



Identificação de Problemas nos Sensores de Estacas Hélice a Partir do Relatório de Monitoramento

Lucas Marques Pires da Silva

Mestrando, COPPE/UFRJ, Rio de Janeiro, Brasil, lucas.silva@coc.ufrj.br

Gustavo Vaz de Mello Guimarães, D.Sc.

Docente, UFRJ, Macaé, Brasil, gvmg@poli.ufrj.br

Daniela Fernandes Tavares

Discente, UFRJ, Macaé, Brasil, daniela.ft@poli.ufrj.com

João Victor Braga Nunes

Discente, UFRJ, Macaé, Brasil, joaobnunes2000@gmail.com

Leonardo Alves da Silva

Discente, UFRJ, Macaé, Brasil, leonardoas@poli.ufrj.com

RESUMO: As estacas tipo hélice contínua monitoradas vem ocupando um espaço cada vez maior no mercado de fundações, principalmente nos locais onde o apoio logístico é favorável. Essas estacas fornecem, devido ao seu monitoramento, uma série de parâmetros obtidos durante sua execução. Por sua vez, é nitido que os algoritmos de inteligência artificial vem tomando conta da indústria a nível global. Em grandes obras, utilizando estacas tipo hélice contínua monitorada, naturalmente, a quantidade de dados a serem analisados é considerável, tornando a análise tradicional muito difícil de ser executada em tempo real. Assim sendo, é interessante e promissor associar a inteligência artificial e o aprendizado de máquinas de modo a auxiliar a engenharia na área de fundações. Desta forma, a ideia é utilizar a inteligência artificial para interpretar os boletins de estacas hélice contínua monitoradas, para em seguida ser capaz de verificar quais estacas poderiam apresentar falhas durante a sua execução. A partir daí, ficaria mais simples saber em quais estacas deveriam ser tomadas providências para saber se deveriam ser condenadas ou aprovadas para uso na fundação de uma determinada obra.

PALAVRAS-CHAVE: Inteligência Artificial. Aprendizado de máquina. Estaca tipo Hélice Contínua Monitorada.

ABSTRACT: The design of construction with foundation executed with auger piles is growing a lot nowadays specially in places with enough technologies and facilities needed on those kind of piles. The auger piles have an effort that brings a better relation of data about the procedure of execution of it. In other hand, it's clear that artificial intelligence algorithms are taking over the industry worldwide. In big constructions, using cast-in-place continuous auger piles, naturally, the amount of data to be analyzed is considerable, making the traditional analysis very difficult to be performed in real time. Therefore, it is interesting and promising to associate artificial intelligence and machine learning in order to help engineering in the area of foundations. Thus, the idea is to use artificial intelligence to interpret the cast-in-place continuous auger piles bulletins, to be able to verify which piles could present failures during their installation. From there, it would be simpler to know which piles should be taken to know whether they should be condemned or approved for use in the foundation of a particular construction.

KEYWORDS: Artificial Intelligence. Teachable Machine. Auger Piles.

1 Introdução

Cada vez mais em grandes centros urbanos a engenharia de fundações vem priorizando a utilização das estacas tipo hélice contínua monitorada (HCM). Este tipo de estaca possui um custo acessível quando comparado a outros tipos e, principalmente, agilidade e facilidade em seu processo executivo. Outra vantagem deste tipo de estaca frente as outras está associado ao possível controle de qualidade devido aos parâmetros obtidos durante o monitoramento em sua execução.



O aprendizado de máquina se tornou uma das tecnologias mais interessantes da atualidade. Grandes empresas, como “Google, Facebook, Apple, Amazon e IBM”, investem pesadamente em pesquisas e aplicativos de aprendizado de máquina. Embora possa parecer que o aprendizado de máquina se tornou a expressão da moda e, certamente, não é uma moda passageira. Este campo estimulante abre caminho para novas possibilidades e se tornou indispensável no dia a dia das empresas.

O século XXI é uma época em que os dados vêm em abundância; usando algoritmos de autoaprendizagem da área de aprendizado de máquina, pode-se transformar esses dados em conhecimento. Graças às muitas bibliotecas de código aberto poderosas que foram desenvolvidas nos últimos anos algoritmos poderosos para detectar padrões nos dados e fazer previsões sobre eventos futuros.

