para desenvolver o sistema, contendo os conceitos e materiais usados, bem como uma breve explicação da rede utilizada. A seguir, na terceira seção é explicada a detecção de pelotas de minério utilizando a Mask-RCNN (*Mask Recurrent Convolutional Neural Network*). Por fim, a quarta seção apresenta os resultados obtidos dos experimentos e as conclusões com trabalhos futuros.

2 Metodologia

2.1 Detecção de pelotas de minério de ferro

Como citado anteriormente, para o problema em questão é necessária a detecção e contagem de pelotas de minério de ferro. Tais itens são necessários para boa performance da aplicação. Com isso, surgiu a necessidade de uso de uma tecnologia que faça a localização visual e a contagem dos objetos alvo. Especificamente neste trabalho, foi utilizada a rede neural Mask R-CNN, explicada na seção 2.2.

2.2 Mask R-CNN

A rede neural convolucional Mask R-CNN foi desenvolvida em 2017 pela FAIR (*Facebook AI Research*) e é uma extensão da Faster R-CNN (*Faster Recurrent Neural Network*), sendo que na primeira foi incluído um ramo para prever e desenhar a máscara de objeto em paralelo com ramo já existente para o reconhecimento e desenho da caixa delimitadora do objeto na imagem (HE et al., 2017). O algoritmo tem um alinhamento pixel-a-pixel, que é uma parte que não existe no algoritmo Fast/Faster R-CNN (*Fast/Faster Recurrent Neural Network*) (HE et al., 2017).

A Mask-RCNN utiliza uma rede que gera propostas de região na imagem. Ela segue o modelo Faster-RCNN, que possui um extrator de características seguido de uma rede que gera as propostas de região. Além disso, a Faster-RCNN adota uma operação conhecida chamada ROI-Pooling (*Region of Interest Pooling*), a qual é responsável por produzir saídas de dimensões corretas para a entrada do classificador.

A *framework* da Mask R-CNN possui três mudanças importantes a respeito da Faster R-CNN. Primeiramente, há uma substituição da ROI-Pooling usada na Faster-RCNN por uma operação chamada ROI-Align (*Region of Interest Align*), a qual permite a construção de máscaras de segmentação de instâncias bem precisas. Além disso, adiciona-se à entrada uma pequena rede neural convolucional a fim de produzir as segmentações de instância desejadas, conforme mostra a Figura 2 (JOHNSON, 2018).

Por fim, as previsões de máscara e classe são separadas e realizadas de forma paralela, em ramos diferentes da rede (HE et al., 2017).

2.3 Treinamento e aplicação da Mask R-CNN

A fim de construir o sistema para identificar as pelotas de minério na imagem, primeiramente foi realizada a aquisição de imagens das pelotas posicionadas sobre o disco através de uma *webcam* modelo **Logitech – Brio Ultra HD Pro Webcam**. Foram feitos vídeos capturando um frame a cada 20 segundos, os quais foram utilizados para posterior treinamento da rede neural. Esses frames foram utilizados para construir os subconjuntos do *dataset*, composto por imagens de treino e de validação em uma proporção baseada no método *holdout* (70%-30%). Este procedimento consiste em dividir o conjunto de dados em dois subconjuntos mutuamente exclusivos

