INTELLIGENT ANALYSIS SYSTEM AND AUTOMATIC CHECKING OF TRAIN FORMATIONS AT EFVM USING COMPUTER VISION

Diogo Henrique da Silva, Gustavo Maia de Almeida

Industrial Automation Coordinator, IFES - Instituto Federal do Espírito Santo

[1]*Av. dos Sabiás, 330, Morada de Laranjeiras, Serra, 29166-630, Espírito Santo, Brasil*
dhs.ddd@gmail.com, gmaia@ifes.edu.br

Abstract. The formation of trains is an inherent step in the railway logistics process, where resources (wagons and locomotives) are joined or separated. This activity involves operational, fiscal and commercial risks, which if not mitigated can result in results ranging from fines applied by government agencies to potentially catastrophic accidents, if a composition is delivered in the wrong sequence to a customer and enters the production line. of some steel mill. Taking a look at the Iron Ore circuit, approximately 7,800 wagons are allocated to the train daily, making checking all formations/compositions a very laborious task. The present work consists of developing a methodology for intelligent analysis and automatic checking of train formations using images generated by cameras installed in iron ore loading terminals along the EFVM (Estrada de Ferro Vitória a Minas).

Keywords: Computer Vision, Character Recognition, Railway, Wagon Conference.

1 Introduction

Characterized by the ability to move large volumes over long distances with energy efficiency and low cost, the railway mode is important in our country's logistics strategy.

Currently, participating with around 21.5% of Brazil's transport matrix (as shown in Figure 1), we can observe that there is still room for increasing the representation of the railway modal, especially when compared with the transport matrices of countries in same territorial size and which are better classified economically, ANTF [1].

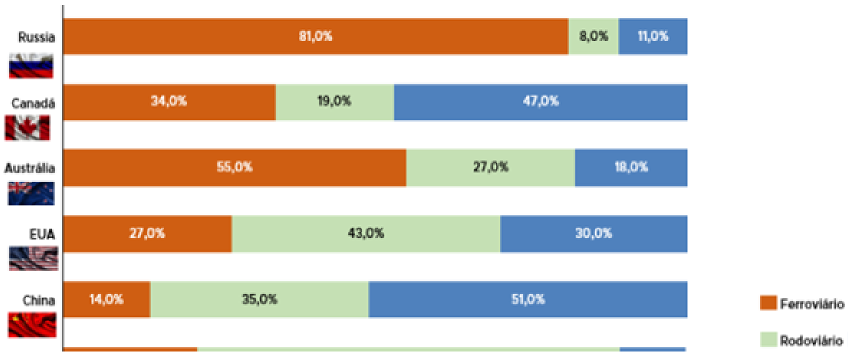

Figure 1. Comparative graph of cargo transport matrices

The Vitória a Minas Railway (EFVM) represents only 3% of the Brazilian railway network in length (approximately 905 kilometers), but in the last 15 years it has responded, on average, for almost 30% of all national railway cargo. Almost 90% of this transport on average is dedicated to iron ore, corresponding to more than 100 million tons annually. Such numbers highlight and place EFVM among the most productive in Brazil, VALE [2].

All movement of this macro volume in the railway mode is carried out through train formations and circulation. Approximately more than 230,000 cars are allocated to trains every month, where this activity is most often carried out without due checking of the compositions in the physical. Therefore, if there is a change in the sequence or the insertion of a wagon or batch (set of 86 wagons) in the composition, it may take days to be noticed. This causes an operational risk, as there may be a wagon in the train that needs some maintenance and should not be traveling. This causes a fiscal risk, as the railway management system is updated manually, and dispatches/billings are made on a wagon-by-carriage basis. Therefore, if a carriage has an associated invoice, but is not actually on the train, a fine may be imposed due to fiscal discrepancy.

Another risk that exists, when there is an unnoticed change in wagons or lots on a train, is that of contamination of a customer's stack. Because, in the steelmaking process, if we mix granulated ore with very fine ore, we can cause damage to the blast furnace with the risk of an accident (explosion).

Bearing in mind that there are thousands of wagons that are part of train formations and circulation every day, exposure to these risks is frequent, as the physical checking of wagon sequences is carried out sporadically and there is no method that guarantees a standardized activity.

