

USO DE UM APLICATIVO DE VISÃO COMPUTACIONAL PARA ANÁLISE DE IMAGENS DE PROCESSOS DINÂMICOS

A. C. C. MIGUEL¹, C. D. GROSSI¹, L. A. CALÇADA¹, C. M. SCHEID¹ e L. A. C. MELEIRO¹

¹Universidade Federal Rural do Rio de Janeiro, Departamento de Engenharia Química
E-mail para contato: andreh.miguel@hotmail.com

RESUMO – O presente trabalho trata da aplicação da visão computacional na análise de um processo dinâmico constituído de um equipamento para alimentação contínua de sólidos, uma câmera para filmagem e um *software* para processamento e análise de imagens. Este sistema foi utilizado para verificar a capacidade do *software* de visão computacional desenvolvido de estimar a taxa de grãos de milho que escoam pela calha do alimentador de sólidos. O resultado obtido foi satisfatório, tendo em vista que a prova de conceito foi realizada e o *software* foi capaz de estimar taxa de grãos que escoam pela calha com erro médio de 17,5%.

1. INTRODUÇÃO

1.1. Definição e Motivação

A visão computacional é um assunto relativamente novo e totalmente conectado com máquinas de aprendizado (*Machine Learning*). Este conceito também é conhecido como visão por computador e é a área da ciência que desenvolve técnicas e métodos voltados à extração automática de informações úteis contidas em imagens. O objetivo dessa ciência é transmitir essas informações às máquinas de forma adequada (FORSYTH e PONCE, 2003).

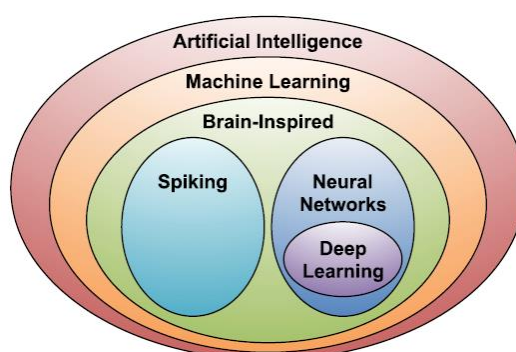
As máquinas de aprendizado são o campo de estudo que usam computadores para fazer o que é natural para os seres humanos, ou seja, aprender com a experiência. Algoritmos de máquinas de aprendizado usam métodos computacionais para “aprender” informações diretamente de dados sem depender de uma equação predeterminada como modelo. Os algoritmos melhoram de modo adaptativo o seu desempenho na medida em que o número de amostras disponíveis para o aprendizado aumenta. As máquinas de aprendizado usam dois tipos de técnicas: *Aprendizado Supervisionado*, que treina um modelo a partir de um conjunto de dados de entrada e saída conhecido de modo que seja possível prever saídas futuras e; *Aprendizado Não Supervisionado*, que encontra padrões escondidos ou estruturas intrínsecas nos dados de entrada (BISHOP, 2016).

A técnica de máquinas de aprendizado supervisionado visa construir um modelo que faz previsões baseadas em evidências na presença de incertezas. Um algoritmo de aprendizado supervisionado usa um conjunto conhecido de dados de entrada e respostas conhecidas para os dados (saída) e treina um modelo para gerar previsões para a resposta aos novos dados. O aprendizado supervisionado usa técnicas de classificação e regressão para desenvolver

modelos preditivos. Por outro lado, a técnica de aprendizado não supervisionado encontra padrões escondidos ou estruturas intrínsecas em dados, onde o *Clustering* é a técnica mais comum de aprendizado não supervisionado e geralmente é usado para análise exploratória de dados para localizar padrões escondidos ou agrupamentos em dados (BISHOP, 2016).

Aprendizagem profunda é um tipo de máquina de aprendizado em que um modelo aprende a executar tarefas de classificação diretamente a partir de imagens, texto ou som. O aprendizado profundo normalmente é implementado utilizando uma arquitetura específica de rede neuronal. O termo *deep* refere-se ao número de camadas na rede, de modo que quanto mais camadas, mais profundo é o aprendizado. Enquanto redes neurais tradicionais contêm apenas duas ou três camadas, as redes para o aprendizado profundo podem ter centenas. Apresentam-se na Figura 1 algumas técnicas de aprendizado no contexto da Inteligência Artificial (GOODFELLOW *et al.*, 2016).

