



Detecção e classificação de objetos em linhas de distribuição de energia utilizando a CNN YOLOv3

Milena Lucas dos Santos
Fundação Parque Tecnológico Itaipu
Universidade Estadual do Oeste do
Paraná
Foz do Iguaçu, Brasil
Email: mii.santos342@gmail.com

Valéria Nines dos Santos
Fundação Parque Tecnológico Itaipu
Foz do Iguaçu, Brasil
Email: valeria.valsnfw@gmail.com

Claudio Roberto Marquetto Mauricio
Fundação Parque Tecnológico Itaipu
Universidade Estadual do Oeste do
Paraná
Foz do Iguaçu, Brasil
Email: crmmauricio@gmail.com

Fabiana Frata Furlan Peres
Universidade Estadual do Oeste do
Paraná
Foz do Iguaçu, Brasil
Email: fabiana.peres@unioeste.br

Abstract—Electric power is essential for all areas of production. Therefore, the condition of the power line distribution equipment is essential for energy quality. However, the power grids extend for miles and even stretches of difficult access, compromising the search for damaged equipment. This article presents a solution for detection and classification of objects in power distribution lines using Convolutional Neural Networks. In the first stage of the project, CNN was trained to detect and classify four types of objects that are part of the power grid. YOLOv3 was the artificial neural network selected to detect and classify the objects. To compose the test and training image bank, a drone was used and image collections were made in 10 different locations. After the training, CNN achieved a IoU of 60.38%.

Keywords— YOLOv3; Convolutional Neural Networks; Energy Distribution.

Resumo —A energia elétrica é algo essencial para todas as áreas de produção. Sendo assim, a condição dos equipamentos é essencial para a distribuição de uma energia de qualidade. Contudo, as redes elétricas se estendem por quilômetros e ainda por trechos de difícil acesso, comprometendo encontrar equipamentos danificados. Neste artigo, apresenta-se uma solução de detecção e classificação de objetos das linhas de distribuição de energia utilizando Redes Neurais Convolucionais. Na primeira etapa do projeto, a CNN foi treinada para detectar e classificar quatro tipos de objetos que fazem parte da rede elétrica. A rede neural artificial selecionada para detectar e classificar os objetos foi a YOLOv3. Para compor o banco de imagens de teste e treinamento, foi utilizado um drone e efetuadas coletas de imagens em 10 locais distintos. Após o treinamento, a CNN alcançou um IoU de 60,38%.

Palavras-chave— YOLOv3; Redes Neurais Convolucionais; Distribuição de energia.

I. INTRODUÇÃO

A distribuição de energia é necessária em todas as áreas, tanto na área rural como na urbana. As redes de distribuição de energia abastecem consumidores industriais, comerciais e residenciais. No Brasil são mais de 140.000 km de linhas de transmissão [1]. Espera-se que estas linhas de transmissão estejam operantes e em bom estado de manutenção para

garantir energia aos consumidores, porém é custoso devido a extensão das linhas de transmissão, pelo difícil acesso a certos trechos e pelos riscos de segurança. Desse modo é muito importante a investigação de novas abordagens explorando tecnologias que possam contribuir com o processo de inspeção, que facilite o acesso aos locais de forma segura. Baseado neste contexto, surgiu um projeto com objetivo de desenvolver uma solução para inspeção dos componentes da rede elétrica de forma automatizada, utilizando drones e Rede Neural Artificial. Deseja-se que a inspeção seja realizada considerando um modelo treinado para detecção e classificação de componentes da rede elétrica. Inclusive deseja-se que o sistema seja capaz de identificar os componentes que apresentam algum defeito. A ideia inicial era a detecção em tempo real utilizando a internet 4G, pois na teoria e na realidade de outros países essa tecnologia tem o suporte necessário, entretanto a capacidade de *streaming* da tecnologia 4G na realidade do país não suporta o envio em tempo real, sendo a latência do *streaming* em torno de 30 segundos, o que não é satisfatório para o contexto atual. Uma saída é a captura das imagens pelo drone e posteriormente feita a análise das imagens adquiridas. Este artigo descreve os resultados alcançados nas etapas iniciais do projeto. Desse modo é descrito os detalhes do treinamento de uma rede CNN para a detecção e classificação de quatro tipos de objetos presentes nas linhas de distribuição de energia.

