

Reconhecimento de Objetos em Linhas de Distribuição de Energia Utilizando IBM Cloud

Raianny Proença de C. de Oliveira Valéria Nunes dos Santos Claudio Roberto Marquetto Mauricio
Bolsista Fundação Parque Tecnológico Itaipu Fundação Parque Tecnológico Itaipu Universidade Estadual do Oeste do Paraná
Universidade Estadual do Oeste do Paraná Foz do Iguaçu, Paraná Foz do Iguaçu, Paraná
Ciência da Computação Email: valeria.valsnfw@gmail.com Email:claudio.mauricio@unioeste.br
Foz do Iguaçu, Paraná
Email: raiannyproenca1@gmail.com

Abstract—O reconhecimento visual de objetos em uma imagem e a habilidade de classificá-los é uma das principais áreas de pesquisa utilizando a inteligência artificial. Atualmente, existem diversos *frameworks* disponíveis para auxiliar nesta tarefa. IBM Cloud é um desses *frameworks* no qual oferece um amplo catálogo de ferramentas na área de inteligência artificial, uma delas é a IBM Watson Visual Recognition. Foi criado um modelo através dessa API que utiliza algoritmos de *Deep Learning* para analisar imagens. O objetivo principal do modelo é o reconhecimento de oito tipos de objetos distintos em linhas de distribuição de energia. Todas as imagens utilizadas no experimento foram coletadas em aproximadamente 10 regiões diferentes da cidade. Os resultados obtidos no modelo criado no artigo foram comparados com o classificador geral da IBM. A taxa de acerto foi de pelo menos um componente reconhecido corretamente em 33 das 40 imagens utilizadas para testes, ou seja, classificou corretamente 82,5% do conjunto de imagens testes, apresentando uma acurácia acima de 50% para o objeto reconhecido. Os serviços utilizados para todo o experimento são do plano *lite*, gratuito. Logo, este artigo demonstra a capacidade da ferramenta mesmo as limitações dessa versão.

Keywords — inteligência artificial, reconhecimento visual, IBM Cloud.

I. INTRODUÇÃO

A inteligência artificial apresenta como uma de suas definições, que foi produzida para imitar o processamento feito pelo cérebro humano, ou seja, se concentra em dar ao computador a habilidade de aprender, compreender, e pensar de forma semelhante aos humanos. Interessa a máquina analisar constantemente os dados, dividindo-os em várias camadas de tal forma que torna desnecessário a intervenção humana de ter que proporcionar explicitamente todo o conteúdo necessário para a máquina [1]. Por ser uma área tão ampla, a inteligência artificial pode ser dividida em partes ou em camadas, como *Machine Learning* e *Deep Learning*. Os dois conceitos foram usados neste artigo.

Machine Learning utiliza algoritmos para extrair informações de dados e representá-los através de algum modelo, possibilitando o uso deste modelo para formar previsões a partir de outros dados [2]. Este artigo é um bom exemplo prático dessa técnica, onde foram fornecidos para a máquina como dados de entrada, centenas de imagens separadas de acordo com o seu conteúdo para que

posteriormente a mesma seja capaz de reconhecer e classificar novas imagens que contenham o conteúdo amostrado. Dentre os vários serviços fornecidos na IBM Cloud na área de Inteligência artificial, a API (*Application Program Interface*) escolhida para o desenvolvimento deste artigo foi a *Watson Visual Recognition*, que permite marcar e classificar imagens usando *Machine Learning*. Os serviços desta API utilizam algoritmos de *Deep Learning* para analisar imagens e classificá-las. Foi criado um modelo na *IBM Watson Studio* com classes customizadas e os resultados foram comparados com a classificação feita pelo modelo geral da IBM fornecido pelo *Watson Visual Recognition*.

II. CONCEITOS GERAIS

A. Aprendizagem Supervisionada

Esta aprendizagem utiliza como entrada dados rotulados, sendo um conjunto de dado de entrada e possíveis saídas, em seguida define um modelo. É uma técnica comumente utilizada em aplicações nas quais dados históricos podem prever possíveis acontecimentos futuros e utiliza padrões para prever os valores dos rótulos, também conhecidos como *label*, em dados adicionais não utilizados, através dos métodos de classificação e regressão [3].