Na segunda metade do século XX, o aprendizado de máquina evoluiu como um subcampo da Inteligência Artificial (IA) que envolvia algoritmos de autoaprendizagem que derivavam conhecimento de dados para fazer previsões, em vez de exigir que os humanos derivem regras e construam modelos manualmente a partir da análise de grandes quantidades de dados. De acordo com Angra e Ahuja (2017), o aprendizado de máquina oferece uma alternativa mais eficiente para capturar o conhecimento em dados para melhorar gradualmente o desempenho de modelos preditivos e tomar decisões baseadas em dados.

O aprendizado de máquina não apenas está se tornando cada vez mais importante na pesquisa em ciência da computação, mas também desempenha um papel cada vez maior no cotidiano. Graças ao aprendizado de máquina, desfruta-se de filtros robustos de “spam” de e-mail, “software” conveniente de reconhecimento de texto e voz, mecanismos de pesquisa na “web” confiáveis, programas desafiadores para jogar xadrez e carros autônomos seguros e eficientes.

Neste artigo, o “Teachable Machine” será utilizado para auxiliar no controle de qualidade do monitoramento de estacas HCM a partir de um banco de dados composto por imagens de boletins de monitoramento.

2 Metodologia

2.1 Programa Utilizado

O “software Teachable Machine” foi escolhido para ser utilizado nesta pesquisa devido a fácil interface com o usuário e por ser um programa gratuito. Este “software” é um programa de classificação, a partir de classes pré-definidas, que possui o “tensorflow” como biblioteca base. O endereço eletrônico para acessar esse programa é <https://teachablemachine.withgoogle.com/>. Ao clicar em “Get Started” aparecem as três opções de projetos que podem ser analisados por esta ferramenta: imagens, audios e poses.

O programa possui uma teoria associada a uma relação matemática. Na verdade, é realizado um ajuste de uma equação matricial, onde cada imagem analisada é transformada em uma matriz. Cabe destacar, que as imagens são analisadas pixel a pixel e, quando é realizado o “upload” de uma imagem no programa, ela é automaticamente transformada em uma imagem quadrada. Assim, esta imagem é comparada, pela I.A., a uma matriz quadrada, onde cada pixel da imagem representa um termo da matriz. A equação 1 representa o modelo matemático que transforma cada pixel, através da sua cor (RGB), em um termo de uma matriz genérica A.

$$A = (a_{ij}) \propto RGB \quad (1)$$

A figura 1a mostra a página inicial do “teachable machine” e a figura 1b apresenta a página seguinte com as opções de projeto: “image project”, “audio project” e “pose project”. Neste projeto utilizou a opção “image project”.

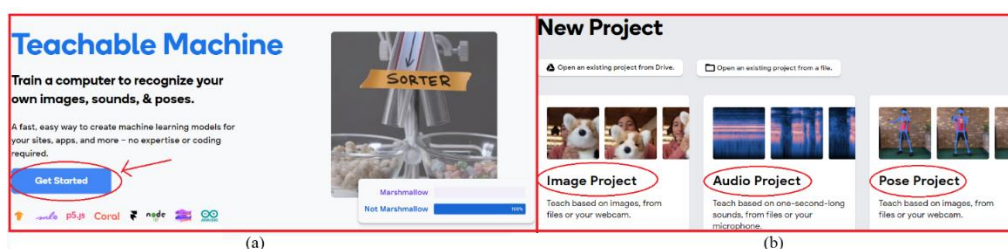


Figura 1 – (a) Página inicial do site de acesso; (b) Possíveis opções de projetos do “software Teachable Machine”, adaptado de <https://teachablemachine.withgoogle.com/>, acessado em 07/11/2021.



2.2 Banco de Dados

O banco de dados deste trabalho foi constituído por 879 imagens de boletins de monitoramento de estacas HCM (Solo Suporte Fundações, 2020). Todas as estacas foram executadas no estado do Rio de Janeiro, sendo que seus diâmetros variam entre 30 e 50 centímetros, e os comprimentos entre 6 e 24 metros, além disso o f_{ck} varia entre 30 e 40 MPa. Destaca-se que os boletins de monitoramento das estacas HCM fornecem seis gráficos elaborados por sensores eletrônicos dispostos no equipamento de execução destas fundações. Três destes gráficos são obtidos por sensores durante o movimento de entrada do trado no solo e dois deles referentes a saída. O último gráfico está associado a um sensor que, de acordo com o volume do concreto injetado, fornece o perfil estimado da estaca.

