

Inżynieria Mechatroniczna

Podstawy Sztucznej Inteligencji i Uczenia Głębokiego: 10: Od uczenia płytkiego do głębokiego

Ziemowit Dworakowski
AGH w Krakowie

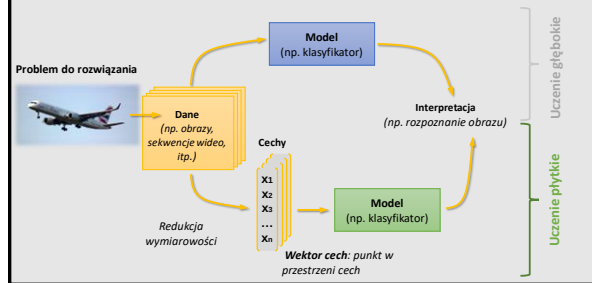
1

Stosowanie metod AI w interpretacji obrazów SD

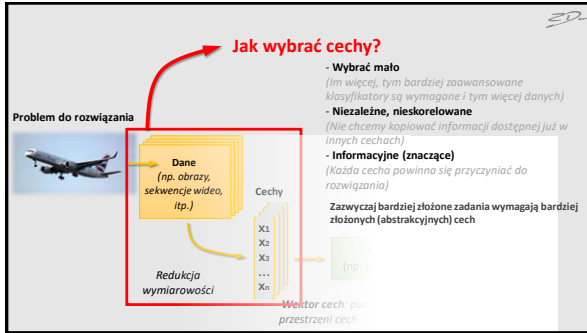
- Korzystamy z metod płytkich lub głębokich
- **Classification / Labeling**
Rozpoznawanie obrazów lub ich fragmentów
 - **Inteligentne przetwarzanie obrazów**
Kalibracja filtrów, filtracja, segmentacja, poprawa jakości
 - **Modelowanie – zaawansowane wnioskowanie**
Rozumienie co się dzieje na obrazie i dlaczego
 - **Generowanie nowych obrazów**
Budowa obrazów z promptów, modyfikowanie obrazów, augmentacja zbiorów obrazów
- Możliwe wyłącznie z zastosowaniem metod głębokich