Faced with so many potential risks involved, it is necessary to create a structured, permanent, standardized and robust routine (due to the number of wagons involved in the process). Some aspects of the current railway process can be explored to structure this routine.

A favorable aspect is that every wagon has on each of its longitudinal faces an alphanumeric code, painted on the upper left end, which uniquely identifies it, thus distinguishing it from any other railway vehicle on the Brazilian network, as shown in the photo presented in Figure 3.

Figure 2. Photo of GDE type wagon loaded with Ore side view

Another favorable condition to be explored is the fact that close to the exit of the Tubarão railway yard, located at Ponta de Camburi, in Vitória - ES, there are cameras installed that help to ensure the safety of train circulation. Such cameras can be repositioned (and/or installed others) laterally and used to monitor the passage of cars after trains form and depart from the yard.

Therefore, in this article, we propose the development of an intelligent solution, based on Computer Vision, that automatically checks train compositions, aiming to guarantee the correct sequence of wagons in each formation of ore trains at EFVM's Tubarão yard.

The remainder of this article is organized as follows. The EFVM ore train formation process is briefly described in section 2. The proposed methodology is presented in section 3. Results obtained in section 4. Finally, conclusions and future work are presented in section 5.

2 Ore Train Formation Process at EFVM

The ore loading process begins with the formation of trains and subsequent shipment to loading points. When serving the flow of iron ore at EFVM, a typical train is made up of three locomotives and 252 wagons (or three batches of 86 wagons) of the GDE type, which have a top opening and are unloaded into wagon dumpers. This composition covers the majority of trains, when there are high transport volumes, with the remainder consisting of two locomotives and 252 wagons (3 lots).

After the train is physically formed in the yard, it is necessary to carry out the process of checking the compositions. This activity, performed manually, consists of visually identifying the wagon codes and recording the sequence (or order) of these vehicles in the train composition.

After checking the train in the field, this train is dispatched into the system, where a form will be registered containing all assets (wagons and locomotives) and their sequences in the composition, which must accompany the train's journey and be delivered to customers for compliance with tax orders and assisting in operational processes.

Ideally, all train compositions would be checked. However, as the number of wagons is very large and the checking process is all done visually directly in the field, some wagons end up passing straight through and in cases where the trains arrive at the loading points with wagons recommended to the workshop, for example, the risk accident rate increases significantly during the circulation of these trains.

To robustly evaluate all allocations of cars on trains, it is essential to use automatic image analysis, as in addition to the exposure of employees to traveling the entire length of the train (approximately 2,000 meters, most of the time in uneven locations, with risks of venomous animals, among others), the activity requires a lot of time and can also lead to errors, as it depends on the level of attention/concentration at the time of the observation made by the employee.

3 Methodology

As mentioned in section 1 of this article, every wagon has an alphanumeric code on each of its longitudinal faces that identifies and distinguishes it from any other railway vehicle on the Brazilian network.

The standardization of this coding is defined by the current standard ABNT NBR 11.691/2023, which establishes the criteria for classification, identification and marking of railway wagons used on Brazilian railway networks, ABNT [3]. This alphanumeric code is made up of 3 letters, followed by 7 Arabic numerals (the last digit being a check digit, which is defined by a pre-defined mathematical formula).

3.1 *Dataset*

Using cameras already installed in some strategic locations on the railway (loading and unloading points), images were obtained with a side view of the wagons. From this angle of view we can capture frames of images of the wagon codes, form an image bank, and through the application of machine learning and optical character recognition methodologies and techniques, we identify the wagons and their sequences in the composition/batch.

To obtain the images, we used the database integrated with the cameras and videos recorded on Vale's video analytics system server, thus generating a dataset of 1,087 images that were properly treated.

In order to increase the quality of the samples, images were collected at different times of the day (day and night) and in different environmental conditions (rain, light).

3.2 *Annotations*

The dataset was prepared for training, where manual annotations of the images were made using a simple tool, “LabelImg” developed by Tzutalin. With this tool, it was possible to make demarcations (bounding boxes), in all images, in the regions of interest (observable in Figure 3), which, in this case, were the places where the wagons' alphanumeric codes appeared, generating the labels (labels) and “.txt” files in the format required for training in YOLOv5, which was the model used to locate the wagon codes.