B. Classificação

Tem como alvo variáveis qualitativas. É o processo de tomar algum tipo de dado como entrada e atribuir um rótulo a ele. Geralmente são usados quando as previsões de naturezas distintas e claras, por exemplo, o paciente está ou não doente. Em casos como este, onde há apenas duas opções, é denominado de classificação binária ou de duas classes. Quando há mais opções é então nomeado como classificação múltipla, pois contém múltiplas classes [3].

C. Base de Dados

O *dataset* foi fornecido para o projeto de reconhecimento de objetos em linhas de distribuição de energia através de uma parceria entre o Centro Internacional de Hidroinformática (CIH) e o Centro Latino-americano de tecnologias abertas (CELTAB), para desenvolvimento de soluções tecnológicas para gestão territorial aplicada à água e energia bem como o uso de inteligência artificial.

III. MATERIAIS E MÉTODOS

A plataforma computacional utilizada para a execução das tarefas foi a IBM Cloud, que oferece um conjunto de serviços em seu catálogo em diversas áreas, entre elas, a área de Inteligência Artificial [4]. Um dos serviços ofertados é o *Watson Visual Recognition* no qual utiliza algoritmos de Deep Learning para analisar imagens e classificá-las [5]. Oferece um conjunto de classes prontas para o uso, ou a opção de criar o próprio classificador, customizando e treinando estes classificadores com uma coleção própria de imagens [3]. As classes que estão prontas para o uso são as classes que reconhecem: rostos, comidas, conteúdos explícitos em geral. O modelo geral da IBM foi utilizado como base de comparação em relação ao modelo treinado desse projeto. Entre os vários planos ofertados pela empresa, o plano utilizado neste experimento foi o Lite, que é a versão gratuita para estudante e que permite o treino de até 2 modelos por conta e é possível treinar 1000 imagens por mês apenas [5].

Para este projeto foram realizadas 6 coletas, totalizando um conjunto com 600 imagens. Destas, 560 foram utilizadas para o treinamento do modelo e 40 compõe o conjunto de testes. Diferente de outras ferramentas de reconhecimento visual, a *API Visual Recognition* não permite que seja feito o reconhecimento do conteúdo da imagem manualmente, demarcando o objeto e classificando-o, usando por exemplo, o conceito de *bounding box*. É necessário que treine as classes de um determinado objeto com imagens dele, de forma que ele seja o principal conteúdo da imagem. Devido a isto, todas as imagens foram recortadas diversas vezes, cada vez destacando um componente de interesse. Isso fez com o número de imagens para treinar o modelo subisse de 560 para 2923 que foram divididas em 8 classes. A Tabela I contém as classes de interesse e o número de imagens utilizadas para treinar cada uma. A Figura 1 demonstra um exemplo de cada um dos componentes na Tabela I.

Tabela I
CLASSES DO MODELO DO ARTIGO

Classes	Número de Imagens
Banco de Capacitor	16
Chave Fusível	381
Espaçador	102
Grampo	493
Isolador	787
Equipamento de Iluminação	459
Para-raios	292
Transformadores	393

Vale ressaltar que nem todas as imagens mantiveram uma boa resolução uma vez que o recorte redimensiona a imagem e faz com que perca qualidade e escopo. Em outros casos a câmera capturou o ângulo oposto ao lado do objeto e isso pode dificultar o reconhecimento do componente. A discrepância de modelos do mesmo objeto também foi um desafio, uma vez que no caso dos pára-raios por exemplo, existe um

grande volume de imagens do mesmo modelo, e uma pequena quantidade de modelos que divergem deste. Uma segunda observação importante sobre os pára-raios é que na maioria das imagens eles estão camuflados possuindo a mesma cor do poste. O banco de capacitor é a classe que possui o menor número de imagens, e suas características físicas lembram a classe de transformadores, logo, em alguns testes foram considerados como tal.

