

Reconhecimento de fechamento do porão.

Francisco Carlos Silva Pimentel

II HACKATHON Softex/IFMA 2023.

Instituto Federal do Maranhão (IFMA) – São Luís – MA – Brasil

1 Introdução

De acordo com a listagem de problemas propostos pela Empresa Maranhense de Administração Portuária - EMAP (Porto do Itaqui) cuja demanda de número onze sugere:

- Como registrar automaticamente o momento exato do fechamento do porão?
- Para validar uma parada operacional, principalmente aquelas que necessitam do fechamento do porão, é importante capturar e registrar esse momento. Para isso, tem-se a ideia de instalar câmeras, com inteligência artificial, que capturem o exato momento que o porão for fechar. A Suzano tem feito algo nesse sentido, e os testes estão em fase avançada.

O referido problema, entre outros, faz parte da segunda edição do concurso Hackthon Softex/IFMA, sendo uma das regras a exigência da aplicação de Inteligência Artificial (IA) na solução proposta. Neste sentido, idealiza-se uma aplicação que valide o modelo de Aprendizagem de Máquina (AM) para reconhecer o porão do navio e sua tampa, com o intuito de criar um balizador das duas situações possíveis.

Assim, a proposta para esta solução, terá como ponto de partida, um modelo que possa fotograficamente coletar a imagem dos porões do navio por meio de câmeras e detectar se ele está fechado ou aberto. Após este balizamento inicial como premissa, a solução poderá evoluir na direção da detecção da transição do estado do porão, por exemplo: de aberto para fechado ou vice-versa.

2 A solução

A proposta da solução do problema de detecção do fechamento do porão vai basear seu funcionamento em conjuntos de soluções com módulos separados, trabalhando em conjunto, conforme Figura 1.

O primeiro módulo do projeto vai contar com um dispositivo de *Internet of Things* (IoT), a ESPCAM 32, uma câmera acoplada a um microcontrolado programável, o qual permite a execução de uma aplicação em C++, bem como a transmissão das imagens por meio da rede local de forma wireless. O objetivo será transformar este dispositivo em um *web service* para *streaming* de vídeo via IP fixo.

Na parte do módulo do cliente, propõe-se uma estação de trabalho, preferencialmente, conectada a rede interna e com placa aceleradora de vídeo. Nesta estação uma aplicação vai receber as imagens da câmera por meio de um IP, onde será

executada a inferência do modelo de IA para a detecção do fechamento do porão, sendo possível o registro do exato horário do fechamento do porão em um banco de dados.

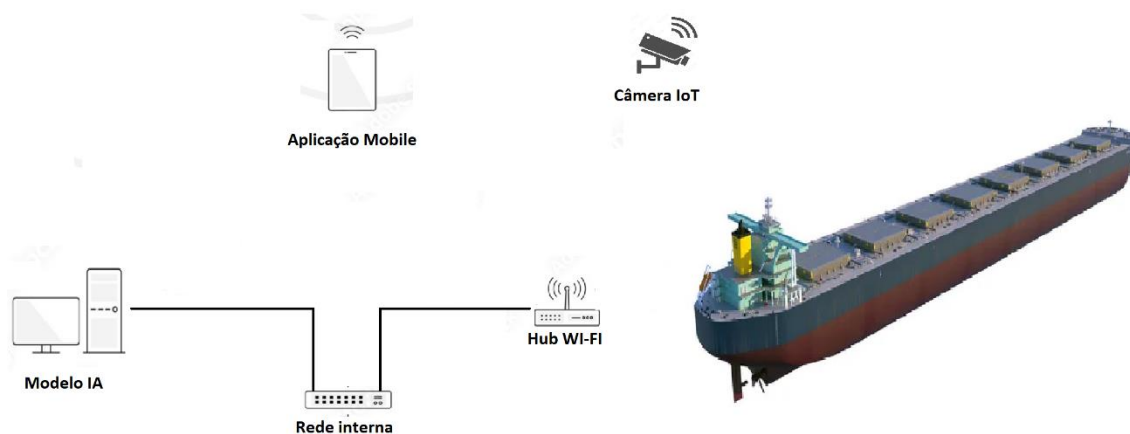


Figura 1 – Proposta da solução do problema da detecção do momento do fechamento do porão.

Facultativamente, como terceiro módulo, a aplicação poderá mandar uma mensagem para uma *app mobile*, informando o fechamento do porão, ou o disparo de uma sirene, acender uma lâmpada ou painel de ledes. Para que esse *app mobile* possa funcionar, faz-se necessário a utilização de mais uma aplicação que possa controlar as mensagerias entre a base de dados e o app dos dispositivos móveis.

