

CONSTRUÇÃO DA BASE DE DADOS PARA APLICATIVO PARA APLICAÇÃO EM PLACAS DE SINALIZAÇÃO DE MATERIAIS PERIGOSOS.

Maurício Nascimento Cunha, Murilo Freitas, Rafael Stubbs Parpinelli

INTRODUÇÃO

O transporte de produtos perigosos exige identificação precisa e rápida da carga em caso de acidentes, pois quando não se identifica quais são os produtos envolvidos, deve-se considerar que uma situação perigosa existe e deve ser tomadas as medidas de segurança e precauções máximas para prevenir qualquer efeito indesejável nos socorristas ou em qualquer outra pessoa na área. Após a identificação do produto, é possível determinar os riscos associados e planejar medidas de controle e segurança mais adequadas, tanto para o pessoal que participa da emergência quanto para o resto das pessoas, com respeito aos riscos a que estão expostas. Com isso, o projeto utiliza aprendizado de máquina e redes neurais convolucionais para identificar automaticamente painéis de segurança e números ONU em veículos. Atualmente, a etapa em execução foca na construção de uma base de conhecimento enriquecida com informações extraídas das Fichas de Segurança de Produtos Químicos (FDS), que irão alimentar o aplicativo móvel para consulta offline. O objetivo desta fase é automatizar a coleta desses dados, padronizá-los e integrá-los de forma prática à aplicação.

DESENVOLVIMENTO

A abordagem de pesquisa utilizada é de caráter exploratório e experimental, começando com uma investigação teórica e revisão de literatura sobre o transporte de substâncias perigosas, reconhecimento pelo número ONU e extração automatizada de informações. A fase prática teve início com a criação de um algoritmo em Python para a leitura por automação de documentos PDF que contém Fichas de Segurança de Produtos Químicos (FDS), disponíveis no site da Labsynth (<https://www.labsynth.com.br/fispq/>).

A maneira de realizar a extração envolve as fases a seguir:

Análise e extração de conteúdo de arquivos PDF utilizando bibliotecas especializadas em manipulação de documentos.

Padrões consistentes para a identificação e extração de dados, como nome da substância, número ONU, categoria de risco, riscos adicionais e medidas de proteção.

A organização das informações coletadas em arquivos CSV possibilita a alteração manual e a conferência antes de serem incorporadas ao banco de dados do aplicativo.

Simultaneamente, foi desenvolvida uma automação para verificar placas utilizando a numeração da ONU no portal da ANTT (http://200.144.30.103/siipp/public/busca_pp.aspx). Isso permitiu a obtenção dos dados no formato JSON e sua conversão para CSV.

Até o presente momento, foram coletadas 682 informações de placas, e atualmente está em curso o tratamento e validação desses dados para garantir a integridade e formatação das informações coletadas. Essa estratégia tem como objetivo uniformizar os dados, diminuir a necessidade de intervenções manuais e reduzir as falhas na criação da base de conhecimento que será empregada pelo aplicativo de reconhecimento.

RESULTADOS

Para alcançar os objetivos propostos, empregou-se uma abordagem exploratória por meio de procedimentos experimentais, utilizando métodos qualitativos e soluções técnicas

práticas (algoritmos de automação) de caráter aplicado. Abaixo estão descritos a justificativa, os métodos e os procedimentos empregados.

Com base nos princípios de extração automatizada de dados e processamento de informações estruturadas, foi criado um algoritmo em Python 3 capaz de analisar Fichas de Segurança de Produtos Químicos (FDS) no formato PDF. Esse algoritmo foi estruturado usando para reconhecer campos específicos, como nome do produto, número ONU, classe de risco, riscos subsidiários e instruções de segurança (primeiros-socorros e medidas de incêndio).

Ademais, foi construída uma automação para consulta no banco de dados público da ANTT, o que possibilita a identificação direta de dados a partir do número ONU e sua conversão em arquivos JSON e CSV. A junção dessas duas fontes garantiu a coleta de 682 registros de placas de identificação de substâncias perigosas.

Os resultados obtidos através da automação acarretaram a diminuição do tempo disposto na coleta e organização dos dados em relação à digitação manual, melhorando a eficiência e reduzindo a possibilidade de falhas humanas. Além disso, essa estratégia facilita a padronização da base, um requisito importante para consultas rápidas e consistentes no aplicativo móvel que foi desenvolvido através dessa pesquisa.

Dentre os benefícios identificados, destacam-se a escalabilidade do método e a oportunidade de atualizações regulares. Como dificuldades encontradas nessa etapa da pesquisa, destacam-se a falta de qualidade dos documentos originais (que podem apresentar variações de formatação de site para site), não ter uma API específica contendo os dados das placas e a necessidade de modificações no algoritmo para tratamento de dados incompletos ou inconsistentes.

CONSIDERAÇÕES FINAIS

A fase concluída até agora, baseada no trabalho Aprendizado de máquina aplicado no reconhecimento de placas de identificação no transporte de substâncias perigosas (Freitas, 2024), possibilitou a automação da extração e estruturação de dados importantes das Fichas de Segurança de Produtos Químicos, bem como do banco de dados público da ANTT. Isso gerou uma base inicial com 682 registros de placas de identificação de substâncias perigosas. Com o uso da linguagem Python e bibliotecas como PyPDF2 para leitura de PDFs, RE para expressões regulares, PANDAS para manipulação e organização de dados, e Requests/BeautifulSoup para coleta em páginas web, foi possível criar um processo eficaz para diminuir o tempo de coleta, reduzir erros e padronizar os dados para integração ao aplicativo móvel, cumprindo os objetivos estabelecidos nesta etapa da pesquisa.

A metodologia utilizada provou ser apropriada para lidar com diversas fontes e formatos de dados, apesar de ainda ser necessário melhorar o manejo de casos com formatações irregulares ou dados incompletos. A continuidade do trabalho incluirá a expansão da base de dados, a introdução de validações automáticas, a integração final ao sistema de identificação no aplicativo móvel, o aprimoramento do algoritmo de detecção e reconhecimento de painéis e dígitos, além da ampliação do conjunto de imagens de placas. Essas medidas fortalecerão a utilidade prática da solução no âmbito da fiscalização e resposta a emergências.

Palavras-chave: Transporte de Produtos Perigosos; Identificação Automática; Aprendizado de Máquina; Redes Neurais Convolucionais; Fichas de Segurança de Produtos Químicos (FDS); Número ONU; Aplicativo Móvel.

REFERÊNCIAS BIBLIOGRÁFICAS

FREITAS, Murilo. Aprendizado de máquina aplicado no reconhecimento de placas de identificação no transporte de substâncias perigosas. 2024. Trabalho de Conclusão de Curso (Graduação em Ciência da Computação) – Universidade do Estado de Santa Catarina, Joinville-SC, 2024.

DADOS CADASTRAIS

BOLSISTA: Maurício Nascimento Cunha

MODALIDADE DE BOLSA: PIBITI/CNPq (IT)

VIGÊNCIA: 09/2024 a 08/2025 – Total: 12 meses

ORIENTADOR(A): Rafael Stubs Parpinelli

CENTRO DE ENSINO: CCT

DEPARTAMENTO: Departamento de Ciência da Computação CCT

ÁREAS DE CONHECIMENTO: Ciências Exatas e da Terra / Ciência da Computação / Metodologia e Técnicas da Computação.

TÍTULO DO PROJETO DE PESQUISA: Aplicação de Inteligência Computacional na Indústria Siderúrgica Parte 2 (NPP2015010003636)

Nº PROTOCOLO DO PROJETO DE PESQUISA: NPP2015020003186