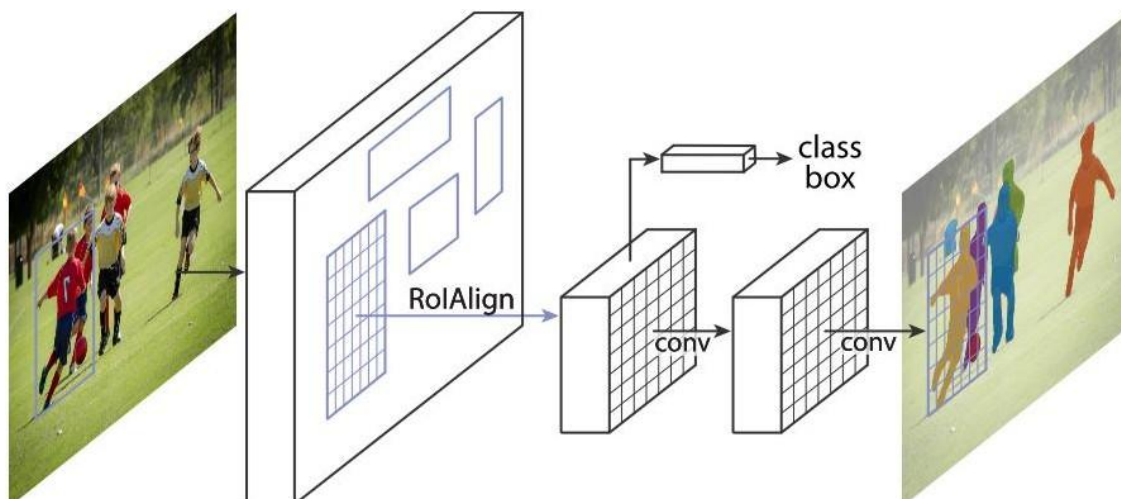


Figura 2: O modelo Mask-RCNN (JOHNSON, 2018)

(Pal & Patel, 2020). Ao finalizar a construção do *dataset* este tinha 184 imagens para treino e 55 para validação.

O segundo passo para a construção do sistema consistiu em trabalhar com a *framework* Mask R-CNN. Essa rede foi inicializada com os pesos relativos a um modelo de detecção de balões. Desta forma, tais pesos foram reutilizados, sendo necessário apenas o retreino da última camada com as imagens das pelotas de minério. Tal processo economiza tempo e poder computacional, visto que foi feito o congelamento das camadas e um ajuste de pesos apenas na última camada (que estava descongelada) para detectar as pelotas de minério, ao invés dos balões. Essa técnica é denominada transferência de aprendizado e é muito efetiva, especialmente em modelos similares, como é o caso de pelotas e balões, além de conseguir economizar tempo e poder computacional, e melhorar as métricas de avaliação da rede.

A redução do tempo necessário para o treino se dá devido ao menor número de pesos que necessitam ser ajustados, podendo-se assim, treinar a rede de forma mais rápida. Isso foi verificado nos experimentos práticos realizados para esse trabalho, sendo que a utilização dessa técnica reduziu o tempo da etapa de treinamento.

A última camada foi, inicialmente, retreinada com apenas uma imagem, mas durante a validação resultou em métricas de avaliação insuficientes. Em seguida, foram utilizadas duas imagens e a rede foi novamente validada. Esse processo foi repetido de forma incremental e quando cinco ou mais imagens para treino eram utilizadas, o resultado das métricas de validação se mantinha constante. Com isso, chegou-se ao resultado lógico de que cinco imagens eram necessárias para realização deste trabalho, visto que, após o treino com essas imagens, a rede já conseguia detectar as pelotas de forma satisfatória. Isso foi possível devido à transferência de aprendizado e à similaridade dos modelos.

Para realizar o treino e validação da rede era necessário gerar, a partir das imagens, um arquivo json (*JavaScript Object Notation*) que continha as localizações das pelotas. Cinco imagens foram utilizadas para gerar o json de treinamento e uma imagem para gerar o de validação. O arquivo com a localização foi construído com o programa VIA. Este software constrói um arquivo json que terá cada posicionamento das respectivas pelotas. Para construção deste arquivo, é

necessário identificar manualmente os objetos que deverão ser reconhecidos pela aplicação (no caso, as pelotas), conforme mostra a Figura 3.

