

Zaawansowane Techniki Sztucznej Inteligencji

Projekt, 2024/25

prof. dr hab. inż. Piotr Cofa

Celem projektu jest systematyczne i przemyślane opracowanie metody klasyfikacji obrazów, na podstawie udostępnionych zbiorów danych. Projekt może wykorzystywać dowolne architektury w dowolnych konfiguracjach.

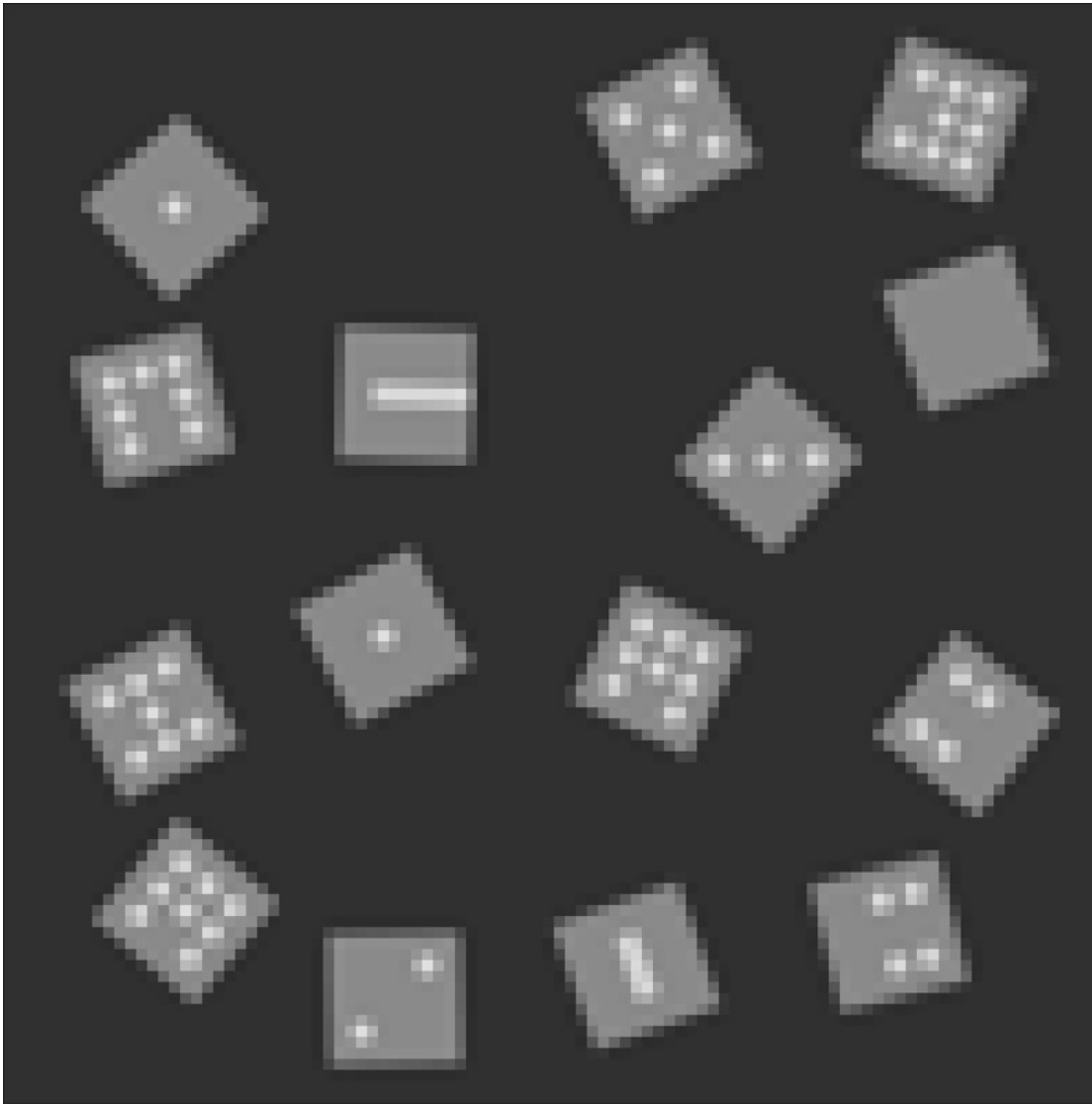
Problem i dane

Problemem jest rozpoznanie sumy liczb oczek poprawnych kostek do gry na podstawie obrazu zawierającego zarówno poprawne jak i niepoprawne kostki do gry (pseudo-kostki).

Podstawowym zestawem danych jest (symulowany) zbiór obrazów z których każdy zawiera piętnaście elementów. Pięć z nich to obrazy poprawnych kostek do gry, a pozostałe to obrazy pseudo-kostek. Całość jest umieszczona na jednolitej powierzchni.

Przykład takiego obrazu jest pokazany poniżej. Znajdują się na nim obrazy poprawnych kostek: dwóch z jednym oczkiem, jedną z dwoma, jedną z trzema i pięcioma. Ponadto znajdują się obrazy pseudo-kostek zawierających więcej oczek, oczka niewłaściwie ułożone, kreski itp. Możliwe jest również wystąpienie m.in. kostki 'pustej', bez oczek oraz inne nieprawidłowości.

