

Análise do teor de óxido de nióbio (Nb_2O_5) na etapa de flotação utilizando imagens de textura

Fábio Lucas de Almeida, Douglas Farias Cordeiro, Núbia Rosa da Silva

Resumo O emprego de técnicas da inteligência artificial está cada vez mais popular em ramos industriais para controle de processos. Este trabalho investiga, através de aplicações que usam aprendizado de máquina, a predição do teor químico Nb_2O_5 nas espumas geradas a partir da etapa de flotação no beneficiamento mineral do nióbio em uma empresa de mineração. A partir de imagens digitais das espumas de flotação de nióbio é proposto um método para agrupar as mesmas em grupos baseados no tamanho das bolhas presentes na espuma. É demonstrado forte relação dos grupos com o teor químico que implica na investigação de um modelo de inteligência artificial com o objetivo de classificar as imagens. Por fim, é demonstrado acurácia entre 70% e 85% pelos modelos propostos e os resultados são brevemente discutidos.

1 Introdução

A mineração é uma atividade econômica e industrial que constitui-se da extração e do beneficiamento de substâncias minerais do subsolo (FARIAS, 2002). Para concentrar o teor mineral, diversas etapas são realizadas, dentre elas a flotação. A flotação geralmente é a última etapa de concentração, nela é realizada uma separação feita numa suspensão em água (polpa) onde as partículas que se deseja flotar são capturadas por bolhas de ar e levadas a superfície formando uma espuma (LUZ; SAMPAIO; ALMEIDA, 2010). Essa espuma apresenta características visuais semelhantes que podem servir de indicativo para controle do processo. O material final gerado nessa etapa deve obedecer um certo limite mínimo de qualidade, de tal forma que se a qualidade estiver acima do limite há potencial de maior produção reduzindo a qualidade até o limite estipulado. Assim como, se a qualidade estiver abaixo do limite é necessário reduzir a produção para entregar um material final dentro da especificação.

Conforme abordado por Fu e Aldrich (2019) a imagem de flotação capturada em diferentes cenários operacionais geram distintos tamanhos de bolhas que podem indicar um valor aproxi-

Fábio Lucas de Almeida
Instituto de Biotecnologia, Universidade Federal de Catalão, Catalão, Goiás, Brasil.
e-mail: fabiolucas@discente.ufg.br

Douglas Farias Cordeiro
Faculdade de Informação e Comunicação, Universidade Federal de Goiás, Goiânia, Goiás, Brasil.
e-mail: cordeiro@ufg.br

Núbia Rosa da Silva
Fábio Lucas de Almeida
Instituto de Biotecnologia, Universidade Federal de Catalão, Catalão, Goiás, Brasil.
e-mail: nubia@ufg.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/>).

mado de concentração do teor analisado. Dessa forma, uma tarefa de aprendizado de máquina que possa classificar a imagem de flotação de acordo com o tamanho de bolha também indicaria um valor aproximado de concentrado, estes dados serviriam como um norteador para a equipe de produção se guiarem nos ajustes de processo.

O uso de recursos computacionais e automação nas indústrias trazem diversos benefícios como monitoramento dos indicadores chaves de processos, antecipação de correções em máquinas e processos, dentre outros (VIEIRA, 2006). Ao monitorar materiais que apresentem características físicas semelhantes e padrões visuais relacionados pode-se aplicar técnicas de visão computacional para identificar comportamentos diferentes, que só seriam percebidos por um técnico especialista, e classificá-los em variados cenários ou classes, ganhando, com isso, agilidade na análise pois máquinas não se cansam (ALDRICH et al., 2010; CAO et al., 2013; ESTRADA-RUIZ; PÉREZ-GARIBAY, 2009; FU, Yihao; ALDRICH, Chris, 2018; FU; ALDRICH, 2019; JAHEDSARAVANI; MASSINAEI; MARHABAN, 2017; KISTNER; JEMWA; ALDRICH, 2013).