É importante observar que, sempre que possível, deve-se preferir um banco de dados mais balanceado possível, pois dessa forma as análises provenientes terão melhor desempenho.

2.3 Classes, hipóteses e conceitos

As classes consideradas na metodologia desta pesquisa para a identificação de problemas nos sensores dos boletins de monitoramento foram: “sem falha” e “com falha”.

A partir destas duas classes as seguintes hipóteses foram adotadas para realizar a classificação dos sensores do monitoramento das estacas HCM:

- Todos os sensores (associados a pressão hidráulica, rotação, velocidade de inserção, pressão do concreto, velocidade de extração e perfil estimado) são classificados “com falha” quando não há nenhuma representação gráfica ou, ainda, existe algum trecho de descontinuidade nos gráficos. Caso contrário, naturalmente os sensores são classificados “sem falha”.
- Especificamente o sensor de velocidade de extração também pode ser classificado “com falha” quando a velocidade medida alcança o fundo de escala em um trecho considerável¹.
- Os erros associados ao perfil de monitoramento da estaca podem acontecer de duas maneiras. No caso, quando em alguma profundidade, que apresente sobre-consumo, a estaca apresenta uma estimativa maior que 25% do diâmetro de projeto é identificada uma possível falha. Ou ainda, no caso de sub-consumo, quando em alguma profundidade a estaca apresenta qualquer estimativa de diâmetro menor que o diâmetro de projeto.

Cabe esclarecer dois conceitos básicos para a utilização da inteligência artificial através de qualquer plataforma: treino e teste.

O treino da IA é o procedimento que vai ensinar a plataforma que está sendo utilizada a se acostumar com padrões das classes pré-determinadas anteriormente. O teste da IA vai verificar se um dado pertence a uma classe ou a outra. Com efeito, os dados que são utilizados para teste não podem ter participado da fase de treino da IA. Nesta pesquisa, de acordo com Raschka e Mirjalili (2017), foram adotados 80% do valor total de imagens do banco de dados para o caso do treino e, conseqüentemente, os outros 20% das imagens foram destinados ao teste.

2.4 Procedimentos

2.4.1 Generalidades

Um procedimento muito trabalhoso que merece destaque foi a transformação de todos os boletins de monitoramento que, originalmente, estavam em formato “.PDF” para imagens com formato “.JPEG”. Para a utilização dos boletins no software escolhido foi necessária a numeração contínua de todos os arquivos .JPEG (1 até 879).

2.4.2 Etapas

¹ Para fins deste trabalho foi considerado um trecho maior que aproximadamente um metro.

Os procedimentos a serem adotados neste trabalho estão associados a três etapas: (i) pré-processamento; (ii) processamento e (iii) pós-processamento. Cabe destacar que nestas etapas algumas ações são realizadas pela equipe e outras podem ser realizadas de forma automática pelo programa que está sendo utilizado. A figura 2 apresenta um organograma com as três principais etapas dos procedimentos e as ações necessárias em cada uma delas.

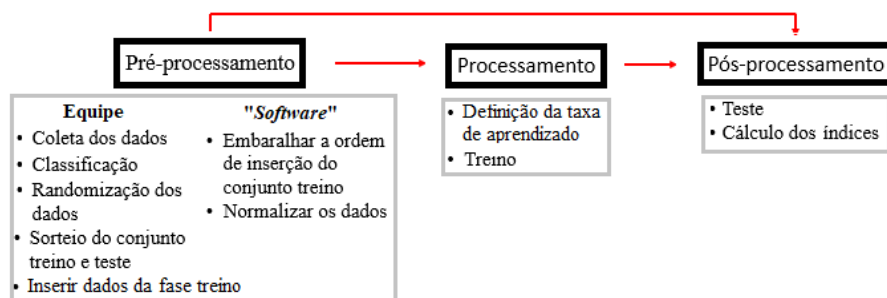


Figura 2 – Organograma das principais etapas para a utilização da IA.

(i) Pré-processamento:

Os boletins do banco de dados foram analisados previamente pela equipe e as imagens foram separadas de acordo com a classificação: “com falha” e “sem falha”. Como resultado, foram obtidos 150 boletins “sem falhas” e 709 boletins “com falhas”. Cada membro da equipe analisou, aleatoriamente, uma parte dos boletins que compunha o banco de dados, assim uma randomização natural dos dados foi realizada.

