

UM MODELO PARA CONSTRUÇÃO DE DATASETS DE COMPONENTES INDUSTRIAIS PARA VISUAL-BASED TRACKING COM DEEP LEARNING

AZEVEDO, Lis da Silva¹; MATOS NETO, Rosalvo²; WINKLER, Ingrid³

¹ Projeto de pesquisa de Iniciação Científica

¹ Graduanda em Engenharia da Computação; lisazevedo3@gmail.com

² Mestrando do Programa de Modelagem Computacional e Tecnologias Industriais (PPG MCTI/ CIMATEC), Salvador, BA, rosalvo.matos@gmail.com

³ Professora Pesquisadora do Programa em Modelagem Computacional e Tecnologias Industriais do SENAI CIMATEC (PPG GETEC/ CIMATEC), Salvador, BA, ingrid.winkler@fiob.org.br

RESUMO

Em contextos industriais, vem sendo aplicadas tecnologias de *deep learning* e *visual-based tracking* para manutenção industrial, porém existem dificuldades encontradas, atualmente, no *visual-based tracking* que dificulta o seu uso. Com isso, o objetivo deste trabalho é propor um modelo para criação de *dataset* de componentes industriais para *visual-based tracking* em ambientes industriais que utilizem *deep learning*. Para tanto, foi realizada uma pesquisa para selecionar as melhores técnicas que resolvam esses problemas. Observou-se que utilizando técnicas de *data augmentation* e segmentação de imagem, no modelo, a precisão do *tracking* aumentou, contudo, ainda é necessário realizar teste em um ambiente industrial.

PALAVRAS-CHAVE: Dataset, Deep Learning, Data Augmentation, Segmentação de Imagem.

1. INTRODUÇÃO

Atualmente, em contextos industriais, vêm sendo aplicadas novas tecnologias, como *Deep Learning*, para predição da vida útil de equipamentos¹ e para alerta prévio de acidentes². Porém algumas tecnologias ainda faltam amadurecimento para utilização. Por exemplo, aplicações de manutenção que utilizam *visual-based tracking* de componentes ainda não estão presentes, pois existem problemas em ambientes industriais que dificultam a aquisição de imagens como, iluminação variada, ambiente modificado com frequência e obstrução³. Nesse contexto, essa pesquisa tem como objetivo propor um modelo para criação de *dataset* de componentes industriais para *visual-based tracking* em ambientes industriais que utilizem *deep learning*.

A pesquisa realizada caracteriza-se como um estudo experimental exploratório, que tem como meta proporcionar maior familiaridade com o problema⁴. Este trabalho está organizado em 4 seções: além desta Introdução, a seção 2 descreve a Metodologia utilizada; a seção 3 descreve os Resultados observados e discute-os e, finalmente, na seção 4 são realizadas as Considerações Finais e indicações de pesquisa futura.

2. METODOLOGIA

Para a construção do modelo proposto, foram constituídos seis passos. Primeiro foi selecionado, classificado e analisado os componentes para realizar o *tracking* (rastreamento ou identificação); segundo, foram capturadas imagens dos componentes com variações de distância, luminosidade, sujeira, angulação e obstrução; terceiro, foi definido a técnica de *deep learning* para saber as limitações da técnica; quarto, segmentado as imagens; quinto, foi definido os métodos de *data augmentation* e os filtros de imagem a serem aplicados; sexto, foi desenvolvido um algoritmo para criação do modelo do *dataset*.

O processo de seleção, classificação e análise dos componentes é importante, pois é necessário fazer uma análise dos componentes, levando em consideração a distância que os componentes de interesse estão entre si e o grau de semelhança e suas características únicas.

Na captura de imagem foi utilizado uma câmera de celular, que é um *hardware* de uso popular, e o modelo escolhido possui *flash*, criando mais possibilidades de imagens.

No processo de definição da técnica de *deep learning* a ser utilizada foi levado em consideração os componentes a serem identificados e sua análise, pois a complexidade das características dos componentes e a quantidade dos componentes por imagem é de suma importância, pois influenciam tanto na complexidade da técnica de *tracking* utilizada, quanto na técnica de *deep learning*.

No processo de segmentação de imagem, tecnologia que possibilita o *tracking* de múltiplos componentes em uma mesma imagem⁵, foi pesquisado um *software* que realiza processamento e auxilia na

criação dos segmentos de imagens no formato aceito pelo *software* de treinamento de *deep learning* com GPU, o Nvidia DIGITS.