O estudo publicado por Weixing e Liangqin (2016) apresenta bons resultados para extrair características das imagens de flotação utilizando descritores de textura a partir de uma matriz em nível de cinza (GLCM) das próprias imagens. Com esses descritores é possível aplicar técnicas de agrupamento como o k-means para particionar os dados em diferentes categorias com características próximas e modelar a base de dados. A partir de um modelo de dados supervisionado é possível aplicar técnicas de aprendizado de máquina supervisionado.

O trabalho de Fu e Aldrich (2019) investiga o uso de algoritmos que englobam o aprendizado de máquina e a visão computacional chamados de Rede Neural Convolutiva (CNN). Uma Rede Neural Convolutiva é um algoritmo preparado para captar uma imagem de entrada, atribuir importância e pesos a vários aspectos da imagem e ser capaz de diferenciar um do outro. Na oportunidade os autores apresentaram um estudo de caso realizado na etapa de flotação de uma indústria de platina. As imagens foram coletadas em quatro regimes operacionais em que a taxa de fluxo de ar teve variação para caracterizar diferentes tamanhos de bolhas da espuma de flotação e gerar diferentes faixas de valores de teor concentrado de platina. Essas imagens serviram de treinamento para um modelo de CNN que obteve êxito em classificar as imagens nas categorias corretas. A correta classificação pelas características observadas na espuma de flotação puderam justificar o método como um bom preditor da taxa de concentração do teor de platina.

Para este caso de estudo, imagens foram obtidas de uma célula de flotação de nióbio na empresa CMOC Brasil, localizada em Catalão/GO. O objetivo deste trabalho é desenvolver uma metodologia utilizando essas imagens de flotação de nióbio e indicar, aproximadamente, a concentração do teor de óxido de nióbio (Nb_2O_5) por meio do uso de técnicas de visão computacional e aprendizado de máquina para ajudar nas tomadas de decisão da equipe de produção.

Este trabalho está organizado no processo de amostragem dos rótulos e aquisição das imagens, nesta etapa uma abordagem utilizando segmentação por tamanho de bolha é aplicado a fim de modelar a base de dados para treinamento supervisionado. O algoritmo genérico de rede neural convolutiva é brevemente explicado bem como a arquitetura ResNet50 empregada no experimento. O descritor de textura *Local Binary Patterns* também é explorado para fins de comparação de resultados. Por fim, os experimentos são realizados sobre a base de dados coletada e os resultados são brevemente discutidos.

2 Aquisição e modelagem da base de dados

Para o processo de aquisição dos dados uma câmera digital instalada no último estágio da flotação de nióbio foi usada para capturar imagens da espuma com um intervalo de trinta segundos. As imagens foram geradas em formato JPG com resolução 256x256 pixels, conforme ilustrado na Figura 7.

Para aquisição dos teores os resultados foram gerados através do método de amostragem da própria empresa, gerando um resultado a cada duas horas que representa todo o material passado nesse intervalo. Para que o resultado represente toda a flotação e não apenas uma parte dela o ponto amostrado foi na descarga do produto flotado, conforme a Figura 1 representa.

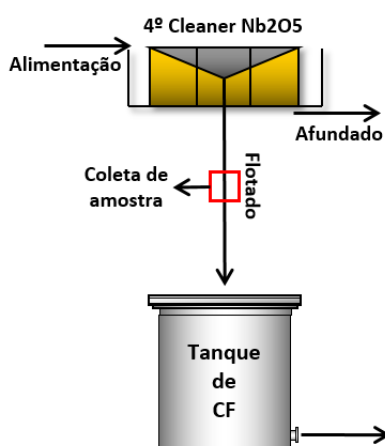


Figura 1: Ponto de amostragem.

3 Arquiteturas de Rede Neural Convolutacional

A rede neural convolutacional (CNN) é um tipo de arquitetura profunda que pertence a classe das redes neurais profundas e vem demonstrando desempenho satisfatório no processamento de dados bidimensionais com topologia em grade, como imagens e vídeos. Esse tipo de rede tem inspiração biológica no córtex visual animal que é formado por um conjunto de neurônios. Ao ver algo, esses neurônios são ativados hierarquicamente, e cada um detecta um conjunto de atributos sobre o que foi visto (LIU et al., 2017). Dessa forma, eles atuam como filtros sobre o espaço de entrada e são adequadas para explorar a forte correlação espacial presente nas imagens naturais.