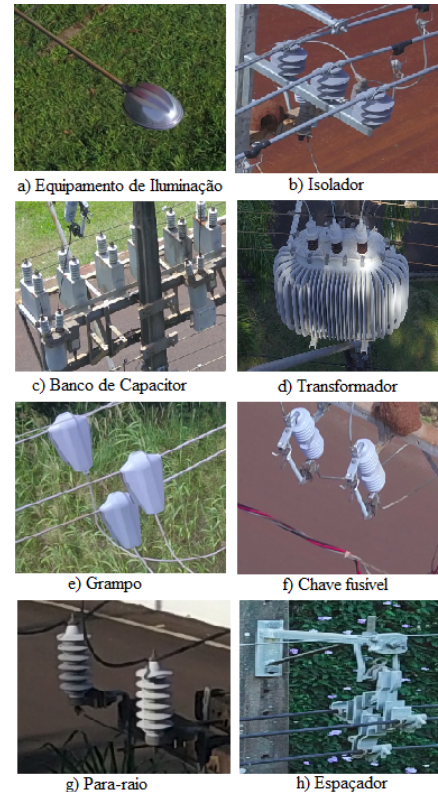


Fig. 1. Exemplo de cada componente.

A classe dos isoladores é a que possui o maior número de imagens e a maior variação também, apresentando vários modelos divergindo bastante de formatos, cores, ângulos e posicionamento. Neste caso por ter muitos modelos e uma grande variação de imagens para cada um, também aparentou interferir no reconhecimento deles.

A classe dos transformadores, assim como a dos isoladores, também possui imagens bem diversificada. A angulação, o modelo e a cor dos transformadores variam bastante e alguns modelos são mais comuns que outros. Notou-se por ser um objeto grande que se destaca na maioria das imagens, o modelo teve uma maior facilidade para reconhecê-lo.

Como dito anteriormente, para treinar as classes do modelo apenas com imagens do objeto de interesse, estas foram recortadas diversas vezes, onde uma imagem que continha todos os objetos forneceu novas imagens para as classes customizadas (que no caso deste artigo são oito). No plano utilizado da plataforma notou-se que uma vez que suas classes e imagens foram treinadas e classificadas não é mais possível removê-las.

IV. EXPERIMENTOS

Para validar o modelo criado, foram testadas as 40 imagens que compõem o conjunto de testes. Abaixo estão alguns exemplos das classificações resultantes utilizando o modelo construído no artigo e modelo geral da IBM. As imagens que apresentam mais de um componente descrito na Tabela I foram marcadas manualmente a fim de facilitar o entendimento do leitor, no entanto vale ressaltar que a ferramenta utilizada não apresenta *bounding box* nos resultados das classificações, como dito anteriormente.



Fig. 2. Imagem original utilizada nos testes dos modelos.

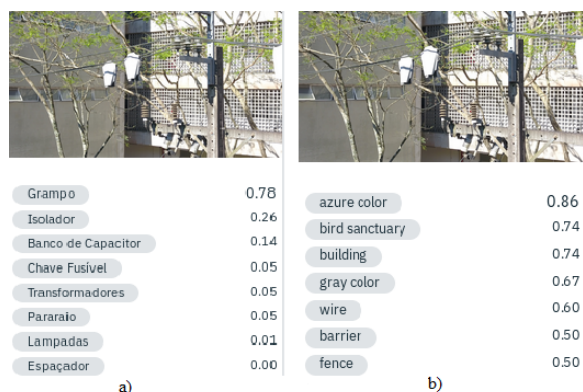


Fig. 3. a) Resultado da classificação do Modelo do Artigo. b) Resultado da classificação do Modelo Geral da IBM.

Na Figura 2 estão presentes os seguintes componentes: grampo, circulado em azul, isoladores em vermelho e para-raios em amarelo e sua classificação em ambos os modelos está exposto na Figura 3. A Figura 3a) apresenta o resultado obtido através do modelo construído neste artigo que reconheceu corretamente os isoladores e os grampos, porém os para-raios estão com uma porcentagem baixa e nota-se que sua identificação realmente está dificultada devido ao fato de estar camuflada com o cenário ao fundo. O terceiro componente que ele afirmar ter na imagem está errado, pois não existe um banco de capacitor na imagem como afirmado pelo modelo. Na Figura 3b) está exposto o resultado do modelo geral da IBM, que reconheceu na imagem a cor azul, um santuário de pássaros, uma construção, a cor cinza entre outros.

