

IO, Grupa I	Karolina Kurowska Dominik Klich	Sprawozdanie I - Analiza celu projektu
-------------	------------------------------------	--

WikiArt Art Movements/Styles

Temat projektu

Celem projektu jest stworzenie modelu sztucznej inteligencji zdolnego do klasyfikacji obrazów na podstawie ich stylu artystycznego. Model ten będzie wykorzystywał dane z zestawu **WikiArt Art Movements/Styles**, który zawiera obrazy przypisane do różnych stylów sztuki. Zastosowaniem naszego projektu może być pomoc w organizacji dzieł w muzeach sztuki, w pomocy rozróżniania stylów artystycznych itp..

Zakres projektu

- Przygotowanie i eksploracja zbioru danych
- Ewentualne wstępne przetwarzanie obrazów (normalizacja)
- Wybór odpowiedniej architektury modelu - CNN
- Trenowanie i walidacja modelu
- Testowanie dokładności na zbiorze testowym
- Optymalizacja i wdrożenie

Wymagania techniczne

1. **Dane wejściowe:** Wylosowane obrazy z różnych stylów artystycznych, 13 kategorii: Academic Art, Art Nouveau, Baroque, Realism, Renaissance (Western), Romanticism, Expressionism, Japanese Art, Neoclassicism, Primitivism, Rococo, Symbolism, Western Medieval.

W niektórych przypadkach style zostały połączone w jedną kategorię (np. sztuka bizantyjska i średniowieczna) z powodu ograniczonej liczby próbek. Autor podkreśla, że nie jest ekspertem, więc zakładamy że w zbiorze uczącym mogą być obrazy niepoprawnie sklasyfikowane. Każda z kategorii ma średnio 2-3 tysiące obrazków (niektóre najmniejsze koło 1000, a największe koło 5000), różnej wielkości, w kolorach (RGB), o rozszerzeniu “.jpg”.

2. **Dane wyjściowe:** Klasa (styl artystyczny) przypisana do obrazu, jeden z wcześniej wymienionych stylów.
3. **Wydajność:** Model powinien osiągać wysoką dokładność (>80%). Dataset ma dużą ilość danych wejściowych (prawie 30 GB danych), co powinno być wystarczające do osiągnięcia wysokiej dokładności.
4. **Skalowalność:** Możliwość uruchamiania na GPU w celu przyspieszenia obliczeń. Wykorzystywana będzie biblioteka PyTorch, z którą będzie można wykonać obliczenia na GPU z użyciem CUDA.
5. **Bezpieczeństwo:** Ograniczenie możliwości błędnej klasyfikacji poprzez monitoring modelu, na przykład generowanie wykresu wyników treningu co kilka epok i zapewnienie

różnorodności danych (30GB), aby model był bardziej odporny na nieprzewidziane warianty obrazów (Data Augmentation).

WikiArt Art Movements/Styles:

<https://www.kaggle.com/datasets/sivarazadi/wikiart-art-movementsstyles>

CZEŚĆ II - Zbiór danych i ich przygotowanie

W ramach przygotowania danych do klasyfikacji stylów artystycznych rozpoczęliśmy od ich wstępnego przetworzenia. Obrazy były zapisane w strukturze katalogów odpowiadającej stylom (niektóre z nich zawierały dodatkowe podkatalogi).