Każdy obraz ma rozdzielczość 100x100 pikseli, każdy piksel opisany jako stopień szarości od zero do 255. Zastosowano konwencję znaną z zestawu danych MNIST, tj. 255 odpowiada kolorowi czarnemu. Kolory podłoża, kostki i oczek kostki (lub elementów je przypominających) są różne, ale wystarczająco kontrastowe. Elementy znajdują się w przypadkowych miejscach obrazu, ale jest gwarantowane, iż cały element jest widoczny na obrazie oraz że elementy nie zachodzą na siebie. Element może być obrócony względem krawędzi obrazu.



W pliku 'dice5.csv' znajduje się 100,000 obrazów, zaetykietowanych sumą liczby oczek na poprawnych kostkach. Każdy obraz jest zapisany w jednym wierszu. Pierwsza pozycja to etykieta (6..30), pozostałe to stopnie szarości pikseli, wierszami, poczynając od lewego górnego.

Dane są dostępne pod adresem

https://utpedupl-my.sharepoint.com/:u:/g/personal/piotr_cofta_o365_pbs_edu_pl/EaWk12ZDznNCsWHy9iJ5TbwB273-pInwk0lisAmx-EGcYQ?e=8tYKAs

Projekt

Projekt jest realizowany w trzech etapach.

Etap 1. – kontrola danych.

Nie ma gwarancji, iż dane w plikach są właściwie sformatowane, ani że są zgodne z opisem. W szczególności zbiór danych 'dice5.csv' jest nowy i nie był jeszcze weryfikowany. W pierwszym etapie należy sprawdzić, czy możliwy jest dostęp do danych, czy jest ich odpowiednia liczba, i czy są one zgodne z opisem.

Etap ten powinien zakończyć się krótkim raportem, obejmującym

- Opis zakresu zastosowanych metod weryfikacji
- Wyniki weryfikacji
- Wydruk jednego z obrazów, o indeksie zgodnym z numerem indeksu studenta modulo liczba danych
- Kod stosowany do weryfikacji (jeżeli weryfikacja wymagała kodu)

Etap 2 – ustalenie minimalnego poziomu jakości (benchmark)

Celem tego etapu jest skalibrowanie oczekiwań poprzez ustalenie minimalnego poziomu jakości, tj. wartości wybranej metryki dla rozwiązania prostego i oczywistego. W kolejnym etapie wprowadzenie bardziej zaawansowanych metod powinno skutkować poprawą tej metryki.

W tym celu należy opracować 'naiwny' model klasyfikujący obrazy zgodnie z sumą liczbą oczek na poprawnych kostkach, niezależnie od innych cech obrazu. 'Naiwność' oznacza, iż można użyć tutaj np. podejścia z poprzedniego semestru, odpowiednio przeskalowanego, lub jakiegokolwiek innego, niekoniecznie związanego z sieciami neuronowymi.

Celem tutaj nie jest opracowanie doskonałego rozwiązania, ale nawet takie 'naiwne' rozwiązanie powinno być zrealizowane starannie tak, aby dostarczyło miarodajnego poziomu odniesienia.

Uwaga. Jeżeli 'naiwne' podejście okaże się dostatecznie dobre, to etap 3. zostanie zrealizowany dla innego zbioru danych o podobnej charakterystyce. Decyzja w tym zakresie zapadnie po otrzymaniu wyników drugiego etapu.

Etap ten powinien zakończyć się raportem, obejmującym

- Opis podejścia: typu wybranego 'naiwnego' rozwiązania oraz wybranej metryki, z uzasadnieniem
- Dobór narzędzi, z uzasadnieniem
- Opis opracowanego modelu, z uwzględnieniem struktury sieci (lub innych parametrów i meta-parametrów), z uzasadnieniem
- Wyniki wraz potrzebnymi wyjaśnieniami; w szczególności wartość odniesienia
- Wytworzony kod

Etap 3 – podejście zaawansowane

Celem tego etapu jest opracowanie zaawansowanego modelu sztucznej sieci neuronowej klasyfikującej obrazy zgodnie z liczbą oczek na kostce, niezależnie od innych cech obrazu. Podstawową miarą jakości jest ta sama metryka która była stosowana w poprzednim etapie.

Poprzez 'zaawansowany' rozumiany jest tutaj model wykorzystujący elementy wiedzy z wykładu. W szczególności może on wykorzystywać segmentację i detekcję, wykrywanie anomalii itp. Być może zajdzie konieczność łączenia kilku modeli, połączenie lub uzupełnienie modeli fragmentami kodu itp.

Etap ten powinien zakończyć się raportem, obejmującym

- Opis zastosowanego podejścia do zagadnienia oraz szczegóły procesu
- Dobór narzędzi, z uzasadnieniem
- Opis opracowanego rozwiązania, z uwzględnieniem struktury sieci, z uzasadnieniem
- Wyniki wraz potrzebnymi wyjaśnieniami
- Wytworzony kod

Ocena

Warunkiem koniecznym zaliczenia jest

- weryfikacja zbioru danych
- realizacja i przetestowanie modeli: 'naiwnego' i 'zaawansowanego'
- dostarczenie raportów w uzgodnionych terminach
- aktywny udział w spotkaniach projektowych zgodnie z programem zajęć
- prezentacja postępu prac na wybranych spotkaniach

Ocena w mniejszym stopniu zależy od jakości modeli, a głównie zależy od dojrzałości podejścia, udokumentowanej w opracowaniu. Oceniane są takie elementy jak

- uzasadnienie wybranego podejścia / podejść
- uzasadnienie doboru narzędzi i metod pracy
- uzasadnienie doboru modeli i ich meta-parametrów
- refleksja nad uzyskanymi wynikami
- jasność prezentacji wyników (opisy, wykresy, tabele)