A Figura 4 contém apenas um componente: o transformador, e sua classificação pode ser observada na Figura 5, sendo



Fig. 4. Imagem original utilizada nos testes dos modelos.

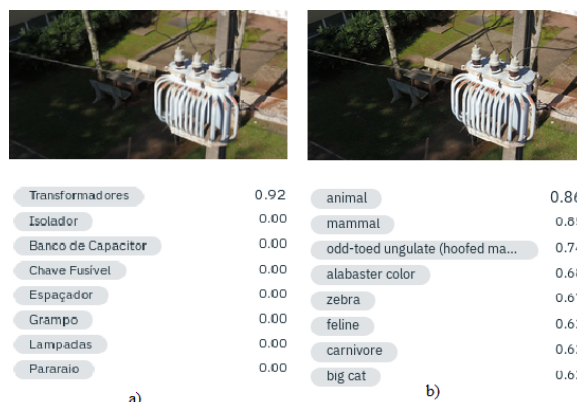


Fig. 5. a) Resultado da classificação do Modelo do Artigo. b) Resultado da classificação do Modelo Geral da IBM.

este um bom exemplo de como o modelo construído neste artigo é bom em classificar um item por vez, pois na imagem o único componente reconhecido foi o transformador e esta informação está correta (Figura 5a)). Já a classificação feita pelo Classificador Geral da IBM (Figura 5b)) afirmou que na imagem há um animal e apontou números altíssimos para este caso.

A Figura 6 é composta por isoladores, destacado na cor amarela, grampos na azul e chave-fusíveis com a caixa delimitadora na cor vermelha. Em sua classificação nota-se que o número de isoladores na imagem realmente é dominante e isto foi bem reconhecido pelo modelo, o mesmo ocorreu para chave fusível, no entanto um espaçador foi reconhecido erroneamente na imagem. Nota-se que não houve reconhecimento dos grampos que também estão na imagem. No entanto, o classificador geral da IBM reconheceu na imagem componentes como antena, dispositivo elétrico, cor cinza entre outros (Figura 7b)). Observa-se que a plataforma mesmo com o número de imagens sendo pequeno em comparação a outros

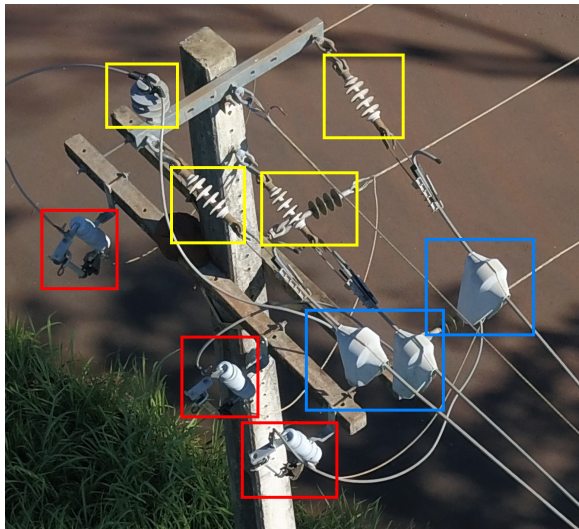


Fig. 6. Imagem original utilizada nos testes dos modelos.

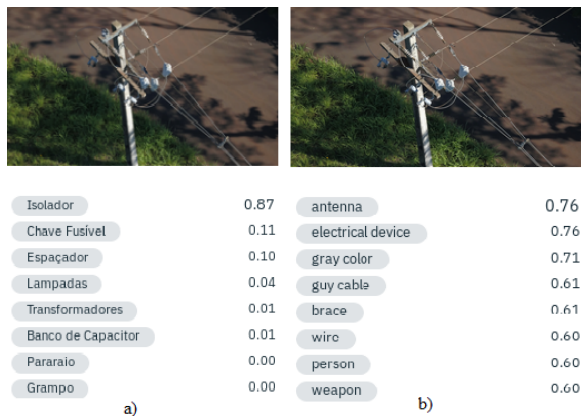


Fig. 7. a) Resultado da classificação do Modelo do Artigo. b) Resultado da classificação do Modelo Geral da IBM.

dataset enormes como MNIST [6], ela teve capacidade de reconhecer os objetos em linhas de distribuição de energia, apresentando uma acurácia da possibilidade de o item estar presente na imagem, sendo alguns resultados melhores que outros.