II. CONCEITOS GERAIS

A. Redes Neurais Artificiais

Rede Neural Artificial (RNA) expõe que a máquina tem a capacidade de aprender sem interferência externa e tem a habilidade de generalizar a informação aprendida [2]. Uma RNA pode ser compreendida ou relacionada com o cérebro humano; direcionada no processamento de informações computacionais, ela busca simular a maneira que o cérebro realiza alguma tarefa, aprendendo através de exemplos. A forma mais simples de RNA é a *perceptron*. Ela possui uma única camada contendo uma unidade neural. As RNAs com várias camadas são chamadas de *multilayer perceptron*. Elas



possuem uma camada de entrada, outra de saída e uma ou mais camadas ocultas. As Redes Neurais Convolucionais são variações das redes *multilayer perceptron* [3]. As Redes Neurais Convolucionais são métodos que fazem parte do subconjunto de algoritmos baseados em Deep Learning. Esses algoritmos extraem intuitivamente características de alto nível essenciais para a classificar [4]. As CNNs se destacam pela facilidade de treinamento e por sua capacidade de generalização; ela é capaz de extrair automaticamente as características de objetos [4]. Recentemente diferentes pesquisas demonstraram a capacidade das CNNs para detecção e classificação de objetos, como por exemplo, a CNN SSD, RetinaNet, entre outras.

A CNN SSD (Single-Shot MultiBox Detector) possui uma arquitetura baseada em *Region Proposal Networks* (RPN), anchor boxes e representação multi-escala [5]. A CNN SSD ajusta as previsões dos *feature maps* de variadas resoluções para trabalhar com objetos de diversos tamanhos. Ela gera um número x de anchor boxes de tamanhos diferentes, que formam um conjunto de *bounding boxes*, inicialmente para a detecção de objetos. Durante o treinamento essas *bounding boxes* são ajustadas para estarem de acordo com a posição do objeto. No final, o algoritmo *Non-Maximum Suppression* reduz as *boxes* para a detecção [5] [3].

A CNN RetinaNet [6] é outra rede de detecção. Esta rede possui um estágio que procura alcançar o desempenho das redes de dois estágios (Faster R-CNN), mantendo a complexidade computacional reduzida. Uma das dificuldades da CNN SSD, que também é de um estágio, está em lidar com a detecção densa de objetos, pois o processamento é feito em uma etapa só. Entretanto, a RetinaNet possui duas sub-redes convolucionais: uma para a classificação das detecções e a outra para a regressão das *bounding boxes* ligadas ao objeto reconhecido. A principal contribuição da CNN RetinaNet foi o desenvolvimento da função de Perda Focal, resolvendo o problema de desbalanceamento de classes [7].

B. Métricas de desempenho

Os algoritmos de detecção de objetos são avaliados através de seu desempenho em relação aos resultados de treinamento. As métricas utilizadas para a avaliação do desempenho do treinamento são *Intersection over Union* (IoU) e *Mean Average Precision* (mAP). *Intersection over Union* é uma métrica que define quanto a predição da localização de um objeto na imagem coincide com sua verdadeira localização [4]. Essa métrica consiste basicamente em dividir a interseção pela união de dois *bounding boxes*. Quanto maior o valor do IoU, mais preciso é a seleção de um objeto na imagem. O mAP é a média sobre o *Average Precision* (AP). Basicamente, o mAP corresponde a acurácia individual de cada classe [4].

III. TRABALHOS RELACIONADOS

Desde a geração até o uso de fato da energia, ocorre um trabalho complexo. A inspeção e a manutenção das redes de transmissão e distribuição de energia elétrica compõem este trabalho. A degradação dos componentes da rede de transmissão provém de ações de agentes ambientais como a temperatura, umidade, poluição e radiação ultravioleta, mas também pode provir de vandalismo, entre outros [8].

Na literatura, há vários estudos e métodos para a detecção de componentes em linhas de distribuição [9] [10] [11], sendo que alguns exploram técnicas de inteligência artificial e processamento digital de imagens, como também radiografias e tomografias. O trabalho apresentado em [9], notabiliza uma maneira de detectar e classificar os defeitos de isoladores poliméricos através de processamento digital de imagens radiográficas. São realizados raios-x em isoladores poliméricos a fim de montar um *dataset* de imagens. Em seguida é realizado um pré-processamento nas imagens, onde a área de interesse é extraída, por fim classificados componente a componente no qual só então são usadas para treinamento de uma RNA.