2

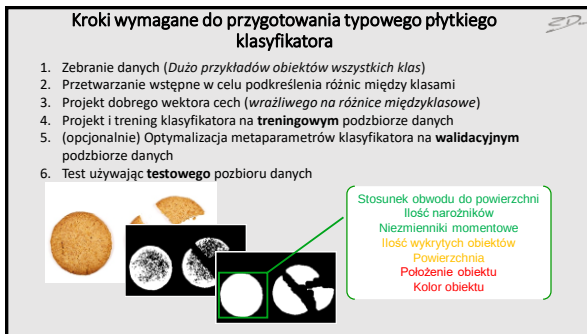
Podstawowa reprezentacja SD



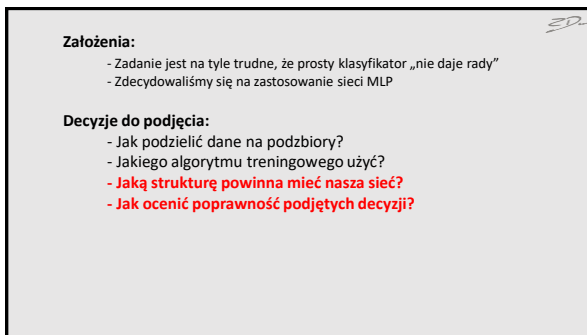
3



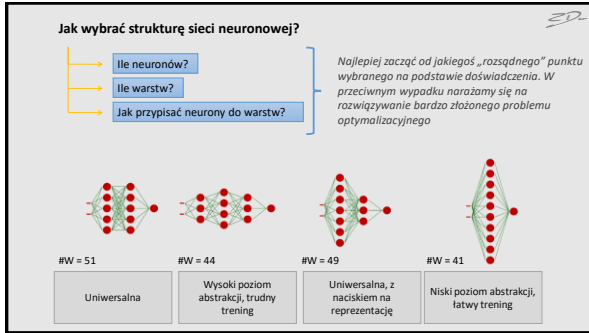
4



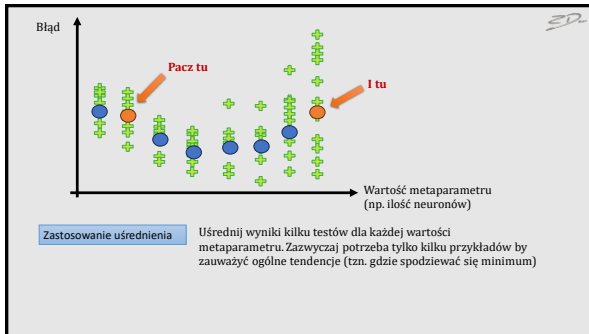
5



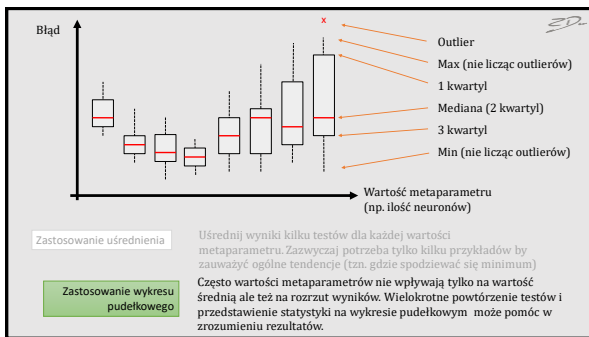
6



7



8



9

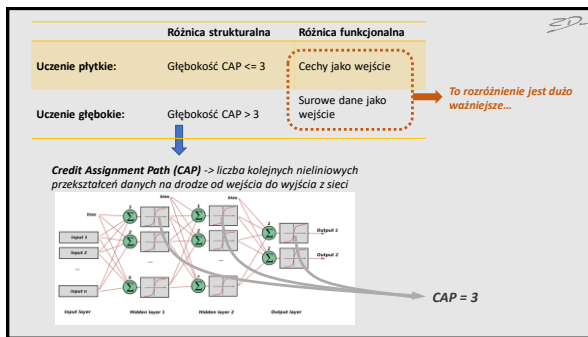
Uwagi ogólne:

- Klasyfikator powinien być wybierany na podstawie naszej wiedzy o przestrzeni cech**
Ile wymiarów? Ile cech? Czy widzimy wyraźne klastry? „Rodzaj zadania”, (np. klasyfikacja warzyw ciasteczek czy zwierząt) jest dużo mniej istotny.
- Podzbiory treningowy, walidacyjny i testowy muszą być rozłączne!**
Zaczynamy od losowego podziału danych na podzbiory lub (lepiej!) zbieramy każdy podzbiór w osobnym eksperymencie.
- Liczba stopni swobody modelu (tj. ustawialnych parametrów) zależy od ilości danych**
Dobłą praktyką jest trzymanie się min. 10 punktów danych na każdy trenowalny parametr modelu
- Jakość cech > Konfiguracja klasyfikatora**
Dobre cechy pozwalają na dobrą separację danych nawet z zastosowaniem słabego klasyfikatora. Słabe cechy nie pozwolą na uzyskanie dobrych wyników pomimo spędzenia setek godzin na optymalizacji modelu. Lepiej w pierwszej kolejności zainwestować czas i zasoby we wstępne przetwarzanie danych

10

Wprowadzenie do uczenia głębokiego

11

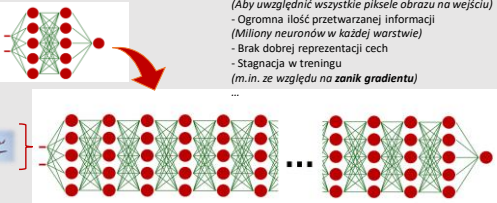


12

Jak **nie** budować głębokiego modelu?

Efekt:

- Miliony neuronów wejściowych (Aby uwzględnić wszystkie piksele obrazu na wejściu)
- Ogromna ilość przetwarzanej informacji (Miliony neuronów w każdej warstwie)
- Brak dobrej reprezentacji cech
- Stagnacja w treningu (m.in. ze względu na **zanik gradientu**)



13

Więc... Co właściwie pozwala na automatyczne nauczanie się przetwarzania surowych danych?

- Aktywacje ReLU i Leaky ReLU
- Kernele konwolucyjne
- Pooling
- Regularyzacja
- Fine-tuning

To podstawowe koncepty. Większość typowych modeli uczenia głębokiego używa ich w jakiejś formie

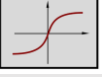
- Adversarial training
- Atencja
- Autoenkodery
- Latent space
- Embeddings
- Uczenie ze wzmocnieniem
- Transfer learning
- Transformatory
- Sieci rekurencyjne (np. Long-Short-Term-Memory)

Te koncepty są bardziej zaawansowane, ich wybór pozwala na zrozumienie zaawansowanych modeli typu Midjourney czy ChatGPT

14

Funkcja sigmoidalna

- + Dobrze działa w przypadku małych sieci
- Może powodować zanik gradientu (powodując stagnację treningu)



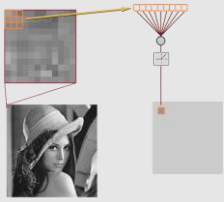
Funkcja aktywacji typu Rectified Linear Unit (ReLU)

- + Przeciwodziła zanikowi gradientu
- + Przyspiesza trening
- /+ Może powodować „umieranie” neuronów



15

Kernel konwolucyjny

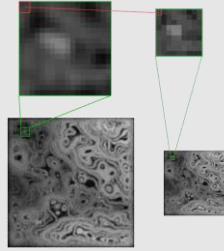


Zapewnia:

- * Filtrację konwolucyjną
- * Filtrację nieliniową
- * Filtrację morfologiczną
- * Wydobycie cech niskiego poziomu
(np. rozpoznawanie narożników)
- * Dzięki **dzieleniu wag** jedna warstwa może realizować wiele zadań w tym samym czasie
- * Wyjściem są **mapy cech**

16

(Max) pooling



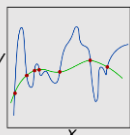
- * Zachowujemy maksymalne wartości mapy cech
- * Zachowujemy **przestrzenne zależności** w cechach niskiego poziomu
- * Zmniejszamy wymiarowość problemu

17

Regularyzacja

Aby uniemożliwić zapamiętywanie danych (przeuczenie) wprowadzamy dodatkowe ograniczenia w treningu sieci

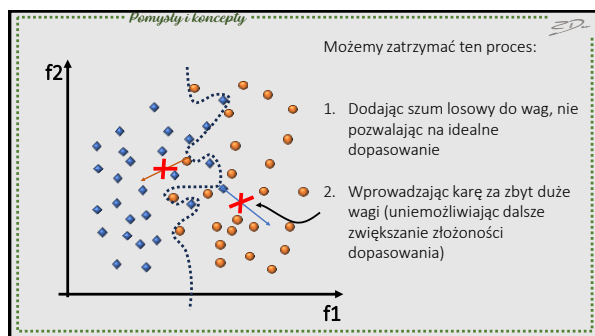
W praktyce zazwyczaj „iteracyjnie psujemy model”. W momencie osiągnięcia generalizacji (gdy klasyfikator zaczyna się dopasowywać do szumów w danych), czynnik regularyzacyjny zaczyna dominować uniemożliwiając dalsze dopasowanie.



Na przykład:

- * W algorytmie gradientowym, poza aktualizacją wag wynikającą z kierunku gradientu dodajemy niewielką losową modyfikację do części lub całości wag
- * Niektóre połączenia między neuronami są usuwane (tzw. metoda „brain damage”)

18



19

Fine tuning (doucewanie)

Każdorazowo, w uczeniu głębokim potrzebujemy **ogromnych ilości danych i ogromnej mocy obliczeniowej**

W przypadkach specjalistycznych tych danych nie będzie wystarczająco dużo, trening modelu może być również nieopłacalny czasowo i energetycznie

Rozwiązanie:
Trening ogólny na dużym zbiorze, **fine-tuning** (doucewanie) na zbiorze specjalistycznym

↓

Robimy to poprzez reinicjalizację wag w końcowych warstwach sieci, często jednocześnie zamrażając wagi w warstwach początkowych

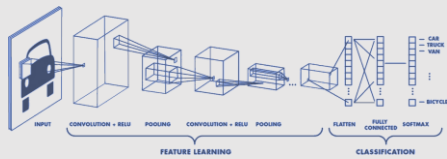
20

Fine tuning - przykłady

Nauczenie dużego modelu językowego na dużym, ogólnym zbiorze tekstów	→	Doucewanie go na twórczości jednego autora, by umożliwić dokładne naśladowanie jego stylu
Nauczenie głębokiej sieci neuronowej rozpoznawania gatunków różnych zwierząt	→	Doucewanie jej na przykładach kilku tysięcy zdjęć wiewiórek dwóch gatunków – do specjalistycznego zliczania wiewiórek w parku narodowym
Nauczenie głębokiego modelu do rozpoznawania mowy na zbiorze nagrań mówiących ludzi	→	Doucewanie go na zbiorze danych ultradźwiękowych (!?) w celu odróżniania struktur uszkodzonych od zdrowych

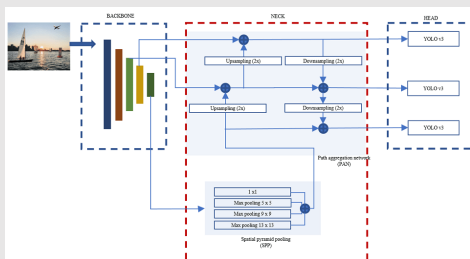
21

Przykłady: Matlab DCNN



22

Przykłady: YOLO v4



23

Podsumowanie (tematy na test):

- 1) Kategorie zadań w inteligentnym przetwarzaniu obrazów (z przykładami)
- 2) Schemat przetwarzania głębokiego i płytkiego (wymagane kroki, z przykładem)
- 3) Funkcjonalna i strukturalna różnica między płytkim i głębokim uczeniem
- 4) Różnice pomiędzy różnymi strukturami sieci MLP
- 5) Jak możemy zoptymalizować sieć MLP? (Uśrednianie, wykres pudełkowy)
- 6) 4 uwagi ogólne w treningu klasyfikatorów
- 7) Wyjaśnij funkcję ReLU, porównaj z sigmoidalną
- 8) Wyjaśnij kernel konwolucyjny
- 9) Wyjaśnij max pooling
- 10) Wyjaśnij regularyzację
- 11) Wyjaśnij fine tuning (doucewanie) – z przykładami

(Szczegółowych schematów głębokich sieci neuronowych nie będzie na teście)

24