Figura 1 - Aprendizagem profunda (*Deep Learning*) no contexto da Inteligência Artificial (Fonte: Sze, V. *et al.*, 2017).



Visão computacional e reconhecimento de imagem são termos frequentemente usados como sinônimos, mas o primeiro possui um conceito mais abrangente do que apenas analisar imagens. O reconhecimento de imagens em si (a análise de pixel e padrão de imagens) é uma parte integrante do processo de visão computacional que envolve tudo, desde reconhecimento de objetos e caracteres até análise de texto e sentimento. Já a visão computacional em seu sentido mais pleno está sendo usada na vida cotidiana e nos negócios para conduzir todos os tipos de tarefas, incluindo identificar doenças, produtos, anúncios dentro de imagens editoriais, entre outros. A visão computacional pode ser usada para digitalizar plataformas de mídia social a fim de encontrar imagens relevantes que não podem ser descobertas por meio de pesquisas tradicionais. A tecnologia é complexa e, assim como todas as tarefas acima mencionadas, requer mais do que apenas reconhecimento de imagem, mas também análise semântica de grandes conjuntos de dados (BACKES e SÁ, 2016).

As tarefas típicas da visão computacional são: Reconhecimento; Identificação; Detecção; Movimento; Reconstrução de cena; Restauração de imagens. Essas tarefas são aplicadas ao processo de visão computacional a fim de realizar a identificação total das imagens de forma robusta e sem esforço humano. A operação desse tipo de tecnologia vem sendo cada vez mais estudada, apesar de ser bastante complexa e, ainda hoje, é uma tarefa bastante desafiadora. Concomitantemente, com os recentes avanços na ciência da computação, observa-se uma explosão da utilização de sistemas que imitam o comportamento

do cérebro com uso de tecnologias de *Machine Learning*, *Big Data*, processamento paralelo em GPU's e técnicas de Inteligência Artificial. Os avanços nestas áreas já são perceptíveis e o futuro da visão computacional é integrar sistemas poderosos e específicos. Isso significa que mesmo em um estágio inicial, a visão computacional já é uma realidade tecnológica em diversas áreas de conhecimento (FORSYTH e PONCE, 2003).

A organização de um sistema de visão computacional depende da aplicação. A implementação específica de tal sistema depende também se sua funcionalidade é pré-especificada ou se existe alguma parte de aprendizagem durante a operação. Funções típicas encontradas normalmente em sistemas de visão computacional são destacadas na Figura 2.

Figura 2 – Esquema de um sistema de visão computacional (Fonte: BACKES e SÁ, 2016).



2. MATERIAIS E MÉTODOS

2.1. MATERIAIS

Sólido a ser analisado: Os sólidos utilizados neste trabalho foram grãos de milho da marca Chinezinho Tipo 1.

Sistema de alimentação de sólidos (Moega): Utilizou-se neste trabalho um dosador vibratório de média capacidade, marca MVL vibradores industriais LTDA, de aço inoxidável acoplado a um silo armazenador de sólidos. Este equipamento foi utilizado para alimentar os grãos com vazão mássica ajustada por meio de um potenciômetro. Os grãos foram coletados em um recipiente de aço inoxidável, do tipo balde industrial.

Sistema de aquisição de imagens: Para a aquisição de imagens foi utilizado um conjunto formado por um tripé e uma câmera filmadora NIKON D5100 DSLR, com sensor óptico CMOS de 16,2 megapixels e capacidade de filmar em *High Definition*.

Sistema de análise de imagens: Os códigos computacionais utilizados neste trabalho para analisar as imagens captadas pela filmadora foram escritos no *software* GNU Octave.