Existem diferentes variações de arquiteturas de CNN que podem ser encontradas, porém, em uma arquitetura básica de uma CNN estão presentes três tipos de camadas consideradas fundamentais que são as camadas de convolução, agrupamento e totalmente conectadas, podendo ser encontradas uma ou mais destas camadas dependendo da arquitetura abordada. Assim como mostrado na Figura 2, uma imagem é inserida diretamente na rede seguido por vários estágios de convolução e agrupamento. Posteriormente, representações dessas operações alimentam

uma ou mais camadas totalmente conectadas alcançando a última camada que gera o rótulo da classe.

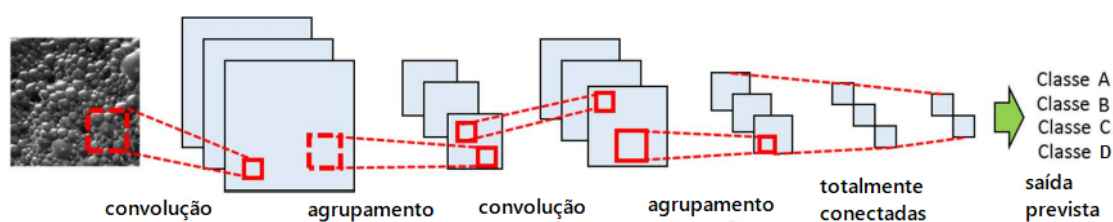


Figura 2: Arquitetura padrão de uma CNN.

As camadas convolucionais servem como extratores de características e, assim, aprendem as representações de características das imagens de entrada (RAWAT; WANG, 2017). Os neurônios nas camadas convolucionais são organizados em mapas de características, ilustrado na Figura 3. Cada neurônio em um mapa de característica (em verde) tem um campo receptivo que é conectado a uma vizinhança de neurônios na camada anterior (em vermelho) por meio de um *kernel* (em azul), que é um conjunto de pesos treináveis. Nas entradas são realizadas operações de convolução sobre o *kernel* para calcular um novo mapa de características, e nos resultados dessa convolução são aplicados uma função de ativação não linear. É importante observar que para gerar cada mapa de recursos o *kernel* é compartilhado por todos os locais da entrada. Todos os neurônios dentro de um mapa de características têm pesos de *kernel* restritos a serem iguais, no entanto, mapas de recursos diferentes na mesma camada convolucional têm pesos de *kernel* diferentes, de modo que vários recursos podem ser extraídos em cada local.

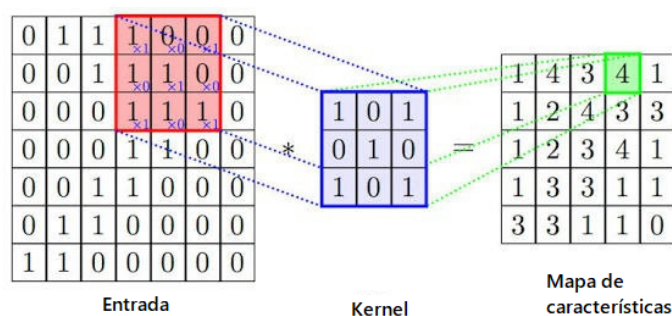


Figura 3: Operação realizada na camada convolucional.

A camada de agrupamento geralmente é aplicada após as camadas de convolução. A função dessa camada é reduzir a dimensionalidade dos dados na rede. Essa redução é importante por questões de agilidade no treinamento, mas principalmente para criar invariância espacial, isto é, não deixar a rede sensível a posição de, por exemplo, um objeto na imagem. Diferentes funções são encontradas na literatura, porém a mais utilizada é a função de máximo (RAWAT; WANG, 2017). Conforme ilustrado na Figura 4, a função de agrupamento máximo é aplicada sobre um mapa de característica de tamanho 4x4 e um campo receptivo de tamanho 2x2 com passo 2 propagando o valor máximo encontrado para a próxima camada, reduzindo a altura e a largura do mapa de característica.