A figura 3a e 3b apresentam um boletim de monitoramento “sem falha” e um “com falha” respectivamente. A figura 3c mostra parte dos 879 arquivos numerados em ordem crescente armazenados em uma pasta.

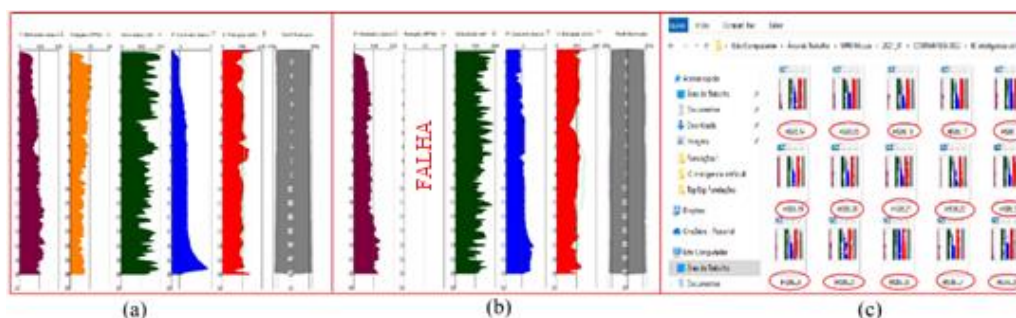


Figura 3 – (a) Boletim “sem falha”; (b) Boletim “com falha”; (c) Arquivos numerados em ordem crescente.

O sorteio dos conjuntos a serem treinados (80%) e testados (20%) foram realizados a partir de uma ferramenta gratuita (<https://sorteador.com.br/>). A figura 4a apresenta a página inicial do “site” utilizado como ferramenta de sorteio dos números e a figura 4b mostra um exemplo de sorteio.



Figura 4 – (a) Página inicial do site escolhido para realizar o sorteio; (b) Exemplo de números sorteados, adaptado de <https://sorteador.com.br/>, acessado em 15/11/2021.



Os boletins referentes a fase de treino da IA foram insridos na plataformapela equipe. Cabe destacar que no caso do “*teachable machine*” as ações referentes a embaralhar a ordem de inserção do conjunto de imagens da fase de treino e normalização dos dados são realizadas automaticamente pela plataforma.

(ii) Processamento:

Antes de se iniciar o treino da IA é necessario a definição da taxa de aprendizado (*learning rate*) que, neste caso, por padrão, foi de 0,001. Para treinar a IA, a partir do programa escolhido, clica-se no botão “*train model*” e a plataforma inicia esta importante fase de aprendizado. A figura 5a apresenta o início do treinamento do programa com a definição da taxa de aprendizado.

(iii) Pós-processamento:

Por fim, com a IA pronta para identificar os padrões ensinados a partir das clases pré-determinadas, inicia-se o teste com as imagens sorteadas anteriormente. É muito importante destacar que, naturalmente, os boletins que são utilizados na fase de teste não são os mesmos utilizados para a fase de treino. Este procedimento visa evitar qualquer viés na análise dos dados. A figura 5b apresenta o resultado de um dos testes realizados em um boletim de monitoramento classificado anteriormente pela equipe como “sem falha” pelo “*teachable machine*”.

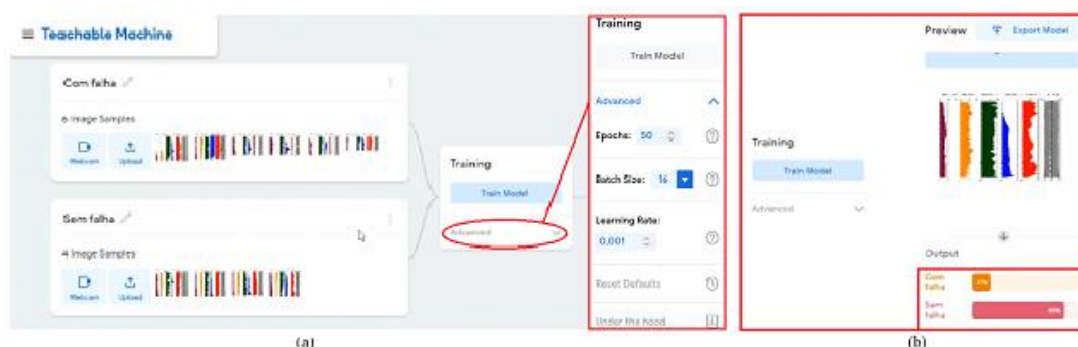


Figura 5 – (a) Fase de treinamento do software; (b) Saída de dados do “*teachable machine*”, adaptado de <https://teachablemachine.withgoogle.com/>, acessado em 07/11/2021.