A termografia infravermelha aplicada no sistema elétrico pode ser muito efetiva para prevenção de falhas e ausência de produção, o trabalho de [10], propõe o uso de algoritmos de RNAs para indicar possíveis falhas em transformadores, através das temperaturas marcadas nas imagens termográficas. A abordagem apresentada em [11] tem o foco em manter a qualidade de energia nas redes de distribuição. O método proposto também utiliza processamento digital de imagens, como o pré-processamento e RNA para a classificação dos eventos, ou seja, ter o monitoramento em barras de subestações, através das oscilografias. Os algoritmos utilizados foram o MultiLayer Perceptron e Support Vector Machines (SVM). A pesquisa de [8] utilizou a radiação ultravioleta e a termografia como parâmetros para determinar o estado da degradação de isoladores poliméricos, auxiliando na tomada de decisões para intervenção e manutenção do componente, utilizando processamento digital de imagens e RNA.

Os estudos apresentados trazem soluções com RNAs assim como este trabalho. No entanto, a diferença deste trabalho para os demais apresentados e principalmente por buscar a detecção e classificação de componentes nas linhas de distribuição de energia elétrica utilizando algoritmos mais avançados.

IV. MATERIAIS E MÉTODOS

Para criar o modelo capaz de detectar e classificar os componentes, foi utilizado a RNA YOLOv3 e o *framework* Darknet. O uso da YOLOv3 se deve por ela se destacar pela sua acurácia e velocidade para treinamento. A YOLOv3 foi proposta por Redmon [12] em 2018. O algoritmo é baseado em Redes Neurais Convolucionais e se propõe a detectar objetos em imagens, em tempo real. Ela possui capacidade de generalização e sua arquitetura contém uma rede convolucional que prevê várias regiões da imagem ao mesmo tempo. Na camada de entrada, a imagem é subdividida em uma grade de tamanho $N \times N$, onde para cada célula da grade é predito um objeto. Para cada célula ela prediz um número limitado de *bounding boxes* entregando um valor de confiança e a probabilidade do objeto ser de uma classe específica. A YOLOv3 [12] possui 53 camadas convolucionais. Para melhorar a detecção de objetos pequenos, é realizado um conjunto de operações de convolução usando filtros 1×1 e 3×3 [13].

O Darknet é um *framework* de código aberto no qual permite ao pesquisador customizar parâmetros de acordo com sua capacidade de processamento [14]. Ela não possui interface porém sua utilização é simples. É possível visualizar o treinamento por iterações: a cada conjunto de imagens que entra na RNA, é exibido um *score* de

informações da RNA em relação à aquele conjunto. Além disso o Darknet permite uma subdivisão sobre o *batch*. Por exemplo, um *batch* de 64 são 64 imagens entrando ao mesmo tempo na RNA para extração de características comuns, e esse valor é multiplicado por cada camada de acordo com suas especificações. O Darknet permite passar um parâmetro para dividir esse *batch* e assim diminuir o custo de processamento; ao final ele realiza uma operação para juntar as partes e retornar o valor total do *batch* inicial.

Para a execução do treinamento da rede e para realizar os testes de detecção e classificação dos componentes, foram coletadas imagens dos objetos em 10 locais diferentes da zona urbana de uma cidade. Foram obtidas ao todo 675 imagens contendo vários objetos de interesse em cada imagem, das quais 615 imagens foram utilizadas para o treinamento e testes e 50 imagens para a validação da CNN. Foram escolhidos quatro tipos de objetos para detectar e classificar: Isolador, chave fusível, transformador e para-raio. A escolha se deve pelo número de componentes no *dataset* serem relevantes para o treinamento. Conforme a limitação do *dataset*, alguns objetos não foram inseridos para treinamento. A figura 1 mostra um exemplo de cada objeto, porém há outros modelos e cores dos objetos inseridos no *dataset*.



Fig. 1. Componentes escolhidos para a execução do projeto.

V. RESULTADOS E DISCUSSÃO

A primeira etapa do processo de detecção e classificação da CNN é o pré treinamento, na qual a ferramenta BBox-Label-Tool é usada para demarcar os objetos em suas respectivas classes, gerando as anotações necessárias. Essa etapa engloba todo o processo anterior ao treinamento. Posteriormente, as 615 imagens foram divididas em 75% para o treinamento e 25% para testes. Essa divisão foi feita de forma aleatória, o que pode ter interferido significativamente nos resultados, uma vez que a quantidade de cada componente ficou desproporcional, gerando um desbalanceamento de classes. A Tabela 1 apresenta o número dos componentes na fase de treinamento, testes e validação.

TABELA I
NÚMERO DE COMPONENTES DAS FASES DE TREINAMENTO, TESTE E VALIDAÇÃO.

