

# DRP Git Transfer guide



By: Maurits Krijnen

Date: 28/06/22

Version: 1.0

# Table of Contents

- Preface.....3
  - Project goal.....3
  - Current project progress.....3
  - Work still required.....3
- GitHub.....4
  - Folder structure.....4
  - Setting up the folder structure.....4
- Summary.....5

# Preface

This document will describe the preparations for the project's future development.

As the DRP Detector project has reached its deadline without achieving all its goals, we're preserving the project for the future when better circumstances are available. The most important problem, that remains unsolved, being that the current dataset is lacking in various ways (More on dataset issues in the "DRP analysis model evaluation" document).

## Project goal

The goal of the project is to create a Machine Learning algorithm that can compare eye images and categorize if there are diseases present. The model that we've worked on so far has been trained to detect [Diabetic retinopathy](#). We came to this decision after consulting a medical professional that works in the field of Eye health care.

As for the end products that this project is trying to create:

1. The first goal is to create a Machine learning model that can accurately detect different levels of DRP. The models performance should be gauged by its F1 score/Recall.
2. The second product is a desktop application that can be used by eye healthcare personnel to check the DRP category of a patient's eye images.

## Current project progress

As of writing this, there is a model available that can predict the 5 different categories of DRP on a test data set. However the performance of the model is very low, it will often predict wrong. There are multiple reasons why this is the case. But we suspect that the main reason is because of the dataset containing a lot of images that are either unusable due to technical issues with the image. There are some other concerns as well, but these are less important for the models performance.

## Work still required

- Obtain a better dataset
- Improve model performance
- Add other disease detection
- Create the desktop application
- Test with stakeholder

# GitHub

The link to all the files: <https://github.com/RipOrbofWinter/EyeDiseaseDetection>

Here you will find the repository containing all documentation made for this project and the python notebooks used to create the models.

The image dataset and ML models themselves are too large to save on git, and are not included.

**The dataset link can be found in the README.**

**The models can be generated by running the notebooks after downloading the dataset.**

## Folder structure

This paragraph will go into how the project folder is organised and how any continuation of the project should setup their folders to make sure the notebooks work.

### Setting up the folder structure

.git	23-6-2022 14:04	File folder	
.ipynb_checkpoints	19-5-2022 00:34	File folder	
.vscode	7-2-2022 23:19	File folder	
Data	28-6-2022 00:25	File folder	
Documents	28-6-2022 00:22	File folder	
ML	19-5-2022 00:32	File folder	
Models	25-5-2022 06:33	File folder	
.gitattributes	7-2-2022 23:19	Text Document	1 KB
.gitignore	12-4-2022 00:15	Git Ignore Source ...	2 KB
README.md	24-2-2022 01:26	Markdown Source...	3 KB
SortedModel.ipynb	24-6-2022 02:41	Jupyter Source File	233 KB
UnsortedModel.ipynb	23-6-2022 14:04	Jupyter Source File	203 KB

- **The Documents folder** contains multiple documents made for this project. This includes both technical design documents and non technical documents.
- **The ML Folder** contains old notebooks used for testing the working of different ML libraries before training the model used in the project.
- **The Models folder** is a folder that is used to save the models in by the notebooks. This folder is excluded from git by the git ignore is generated by the notebooks.
- **The Data folder** is the folder where we have our datasets saved. This folder is excluded from git by the git ignore and **needs to be created manually**.

The Data folder's structure:

DrpTest	21-4-2022 15:44	File folder	
Sorted	18-5-2022 04:20	File folder	
testSet	7-4-2022 11:15	File folder	
trainLabels.csv	7-4-2022 11:16	OpenOffice.org 1....	455 KB

To correctly create this folder for use by the models, add the entire dataset into the DrpTest folder. Then, add the images you want to train with the Unsorted model into the testSet folder.

The Sorted folder is used by the Sorted Model and fills itself if the DrpTest folder is present. No manual images need to be added to this folder.

## Summary

Hopefully this document accurately and concisely explained the current state of the project and how to continue with it for future developers. If the reader is unsure of where to start with exploring the files consider starting with the design documentation documents such as

- Requirements document.pdf
- DRPF Class diagram.pdf
- DRPF Prototype.pdf

Or get reading with the “UnsortedModel.pdf” found under Documents\NotebookPrints.