Neste caso², a IA foi capaz de classificar o boletim com 83% de chance de ser “sem falha” e com apenas 17% de chance de apresentar falha.

3 Avaliação do Desempenho

Diversos índices podem ser utilizados para avaliar o desempenho de uma análise realizada através de um algoritmo, como por exemplo: acurácia, erro, razão de verdadeiros positivos, razão de falsos positivos e precisão.

A acurácia mede quantas previsões foram corretas no conjunto total de previsões, por sua vez, o erro, mede quantas previsões foram erradas no conjunto total de previsões. A razão de verdadeiros positivos mede o percentual de valores positivos reais que são previstos como verdadeiros positivos e a razão de falsos positivos mede o percentual de valores negativos reais que são previstos como falsos positivos. Já a precisão mede o percentual de previsões positivas que correspondem a verdadeiros positivos.

Um conceito muito interessante e facilitador das análises utilizando esses índices é a matriz confusão. Para esta pesquisa, definiu-se o verdadeiro positivo quando o algoritmo acerta que o boletim é “sem falha”; o falso negativo quando o algoritmo prediz que o boletim é “com falha”, mas na realidade o boletim é “sem falha”; o verdadeiro negativo quando o algoritmo acerta que o boletim é “com falha”; e o falso positivo quando o algoritmo prediz que o boletim é “sem falha”, mas na realidade é “com falha”. A figura 6 apresenta a matriz confusão e as equações 2, 3, 4, 5 e 6 que definem os índices acurácia (ACC), erro (ERR), razão de verdadeiros positivos (TPR), razão de falsos positivos (FPR) e precisão (PRE), respectivamente.

² Destaca-se que este caso foi realizado para familiarização do “software” escolhido.



Valor real	Positivo	Verdadeiros positivos (VP)	Falsos negativos (FN)
	Negativo	Falsos positivos (FP)	Verdadeiros negativos (VN)
		Positivo	Negativo
		Valor previsto	

$$ACC = \frac{VP+VN}{VP+VN+FP+FN} \quad (2)$$

$$ERR = 1 - ACC \quad (3)$$

$$TPR = \frac{VP}{FN+VP} \quad (4)$$

$$FPR = \frac{FP}{FP+VN} \quad (5)$$

$$PRE = \frac{VP}{VP+FP} \quad (6)$$

Figura 6. Matriz confusão.

Outra avaliação que pode ser feita é comparar a acurácia da IA com a de um “dummy test”. O “dummy test” é o método mais básico, sendo utilizado como comparação com outros métodos. Basicamente, esse método consiste em analisar o banco de dados e verificar qual das classes é a predominante e, assim, estimar que todo o banco de dados se comporta dessa forma. Cabe ressaltar que o “dummy test” deve sempre possuir um tempo de aplicação menor que a IA. Com efeito, pode-se medir uma certa eficácia da IA em relação ao “dummy test”.

4 Estudo de Caso – Análises e Resultados

Fora a tentativa inicial de familiarização da plataforma, outras três tentativas de utilização do programa para classificação dos boletins de monitoramento nas classes “sem e falha” e “com falha” foram realizadas. Em cada uma delas o número de boletins, associado ao banco de dados, foi diferente.

4.1 Primeira Tentativa

A primeira tentativa contou com 250 relatórios, sendo que, 208 imagens classificadas pela equipe “com falha” e 42 imagens “sem falha”. Para a fase de treino do programa foram utilizados 200 boletins de monitoramento (80% de todas as imagens do banco de dados): 162 imagens “com falha” e 38 imagens “sem falha”. Para a fase teste, por sua vez, foram utilizados os 50 boletins restantes (20% de todas as imagens do banco de dados). Destes boletins, 47 eram “com falha” e 3 “sem falha”.

A inteligência artificial foi capaz de acertar a classificação de todos os 3 boletins sem falhas. Já na classificação dos 47 boletins com falha, a IA errou somente um deles. A figura 7 apresenta o resultado da primeira tentativa representado pelo conceito da matriz confusão. A tabela 1 mostra os parâmetros calculados a partir da matriz confusão para a primeira tentativa.

