

Biblioteki cyfrowe metadane

Jakub Wierzbą 184834

15 kwietnia 2024

1 Wstęp

Klasą obiektów wybranych w ramach projektu są znaki drogowe.

2 Implementacja

Do wydobywania exifów ze zdjęć posłużono się pakietem `exif`, z którego wykorzystano obiekt `Image`, który pozwala wydobyć właściwości zdjęcia.

W celu generacji podpisów wykorzystano pakiet `transformers`, a z niego obiekty `BlipProcessor`, `BlipForConditionalGeneration`, które wykorzystują gotowe generatory opisów ze strony:

<https://huggingface.co/Salesforce/blip-image-captioning-base>

3 Wydobywanie exifów

```
with open(os.path.join(images_directory, imageFile), 'rb') as image_file:
    image = Image(image_file)
    # add exif
    for key in dir(image):
        if key != '_segments':
            val = image.get(key)
            # check if it's a method
            if not hasattr(val, '__self__'):
                exif = SubElement(exif_node, key)
                exif.text = str(val)
```

4 Generowanie podpisu

```
# utworzenie modelu i procesora wejścia
processor = BlipProcessor.from_pretrained("
----Salesforce/blip-image-captioning-base")
model = BlipForConditionalGeneration.from_pretrained("
----Salesforce/blip-image-captioning-base")

# Wczytanie
raw_image = ImageMod.open(os.path.join(images_directory , imageFile))
    .convert('RGB')

# transformacja do inputu modelu
inputs = processor(raw_image , return_tensors="pt")

# generacja podpisu
out = model.generate(**inputs)
```

5 Skuteczność generacji podpisów

Generator podpisów nie był w stanie na żadnym zdjęciu rozpoznać jaki znak znajduje się na zdjęciu. Co najwyżej potrafił określić kolor znaku, bądź na czym wisi. Na kilku zdjęciach rozpoznał samochód zamiast znaku drogowego. Wygenerowane opisy są bardzo ogólne.