2.1. MÉTODOS

Procedimento experimental de alimentação de sólidos: Os grãos de milho foram alimentados no silo do sistema de alimentação de sólidos, conforme ilustrado na Figura 3 (a),

e, em seguida, a intensidade de vibração do alimentador foi ajustada para diferentes valores por meio da manipulação do potenciômetro, conforme ilustrado na Figura 3 (b). Várias corridas foram feitas para diferentes vazões volumétrica e as imagens do processo de alimentação de grãos foram coletadas utilizando o sistema de aquisição de imagens. A filmagem foi feita posicionando a câmera em um tripé na frente do alimentador de sólidos em um ângulo aproximadamente de 45°, de forma que o foco da imagem seja a calha por onde escoam os grãos. As imagens adquiridas são utilizadas para alimentar o programa computacional feito no *software* GNU Octave, cujo objetivo é estimar a taxa de grãos que escoam pela calha para uma determinada potência de agitação.

Figura 3 – Unidade de alimentação composta pelo silo, calha e coletor (a) e detalhe do potenciômetro para controle da vazão de alimentação (b).



(a)



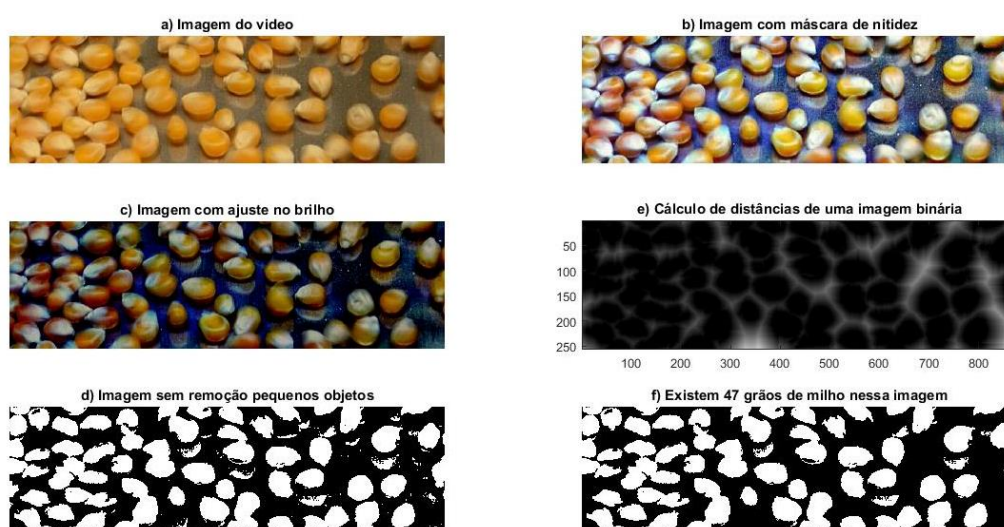
(b)

Procedimento computacional para análise das imagens dos sólidos: O algoritmo do programa de visão computacional consiste em carregar o vídeo do escoamento de grãos pela calha do alimentador para cada potência e determinar o número de *frames* correspondente ao tempo de cada filme. Conhecido o tempo de filmagem e o número de *frames* do vídeo, foi determinado, por meio da análise visual das imagens, o número real de grãos que escoava na calha por segundo (NRG/s). Após a contagem dos *frames*, o programa executa uma série de comandos para aplicar filtros de contraste, com objetivo de realçar as bordas com equalização histográfica. Este recurso tem como objetivo melhorar a nitidez e ajustar o brilho, minimizando o reflexo da imagem do grão na calha de metal e evitando a contagem equivocada de grãos. Na sequência, a imagem é binarizada e são traçadas as regiões limites de interesse, após o que as características estatísticas são calculadas e com essas informações o programa é capaz de identificar e realizar a contagem do número de grãos por *frame* e assim, estimar quantos grãos passam por segundo na calha (NEG/s).