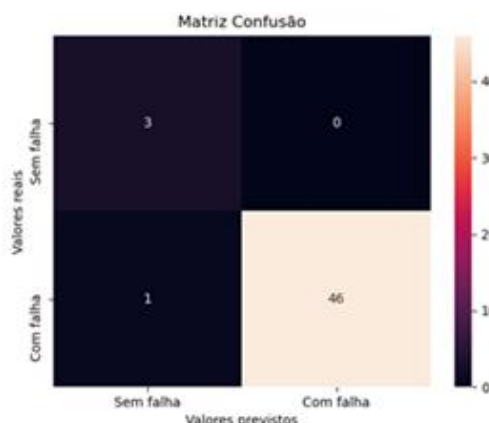


Tabela 1. Parâmetros calculados a partir da matriz confusão para a 1ª tentativa.

Parâmetros	1ª tentativa
Acurácia	98,0%
Erro	2,0%
Razão de verdadeiros positivos	75,0%
Razão de falsos positivos	0,0%
Precisão	100,0%

Figura 7 – Matriz confusão para a 1ª tentativa.



4.2 Segunda Tentativa

A segunda tentativa contou com mais que o dobro de boletins de monitoramento que na tentativa anterior, 565 imagens no total, sendo que 515 imagens de imagens “com falha” e 50 “sem falha”. Para a fase de treino do programa foram utilizados 452 boletins de monitoramento (80% de todas as imagens do banco de dados): 408 imagens “com falha” e 44 imagens “sem falha”. Para a fase teste, nesta tentativa, foram utilizados os 113 boletins restantes (20% de todas as imagens do banco de dados). Destes boletins, 107 eram “com falha” e 6 “sem falha”.

A inteligência artificial também foi capaz de acertar a classificação de todos os boletins sem falhas. Porém, na classificação dos 113 boletins “com falha”, a IA se equivocou apenas 4 vezes. A figura 8 apresenta o resultado da segunda tentativa representado pelo conceito da matriz confusão. A tabela 2 mostra os parâmetros calculados a partir da matriz confusão para a segunda tentativa.

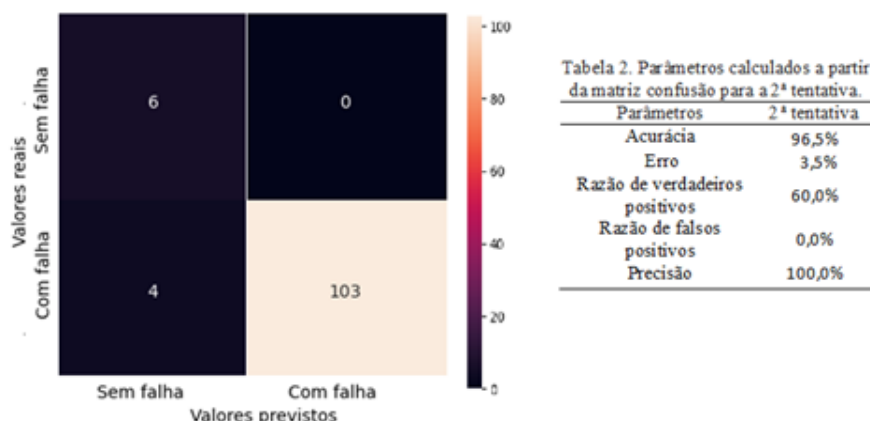


Figura 8 – Matriz confusão para a 2ª tentativa.

4.3 Terceira Tentativa

A última tentativa contou com todos os de boletins de monitoramento do banco de dados disponível, 859 imagens no total, das quais 709 boletins de monitoramento “com falha” e 150 “sem falha”. Para a fase de treino do programa foram utilizados 687 boletins de monitoramento (80% de todas as imagens do banco de dados): 573 imagens “com falha” e 114 imagens “sem falha”. Para a última fase teste foram utilizados os outros 171 boletins (20% de todas as imagens do banco de dados). Destes boletins, 140 eram “com falha” e 31 “sem falha”.

Neste caso, a IA se equivocou em 8 boletins “sem falha” e 2 boletins “com falha”. A figura 9 apresenta o resultado da terceira tentativa representado pelo conceito da matriz confusão. A tabela 3 mostra os parâmetros calculados a partir da matriz confusão para a última tentativa.