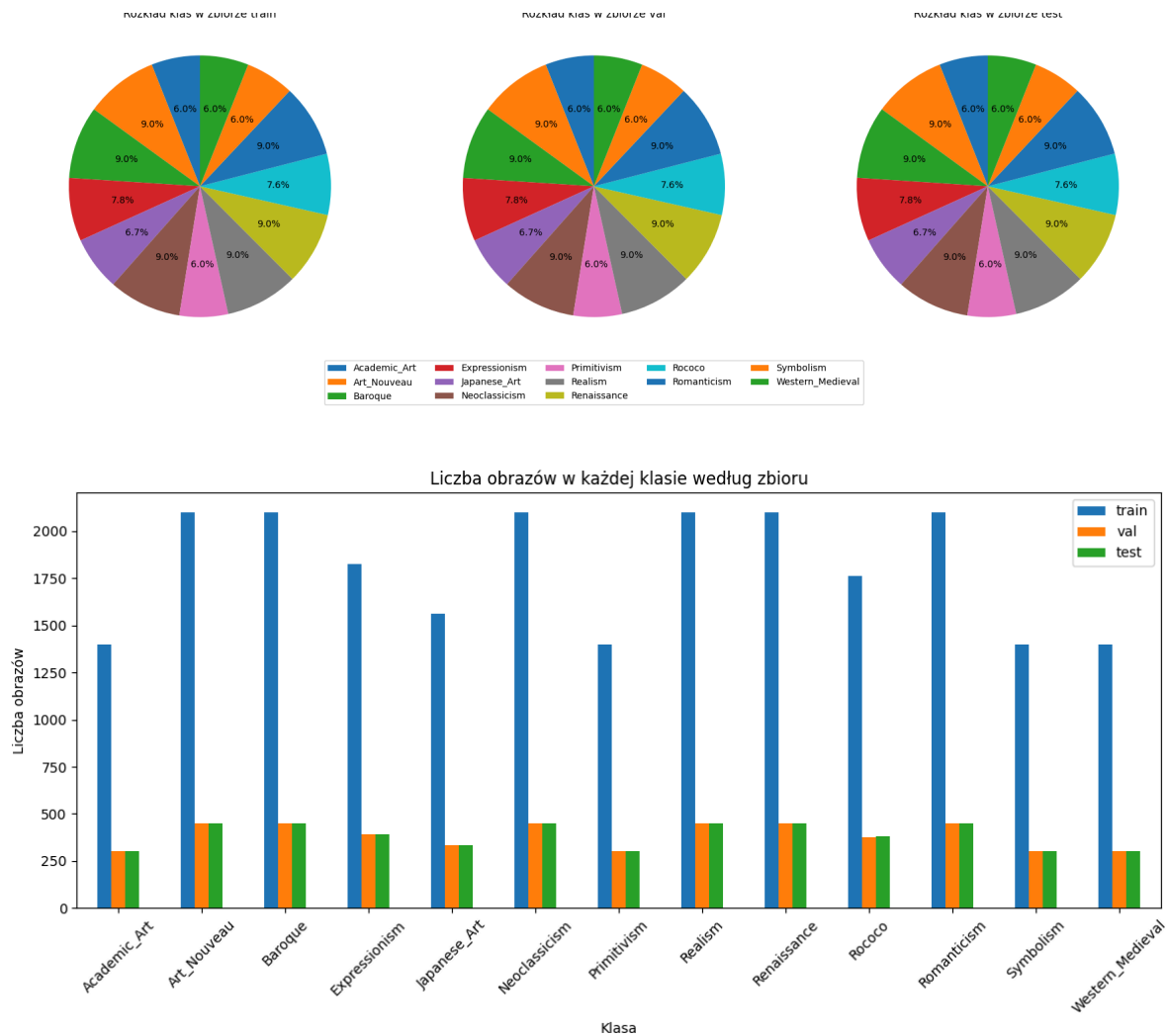
Aby ułatwić dalsze przetwarzanie, zdecydowaliśmy się na ujednolicenie rozmiaru obrazów. Każdy obraz został przeskalowany do wymiaru 512x512 pikseli z zachowaniem proporcji i dodaniem czarnego tła (paddingu), co pozwoliło na zachowanie spójności.

W kolejnym kroku skupiliśmy się na przygotowaniu zbiorów treningowego, walidacyjnego i testowego. Podział ten wynosił odpowiednio 70%, 15% i 15%. Aby unormować dane, postanowiliśmy, że liczba obrazów w każdej klasie (czyli folderze stylu) powinna mieścić się w przedziale od 2000 do 3000. Foldery zawierające zbyt dużą liczbę zdjęć zostały zredukowane – losowo wybraliśmy określoną liczbę przykładów. Natomiast w przypadku klas zbyt ubogich w dane, zdecydowaliśmy się na prostą augmentację – poprzez rotację wybranych obrazów pod kilkoma kątami utworzyliśmy nowe przykłady, które zwiększyły liczebność klasy.

W trakcie przetwarzania wprowadzono też dodatkowe przekształcenia. Dla zbioru treningowego zastosowaliśmy klasyczne techniki augmentacji, takie jak obrót czy odbicie lustrzane. Celem było zwiększenie różnorodności danych, a tym samym poprawienie uogólnienia modelu. Dla zbiorów walidacyjnego i testowego ograniczyliśmy się wyłącznie do przeskalowania i centralnego przycięcia, unikając w ten sposób przypadkowego zniekształcenia obrazów.

Aby móc efektywnie trenować modele, przygotowaliśmy klasę Dataset, która odpowiada za ładowanie obrazów oraz przypisywanie im etykiet. Do ładowania danych wykorzystaliśmy DataLoader z odpowiednią konfiguracją – zbiór treningowy był mieszany (shuffle=True), podczas gdy walidacyjny i testowy pozostawały w ustalonej kolejności.

W celu oceny jakości podziału i weryfikacji zbalansowania danych, wygenerowaliśmy wykresy kołowe oraz słupkowe przedstawiające liczebność poszczególnych klas w każdym ze zbiorów:



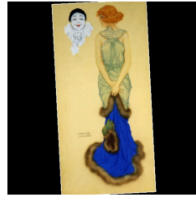
Dodatkowo przygotowano wizualizacje przykładowych obrazów wraz z przypisanymi etykietami:



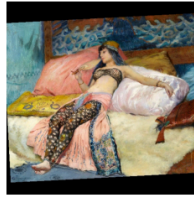
245438.jpg



252369.jpg



Przykłady klasy: Art_Nouveau
9223372032559860615.jpg



203527.jpg



262837.jpg



232486_aug_r3.jpg



295684.jpg



Przykłady klasy: Academic_Art
9223372032559879371.jpg



9223372032559874381.jpg



340348.jpg