3. RESULTADOS E DISCUSSÃO

Para ilustrar o procedimento de identificação e contagem automática do número de grãos que escoam pela calha do alimentador por segundo, foi selecionado o experimento correspondente à potência de 40% para o agitador (Imagem 1), ilustrado na Figura 4. A sequência de imagens na Figura 4 ilustra as etapas do algoritmo utilizado neste estudo de caso, onde é possível verificar que o *software* desenvolvido conseguiu estimar a quantidade de grãos de milho no *frame* ilustrado. Contudo, em alguns casos, foram verificados erros na contabilização, visto que devido ao reflexo dos grãos na calha ou à sua aglomeração durante o escoamento, a contabilização dos grãos foi equivocada. O mesmo procedimento foi repetido para as 11 imagens processadas neste vídeo.

Figura 4 – Tratamento da Imagem 1 e procedimento para contagem de sólidos.

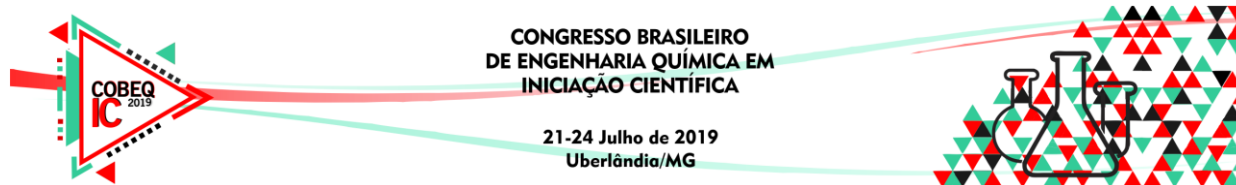


A partir das respostas fornecidas pelo *software* e por meio da contagem visual dos grãos, foi possível avaliar o desempenho do *software* de visão computacional desenvolvido neste trabalho, sendo os respectivos resultados compilados na Tabela 1.

Tabela 1 – Resultado da contagem de grãos nas imagens processadas, em que NRG/s representam o número real de grãos e NEG/s o número estimado de grãos.

Imagem	NRG/s	NEG/s	Erro (%)	Imagem	NRG/s	NEG/s	Erro (%)
1	56	47	16,1	7	55	44	20,0
2	57	45	21,1	8	58	46	20,7
3	53	45	15,1	9	48	39	18,8
4	54	46	14,8	10	36	36	0,0
5	59	44	25,4	11	40	32	20,0
6	59	47	20,3	Média	52	43	17,5

De acordo com os dados apresentados na Tabela 1, é possível verificar que o *software* de visão computacional estimou que, em média, 43 grãos de milho escoaram na calha por segundo e que, quando comparado com a análise visual, apresentou um erro médio de 17,5%. Tal erro provavelmente se deve à aglomeração de grãos observada durante a maior parte dos experimentos. Este erro será corrigido nas próximas versões do *software* por meio da



utilização de filtros mais eficientes, assim como na implementação de algoritmos capazes de estimar o tamanho médio do grão e, desse modo, identificar o fenômeno de aglomeração nas imagens.

4. CONCLUSÃO

O presente trabalho apresenta os resultados iniciais da aplicação de um *software* de visão computacional utilizado na identificação e na análise de sólidos em movimento que, nesse caso, correspondem a grãos de milho escoando em na calha de um alimentador de sólidos. De acordo com os resultados apresentados, o *software* de visão computacional foi capaz de estimar a taxa de escoamento de grãos com erro médio de 17,5%. Apesar das falhas apresentadas, o *software* tem potencial para fornecer resultados mais precisos e, em tese, pode ser empregado na análise dinâmica de qualquer tipo de sólido, desde que possam ter alguma característica, como por exemplo seu contorno, detectado pelo vídeo.

5. REFERÊNCIAS

- BACKES, A.R.; SÁ Jr, J.J.M. Introdução à Visão Computacional Usando MATLAB. Alta Books Editora, Rio de Janeiro, 2016.
- BISHOP, C. M. *Pattern Recognition and Machine Learning*. Springer New York, 2016
- FORSYTH, D. A. and PONCE, J., *Computer Vision. A Modern Approach*, Prentice Hall, 2003.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning - Adaptive Computation and Machine Learning*. MIT Press, 2016.
- SZE, V.; CHEN Y.-H; YANG, T.-J; EMER, J. *Efficient Processing of Deep Neural Networks: A Tutorial and Survey*. Proceedings of the IEEE, 2017.