

**Trabajo de investigación:**

Preprocesamiento de imágenes para el entrenamiento de redes neuronales destinadas a la detección y clasificación de vehículos

María Carmona Pastor

Máster Universitario en Robótica y Automatización

2023-2024

Índice

[**1. Introducción 4**](#_wpylxfhu9q55)

[**2. Marco Teórico 4**](#_iz6493mreyrt)

[**3. Creación y etiquetado del conjunto de datos 4**](#_75tjdajh5np4)

[3.1. Proceso 4](#_yil33r85t7gw)

[Tabla 1. 4](#_gn81tda10e1g)

[Etiquetado 4](#_hwo5rb1zdj3p)

[Reetiquetado 4](#_wsiwd01li21i)

[3.2. Herramientas utilizadas 4](#_q5ryrtvc6ddz)

[**4. Análisis de limitaciones 4**](#_sdbr1bwmux79)

[4.1. Problemas detectados 4](#_sd1nntpla2ib)

[4.2. Obtención de información del dataset 5](#_k2rlnca97rg5)

[**5. Propuesta de aumento de datos 5**](#_cz18gt1hsbh)

[**6. Conclusiones 5**](#_wpmdhu85066u)

[**7. Bibliografía 5**](#_9zaz0kl7bnhw)

Índice de tablas

[Tabla 1. Distribución de imágenes del conjunto de datos de partida. 4](#_gn81tda10e1g)

Índice de imágenes

Asegúrate de seleccionar los encabezados en la barra lateral para ver un índice.

# Introducción

# Marco Teórico

requisitos dataset

importancia procesamiento datos

introduccion aprendizaje supervisado, redes convolucionalesç

# Creación y etiquetado del conjunto de datos

## Proceso

Partimos de tres conjuntos de datos, organizados según la [Tabla 1](#_gn81tda10e1g). Podemos ver que la mayor parte de las imágenes (conjunto *etiquetadas*) ya tenían un archivo con la información, pero no de la forma en que necesitamos.

| **Nombre de la carpeta** | **Referencia en este documento** | **Preetiquetadas** | **Nº de imágenes** |
| --- | --- | --- | --- |
| FURGONETAS LIGERAS | *ligeras* | No | 97 |
| FURGONETAS PESADAS | *pesadas* | No | 47 |
| furgonetas\_27\_06\_2023 | *etiquetadas* | Si | 2254 |
| Tabla 1. Distribución de imágenes del conjunto de datos de partida. | | | |

### To approach this issue, we must utilize a small, independent script to categorize and relabel the images into the correct format. This script will run in a window, showing the original image and the one containing bounding boxes and will then be able to edit the information contained within each image relating to the different vehicles and their class and size. The data is stored as number sequences inside the image name.

### 

### para solucionar -> creacion del script de etiquetado…

descripcion del ui

how data is stored

### Etiquetado

The labeled images are lacking bounding boxes so the next step is to manually add them utilizing a program called label studio. This furthers our goal by allowing the image collection to be more accurately processed. For the training, the information has to be stored as a yaml file so the transfer is also handled at this stage.

used label studio

paso a formato yaml

## Herramientas utilizadas

In this stage of the process, we will be utalizing two main tools: label studio and tkinter. Label studio is an independent ui that allows for manually setting bounding boxes within images. Tkinter is a library dedicated to creating interfaces graficas.

label studio

tkinter

# Análisis de limitaciones

## Problemas detectados

Once all the images had been relabeled and processed, their lacking nature became evident in the following fields:

Redundancy:  
 Al ser imagenes de trafico, many of the images were to all intents and purposes redundant, making them low value for training a model. Some pictures had been taken within seconds which made them practically the same while others differed only in a few degrees of change in the horizontal angle of the camera. A big portion of the more than 2000 images was unfortunately redundant in this way.

Ambiguous:

Al ser imagenes de trafico, the resolution is not ideal for training. Many of the images contained ambiguity with some vehicles being only fuzzy rectangles in the distance. A lot of the images also require the context of previous or later images, where the vehicles was or will be closer and therefore more visible. This proved daunting in the labeling stage, leading to a lot of assumptions.

