

# Universidade do Minho

Mestrado em Engenharia Informática



Visão por Computador e Processamento de Imagem

1º Ano, 2º Semestre

Ano letivo 2021/2022

1 de junho de 2022

Cláudio Moreira - PG47844

Filipe Fernandes - A83996

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b>Arquitetura da solução</b>	<b>3</b>
2.1	Análise do <i>Dataset</i> . . . . .	3
2.2	Estratégias tomadas . . . . .	4
<b>3</b>	<b>Resultados obtidos</b>	<b>5</b>
3.1	Comparações . . . . .	5
<b>4</b>	<b>Conclusão</b>	<b>6</b>

# 1 Introdução

Neste relatório, no âmbito da unidade curricular de Visão por Computador e Processamento de Imagem, pretende-se explorar modelos de deep learning aplicados ao dataset GTSRB para obter a melhor accuracy possível. Dataset esse que é constituído por um conjunto de imagens para treino e teste.

Para isso, numa fase inicial, irão ser aplicados algoritmos de treino aplicando data augmentation, não só em pré-processamento como também de uma forma dinâmica. Pretende-se assim, avaliar os resultados obtidos através das diferentes transformações, mas também explorar os diferentes filtros e métodos de processamento de imagem que podem ser aplicados nos algoritmos.

Numa fase posterior do trabalho, foi-nos proposto a exploração do potencial dos ensembles de redes através das redes construídas na primeira fase.

Assim sendo, iremos analisar em particular os diferentes métodos de data augmentation e ensemble de redes de modo a obter conclusões nos efeitos que os mesmos provocam no dataset definido.

## 2 Arquitetura da solução

### 2.1 Análise do *Dataset*

Foi realizado um script (ppmToPng.py) para a conversão das imagens de ppm para png. De forma a ter uma melhor familiarização com o dataset foi feita uma análise deste, onde não foram encontradas imagens fora das suas classes, nem imagens com label erradas. No entanto, foi detectada um grande desbalanceamento do dataset.

De modo a balancear, foi criado o notebook imgesCreate que contém várias funções de transformação (brightness, contrast, shear, blur, zoom, centershear, randomRotation, rotation, horizontalShift, VerticalShift, translation, box, e flip). Posteriormente, para um threshold pré-estabelecido - neste caso 1500 - são verificadas as classes que têm um número inferior a este threshold e irão ser aplicadas as transformações às imagens dessas classes de forma random. Para os pares de classes 19-20, 33-34, 36-37, 38-39 foi aplicada a função de flip, antes da aplicação das transformações, para gerar novas imagens. E assim foi balanceado o dataset.