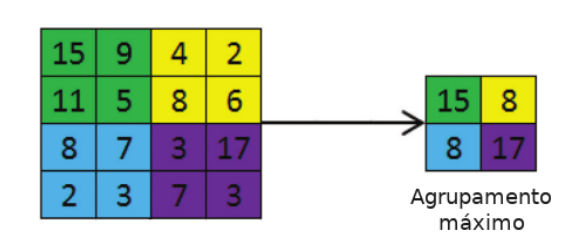


Figura 4: Função de agrupamento máximo.

3.1 ResNet50

O trabalho de He et al. (2016) introduziu a estrutura de aprendizado profundo residual ao campo das redes convolucionais profundas e conquistou o primeiro lugar na competição de classificação da *Large Scale Visual Recognition Challenge* de 2015. Segundo os autores, ao se projetar redes cada vez mais profundas em algum momento a acurácia ficará saturada, atingindo seu limite. Para conseguir contornar o problema e construir arquiteturas mais profundas apresentam a estrutura residual.

Assim como ilustrado na Figura 5, a ResNet50 recebe como entrada uma imagem no formato $224 \times 224 \times 3$ e aplica uma primeira camada convolucional com 64 unidades de *kernel* com tamanho 7×7 e passo 2. Em seguida uma camada de agrupamento com a função de agrupamento máximo com tamanho 3×3 e passo 2 é executada. Nas estruturas residuais 3 camadas convolucionais empilhadas formam um bloco: a primeira com *kernel* de tamanho 1×1 é responsável por reduzir a dimensão; a segunda com *kernel* de tamanho 3×3 executada o mapeamento sobre a dimensão reduzida; a terceira camada com *kernel* de tamanho 1×1 é responsável por aumentar a dimensão.

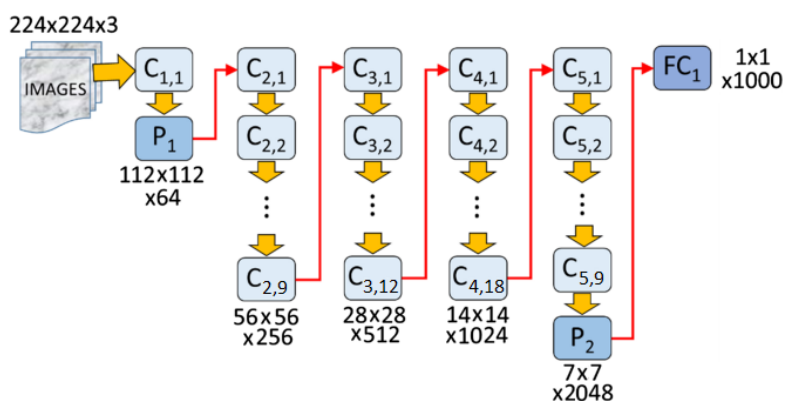


Figura 5: Arquitetura da CNN ResNet50 com 49 camadas convolucionais ($C_{1,1}$, $C_{2,1}$, ..., $C_{5,9}$), 2 de agrupamento (P_1 e P_2) e 1 totalmente conectadas (FC_1).

4 Descritor de textura *Local Binary Patterns*

Local Binary Patterns (LBP) é um descritor de textura que se tornou bastante popular pela publicação do estudo de Ojala, Pietikainen e Maenpaa (2002). A ideia do descritor é comparar cada pixel de uma imagem com sua vizinhança de pixels ao redor. O método é aplicado sobre uma imagem em nível de cinza, onde o valor de cada pixel da imagem é comparado com os valores de uma vizinhança de pixels de tamanho r , se o valor do pixel central for maior ou igual que seu vizinho o valor 1 é definido, caso contrário, o valor 0 é definido na matriz de saída conforme ilustrado na Figura 6.