Na definição das técnicas de *data augmentation*, técnica usado para aumentar a quantidade de dados em um *dataset*⁴, e filtros de imagens foram levadas em consideração as análises iniciais dos componentes, pois a técnica de *data augmentation* e aplicações dos filtros pode gerar alteração na cor e no formato dos componentes, podendo alterar ou remover as singularidades de cada componente.

No desenvolvimento do algoritmo de criação do *dataset* foram pesquisadas tecnologias, *frameworks* e *bibliotecas* apropriados para manipulação de imagens e aplicação de técnicas de *data augmentation*.

3. RESULTADOS E DISCUSSÃO

No processo de seleção dos componentes foram escolhidos dois componentes semelhantes e um distinto dos demais, além de levar em consideração a distância entre os componentes, para assim capturar imagens com mais de um componente.

O processo de captura de imagens foi realizado em um ambiente controlado com a tentativa de simular um ambiente industrial. Foram capturadas imagens que possuíam mais de um componente, variando a luminosidade com o *flash* da câmera, simulando obstrução com objetos e simulando sujeira, passando borra de café nos componentes. Foram capturadas trezentas imagens, sendo cem de cada componente escolhido e dentre as trezentas imagens, algumas possuíam mais de um componente.

A técnica de *Deep Learning* escolhida foi a ImageNet, pois é uma técnica que possibilita o *tracking* de múltiplos componentes em uma mesma imagem, além de ser disponibilizada na plataforma Nvidia DIGITS, que foi a plataforma escolhida para aplicação da técnica de *deep learning*. Para utilizar essa técnica as imagens precisam passar pelo método de segmentação de imagem por *bounding box*, que é um retângulo utilizado para representar uma região de interesse.

Devido à necessidade da segmentação por *bounding box*, foi necessário pesquisar um software de processamento de imagem especializado. Foi escolhido o ImageJ/Fiji, pois a estrutura do *file*, que contém as *labels* dos *bounding boxes*, gerada por esse *software* é compatível com a ImageNet. Foi realizado o processo de segmentação das imagens antes de aplicar *data augmentation*, pois é um processo manual, sujeito a falhas humanas, que demanda tempo e atenção, podendo gerar mais erros no *dataset* realizar o *bounding box* de um número maior de imagens.

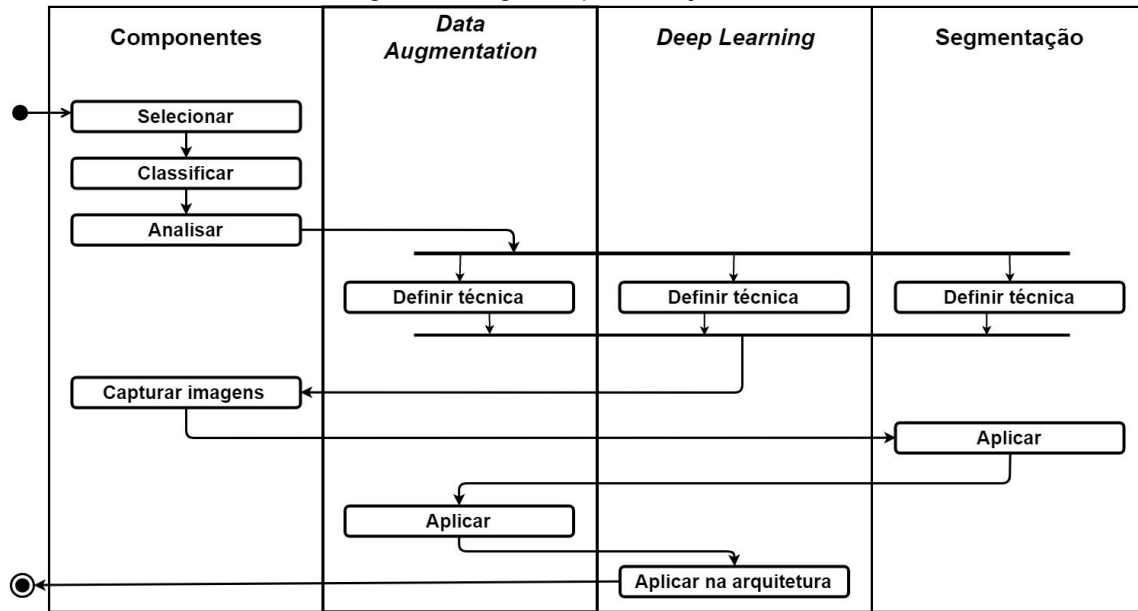
A aplicação da técnica de *data augmentation* foi essencial para o modelo pois ela diminui os problemas como, dificuldade de aquisição de imagens, iluminação variada, obstrução e *overfitting* (sobreajuste)⁴ presente nas técnicas de *deep learning*. No algoritmo desenvolvido ao aplicar a técnica de *data augmentation* e os filtros, uma imagem é representada em diferentes posições, ângulos, iluminação, distorções e filtros de cor, gerando aleatoriamente cinco variações da mesma imagem, aumentando significativamente a quantidade de dados para a criação do *dataset*. Com a aplicação de *data augmentation* nas imagens segmentadas, ocorreram alterações nas imagens, sendo necessário criar novas *labels* das imagens geradas para atender as alterações feitas.