Classes	Treino	Teste	Validação
Transformador	170	150	21
Isolador	1364	288	152
Chave Fusível	549	172	56
Para-raio	424	214	37

Para o treinamento, foram utilizados os pesos sinápticos já pré-treinados, sendo o modelo *darknet53*. Os seguintes parâmetros do *darknet53* foram alterados com o objetivo de melhorar a detecção e classificação da CNN: o *Batch*, o *Subdivisions*, o *Width*, o *Height*, o *Filters* e o *Random*. O parâmetro *Batch* define o número de imagens no lote que será aplicado a normalização; *Subdivisions* permite dividir o valor do *Batch* para reduzir o número de imagens no lote; *Width* e *Height*, são usados para definir a dimensão das imagens de treinamento; *Filters* corresponde ao filtro da última camada de convolução; esse valor é calculado pela fórmula: $(classes+5)*3$; por fim o parâmetro *Random* se alterado para 1, redimensiona as imagens a cada 10 iterações [15].

TABELA II
RESULTADOS OBTIDOS COM O PRIMEIRO E SEGUNDO TREINAMENTO..

Classes	AP T1	AP V1	AP T2	AP V2
Transformador	80,22%	62,5%	74,87%	73,06%
Isolador	27,45%	32,66%	29,49%	38,12%
Chave Fusível	34,0%	18,02%	31,73%	20,85%
Para-raio	11,29%	8,71%	13,22%	3,6%

O primeiro treinamento resultou em um mAP de 17,41% e média do IoU de 51,99%. No segundo treinamento o valor resultante do mAP foi de 34% e média do IoU de 60,38%. Observa-se que o segundo treinamento gerou o melhor resultado. A Tabela II mostra a precisão média (AP) de cada classe para os dois treinamentos. As colunas "AP T1" e "AP V1" apresentam a AP obtida no teste e na validação do primeiro treinamento e as colunas "AP T2" e "AP V2" a AP obtida no teste e na validação do segundo treinamento. A Figura 2 e a Figura 3 mostram os resultados obtidos. Pode-se observar que os *bounding boxes* coincidem aproximadamente com sua verdadeira localização. Na Figura 2 a CNN detecta o transformador e os três isoladores com uma confiança de 100%, mas a CNN classifica o objeto para-raio como chave fusível. Na Figura 3, o componente para-raio é detectado duas vezes, entretanto o componente mais distante do transformador é uma chave-fusível. Este erro em ambos treinamentos é conhecido como falso positivo. Vários fatores podem ter influenciado o mAP resultante, como a distância usada para obter as imagens, que pode ter dificultado o aprendizado da CNN; os diferentes modelos dos componentes; os parâmetros utilizados para o treinamento e/ou a divisão das imagens para treinamento e teste. Ou seja, vários fatores podem ter influenciado o resultado final. É importante ressaltar que os dois

treinamentos foram realizados com o mesmo número de imagens, pelo fato de não terem sido realizadas mais coletas de imagens, no entanto o ideal seria aumentar o número de imagens a cada treinamento.

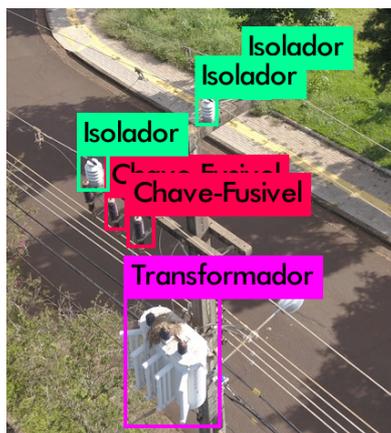


Fig. 2. Exemplo de uma imagem do segundo treinamento.

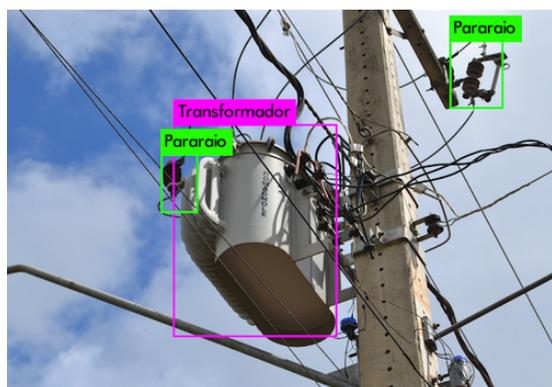


Fig. 3. Exemplo de uma imagem do segundo treinamento.

VI. CONCLUSÕES

Conforme os resultados e testes mostrados, a classe transformador é o componente com maior detecção pela CNN YOLOv3. Analisando o número de componentes, a classe transformador é o objeto com menor número de componentes para o treinamento, mas pelo fato do objeto ser maior, a CNN tem mais facilidade de reconhecer. Os demais componentes são objetos menores, dificultando o aprendizado da CNN.