No próximo passo a matriz é representada em valor decimal, para isso cada valor vizinho definido na etapa anterior é representado em função de 2, iniciando no sentido horário, ilustrado na Figura 6. Com o valor decimal representando o pixel central inicial, o valor é então armazenado na matriz LBP de saída e outro pixel central é escolhido até que a imagem seja completamente percorrida. A abordagem usando descritor LBP é um poderoso recurso para classificação de textura.

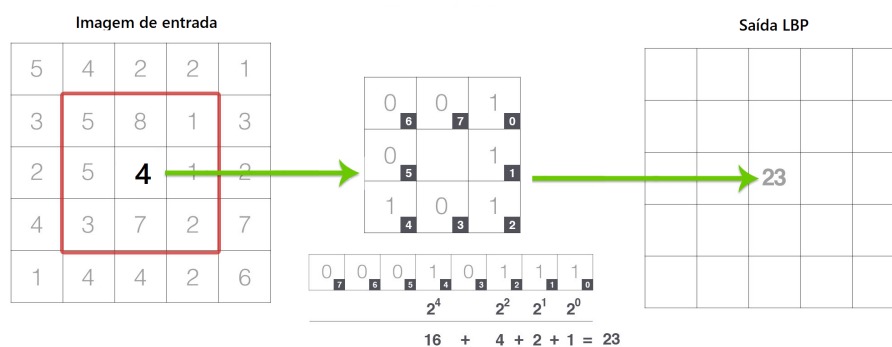


Figura 6: Método LBP de comparação de pixels.

5 Experimentos

A fim de agrupar os dados em categorias para posterior aplicação dos algoritmos de aprendizado de máquina com uma abordagem supervisionada em um modelo de classificação, os dados de teores foram agrupados através do método tamanho de bolha. O método de agrupamento por tamanho de bolha é proposto por Weixing e Liangqin (2016) que a partir de quatro imagens, escolhidas manualmente com base no tamanho da bolha, em escala de cinza aplica o método *threshold* de segmentação de imagens para criar uma imagem binária. Conforme ilustrado na linha 2 da Figura 7, essa segmentação consegue definir bem as áreas brancas do topo das bolhas e é possível observar que a medida que as bolhas aumentam de tamanho as áreas brancas ficam mais densas, isto é, mais pixels brancos agrupados em uma região menor e separados por grandes espaços pretos. Dessa forma, analisando a propriedade de área dessas regiões brancas é possível identificar as imagens da classe tamanho de bolha super grande.

Para explorar as demais classes, o recurso de textura de matriz de co-ocorrência de nível de cinza (GLCM) é sugerido por Weixing e Liangqin (2016) para descrever a propriedade de