V. CONCLUSÃO

O experimento apenas utilizou os resultados da classificação do modelo geral da IBM para comparações porque foi uma das formas encontradas para demonstrar a eficiência do modelo criado. É importante observar que o modelo geral da IBM, tem a capacidade de reconhecer o conteúdo geral das imagens e não características específicas.

O reconhecimento feito pelo modelo construído neste artigo está diretamente ligado às imagens utilizadas para o treino do mesmo, no qual depende do número e da qualidade das imagens e que o *dataset* em geral expunha o máximo possível de diversificações do componente, no qual pode apresentar diferentes posições, ângulos, tamanho e cores, uma vez que

como o exemplo dos transformadores, nem sempre terão o mesmo modelo, cor ou tamanho.

Um ponto relevante é que a ferramenta apenas mostra a acurácia do que ela acredita estar na imagem, não evidenciando a detecção do componente com *bounding box*, isso impede de afirmar com absoluta certeza que o isolador reconhecido por ela, por exemplo, realmente seja o isolador e não outro componente parecido.

Outro fato importante a se considerar é que como as classes foram criadas com imagens que continham no máximo um objeto, a ferramenta apresentou tendência em classificar melhor um componente na imagem, geralmente o primeiro e é identificado com acurácia acima de 50%. Isto é, não classifica todos de forma igualitária. Logo, sempre um componente contém uma acurácia alta e os outros, na maioria dos casos, possuem uma acurácia igual ou abaixo de 30%. Porém, isso pode ser observado apenas no modelo de classificação customizado deste artigo. Conclui-se então, que neste caso, ele é melhor para reconhecer um único objeto e não vários. Isto pode ocorrer devido ao fato do modelo do artigo ser treinado apenas com imagens que contém um único componente.

Por fim, devido a ferramenta não dispor de métricas para análise de desempenho geral da rede neural artificial treinada, a forma encontrada de entender melhor os resultados foi através dos testes apresentados neste artigo. O conceito de classificação errada é usado quando o modelo aponta a maior acurácia para uma classe incorreta, contudo, em alguns casos o componente foi reconhecido corretamente porém não apresentou a maior acurácia. Portanto, mesmo com as limitações da versão utilizada e o pequeno número de imagens do *dataset*, o modelo reconheceu corretamente um objeto alvo em 33 das 40 imagens com acurácia acima de 50%.

AGRADECIMENTO

Agradeço a Fundação Parque Tecnológico itaipu (FPTI) e também a Itaipu pelo incentivo ao projeto, ao Centro Internacional de Hidroinformática (CIH) por propor e financiar a ideia, e também ao Centro Latino Americano de Tecnologias Abertas (CELTA) por conceder material e conhecimento para desenvolvimento do projeto.

REFERENCES

- [1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [2] S. S. Damaceno and R. O. Vasconcelos, "Inteligência artificial: uma breve abordagem sobre seu conceito real e o conhecimento popular," *Revista periódica Caderno de Graduação das Ciências Exatas e Tecnológicas – CGCET*, vol. 5, no. 1, pp. 11–16, 2018.
- [3] R. L. Stange, "Adaptatividade em aprendizagem de máquina: Conceitos e estudos de caso," Mestrado em Engenharia Elétrica, Adaptatividade em Aprendizagem de Máquina: Conceitos e Estudos de Caso, São Paulo, 2011.
- [4] I. Cloud. Catálogo. [Online]. Available: <https://cloud.ibm.com/catalog?category=ai>
- [5] I. W. Studio. Ibm watson visual recognition. [Online]. Available: <https://www.ibm.com/br-pt/cloud/watson-visual-recognition>
- [6] Y. LeCun and C. Cortes, "MNIST handwritten digit database," 2010. [Online]. Available: <http://yann.lecun.com/exdb/mnist/>