Ao finalizar o modelo de criação de *dataset* foi desenvolvido um algoritmo para automatizar a criação de *datasets* em que ele aplica *data augmentation* e os filtros nas imagens, cria as pastas *Training* e *Validation*, conforme a arquitetura aceita pela ImageNet, e divide as imagens entre as pastas criadas em 80% e 20%, respectivamente.

Ao finalizar o algoritmo foram criados *datasets*, com as imagens capturadas anteriormente, para realizar os primeiros testes. Utilizando os *datasets* criados foi possível observar que os *datasets* que não possuíam *bounding box* só conseguiam realizar o *tracking* de um componente por imagem, sem possibilidade de escolha de qual componente. Os *datasets* que não possuíam *data augmentation* apresentavam uma precisão muito baixa no *tracking*, além de possuírem um alto *overfitting*, uma vez que esses *datasets* possuíam trezentas imagens. Ao aplicar as técnicas de *data augmentation* e segmentação por *bounding box*, observou-se que a precisão de *tracking* aumentou, em relação aos resultados anteriores, sendo possível reconhecer múltiplos componentes.

Com os resultados obtidos, foi possível desenvolver um modelo para criação de *datasets*, representado na figura 1. No diagrama, é possível visualizar, passo a passo, o processo de criação de um *dataset*, isso é, selecionar, classificar e analisar os componentes que se tem interesse de criar o *dataset*, logo após definir as técnicas de *data augmentation*, *deep learning* e segmentação, depois capturar as imagens dos componentes, aplicar a técnica de segmentação, definida anteriormente, nas imagens capturadas, aplicar a técnica de *data augmentation* e por último aplicar as imagens na arquitetura da técnica de *deep learning* escolhida.

Figura 1: Diagrama para criação de *datasets*



4. CONSIDERAÇÕES FINAIS

Observou-se que utilizar técnicas de *data augmentation* e segmentação de imagens na criação do modelo, ajuda a suprir as dificuldades do *tracking*, atualmente, encontradas em ambiente industrial.

Em pesquisas futuras, recomenda-se aplicar outras técnicas de segmentação de imagem, como, segmentação em grafos, além de testar mais métodos de *data augmentation*, para assim criar um *dataset* com maior quantidade de dados e testar o modelo com outras técnicas de *Deep Learning*. Também é necessário testar o *visual-based tracking* em um ambiente industrial, pois o mesmo não foi testado.

Agradecimentos

Os autores agradecem ao Grupo de Pesquisa em Realidade Aumentada e Realidade Virtual para Inovação na Indústria, Saúde e Educação (CNPQ) e à FAPESB pelo auxílio financeiro da bolsa de Iniciação Científica concedida.

5. REFERÊNCIAS

- ¹ H. Yan, J. Wan, C. Zhang, S. Tang, Q. Hua and Z. Wang, **Industrial Big Data Analytics for Prediction of Remaining Useful Life Based on Deep Learning**. *IEEE Access*, vol. 6, pp. 17190-17197, 2018.
- ² Y. Zheng, S. Chen, Y. Xue and J. Xue, **A Pythagorean-Type Fuzzy Deep Denoising Autoencoder for Industrial Accident Early Warning**. *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 6, pp. 1561-1575, Dec. 2017.
- ³ PALMARINI, R., ERKOYUNCU, J. A., ROY, R., TORABMUSTAEDI, H. **A systematic review of augmented reality applications in maintenance: Robotics and Computer-Integrated Manufacturing**, v. 49, p. 215-228, 2018.
- ⁴ WANG, Jason, PEREZ, Luis. **The Effectiveness of Data Augmentation in Image Classification using Deep Learning**. CoRR, 2017.
- ⁵ KANG, Kai *et al.*, **Object Detection in Videos with Tubelet Proposal Networks**. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017.