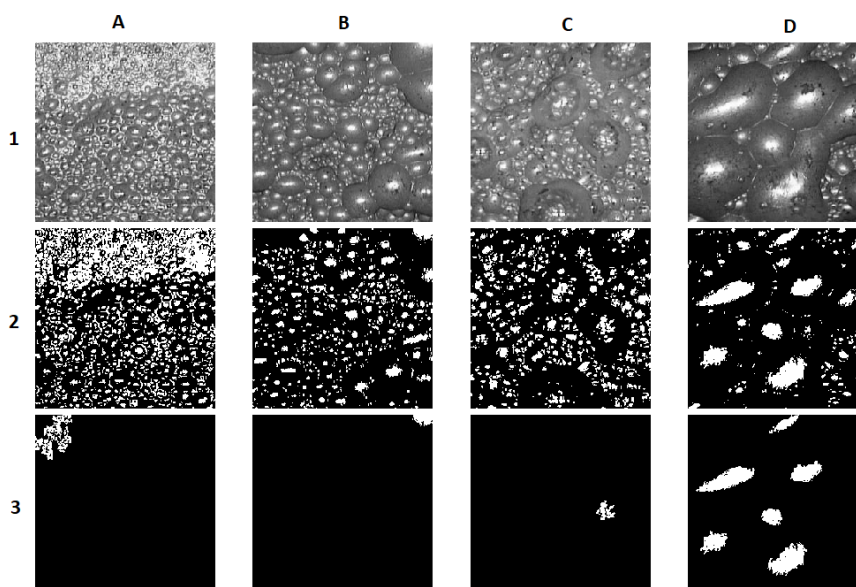


Figura 7: Segmentação de imagens baseado no tamanho da bolha. Classes de bolhas (A) pequeno, (B) médio, (C) grande, (D) super grande apresentadas em (1) escala de cinza, (2) regiões brancas extraídas e (3) regiões brancas com áreas de interesse.

contraste sobre as imagens. A partir das imagens da Figura 7 é calculado a matriz GLCM nos ângulos de 0° , 45° , 90° e 135° e extraído a propriedade de contraste de cada ângulo. Com a média dos valores de contraste, exibidos na Tabela 1, é possível observar diferentes faixas de valores encontrados para as classes de imagens propostas.

Tipo de imagem	Pequeno	Médio	Grande	Super grande
Contraste	1,58	0,65	0,57	0,39

Tabela 1: Propriedade de contraste extraída da GLCM.

Sendo assim, um método que combina o tamanho da área da região branca e a propriedade de textura é empregado para classificar as imagens com base no tamanho da bolha das espumas de flotação de nióbio e agrupar as imagens do conjunto de dados em classes definidas. Um total de 3656 imagens de espuma de flotação de nióbio foram coletadas e agrupadas pelo método descrito, a distribuição final das imagens podem ser visualizadas na Figura 8.

A partir das imagens agrupadas é investigado a relação do teor de concentrado de óxido de nióbio Nb_2O_5 nas classes formadas. A Figura 9 exhibe as estimativas da tendência central para o valor de teor concentrado de Nb_2O_5 em cada classe e plota todas em um eixo para fins de comparação entre as diferentes classes existentes. Analisando as classes formadas acima e a Figura 9, é possível observar uma relação entre as classes e os teores, em que a medida que o tamanho de bolha aumenta o valor de Nb_2O_5 também é maior.

Partindo do princípio onde as imagens identificam as classes, e as classes representam um valor de teor aproximado, um modelo de classificação é proposto usando o descritor de textura LBP e a arquitetura de CNN ResNet50.

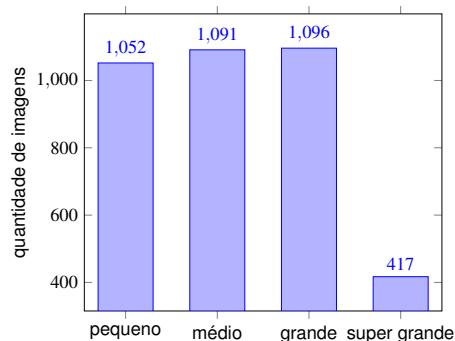


Figura 8: Distribuição das imagens entre as classes do *dataset*.

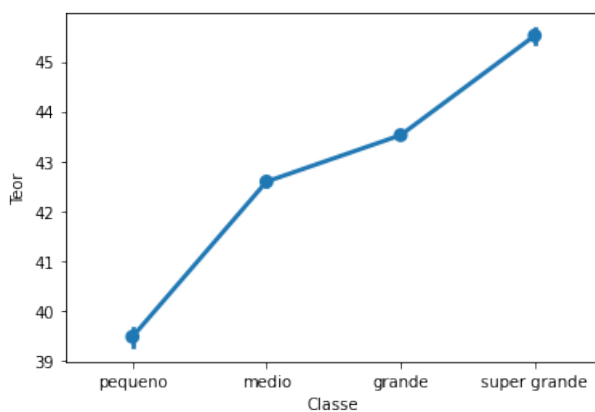


Figura 9: Tendência central dos valores de teor de Nb_2O_5 dentre as classes.

