

REALIDADE AUMENTADA PARA A INDÚSTRIA: UMA PROPOSTA DE MODELO COMPUTACIONAL BASEADO EM DEEP LEARNING

MATOS NETO, Rosalvo¹; WINKLER, Ingrid²; SENNA, Valter³

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

2 Professora Pesquisadora do Programa de Gestão e Tecnologias Industriais do SENAI CIMATEC (PPG GETEC/ CIMATEC), Salvador, BA, ingrid.winkler@fieb.org.br;

3 Professor Pesquisador do Programa de Modelagem Computacional e Tecnologias Industriais do SENAI CIMATEC (PPG MCTI/ CIMATEC), Salvador, BA, senna@fieb.org.br.

RESUMO

Soluções de Realidade Aumentada vem sendo propostas a fim de auxiliar a área de Manutenção Industrial. Porém, as técnicas atuais de *tracking* ainda não são consideradas confiáveis, e uma solução em potencial é a adoção de técnicas de *Deep Learning*. Nesse contexto, o objetivo deste trabalho é propor um modelo computacional baseado em um dispositivo de Realidade Aumentada que aplique técnicas de Visão Computacional para o *tracking* de componentes industriais de forma confiável. Para tanto, foi realizada uma pesquisa exploratória experimental, onde os resultados apresentam o uso de diversas tecnologias e técnicas, onde taxas de precisão de até 99,9% são encontradas, entretanto, ainda são necessários mais estudos sobre estas técnicas. No final é proposto um modelo de arquitetura UML.

PALAVRAS-CHAVE: Realidade Aumentada, Deep Learning, Tracking, Classificação de Objetos, Manutenção

1. INTRODUÇÃO

Nas últimas 5 décadas, soluções com Realidade Aumentada (RA) para apoiar a manutenção vêm sendo investigadas do ponto de vista da pesquisa, porém, ainda existem desafios que inviabilizam ou dificultam a adoção em larga escala dessas soluções no contexto industrial. Um dos desafios está relacionado ao rastreamento ou *tracking*.

Como possíveis soluções para os problemas de *tracking*¹¹, técnicas de Visão Computacional (VC) vem sendo aplicadas, porém, ainda não tornam tais soluções confiáveis por conta de variáveis de ambiente existentes na indústria, como iluminação, obstrução ou modificação constante.

Diante deste problema, têm sido propostas soluções que utilizam técnicas de *Machine Learning* (ML) ou *Deep Learning* (DL) para superar tais brechas. Diante do exposto, este trabalho tem como objetivo propor um modelo computacional baseado em um dispositivo de RA que aplique técnicas de DL para o *tracking* de componentes industriais de forma confiável.

2. MATERIAIS E MÉTODOS

Esta pesquisa foi dividida em três etapas: uma caracterização do conhecimento atual sobre a temática da utilização de ML ou DL para realizar *tracking* em ambientes industriais, a proposta de um modelo e testes preliminares para avaliar o potencial desse modelo.

Para a consecução da primeira etapa, foi realizada uma revisão integrativa que proporciona a síntese de conhecimento. As bases científicas investigadas foram a IEEE Xplore, Science Direct e Scopus. Os critérios de inclusão e exclusão aplicados foram artigos publicados entre 2008 e 2018, string de busca recorrente no título, abstract ou palavras-chave e tipos de documento para revistas e artigos.

Para a consecução da segunda etapa, o modelo computacional é proposto através de um diagrama de implementação, baseado em *Unified Modeling Language* (UML), utilizando a ferramenta *Enterprise Architect* (EA). Para tal diagrama, realizou-se um estudo sobre sistemas de RA, ML e DL^{1, 2, 3, 4}, considerando que as arquiteturas de ML e DL são similares, onde a arquitetura proposta pelos autores sugestionou a forma da proposta neste trabalho.

Na terceira etapa, para realizar os treinos de DL foi utilizada a plataforma *NVIDIA Digits* com os frameworks suportados pela mesma, *Caffe*, *Keras* e *TensorFlow*. Foram selecionados três componentes distintos em um corpo para comporem o *dataset*. Como existem fatores que dificultam a coleta de imagens em ambientes industriais, como custo de parada e má visibilidade, técnicas de *Data Augmentation* (DA) foram

aplicadas, permitindo que tamanho do *dataset* fosse elevado, também reduzindo a chance de *overfitig*. As modificações utilizadas em DA foram rotação e escala, filtro gaussiano e coloração preto e branco.

Como o ambiente onde se encontravam os componentes utilizados no *dataset* era laboratorial, alguns ruídos foram adicionados manualmente para simular ambientes industriais, sombras e objetos foram postos próximos aos componentes para gerar obstruções, colocados pós e alguns objetos ao redor foram movidos para simular a modificação constante presente no ambiente.