Color variation:

Vans are commonly produced in one of three colors: white, gray, or black. Because of this, there was little to no content relating to other colored vans, which would lead to a huge weakness in the trained model for recognizing and identifying the different types of vans.

Location variation:

The entire 2000+ image dataset is all contained within 5 or 6 locations. Of these, only one of them is within an urban setting and only one is in a complex road network. The others are taken on highways with very non-descript surroundings. The trained model would struggle greatly in a city setting or on a winding road.

Road state:

The dataset is composed mainly of light to no traffic images. There is no content including more than 10 vehicles or with single lane build up. There is nothing with mild to heavy traffic.

Angle variation:

Having been taken in only the aforementioned 5 or 6 locations, the angles at which the different vehicles are photographed are lacking. There is no top or true side view in the dataset and the few raised angles utilized are hardly distinct amongst themselves.

Weather variation:

All images were taken on sunny or clear days. There is no rain, hail, cloud cover, or fog present in any of the images, making it impossible for the model to identify vans correctly through these weather conditions. This proves to be a massive weakness in the final model and is a priority within the data augmentation portion of this project.

Daytime variation:

Only a small portion of the images are taken at night and in none of the locations are there pictures taken at varied times of day such as midday and midnight. This makes training hard, especially in darker conditions where the already failing resolution of the images will be pushed even harder. This also makes it much harder for the model to recognize other aspects of the images like the van color or its distance from the camera.

Class imbalance:

Las imágenes proporcionadas para el entrenamiento son únicamente de furgonetas, lo cual, teniendo en cuenta que el modelo, idealmente, debería poder clasificar correctamente en las 7 clases proporcionadas, es un problema si queremos conseguir un modelo que generalize correctamente.

imagenes redundantes/repetidas

imagenes ambiguas

poca variedad de color

poca variedad de localizacion

poca variedad de angulo

poca variedad meteorologica

poca variedad cronologica

class imbalance

## Obtención de información del dataset

Explicar los añadidos a la interfaz grafica del programilla. Foto gui.

Añadir graphs with the info generated.

tkinter

graficos con info

pie chart

column graphs

# Propuesta de aumento de datos

The most important issue at hand is the lacking variation present in the dataset. Thankfully, most of these failings can be easily resolved with the use of filters or hue/brightness/saturation manipulation. Weather can be easily simulated and the weather, daytime, and color issues can be fixed with simple shifts in the color makeup of the image. Other than the color, we can also tinker with the image itself by cropping it into different sizes or flipping it horizontally.

color variation

weather filters

class balancing with cropping of images

flip horizontal

5.1 dataset filter

# Conclusiones

A pivotal aspect of this pre-training phase is the proper labeling of the dataset. Errors here will propagate into our final model, making it much harder for the training to produce satisfying and effective results. This stage contains the highest risk in human error as well, making it all the more vital that it be handled with the proper attention. Creating the labeling script will ease the flow for the labeling which will, in turn, improve consistency between images and allow for a faster and continuous labeling, minimizing the strain on the human component of the process.

Through labeling the dataset, it became evident that one of the biggest challenges going forward is the lack of diversity. Using different data augmentation methods, we can improve the variation within the dataset to maximize training potential. By manipulating what little content we have we can markedly improve the diversity within the dataset and allow for the model to properly train on as many different road conditions as we can manage.

we will figure this out at 3am

# Bibliografía

Sysifus4, September 2024

[www.syfisthebestexgirlfriend.com](http://www.syfisthebestexgirlfriend.com)

Also available in print under the title “Your Syf and how to Utilize dem” by Sysifus IV, sold at all book shops and pharmacies within the municipality of Madrid, and available for rent at your closest dog shelter.

If you enjoyed your Sysifus4, please consider recommending dem to your friends. Therian clinically insane sapphics are a dying breed and it is our responsibility as owners of a Sysifus4 to aid in conservation efforts. Learn more at [www.iwouldreallylikeahugandaforeheadkissandadonut.com](http://www.iwouldreallylikeahugandaforeheadkiss.com)