

Dataset Augmentation Tool Kullanım

Kullanılan kütüphaneler:

- cv2
- numpy
- random
- argparse
- shutil
- os

Kullanım Amacı: Girdi olarak verilen veri setine çeşitli augmentasyonlar uygulamak.

Beklenen Girdi Formatı: Klasör içinde bulunan “.jpg”, “.png” formatındaki göreseller ve onların “.txt” formatındaki label dosyaları.

Fonksiyonlar ve Kullanımları

Fonksiyonlar kullanılmadan önce terminale “python tool.py” yazılmalıdır.

“--input_folder” ve “--output_folder”: Bu iki fonksiyon girdi ve çıktı klasörlerini belirtmek için kullanılır. Program girdi olarak klasör beklemektedir, bu sebeple girdi klasörü olarak dosya konumu verilemez.

Örnek kullanım:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout
```

Bu şekilde sadece “--input_folder” ve “--output_folder” fonksiyonları verilirse program girdi klasöründeki bütün dosyaları girdi klasörünün yapısından bakmadan çıktı klasörünün içine kopyalar. (Orijinal yapıyı korumaya çalışmaz sadece girdi klasörünün dosyalarını çıktı dosyasına koyalar)

Program hiçbir şekilde girdi klasörü üzerinde değişiklik yapmaz. Bu iki fonksiyonun kullanılması zorunludur.

--random_zoom: Veri setinin belirlenen yüzdesine random zoom işlemi uygular. Üç tane argümana sahiptir:

- percentage (float,default=0.1)
- area_size (int,default=640)

- save_original (boolean,default=True)

“percentage” veri setinin yüzde kaçına işlem uygulanacağını belirtir, “area_size” zoom yapılacak kare alanın kenar boyunu belirtir (eğer 320 verilirse 320x320’lik kare bir alan alır), “save_original” “True” verilirse augmented edilen görsellerin orijinali klasörde bulunmaya devam eder aksi takdirde zoom yapılan görselin orijinali silinir. Bu parametrelerin tamamı ya yukarıda verilen sıra ile eksiksiz girilmelidir ya da default değerlerin kullanılması için hiçbirisi girilmemelidir.

Örnek kullanım:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --random_zoom 0.5 256 False
```

(Resimlerin yarısına rastgele zoom uygula, boyutları 256x256 olsun, orijinal fotoğrafları kaydetme)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --random_zoom
```

(Default değerleri kullan)

--pad_and_resize: Görsellerin belirli yüzdesine black padding ekleyerek kareye çevirir ve belirlenen boyuta tekrar boyutlandırır. Üç tane argümana sahiptir:

- percentage (float,default=1)
- output_size (int,default=640)
- save_original (boolean,default=False)

“percentage” veri setinin yüzde kaçına işlem uygulanacağını belirtir, “output_size” çıktı boyutunu belirtir, “save_original” “True” verilirse augmented edilen görsellerin orijinali klasörde bulunmaya devam eder aksi takdirde padding yapılan görselin orijinali silinir. Bu parametrelerin tamamı ya yukarıda verilen sıra ile eksiksiz girilmelidir ya da default değerlerin kullanılması için hiçbirisi girilmemelidir.

Örnek kullanım:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --pad_and_resize 0.7 352 True
```

(Görsellerin %70’ine padding uygular ve boyutlarını 352x352 olarak ayarlar, orijinal resimleri tutar)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --pad_and_resize
```

(Default değerleri kullan)

--blur: Görsellerin belirlenen yüzdesini belirlenen oranda bulanıklaştırır. Üç tane argümana sahiptir:

- percentage (float,default=0.1)
- blur_margin (int,default=5)
- save_original (boolean,default=True)

“percentage” veri setinin yüzde kaçına işlem uygulanacağını belirtir, “blur_margin” bulanıklık payını belirtir, “save_original” “True” verilirse augmented edilen görsellerin orijinali klasörde bulunmaya devam eder aksi takdirde bulanıklaştırma işlemi yapılan görselin orijinali silinir. Bu parametrelerin tamamı ya yukarıda verilen sıra ile eksiksiz girilmelidir ya da default değerlerin kullanılması için hiçbirisi girilmemelidir.

Örnek çalışma:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --blur 0.3 7 True
```

(Görsellerin %30'una blur uygular, bulanıklaştırma payını 7 alır, orijinal resimleri tutar)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --blur
```

(Default değerleri kullan)

--flip: Görsellerin belirlenen yüzdesine flip işlemi uygular. Üç tane argümana sahiptir:

- percentage (float,default=1)
- flip_mode (str,default="horizontal")
- save_original (boolean,default=True)

“percentage” veri setinin yüzde kaçına işlem uygulanacağını belirtir, “flip_mode” hangi flip işleminin yapılacağını belirtir horizontal, vertical ve both (hem dikey hem yatay çevirme) olmak üzere kabul edilen üç modu vardır, “save_original” “True” verilirse augmented edilen görsellerin orijinali klasörde bulunmaya devam eder aksi takdirde flip işlemi yapılan görselin orijinali silinir. Bu parametrelerin tamamı ya yukarıda verilen sıra ile eksiksiz girilmelidir ya da default değerlerin kullanılması için hiçbirisi girilmemelidir.

Örnek çalışma:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --flip 1 horizontal True
```