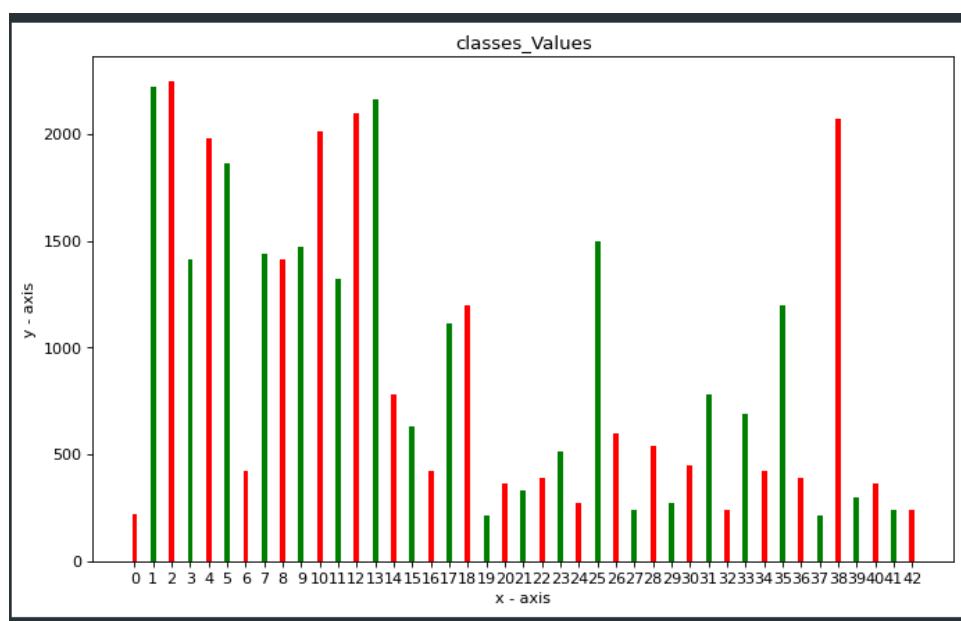


Figura 1: Dataset Desbalanceado

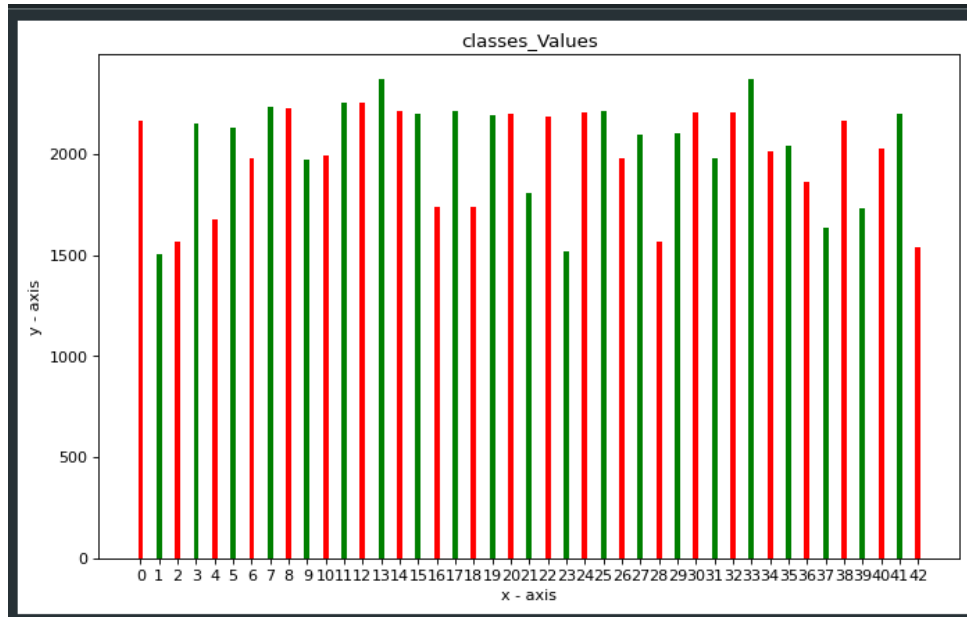


Figura 2: Dataset Balanceado

## 2.2 Estratégias tomadas

Para o presente trabalho, foi proposta a exploração de modelos de Deep Learning. Numa primeira parte, é suposto treinar modelos aplicando data augmentation, tanto em pré-processamento como dinâmico. Começou-se por executar uma exploração do dataset sem realizar nenhum tipo de alterações, com o intuito de perceber a qualidade do mesmo. Esta execução encontra-se no notebook gtsrb-no-augmentation.

De seguida, aplicaram-se alterações ao nível do pré-processamento. A utilização deste tipo de funções tem como objetivo o aumento significativo do número de imagens através de uma implementação de funções simples. Para isso utilizaram-se funções de rotação, translação e coloração (aplicação de alterações do brilho, contraste, saturação e hue). A execução descrita anteriormente encontra-se no dataset gtsrb-Static-augmentation.

Seguidamente, passou-se para a fase de implementações de alterações dinâmicas. Para isto, utilizou-se uma função de processamento de imagem fornecida pelo docente da unidade curricular, de modo a realizar uma transformação do dataset de forma dinâmica. Dentro da função de processamento são realizadas alterações ao nível da rotação, translação e coloração do dataset. Esta execução encontra-se no notebook gtsrn-Dynamic-Augmentation. Finalmente, abordou-se a massive data augmentation que realiza um processamento bastante mais complexo das imagens. Para isto, foram utilizadas funções de ajuste do brilho, contraste, saturação, hue, rotação, rotação, translação e redimensionar as imagens. A execução descrita anteriormente encontra-se no dataset gtsrb-Massive-augmentation.