5.1 Resultados com descritor LBP

O método LBP invariante à rotação, da biblioteca *skimage*, foi inicializado com 24 pontos vizinhos e tamanho de raio 8. Todas as imagens foram descritas gerando 26 características para cada imagem. Os modelos *Linear Support Vector Classification* (SVC), *k-nearest neighbors* (KNN) e *Random Forest Classifier* foram utilizados neste experimento. Os resultados de acurácia foram calculados pelo método *cross-validation*, da biblioteca *sklearn*, com os dados de entrada sendo as características e rótulo das imagens. O tempo de execução do algoritmo para descrever as imagens e classifica-las com os modelos acima foi de 02:24 minutos. As médias dos resultados de cada método de classificação podem ser visualizados na Tabela 2.

	SVC	KNN	Random Forest
Acurácia	71%	76%	79%

Tabela 2: Resultados usando descritor LBP.

5.2 Resultados com ResNet50

No experimento o conjunto de dados foi dividido em três outros conjuntos através da biblioteca *sklearn*: treino com 70% treino, 20% validação e 10% teste. Os conjuntos de treino e validação foram carregadas sem a aplicação da técnica de *data augmentation* e o modelo, construído com a biblioteca Keras, foi iniciado sem a utilização de transferência de aprendizado. O treinamento levou um tempo de 11:43 minutos e a métrica final do modelo avaliada sobre o conjunto de dados de validação foi de 81,69% de acurácia e valor de erro 0,44. Isso significa que o modelo, a partir de uma imagem de flotação de nióbio qualquer, consegue prever a classe, definida na metodologia, com no mínimo 81% de precisão. A Figura 10 mostra o resultado de previsão sobre o conjunto de teste através da matriz de confusão com 85% de precisão.

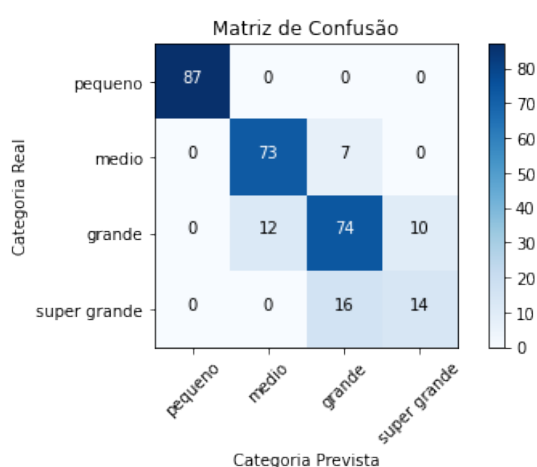


Figura 10: Matriz de confusão prevista do ResNet50 sobre os dados de teste.

6 Discussão final

Para resumir, nesta investigação, foi explorado arquiteturas de rede neural convolucional ResNet50 e descritores de textura como o *Local Binary Patterns* para o desenvolvimento de sensores de imagem de espuma de flotação de nióbio. As seguintes conclusões podem ser tiradas deste estudo:

- A forte relação do teor com as classes pode ser explicada pelo método que classifica as imagens com relação ao tamanho das bolhas na espuma de flotação. Essas bolhas são responsáveis por capturar as partículas minerais de Nb_2O_5 e arrastá-las à superfície (LUZ; SAMPAIO; ALMEIDA, 2010). Portanto, essas classes podem ser vistas como cenários operacionais diferentes que geram diferentes faixas de valores do teor concentrado mineral estudado.
- A aplicação de redes neurais profundas mostrou uma pequena melhora na acurácia quando comparado ao descritor LBP. Na CNN o custo computacional é maior devido as 49 camadas convolucionais que iteram sobre o espaço de entrada das imagens resultando em aproxi-

madamente 25 milhões de parâmetros treináveis. Por outro lado, o descritor LBP compara cada pixel com sua vizinhança gerando valores da mesma dimensão da imagem de entrada. A diferença é perceptível no tempo de treinamento dos algoritmos, porém ambos os modelos depois de treinados podem realizar predições rápidas sobre uma imagem qualquer. Como o experimento é realizado sobre uma quantidade limitada de imagens, uma abordagem utilizando CNN seria mais preferível devido a melhor acurácia e ao tempo de treinamento relativamente baixo.