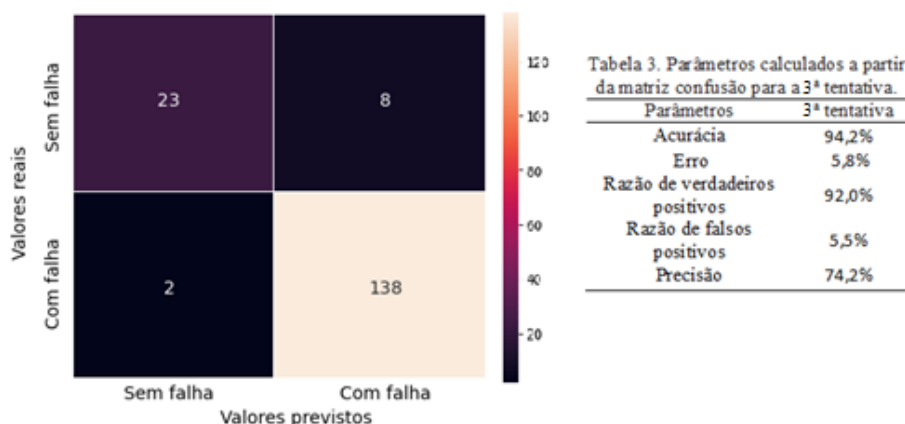


Figura 9 – Matriz confusão para a 3ª tentativa.



Para entender o desempenho da análise compara-se a acurácia da IA com a acurácia do “*Dummy test*”. A figura 10 mostra a eficácia para cada uma das três tentativas.

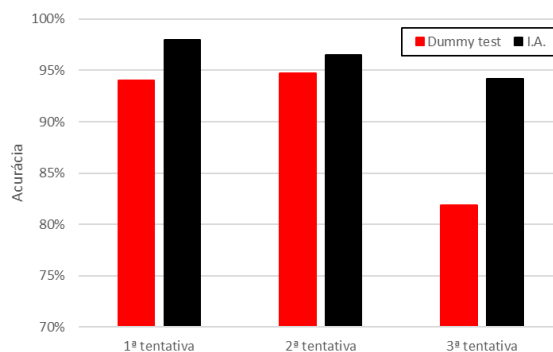


Figura 10 – Desempenho da acurácia da IA e do “*Dummy test*”.

Pode-se verificar que apesar da acurácia da IA reduzir com o aumento do banco de dados, a diferença da acurácia da IA em relação a acurácia do “*Dummy test*” aumenta. Assim, pode-se concluir que o processo de interpretação dos boletins pela IA se torna mais eficaz.

5 Considerações Finais e Discussões

Os procedimentos adotados e a plataforma “*teachable machine*” utilizada para a identificação de problemas nos sensores dos boletins de monitoramento de estacas HCM se mostram adequados. Cabe destacar que o programa é de fácil usabilidade e possui uma interface lógica e amigável.

Os índices utilizados a partir do conceito da matriz confusão se demonstraram muito úteis na avaliação do desempenho da tecnologia com a inteligência artificial. Cabe salientar que os índices apresentaram um desempenho satisfatório, todavia o banco de dados utilizado não possui um balanceamento adequado, pois existem muito mais boletins com falha do que sem falha.

Este trabalho é uma prova de conceito de tecnologia aplicada ao caso estudado. Pode-se dizer que a inteligência artificial foi capaz de acertar boa parte dos testes realizados. No meio da ciência da computação, costuma-se dizer que a IA aprendeu a classificar os dados (adaptado de Raschka e Mirjalili, 2017).

Com efeito, nas futuras pesquisas a evolução natural é ensinar a IA a distinguir os tipos de falhas nos sensores dos boletins de monitoramento, além de identificar possíveis falhas na execução da fundação.

REFERÊNCIAS BIBLIOGRÁFICAS

Angra, S. and Ahuja, S. (2017). Machine learning and its applications: a review. In Big Data Analytics and Computational Intelligence (ICBDAC), International Conference on. IEEE, pp. 57-60.

Raschka S. e Mirjalili V. (2017). *Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow*, 2nd ed., Packt, Mumbai, India, 622 p

Teachable Machine. Disponível em: <<https://teachablemachine.withgoogle.com/>> Acesso em: 07/11/2021.

Sorteador. Disponível em: <<https://sorteador.com.br/>> Acesso em: 15/11/2021.

Solo Suporte Fundações (2020) Banco de dados de boletins de monitoramento de estacas hélice continua executadas no Estado do Rio de Janeiro – comunicação por correio eletrônico.