Figure 3. Demarcation in night photo camera image frame

3.3 *Network training to locate wagon codes*

The formed dataset was randomly subdivided into training and validation datasets (with the insertion of some background images).

Google Colab was used as a development environment to train the network to locate the wagon codes. In the configuration file, “FOTOS_GDE.yaml”, the single class named as “boards”. This choice was made because Google Colab offers high-performance GPU (Graphical Processing Unit) processing, running in the cloud.

The model architecture was developed in Python language based on the code and structure of the YOLOv5 neural network. With 270 layers and more than 7 (seven) million parameters, the model was trained on different network versions, with the YOLOv5s version being chosen, as it presented better performance than the other versions, with close levels of accuracy in shorter processing time.

Executed in 300 epochs each, the 17th round (“/runs/train/exp17”) of training generated satisfactory results, observable in Figure 4, locating the wagon codes in different image conditions, and compiling the weights file. “best.pt”, which was used in the next stage of project development.

Figure 4. Predictions obtained with the “best.pt” file

3.4 *Character identification algorithm*

After training the wagon code location model (“plates”), we carried out a survey of 400 images that were legible and identifiable to the human eye, and thus a validation dataset was created for the application of the character reading algorithm. This dataset contains images with different environmental conditions (day, night, rain), lighting, carriage painting, dirt, among other aspects that influence code reading.

Next, we developed the algorithm for recognizing text in images, that is, reading the wagon plates. For this process, we searched for some optical character recognition (OCR) libraries, and based on preliminary comparative studies, we chose the EasyOCR model (from Jaided AI), as we consider it more suitable for application to the types of images that appear in the dataset. our project.

Using Visual Studio Code, developed by Microsoft, we adapted an OpenCV code with EasyOCR packages. We import and install the necessary auxiliary libraries and pass as a parameter the file “best.pt” (renamed to “best_gde.pt”), which contains the best training weights for locating the wagon “plates” in the images, reading of the characters in the captured images, as shown in figure 5.

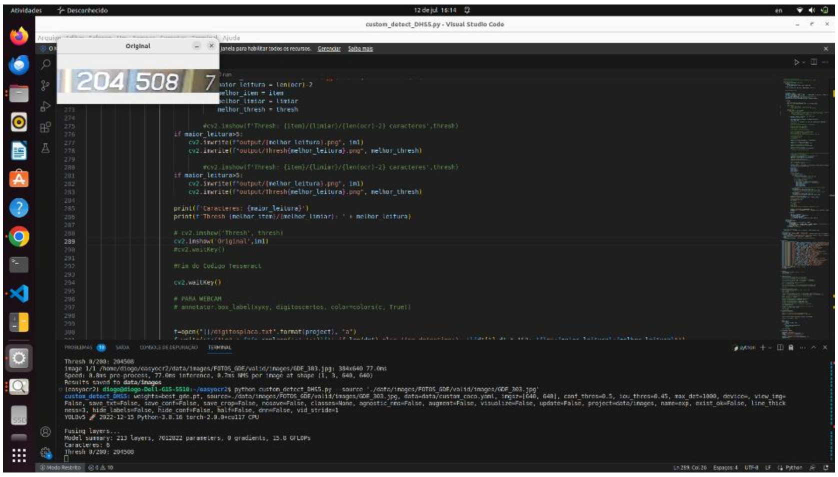

Figure 5. Modal displaying the wagon “plate” segmentation and reading

Based on the ABNT coding standard, we developed a check digit calculation algorithm to evaluate the OCR reading.

Considering that all wagons are necessarily registered in Vale's internal system, we obtained a list of active wagons. With this list, we created a function to validate the reading carried out by the OCR, comparing the result with the list of active wagons in the company's system. The results of readings that are not included in the list of active cars are rejected directly in the code. Therefore, the results obtained only contain codes for existing and active wagons.

4 Results

4.1 *YOLO Results (Location of Wagon Codes)*

The results obtained in the training and validation stages of the YOLOv5 model for the location of the plates can be seen in Figure 51, in the format of the graphs of train/box_loss, val/box_loss, train/obj_loss, val/obj_loss, metrics/precision, metrics/recall, metrics/mAP_0.5 and metrics/mAP_0.5:0.95).