3. RESULTADOS E DISCUSSÃO

3.1 Caracterização do conhecimento atual sobre a temática da utilização de ML ou DL para realizar tracking em ambientes industriais

Os vinte e três estudos identificados foram classificados segundo a técnica utilizada (machine ou deep learning) e agrupados pela área de atuação, permitindo apontar qual das técnica e qual o campo de aplicação vem sendo mais utilizados. As discussões realizadas nos estudos foram agrupadas em quatro temas: Classificação de Qualidade, Monitoramento e Controle, Aprimoramento de Técnicas e Inspeção e Reconhecimento. Também foram analisadas a precisão alcançada nos experimentos em cada estudo (quando informada) e os métodos aplicados.

A análise crítica dos estudos possibilitou identificar alguns padrões. Foi percebida uma tendência a combinação dos métodos de ML e DL para solucionar particularidades encontradas em alguns problemas. Zhang *et al.* ⁵, Na ⁶ e Wang *et al.* ⁷ são exemplos disso, onde problemas específicos são resolvidos com a técnica que melhor se encaixa e os resultados são utilizados em conjunto.

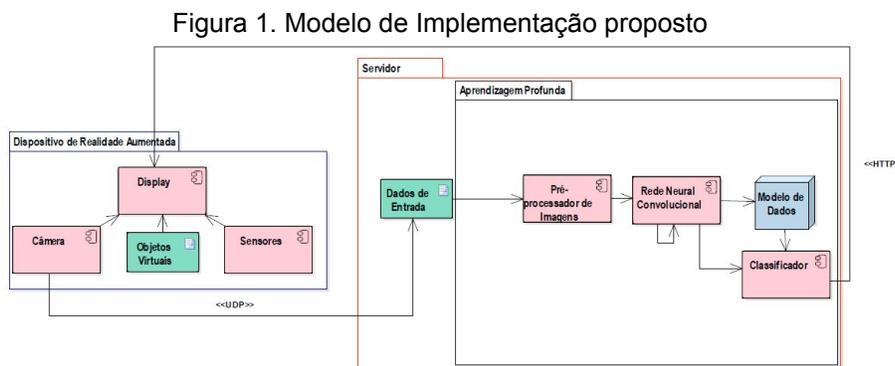
Os resultados da pesquisa demonstraram soluções com taxas de precisão de até 99,9% ⁸, entretanto, ainda são apontadas como não confiáveis ¹. Um possível aspecto para tal suspeita se dá ao fato de que tais soluções não são aplicáveis ou escaláveis, em sua maioria funcionam somente no ambiente ao qual foram projetadas inicialmente, impossibilitando que outra pessoa as utilize em outro ambiente, e caso sejam aplicadas, a precisão obtida pode ser diferente da encontrada pelo criador.

Alguns trabalhos buscam melhorar aspectos de técnicas, visando otimizar o processo, de forma que técnicas já existentes se tornem mais aderentes a ambientes industriais, permitindo um uso mais confiável e de maneira que o processo já existente não seja impactado negativamente em algum aspecto, como o Zhang *et al.* ⁹ que reduz o tempo de processamento e Ahmad *et al.* ¹⁰ que aperfeiçoa uma técnica para que uma precisão maior seja obtida.

Nas demais categorias listadas durante a revisão, cada autor utiliza de combinações entre técnicas, filtros e outros artifícios para solucionar seus obstáculos. Para esta pesquisa, foram observadas soluções e selecionadas as técnicas aplicadas para minimizar os distúrbios gerados pelas variáveis criadas em ambientes industriais. Neste ponto da pesquisa, foram selecionadas as variáveis: oclusão, iluminação deficiente e modificações constantes do ambiente.

3.2 Modelo Proposto

Inicialmente são criados dois grandes módulos, o módulo Servidor e o módulo Dispositivo de Realidade Aumentada. A Figura 1 demonstra o modelo proposto.



3.2.1 Módulo Servidor

Representa um computador robusto, responsável por manter o serviço de aprendizado de máquina (AM), neste caso DL, em funcionamento e realizará a comunicação entre este serviço e o Dispositivo de RA. O serviço de AM não estará alocado diretamente no Dispositivo de RA, pois os dispositivos atuais de RA não dispõem de recursos de hardware robustos o suficiente para uma rede neural grande, ou seja, com muitas camadas e classes, e levando em conta como a quantidade de componentes que a rede poderá abrigar é imensurável, foi optado por mantê-la em um servidor.

3.2.2 Módulo Dispositivo de Realidade Aumentada

Este módulo detém componentes comuns dentre a maioria dos dispositivos encontrados no estado da arte. Neste exemplo, como o serviço de AM realizará classificação de imagens, a comunicação do Dispositivo de Realidade Aumentada com o Servidor parte da câmera e funcionará através do protocolo *User Datagram Protocol* (UDP) para que as imagens sejam enviadas de forma ágil e receberá as respostas por meio do protocolo *HypeText Transfer Protocol* (HTTP). No caso de outra solução de AM, o envio de dados poderia partir do próprio dispositivo ou dos sensores, utilizando outro protocolo. A caixa nomeada Processamento de Imagens será preenchida por DA, onde filtros e transformações serão aplicados nas imagens.