Importante ressaltar que a YOLOv3 melhorou em relação a detecção de objetos pequenos, mas ainda é um desafio para CNNs de um estágio. A diferença da precisão média entre a classe transformador e a classe isolador (segunda classe mais detectada, segundo AP) é muito alta, diferença de 35%.

Um dos fatores importantes de serem citados que interferem na detecção e classificação dos componentes são o modelo, a cor e o ângulo variados nas imagens. Por exemplo, no banco de imagens contém figuras com diferentes modelos de isolador, sendo poliméricos, de vidro ou porcelana. Da mesma forma, acontece com os demais componentes das linhas de distribuição de energia.

Sabendo que o valor de IoU acima de 50% é considerado uma boa predição, o valor resultante do IoU é satisfatório. Acredita-se que a ampliação do *dataset* pode influenciar positivamente nos próximos treinamentos realizados com a CNN.

Os próximos passos deste trabalho, serão melhorar a CNN de acordo com o conhecimento e experiência que foram adquiridos e classificar os componentes como regulares ou com defeitos. Para isso o *dataset* será ampliado e será aplicado novos métodos de detecção de objetos nas imagens.

AGRADECIMENTOS

Os autores gostariam de agradecer a Fundação Parque Tecnológico Itaipu (FPTI), ao Centro Internacional de Hidroinformática (CIH) por propor e financiar o projeto, ao Centro Latino Americano de Tecnologias Abertas (CELTAB) pela disponibilização de material e conhecimento, também a Itaipu pelo incentivo ao projeto.

REFERÊNCIAS

- [1] ONS, "Resultados da operação histórico da operação: Extensão das linhas de transmissão," <http://www.ons.org.br/Paginas/resultados-da-operacao/historico-da-operacao/extensaolinhastransmissao.aspx>, 2020, online; acessado em 28/07/2020.
- [2] H. Pedrini and W. R. Schwartz, *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*, 1st ed. Sao Paulo, SP, Brasil: Cengage Learning, Agosto 2007.
- [3] I. M. Soares, "Uma abordagem bottom-up completa para reconhecimento de atividades humanas em imagens através da pose estimada com redes convolucionais," *M.S. thesis*, UFG, Goiânia, Brasil, 2019. [Online]. Available: <http://repositorio.bc.ufg.br/tede/handle/tede/10191>.
- [4] B. H. P. Marques, "Avaliação de algoritmos baseados em deep learning para localizar placas veiculares brasileiras em ambientes complexos." Pernambuco, Brasil, 2019. [Online]. Available: <http://hdl.handle.net/123456789/833>.
- [5] M. Mazzetto, "Detecção e classificação de múltiplos componentes em linha de montagem automotiva usando deep learning," *M.S. thesis*, UTFPR, Pato Branco, PR, Brasil.
- [6] T.-Y. LIN, "Focal loss for dense object detection," *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.
- [7] R. A. P. Martins, "Aplicação de redes convolucionais profundas para detecção de massas em mamografias," Florianópolis, SC, Brasil.
- [8] G. Ribeiro, "Aprimoramento da classificação de isoladores poliméricos por medições termográficas e radiação uv usando processamento de imagens e rna," *M.S. thesis*, UFCG, Campina, PB, Brasil, 2017. [Online]. Available: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/494>.
- [9] W. C. Godoi, R. R. Silva, V. Swinka-Filho, and T. Loddí, "Detecção de defeitos em isoladores poliméricos por meio da radiografia digital e reconhecimento de padrões," *Espaço Energia*, Abr. 2005.
- [10] G. Santos, "Termografia e inteligência artificial na

- detecção de falhas em transformadores,” *M.S. thesis*, UFPE, Recife, PE, Brasil.
- [11] F. Vitor, “Classificação de eventos em redes de distribuição de energia utilizando transformada wavelet e modelos neurais autônomo,” *Br. Society Neural Net.*, vol. 8, pp. 93–99, Jan. 2010.
- [12] J.Redmon and A. Farhadi, “Yolov3: An incremental improvement.” *ArXiv*, abr 2018. [Online]. Available: arXiv:1804.02767.
- [13] M. H. Scalabrin, “Aplicação de redes neurais artificiais profundas na detecção de placas de pare,” Francisco Beltrão, PR, Brasil, 2019. [Online]. Available: <http://repositorio.roca.utfpr.edu.br/jspui/handle/1/13452>.
- [14] J. Redmon, “Darknet: Open source neural networks in c,” <http://pjreddie.com/darknet/>, 2013–2016.
- [15] V. N. Santos, “Reconhecimento de objetos em uma cena utilizando redes neurais convolucionais.” Medianeira, Brasil, 2018.