3 O protótipo

O protótipo se limita apenas a viabilidade de desenvolvimento da solução. Uma pequena parte será desenvolvida, dando ênfase ao teste funcional da aplicação do modelo de artificializa, cujo objetivo será certificar o uso da AM e avaliar os seus resultados.

Para tal feito, um experimento será realizado com um modelo de detecção de imagens de navios cargueiros. Inicialmente, todas as imagens coletadas e utilizadas, serão de domínio público. Uma vez que o propósito será detectar quando o porão está aberto ou fechado, satisfazendo a prova de conceito e a viabilidade do desenvolvimento da solução do problema de número onze.

Havendo sucesso do experimento, uma segunda etapa, já na fase de desenvolvimento da solução, a qual será realizada com imagens reais e extraídas a partir das câmeras de circuito interno, por fim, anexadas a base já existe. O modelo IA será retreinado para adaptar-se à realidade do Porto do Itaquí.

Sendo importante enfatizar que o referido protótipo não contempla toda a proposta de resolução do problema, tendo seu escopo limitado apenas a prova de conceito e a viabilidade do desenvolvimento da solução com o modelo de IA.

4 Arquitetura

Na arquitetura da solução de cada módulo serão empregadas algumas tecnologias de desenvolvimento de aplicações e IA. No desenvolvimento de aplicações, a linguagem *Python* será usada para construção da aplicação principal, a qual recebe o *streaming* de vídeo e realiza a detecção do fechamento do porão. Já o do *streaming* da câmera, utilizará a linguagem C++ para desenvolvimento da aplicação de captura de vídeo e transmissão via rede interna de forma wireless. Esta aplicação será um *web service* de *streaming* por IP fixo. A princípio, a proposta será a utilização de uma device ESPCAM 32, conforme Figura 2.



Figura 2 – Dispositivo ESPCAM 32.

A escolha do dispositivo ESPCAM 32, em detrimento dos demais equipamentos de *streaming* ofertados no mercado, consiste em seu baixo custo (menos de US\$ 5,00), sua flexibilidade em ser programável, sua conectividade com redes internas e recursos wireless, as quais atendem as necessidades da fase de teste da solução.

O algoritmo de detecção de imagens utilizado no desenvolvimento da aplicação, será o YOLO versão 4.0, por se tratar de um modelo de rápida detecção e por reconhecer imagens em diferentes ângulos e formas, bem como, pela sua capacidade de reconhecer imagens em situações de ruídos, tais como: objetos na frente da imagem (guindastes, veículos, ferragens) e outros tipos de ruídos. O ambiente operacional ideal para executar detecções com o YOLO seria o sistema operacional *Linux*, mas optou-se por usar o *Windows* 10 ou 11 como hospedeiro da aplicação.

A aplicação principal será um programa em linguagem *Python*, que será executado em um computador ou notebook equipado com placa aceleradora de vídeo. Esta aplicação será capaz de receber o *streaming* da câmera. A imagem da câmera será tratada a nível de frame para inferência no modelo de IA.

5 Resultados esperados

Antes que a solução possa ser tida por completa e funcional, será necessário a aplicação de certos experimentos para obtenção de resultados específicos. Tais resultados são de suma importância para certificar o uso da solução, pois o foco será o uso de IA e a maioria dos resultados esperados se concentrarão no desempenho de tarefas de reconhecimento de imagens.

Portanto pode-se destacar como os principais, os que seguem:

- *Streaming* de vídeo;
- Coleta de imagens para testes de detecção de imagens;
- Rotulação das imagens para testes;
- Aplicação de *data augmentation* nas imagens;
- Instalação do framework YOLO para testes de detecção de imagens;
- Obtenção de boa acurácia na detecção de imagens;
- Capacidade de reconhecimento do porão aberto, fechado e fechando;
- Capacidade de reconhecimento do momento exato do fechamento;
- Comunicação com banco de dados ou *App mobile*.

5.1 Resultados preliminares

A primeira iteração de parte do protótipo já foi desenvolvida utilizando uma coleta de imagens públicas obtidas na internet, as quais foram rotuladas e submetidas a um processo de *data augmentation*, que elevou a base para 75 mil amostras. Com tal base, instalou-se o *framework* YOLO e fez-se os testes preliminares com sucesso no reconhecimento do porão aberto e fechado.



Figura 3 – Reconhecimento da imagem com acurácia do modelo em 62%.

Há algumas considerações que devem ser comentadas sobre os resultados obtidos. Uma delas é quanto a acurácia de reconhecimento do modelo que no primeiro treinamento, com 25 mil imagens, teve baixos percentuais de certeza no reconhecimento do porão fechado e aberto.