## 3 Resultados obtidos

### 3.1 Comparações

Após a aplicação do modelo e respetivo treino obtiveram-se os seguintes resultados:

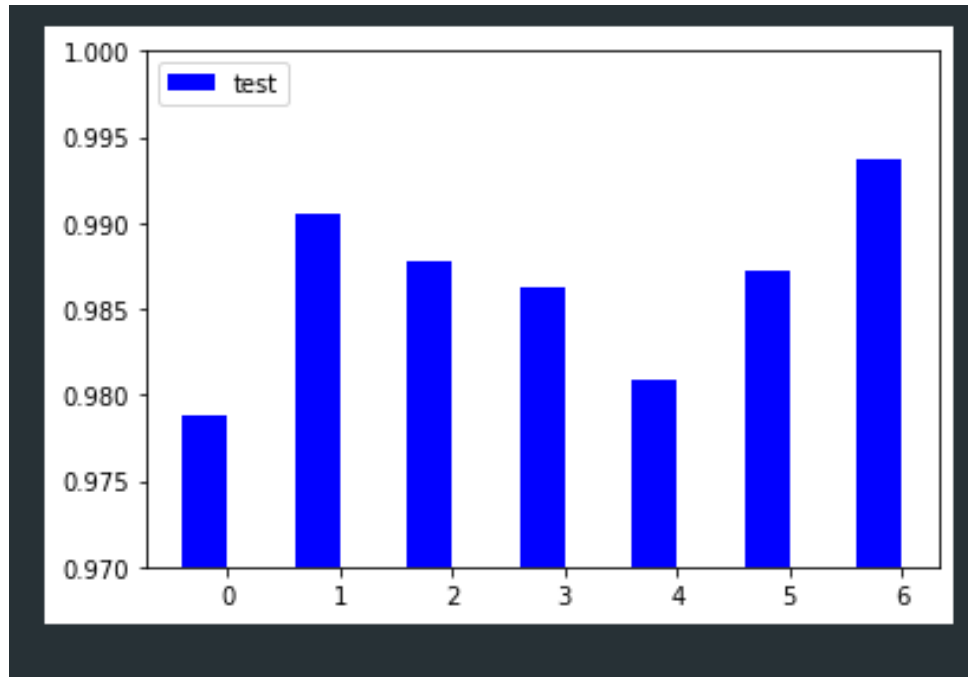


Figura 3: Gráfico de comparações

0. noAugmentation = 0.9788598418235779

1. noAugmentation (Utilizando datasetBalanceado) = 0.9905779957771301

2. static-Augmentation = 0.9878067970275879

3. static-Augmentation (Utilizando datasetBalanceado) = 0.9862232804298401

4. static-Augmentation (Utilizando mais tipos de transformações) = 0.9808392524719238

5. Dynamic-Augmentation = 0.9872525930404663

6. Massive-Augmentation = 0.993665874004364

Sendo assim, podemos verificar que houve uma clara melhoria entre o modelo sem data augmentation e dataset normal e o dataset data augmentation com o dataset balanceado, obtendo melhores resultados que os dataset com static e dynamic augmentation. Por fim, podemos ver que o Massive augmentation obteve a melhor classificação.

No geral é possível verificar que os modelos falham

Em relação ao ensemble de redes: os resultados obtidos foram muito abaixo do esperado com uma average accuracy de 0.612 o que nos fez querer que existiu um erro de implementação nesta secção. Não existe grande sentido em realizar uma análise ao resultado obtido devido ao valor da accuracy ser tão disperso de o que é normal.

## 4 Conclusão

Podemos concluir que data augmentation através de pré-processamento os resultados obtidos em todas as formas foram positivos, obtendo accuracy scores sempre acima dos 0.977. Como era de esperar, devido à complexidade e ao número de funções implementadas, a implementação com maior accuracy-score foi a Massive-Augmentation. Para além disso, podemos concluir que a aplicação de métodos de data augmentation tanto de pré processamento como dinâmicos implica um aumento considerável da qualidade das imagens (subida de noAugmentation para qualquer um dos outros métodos de processamento).

Em relação ao ensemble de redes, podemos concluir que os resultados obtidos não foram os esperados devido a serem tão baixos.

No entanto, pensamos que o objetivo do trabalho foi alcançado devido aos accuracy scores obtidos na 1<sup>o</sup> parte do projeto.