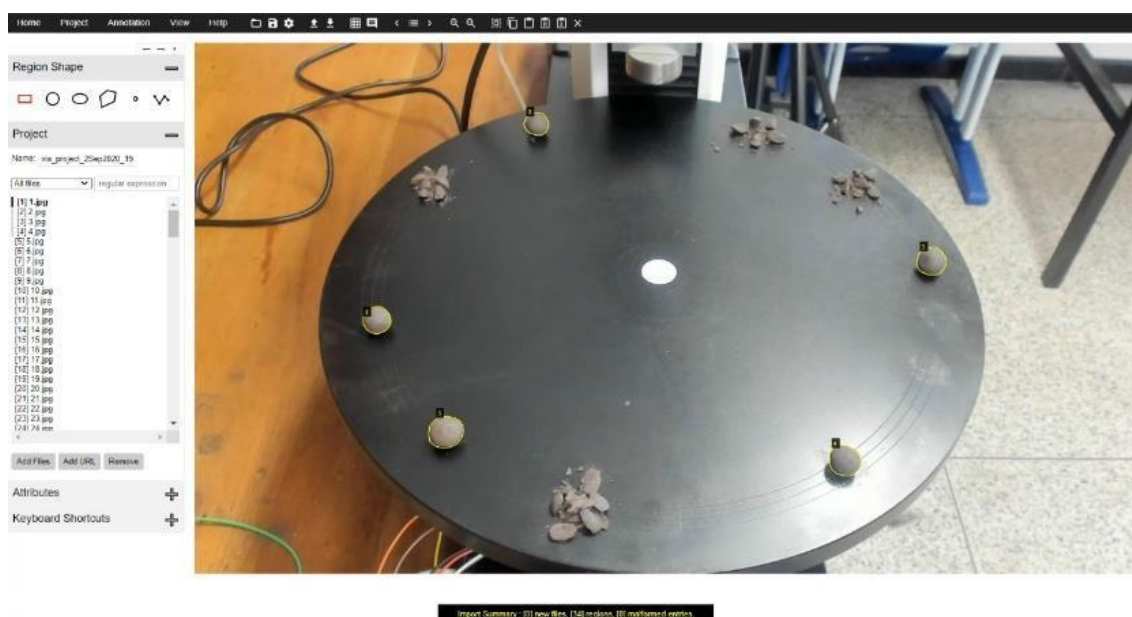


Figura 3: Software VIA utilizado para criação do arquivo de treinamento.

Feito esse processo de treinamento, o modelo resultante do processo foi salvo. Em seguida, foi desenvolvido um programa para avaliar as imagens e, por meio da aplicação da rede neural treinada, detectar as pelotas. Para a validação, foi utilizado um vídeo gravado das pelotas sobre o disco pelotizador.

A Figura 4 mostra um fluxograma com os passos realizados para o treinamento da rede neural e sua validação.

3 Resultados

Para verificação dos resultados, um vídeo contendo a mesma quantidade de pelotas foi analisado com o algoritmo de detecção de objetos treinado. A Figura 5 mostra o teste realizado sobre o vídeo.

Uma amostra de 1000 frames foi submetida ao algoritmo de detecção e os resultados observados e apresentados na Tabela 1. O item BG (*Background*) se refere a não-detecção da pelota, enquanto o item pelota, se refere a detecção da pelota.

	BG	Pelota
BG	1000	100
Pelota	144	5000

Tabela 1: Matriz de confusão do algoritmo de detecção.

Após a execução do teste obteve-se as seguintes informações:

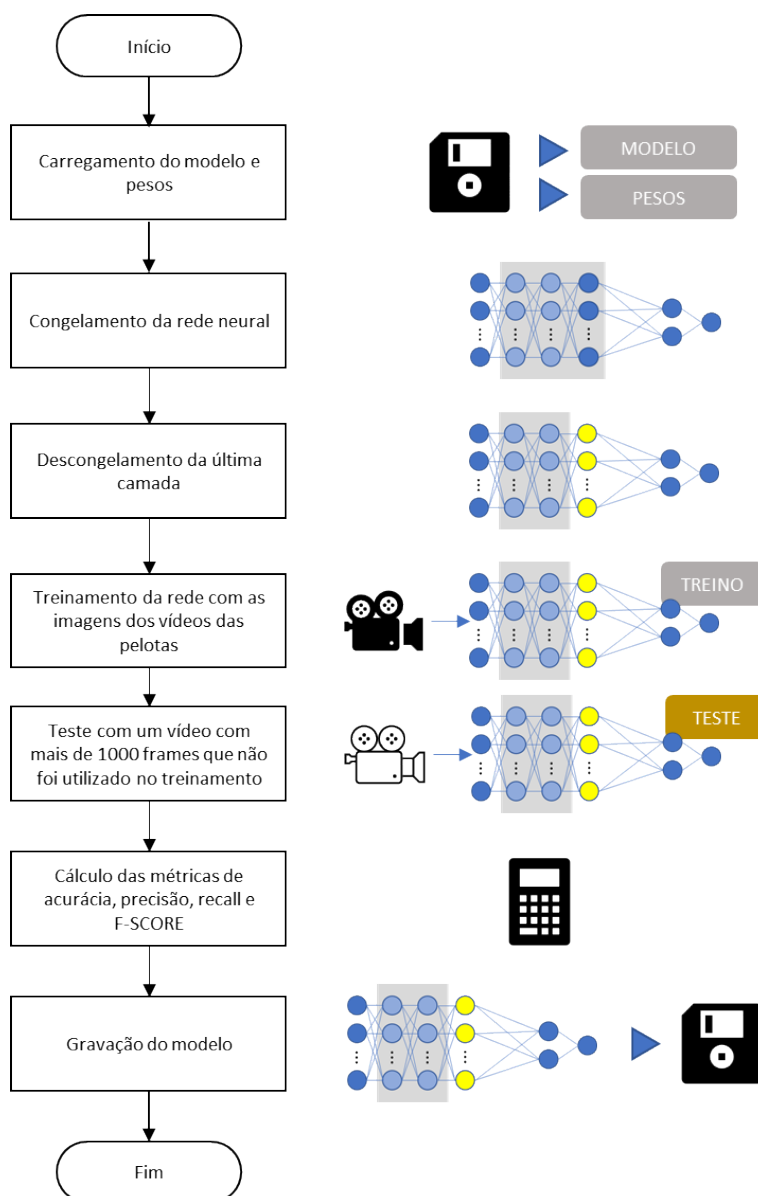


Figura 4: Fluxograma do treinamento e teste da rede neural utilizada.

- Tempo Total de execução: 937.251 segundos;
- Tempo médio de Detecção de objetos: 0.937 segundos;
- Número Total de pelotas detectadas: 5100;
- Número Total de BG detectados: 1144;
- Número Total de pelotas detectadas como BG: 100;
- Número Total de BG detectados como Pelota: 144;
- Média dos Scores do detector das pelotas: 0.9941 ou 99.41%.

Os resultados da Matriz de confusão indicam que, na coluna BG, foram classificados corretamente 1000 objetos da classe BG e foram identificados incorretamente 144 objetos Pelota, aos quais pertenciam a classe BG (falso negativo). Já na coluna Pelota, a primeira linha indica o número de objetos que foram identificados como Pelota que somam 100 objetos, e que na verdade eram objetos BG (falso positivo). Na segunda linha, tem-se o número de pelotas

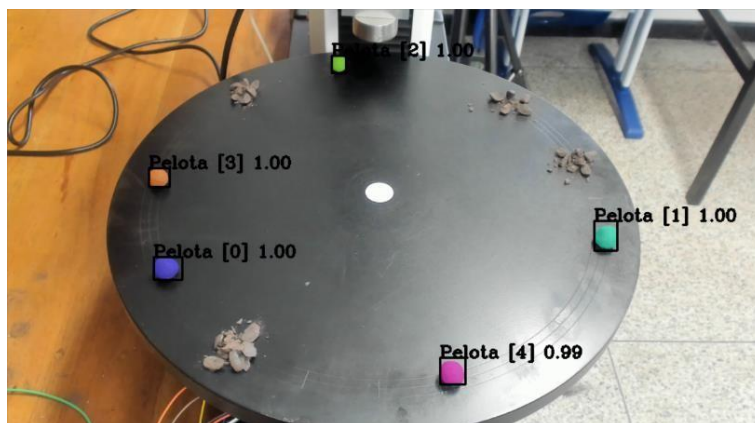


Figura 5: O vídeo sendo analisado pelo algoritmo de detecção de objetos.

corretamente identificadas. Com estes dados, é possível conseguir alguns índices de performance da Mask-RCNN (HASHMI; TABASSUM, 2020). Duas das métricas mais fundamentais de classificação são a acurácia e a precisão. Para realizar o cálculo referente a acurácia basta realizar a seguinte conta: proporção de classificações corretas pelo total de classificações naquele domínio, com mostrado na Equação 1.