O referido modelo, em seu primeiro treino obteve uma acurácia de 62%, mas com um percentual de certeza muito baixo na detecção do porão fechado e aberto, conforme Figura 3.



Figura 4 – Reconhecimento da imagem do porão com erros de detecção.

Na Figura 4, pode-se observar que nem todos os porões fechados foram detectados, apesar de um bom percentual de certeza em apenas três porões. Já na Figura 5, demonstra que mesmo com percentual de acurácia baixo, ainda foi possível detectar o porão aberto em imagem com close.



Figura 5 – Reconhecimento da imagem do porão em close.

Após alguns ajustes no banco de dados, ter aumentado o número de amostras de 25 para 75 mil imagens e todas as rotulações de imagens refeitas, obteve-se um treinamento com uma acurácia de 91%.



Figura 6 – Reconhecimento da imagem com acurácia do modelo em 91%.

Nesta segunda iteração, percebe-se na Figura 6, comparando com a mesma imagem (Figura 3) com baixos percentuais de certeza, agora obteve-se 100%. Na Figura 7, a imagem anterior (Figura 4) que teve dificuldade de localizar todos os porões fechados, após o novo treino com a acurácia maior, conseguiu detectar todos os porões.



Figura 7 – Reconhecimento da imagem do porão com mais acertos.

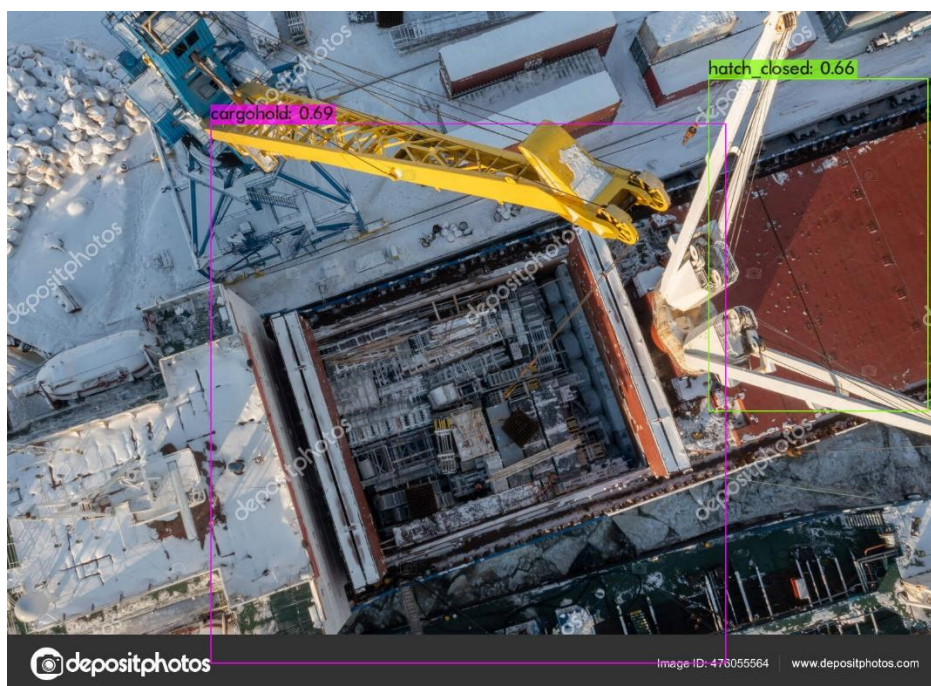


Figura 8 – Reconhecimento do porão aberto e fechado.

Entre os melhores resultados, pode-se destacar o reconhecimento tanto do porão aberto quanto do fechado na mesma imagem, conforme Figura 8. Deste modo, infere-se que o modelo de reconhecimento de imagens já está apto para evoluir. Embora exista um desafio maior que é detectar o exato momento em que o porão muda do *status* de aberto para fechado, ainda há muitos outros experimentos a serem feitos até chegar neste patamar.

6 Conclusão

Após todo o desenvolvimento e testes do algoritmo, alguns pontos interessantes devem ser observados a respeito do desenvolvimento deste trabalho, que são:

1. Considera-se que o protótipo atendendo as todas as expectativas com os seus resultados preliminares.
2. Os experimentos durante o desenvolvimento do protótipo demonstraram a viabilidade da possível solução para o problema número 11 da EMAP.
3. Entende-se que as próximas etapas e experimentos futuros tendem a evoluir para a solução pretendida por meio deste protótipo como ponto de partida.