

Ficha Prática 2

Experimental Setup

1. **Google Colab** – Crie um script, utilizando o Google Colab, capaz de ler o conjunto de imagens “GTSRB_Classification.zip” e dividir o conjunto de dados em três sub-conjuntos disjuntos, para “Aprendizagem”, “Validação” e “Teste”, com base num valor-proporção (para treino) desejado.

Por exemplo, ao especificar o valor “0.6, 0.2” estamos a determinar que o conjunto de aprendizagem fique com 60% dos dados, o conjunto de validação terá 20% dos dados, enquanto o conjunto de teste ficará com os restantes 20%.

O script deverá gerar um ficheiro “.csv” com o seguinte formato:

Imagem_path.jpg, <CLASSE>, <TREINO|VALIDAÇÃO|TESTE>

Em que <CLASSE> consiste num valor inteiro entre 0 e o total de sinais disponíveis e <TREINO|VALIDAÇÃO|TESTE> consiste numa *flag* a indicar o respetivo conjunto onde o elemento será integrado (0=Treino; 1=Validação; e 2=Teste).

Exemplo:

```
Data/Imagem_1.jpg, 0, 0  
Data/Imagem_2.jpg, 1, 1  
Data/Imagem_3.jpg, 1, 0  
Data/Imagem_4.jpg, 2, 2  
Data/Imagem_5.jpg, 2, 0  
Data/Imagem_6.jpg, 0, 1  
Data/Imagem_7.jpg, 3, 0  
Data/Imagem_8.jpg, 3, 0  
...
```