3.3 Teste do potencial do Modelo - SubMódulo Deep Learning

A partir dos *Frameworks* de redes neurais selecionados, serão executados treinos e seus respectivos resultados serão comparados. Para realizar os treinos, as imagens capturadas de componentes industriais serão transformadas utilizando DA, de forma que para cada imagem serão geradas mais cinco, totalizando 600 imagens por componente e um dataset com 1800. A quantidade de imagens dedicada a cada etapa foi selecionada em 80% para a etapa de treino, 10% para validação e 10% para testes. Esta divisão foi realizada de forma aleatória para que a dimensionalidade dos pacotes fosse grande e que durante o treino não houvesse algum tipo de tendência.

Um algoritmo foi desenvolvido para realizar toda estas configurações de forma automática no *dataset* e permitindo que vários testes fossem executados de forma ágil. Até esta fase do projeto, os testes foram executados no Caffe, obtendo uma precisão média de 80% de acurácia. O tempo de consulta obtido foi inferior a um segundo. Alguns treinos foram realizados a fim de definir melhores valores para os parâmetros, com um tempo médio de 30 minutos.

4. CONSIDERAÇÕES FINAIS

Após a leitura e análise dos trabalhos identificados durante a revisão da literatura, foi observado que existem diversas soluções para a Indústria que utilizam artifícios computacionais para auxiliar em processos, sejam estes de produção, segurança entre outros. É percebida uma alta precisão, entretanto tais soluções são aplicadas somente em um único ambiente, sem conseguir funcionar bem em outros. Também é percebido que a maior parte das soluções utiliza combinações entre técnicas para solucionar alguma problemática com maior eficiência. Como proposta de solução para este problema, é necessário que os autores explorem mais ambientes industriais e consigam criar soluções aplicáveis a ambientes distintos sem que a acurácia seja afetada.

As transformações aplicadas nas imagens e a criação de algoritmos que tornam esta ação e a montagem do *dataset* automáticas reduziram significativamente o tempo de preparo dos testes e beneficiou os resultados obtidos. Realizada a análise dos resultados através treino executado no Caffe, foram percebidos valores satisfatórios para precisão e tempo de resposta para classificação. Ainda são necessários treinos em *TensorFlow* e *Torch* para que os resultados sejam comparados e o que se obtiver melhores valores seja definido e aplicado no modelo que será proposto

Através de um estudo sobre o estado da arte em soluções de RA e ML, foi proposto um diagrama de implementação, que visa padronizar a distribuição lógica e física dos componentes, permitindo que um dispositivo de realidade aumentada possa se comunicar com um serviço de DL de maneira confiável e eficaz. Como pesquisas futuras, é sugerido que tal arquitetura seja implantada e que sua eficácia seja testada.

Agradecimentos

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

5. REFERÊNCIAS

- ¹ AZUMA, R. BAILLOT, Y. BEHRINGER, R. FEINER, S. JULIER, S. MACINTYRE, B. **Recent Advances in Augmented Reality**. IEEE Computer Society, 2001.
- ² HARMAYER, K., & DOSWELL, J. **Extending the ‘Serious Game’ Boundary: Virtual Instructors in Mobile Mixed Reality Learning Games**. Situated Play, Proceedings of DiGRA 2007 Conference, 2007.
- ³ YI, B. LEE, J. G., D. & RIM, H. C. **The Effects of Feature Optimization on High-Dimensional Essay Data Mathematical Problems in Engineering**. Mathematical Problems in Engineering, 2015.
- ⁴ CHEN, Chih-Ming, TSAI, Yen-Nung. **Interactive augmented reality system for enhancing library instruction in elementar schools**. Computers & Education, 2012.
- ⁵ WANG, Jason, PEREZ, Luis. **The Effectiveness of Data Augmentation in Image Classification using Deep Learning**. CoRR, 2017.
- ⁶ ZHANG, Z. JIANG, T. LI, S. et al. **Automated feature learning for nonlinear process monitoring – An approach using stacked denoising autoencoder and k-nearest neighbor rule**. Journal of Proces Control, 2018.
- ⁷ NA, Jonggeol. JEON, Kyeongwoo. LEE, Won Bo. **Toxic gas release modeling for real-time anayis using variational autoencoder with convolutional neural networks**. Chemical Engineering Science, 2018.
- ⁸ WANG, Zhenyu. SOUNG, Chufeng. CHEN, Tao. **Deep learning based monitoring of furnance combustion state and measurement of heat release rate**. Energy, 2018.
- ⁹ ZHANG, Zehan. JIANG, Teng. LI, Shuanghong. YANG, Yupu. **Automated feature learning for nonlinear process monitoring - An approach using stacked denoising autoencoder and k-neares neighbor rule**. Journal of Process Control, 2018.
- ¹⁰ AHMAD, J. MUHAMMAD, K. LLORET, J. BAIK, S. W. **Efficient Conversion of Deep Features to Compact Binary Codes Using Fourier Decomposition for Multimedia Big Data**. IEEE Transactions on Industrial Informatics, 2018.
- ¹¹ AKGUL, Omer. PENEKIL, H. Ibrahim. GENC, Yakup. **Appling Deep Learning in Augmented Reality Tracking**. International Conference on Signal-Image Technology & Internet-Based Systems, 2016.