- Assim como no estudo de caso de platina no trabalho de Fu e Aldrich (2019) este estudo mostra potencial em investigar a aplicação de sensores em imagens de flotação de um domínio diferente, como o de nióbio.
- É comprovado que a classificação de uma imagem qualquer de espuma de flotação de nióbio, entre as classes indicadas neste estudo, também indica um valor aproximado do teor concentrado mineral de Nb_2O_5 . Essa indicação pode ser usada para auxiliar na tomada de decisão no controle de processo, em um cenário onde a produção é inversamente proporcional a qualidade, e onde o resultado da qualidade do produto final só é entregue em grandes intervalos de tempo.

Referências

- ALDRICH, Chris et al. Online monitoring and control of froth flotation systems with machine vision: A review. **International Journal of Mineral Processing**, Elsevier, v. 96, n. 1-4, p. 1–13, 2010.
- CAO, Binfang et al. Integrated prediction model of bauxite concentrate grade based on distributed machine vision. **Minerals engineering**, Elsevier, v. 53, p. 31–38, 2013.
- ESTRADA-RUIZ, RH; PÉREZ-GARIBAY, R. Neural networks to estimate bubble diameter and bubble size distribution of flotation froth surfaces. **Journal of the Southern African Institute of Mining and Metallurgy**, The Southern African Institute of Mining e Metallurgy, v. 109, n. 7, p. 441–446, 2009.
- FARIAS, Carlos Eugênio Gomes. Mineração e meio ambiente no Brasil. **Relatório do CGEE/PNUD**, v. 76, p. 2, 2002.
- FU, Y; ALDRICH, C. Flotation froth image recognition with convolutional neural networks. **Minerals Engineering**, Elsevier, v. 132, p. 183–190, 2019.
- FU, Yihao; ALDRICH, Chris. Froth image analysis by use of transfer learning and convolutional neural networks. **Minerals Engineering**, Elsevier, v. 115, p. 68–78, 2018.
- HE, Kaiming et al. Deep residual learning for image recognition. In: PROCEEDINGS of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.], 2016. p. 770–778.
- JAHEDSARAVANI, A; MASSINAEI, M; MARHABAN, MH. Development of a machine vision system for real-time monitoring and control of batch flotation process. **International Journal of Mineral Processing**, Elsevier, v. 167, p. 16–26, 2017.
- KISTNER, Melissa; JEMWA, Gordon T; ALDRICH, Chris. Monitoring of mineral processing systems by using textural image analysis. **Minerals Engineering**, Elsevier, v. 52, p. 169–177, 2013.
- LIU, Weibo et al. A survey of deep neural network architectures and their applications. **Neurocomputing**, Elsevier, v. 234, p. 11–26, 2017.
- LUZ, Adão; SAMPAIO, João; ALMEIDA, Salvador. **Tratamento de minérios**. 5. ed. Rio de Janeiro: CETEM/MCT, 2010.
- OJALA, Timo; PIETIKAINEN, Matti; MAENPAA, Topi. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 24, n. 7, p. 971–987, 2002.
- RAWAT, Waseem; WANG, Zenghui. Deep convolutional neural networks for image classification: A comprehensive review. **Neural computation**, MIT Press, v. 29, n. 9, p. 2352–2449, 2017.
- VIEIRA, Guilherme E. Uma revisão sobre a aplicação de simulação computacional em processos industriais. **Simpósio de Engenharia de Produção, XIII, Bauru, Anais**, p. 1–10, 2006.
- WEIXING, Wang; LIANGQIN, Chen. Mineral Froth Image Classification and Segmentation. In: RECENT Advances in Image and Video Coding. [S.l.]: InTech Rijeka, Croatia, 2016.

Grupo da Automação Industrial – Instituto Federal do Espírito Santo (Ifes) – 29173-087– Rodovia ES-010 - Km 6,5 – Manguinhos – Serra – ES – Brasil