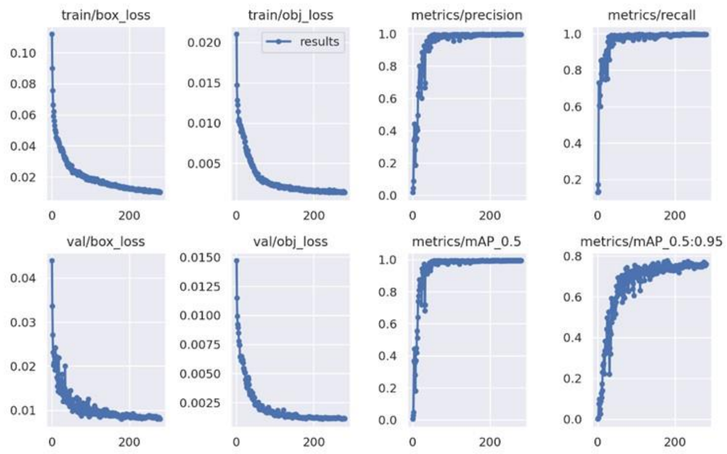

Figure 6. Results obtained with the YOLOv5 network

The metrics/precision graph, calculated by the formula: “true positives” / (“true positives” + “false positives”), evaluates how often the model correctly classifies an image into a specific category. The results reached maximum values of 99.86% and an average of 93.84%.

The metrics/recall graph, resulting from the calculation: “true positives” / (“true positives” + “false negatives”), shows the percentage of images that were classified in a category, compared to all images that should have been classified in this category. The results show maximum and average values of 99.86% and 96.33%, respectively.

The mAP (Mean Average Precision) indicator measures how accurate the model is, that is, the percentage of classified objects that are correct.

As we can see in the metrics/mAP_0.5 graph, the results show maximum and average values of 99.47% and 94.62%, respectively.

4.2 *Easyocr results (character identification)*

The results obtained when applying Easyocr can be seen in Table 1.

Correctly identified characters	**Number of wagons**	**Percentage (%)**
6 characters (all)	356	89%
5 characters	25	6%
Less than 5 characters	19	5%
TOTAL	**400**	**100%**

Table 1. Results obtained with Easyocr

We can see that 356 wagons had their codes completely identified, representing approximately 89% of the total dataset.

If we consider that in the case of reading numerical sequences, each digit is identified independently and separately, the number of characters that were identified in the correct positions totaled 2,342 characters out of a total of 2,400 possible characters (400 characters x 6 digits). Thus, we can determine an accuracy of reading independent characters of approximately 98%.

5 Conclusions

5.1 *Conclusions*

The main technological advances in computer vision in recent years enable the creation and application of tools in industrial processes, which, due to resource limitations and particularities, were not economically and/or technically viable for most companies. The formation of trains at EFVM is an example of this unfeasibility, as the need for scalability, the activity operated with risks and was deprioritized in improvement forums.

The results obtained in the project by the algorithm show that the model is capable of carrying out the conference activity through cameras. The satisfactory results verified in the metrics (precision, recall, mAP) of YOLO (location of wagon codes) and in Easyocr (character identification) show that the application can surpass human capacity in validation and reading speed. In relation to ease of replicability and cost, full coverage of the locations where this activity takes place can be envisaged.

5.2 *Future Work*

In future work, we intend to carry out experiments in real time and integrate the solution with the company's internal system.

References

[1] ANTF. 2022. Informações gerais / O setor ferroviário de carga brasileiro. ASSOCIAÇÃO NACIONAL DOS TRANSPORTADORES FERROVIÁRIOS. [Online] 15 de Janeiro de 2022. https://www.antf.org.br/informacoes-gerais/.

[2] VALE S/A. 2023. Espaço Memória - VALE. Logística VALE. [Online] 20 de Maio de 2023. https://www.vale.com/pt/espaco-memoria.

[3] ABNT. 2019. ABNT NBR 11691. Vagão Ferroviário - Classificação, Identificação e Marcação. s.l. : Associação Brasileira de Normas Técnicas, 2019.