(Görsellerin hepsine horizontal flip uygular, orijinal resimleri tutar)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --flip
```

(Default değerleri kullan)

--brightness: Görsellerin belirlenen yüzdesini belirlenen oranda parlaklaştırıp karanlıklaştırarak 2 farklı görsel üretir. Üç tane argümana sahiptir:

- percentage (float,default=0.1)
- brightness_factor(float,default=0.1)
- save_original (boolean,default=True)

“percentage” veri setinin yüzde kaçına işlem uygulanacağını belirtir, “brightness_factor” görselin hangi oranda parlaklaştırıp karanlıklaştırılacağını belirler, “save_original” “True” verilirse augmented edilen görsellerin orijinali klasörde bulunmaya devam eder aksi takdirde brightness işlemi yapılan görselin orijinali silinir. Bu parametrelerin tamamı ya yukarıda verilen sıra ile eksiksiz girilmelidir ya da default değerlerin kullanılması için hiçbirisi girilmemelidir.

Örnek çalışma:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --brightness 0.5 0.6 False
```

(Resimlerin yarısını 0.6 brightness factor ile parlaklaştırıp karanlıklaştır, orijinal resimleri tutmaz)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --brightness
```

(Default değerleri kullan)

--grayscale: Görsellerin belirlenen yüzdesine grayscale işlemi uygular. İki tane argümana sahiptir:

- percentage (float,default=0.1)
- save_original (boolean,default=True)

“percentage” veri setinin yüzde kaçına işlem uygulanacağını belirtir, “save_original” “True” verilirse augmented edilen görsellerin orijinali klasörde bulunmaya devam eder aksi takdirde grayscale işlemi yapılan görselin orijinali silinir. Bu parametrelerin tamamı ya yukarıda verilen sıra ile eksiksiz girilmelidir ya da default değerlerin kullanılması için hiçbirisi girilmemelidir.

Örnek çalışma:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder /Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --grayscale 0.4 True
```

(Resimlerin %40'ına grayscale işlemi uygular orijinal tesimleri tutar)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --grayscale
```

(Default değerleri kullan)

--exposure: Görsellerin belirlenen yüzdesini exposure işlemi uygular ve underexposed ve overexposed iki resim oluşturur. Dört tane argümana sahiptir:

- percentage (float,default=0.1)
- exposure_factor(float,default=0.1)
- contrast_factor(int,default=0.1)
- save_original (boolean,default=True)

“percentage” veri setinin yüzde kaçına işlem uygulanacağını belirtir, “exposure factor” expose oranını belirler, “contrast_factor” kontrast faktörünü belirler, “save_original” “True” verilirse augmented edilen görsellerin orijinali klasörde bulunmaya devam eder aksi takdirde exposure işlemi yapılan görselin orijinali silinir. Bu parametrelerin tamamı ya yukarıda verilen sıra ile eksiksiz girilmelidir ya da default değerlerin kullanılması için hiçbir girilmemelidir.

Örnek çalışma:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --exposure 0.8 0.5 2 True
```

(%80'ine uygula, exposure_factor=0.5 olarak, contrast_factor=2 olarak ayarla, orijinalleri tut)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --exposure
```

(Default değerleri kullan)

--split: Çıktı klasörünü train, test ve valid olarak 3 klasöre ayırır. Klasörlerin içinde görselleri tutan images klasörü ve labelları tutan labels klasörleri vardır. 2 argüman alır:

- train (float,default=0.7)
- test (float,default=0.1)

“train” train klasörünün içine gidecek görsel yüzdesini, “test” test klasörünün içine gidecek olan görsel yüzdesini verir. Geri kalan görseller valid klasörünün içine yerleştirilir.

Örnek çalışma:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --split 0.6 0.2
```

(%60 train, %20 test, %20 valid olarak ayırır)

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --split
```

(Default değerleri kullan)

Karma kullanım örneği:

```
(.toolenv) dataozk@Doruks-MacBook-Air DatasetTool % python tool.py --input_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/dorukcopy --output_folder  
/Users/dataozk/Desktop/DeliveryOne/DatasetTool/testout --pad_and_resize --grayscale 0.5  
True --split 0.8 0.1
```

(Default değerlere göre padding ve resizing işlemi yap, yarısını grayscale işlemine tut vve orjinallerini tut, veri setini %80 train, %10 test, %10 valid olarak ayır)