

PROJEKT PyTorch - “Znajdź samojedę”



Opis:

Celem projektu jest opracowanie narzędzia umożliwiającego analizę, etykietowanie i klasyfikację obrazów, z wykorzystaniem technik uczenia maszynowego i głębokiego uczenia. Narzędzie ma przetwarzać co najmniej 300 obrazów, oceniać ich jakość, identyfikować i etykietować wybrane klasy obiektów, a następnie przygotowywać dane do dalszej klasyfikacji za pomocą biblioteki PyTorch.

Funkcje dodatkowe:

- Etykietowanie Obiektów (etykietowanie.py) : Detekcja i etykietowanie wybranych klas obiektów na zdjęciach.
- Przygotowanie Danych dla PyTorch (custom_dataset.py): Konwersja zbioru danych do formatu kompatybilnego z PyTorch Dataset.
- Klasyfikacja Obrazów: Wykorzystanie wstępnie wytrenowanych modeli PyTorch oraz opcjonalnie trenowanie własnych modeli do klasyfikacji obiektów.

Technologie:

- Python
- PyTorch
- Numpy
- Pillow
- PIL (Python Imaging Library)
- OpenCV / TensorFlow (do detekcji obiektów)

Opis modeli:

Model resnet50 → patrzy na obrazy, ucząc się z nich, a potem opowiada nam, co dostrzega. Korzysta z doświadczeń innych, aby stać się bardziej zręcznym. Jego struktura przypomina zestaw klocków, gdzie każdy element specjalizuje się w rozpoznawaniu różnych detali obrazu. Doskonale radzi sobie z rozpoznawaniem skomplikowanych obiektów i zwraca uwagę na drobne szczegóły. Kiedy popełni błąd, wyciąga z tego wnioski, dążąc do doskonalenia w przyszłości. Działa bardzo szybko, błyskawicznie identyfikując, co znajduje się na obrazie.

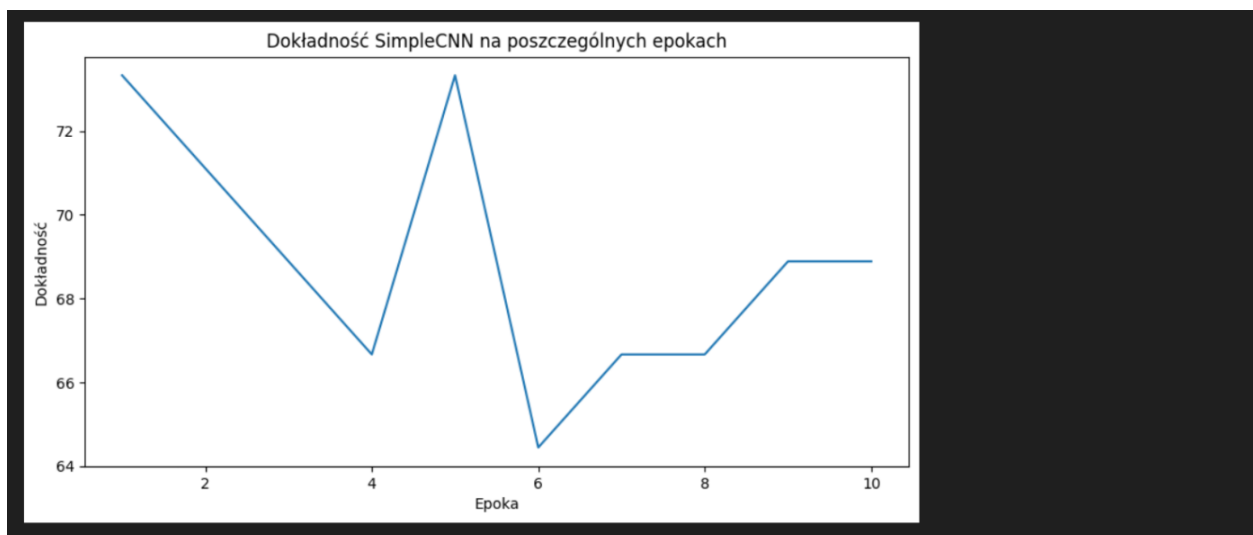
Model AlexNet → rozpoznaje podstawowe kształty, takie jak linie czy tekstury, używając różnych warstw. Stopniowo, analizując obraz w coraz większym stopniu szczegółów, model stara się zrozumieć bardziej skomplikowane obiekty. Ważną cechą jest nauka na błędach - jeśli popełni błąd w rozpoznawaniu, dostosowuje się, aby lepiej radzić sobie z podobnymi sytuacjami w przyszłości. Na koniec, gdy zgromadzi wystarczającą ilość informacji, model podejmuje decyzje na podstawie prawdopodobieństwa, określając, co widzi na obrazie i przypisując to do różnych kategorii.