$$acc = \frac{VP + VN}{VP + VN + FP + FN}, \quad (1)$$

onde acc é o valor da acurácia, VP é o total de verdadeiros positivos, VN , o total de verdadeiros negativos, FP , o total de falsos positivos e FN , o total de falsos negativos. Com isso, o resultado da Acurácia é de 6000 dividido por 6244, o que resulta em um índice de 0,9609, ou seja 96,09%. Já a precisão, cuja fórmula é mostrada na Equação 2, calcula quantos dos resultados positivos são realmente positivos.

$$prec = \frac{VP}{VP + FP}, \quad (2)$$

onde $prec$ é a precisão, VP , o total de verdadeiros positivos e FP , o total de falsos positivos. Assim, o seu cálculo é o total da classificação correta positiva sobre a classificação total positiva. O cálculo desse índice é de 5000 sobre 5100 que resulta em 0,9804 ou seja 98,04%.

Sabendo-se que acurácia é o quanto foi acertado nas previsões, e que precisão é o quão os resultados não constituíram falsos positivos, abordamos o *recall*, ou sensibilidade, mostrado na Equação 3.

$$recall = \frac{VP}{VP + FN}. \quad (3)$$

Este índice mostra a medida do quanto todos os resultados positivos em teste foram encontrados pelo modelo. O seu cálculo é realizado pelo número de verdadeiros positivos, 5000, sobre o total de itens detectados como pelotas, totalizando 5144, o que dá um resultado de 0,972 ou seja 97,2%.

Finalmente, pode-se calcular o F-score, mostrado na Equação 4.

$$F - score = 2 \frac{recall}{recall} \quad (4)$$

Onde, F-score é o valor da métrica, *prec* é a precisão e *recall* é a sensibilidade, como dito anteriormente. Essa métrica reúne em um só valor os valores de precisão e *recall*. É útil porque duas métricas têm idêntica relevância, então é simplesmente a média harmônica das duas métricas, ou seja, é a precisão vezes o *recall* sobre precisão mais *recall* vezes dois que resulta em 0,9761 ou 97,61%.

4 Conclusões

O uso do algoritmo Mask-RCNN para detectar pelotas de minério foi eficaz em seu propósito. A *framework* facilitou o desenvolvimento da aplicação devido a reutilização de código. Além disso, a utilização da técnica de transferência de aprendizado otimizou a etapa de treinamento, fazendo com que fossem usadas poucas imagens nesta etapa. Portanto, o resultado foi bastante significativo e se o algoritmo for treinado com mais imagens tende a melhorar os índices das métricas de avaliação. Para trabalhos futuros, será implementado um algoritmo de *tracking* simples para os objetos detectados conforme proposto por Bewley et al. (2016).

Referências

- BEWLEY, A. et al. Simple online and realtime tracking. In: PROCEEDINGS of 2016 IEEE International Conference on Image Processing (ICIP). [S.l.: s.n.], 2016. p. 3464–3468.
- CASTELLANOS, R. M. et al. Characterization of iron ore pellets by multimodal microscopy and image analysis. **International Engineering Journal**, v. 71, n. 2, p. 209–215, 2018.
- GUSTAFSSON, G. et al. Determination of bulk properties and fracture data for iron ore pellets using instrumented confined compression experiments. **Powder Technology**, v. 241, p. 19–27, 2013.
- HASHMI, A.; TABASSUM, S. Analysis of Metrics of Occupational Stress Classification by CNN with Softmax Classifier. **International Journal of Applied Exercise Physiology**, v. 9, n. 4, p. 229–239, 2020.
- HE, K. et al. Mask R-CNN. In: PROCEEDINGS of 2017 IEEE International Conference on Computer Vision (ICCV). [S.l.: s.n.], 2017. p. 2980–2988.
- JOHNSON, J. Adapting Mask-RCNN for Automatic Nucleus Segmentation. **ArXiv**, abs/1805.00500, 2018.
- LIU, L. et al. Deep learning for generic object detection: a survey. **International journal of computer vision**, v. 128, n. 2, p. 261–318, 2020.
- MEYER, K. **Pelletizing of Iron Ore**. Berlin: Springer-Verlag, 1980.
- ZHU, D. et al. 15 - Iron ore pelletization. In: LU, Liming (Ed.). **Iron Ore**. [S.l.]: Woodhead Publishing, 2015. p. 435–473.