Model VGG16 → operuje w sposób analogiczny do systemu układania klocków, gdzie każdy z elementów modelu pełni określoną specjalizację, obejmującą m.in. identyfikację linii czy rozpoznawanie kształtów. Model uczy się, analizując wiele różnych obrazów, co pozwala mu szybko zauważać podobieństwa na zdjęciach. W razie błędów, model się uczy, dążąc do lepszego rozpoznawania w przyszłości - to swoista nauka na błędach. VGG16 jest szczególnie efektywny w rozpoznawaniu różnorodnych obiektów na zdjęciach, nawet tych bardziej skomplikowanych, przypominając nam, co dokładnie widzi na obrazie.

Model SqueezeNet → zaprojektowany z myślą o oszczędzaniu miejsca, jednocześnie zachowując spryt w rozpoznawaniu obrazów. Jego mniejszy rozmiar sprawia, że jest przyjazny dla urządzeń z ograniczonymi zasobami, takich jak smartfony. To model, który mimo niewielkich wymiarów, szybko analizuje obrazy, identyfikuje podstawowe kształty, jak okręgi czy prostokąty, i efektywnie pracuje nawet przy ograniczonej ilości informacji.

Model DenseNet → wyróżnia się doskonałą współpracą między swoimi częściami, przypominając zespół, który świetnie się ze sobą komunikują i wspólnie analizują obrazy. Jego unikalna cecha to wzajemne dzielenie informacji między poszczególnymi elementami, co sprawia, że razem stają się silniejsze. Dodatkowo, model ten jest elastyczny w procesie uczenia, potrafiąc dostosować się do różnych zadań. Pomimo mniejszej ilości parametrów niż inne modele, DenseNet jest skuteczny i efektywny, co pozwala mu pracować sprawnie nawet na komputerach z ograniczonymi zasobami.

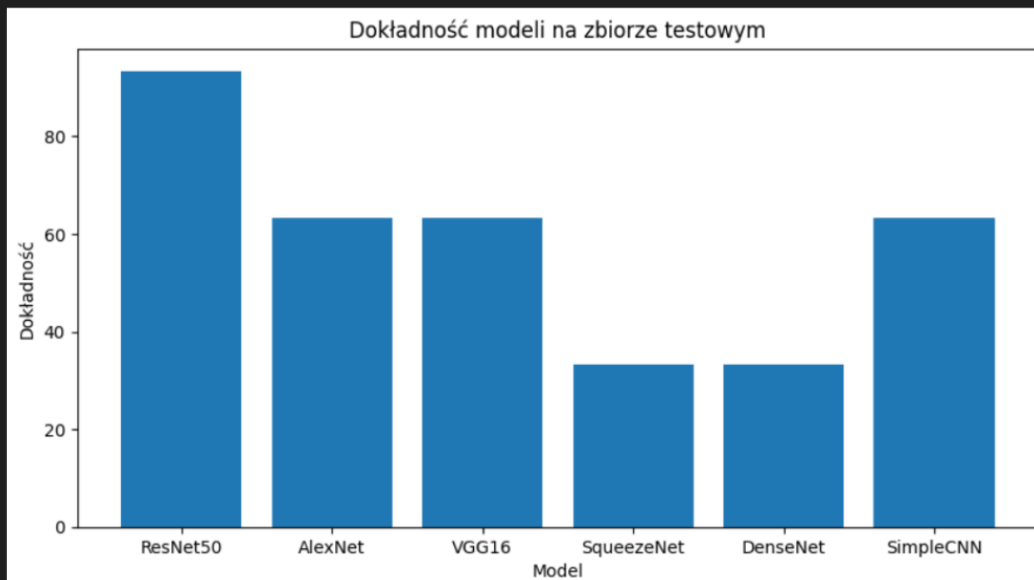
Wyniki (zdjęcia):



```

Model: ResNet
Dokładność: 93.33%
Model: AlexNet
Dokładność: 63.33%
Model: VGG
Dokładność: 63.33%
Model: SqueezeNet
Dokładność: 33.33%
Model: DenseNet
Dokładność: 33.33%
Model: SimpleCNN
Dokładność: 63.33%
[93.333333333333, 63.3333333333336, 63.3333333333336, 33.3333333333336, 33.3333333333336, 63.3333333333336]

```



Wniosek:

Model ResNet osiągnął najwyższą dokładność wynoszącą 93.33%. Jego zdolność do szybkiego uczenia się z błędów, elastyczność w przetwarzaniu obrazów oraz błyskawiczne rozumienie powodują wysoką skuteczność tego modelu.

AlexNet i Simple CNN uzyskały identyczną dokładność na poziomie 63.33%. Oba modele skupiają się na rozpoznawaniu podstawowych kształtów, jednak AlexNet dodatkowo korzysta z nauki na błędach.

VGG16 i DenseNet osiągnęły dokładność na poziomie 33.3%, mniejsza skuteczność w porównaniu do innych modeli. Wyniki mogą sugerować pewne wyzwania związane z bardziej złożonymi danymi.

SqueezeNet, model przyjazny dla urządzeń z ograniczonymi zasobami, uzyskał 63.33% dokładności, co może wskazywać na jego skuteczność w przetwarzaniu obrazów przy mniejszych wymaganiach zasobowych.