

Podsumowanie projektu – Klasyfikacja obrazów CIFAR-10 z użyciem sieci neuronowych

Bartosz Smolibowski, Michał Safuryn, Wojciech Balcer

Cel projektu

Celem projektu było porównanie skuteczności różnych architektur sieci konwolucyjnych (CNN) w zadaniu klasyfikacji obrazów z zestawu danych CIFAR-10. W tym celu zaimplementowano i przetestowano cztery modele:

- Simple CNN – prosta sieć konwolucyjna z dwiema warstwami Conv2D.
- Deep CNN z Batch Normalization – pogłębiona sieć z normalizacją wsadową i dropoutem.
- VGG-like CNN – inspirowana architekturą VGG z większą liczbą filtrów i warstw.
- MobileNetV2 – transfer learning z wykorzystaniem gotowej architektury MobileNetV2 (wstępnie wytrenowanej na ImageNet).

Przygotowanie danych

Dane wejściowe zostały pobrane z zestawu CIFAR-10. Obrazy przeskalowano do przedziału [0, 1], a następnie poddano normalizacji z użyciem z-score (na podstawie średniej i odchylenia standardowego z kanałów RGB). Etykiety zostały zakodowane jako one-hot encoding. W celu poprawy generalizacji modeli zastosowano augmentację danych, m.in. rotację, przesunięcia i odbicie poziome.

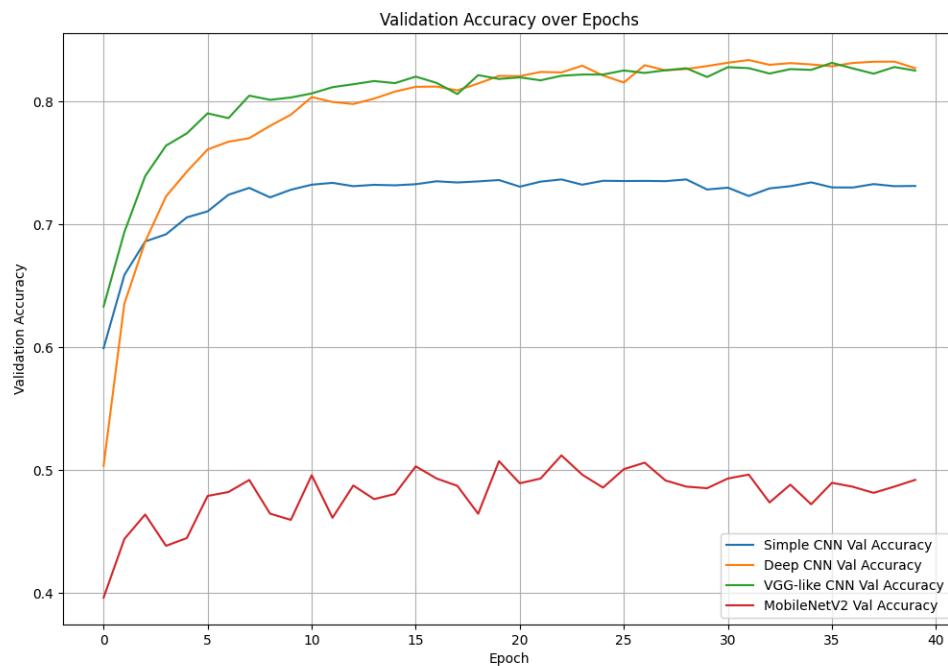
Szkolenie modeli

Wszystkie modele zostały wytrenowane przez 80 epok z użyciem optymalizatora Adam i funkcji straty categorical_crossentropy. Trenowanie przeprowadzono z użyciem ImageDataGenerator, aby zapewnić losową augmentację danych w czasie rzeczywistym. Po zakończeniu trenowania każdego modelu, przeprowadzono predykcję i obliczono macierz pomyłek oraz dokładność walidacyjną.

Wyniki i obserwacje

1. Dokładność walidacyjna w czasie (Validation Accuracy)

Poniżej znajduje się wykres prezentujący zmiany dokładności walidacyjnej w czasie treningu dla wszystkich modeli:



2. Macierze pomylek dla każdego modelu

Model

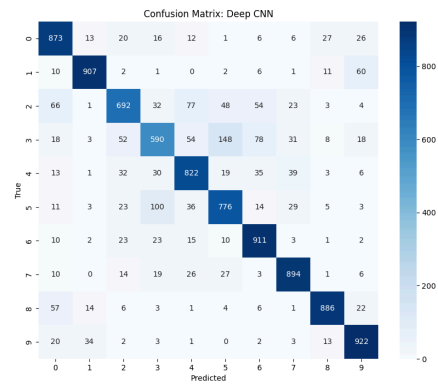
Simple CNN

Macierz pomylek

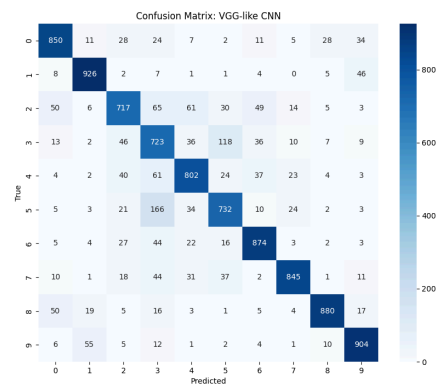
Confusion Matrix: Simple CNN

	0	1	2	3	4	5	6	7	8	9
0	749	24	45	21	17	3	9	20	62	50
1	22	871	6	5	0	2	5	3	12	74
2	63	7	604	53	83	63	61	45	11	10
3	25	7	65	528	61	183	61	38	12	20
4	21	3	58	59	696	34	53	64	8	4
5	12	8	47	167	47	644	16	39	5	15
6	4	4	42	70	49	30	784	10	6	1
7	13	6	41	23	57	58	3	787	2	10
8	59	49	9	13	7	6	6	6	821	24
9	19	89	8	9	5	5	4	15	16	830

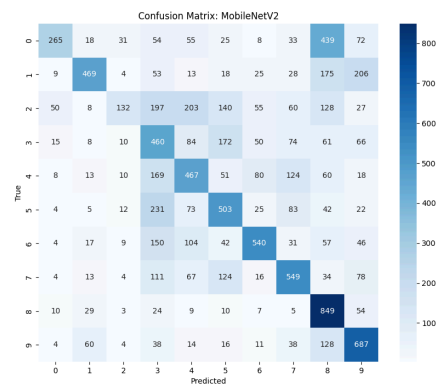
Deep CNN w/ BN



VGG-like CNN



MobileNetV2



3. Kluczowe wnioski

- Najniższą dokładność walidacyjną osiągnął model MobileNetV2
- Modele Deep CNN oraz VGG-like CNN osiągnęły zbliżone wyniki, lecz lepsze od najprostszego modelu Simple CNN.
- Simple CNN, mimo swojej prostoty, wykazał się rozsądną skutecznością, jednak nie dorównał bardziej złożonym architekturom.
- Augmentacja danych oraz normalizacja z-score przyczyniły się do poprawy uogólniania modeli.

Wnioski końcowe

Projekt pokazuje, jak różne architektury sieci konwolucyjnych wpływają na skuteczność klasyfikacji obrazów. W szczególności, wykorzystanie gotowych modeli głębokich może znacznie poprawić wyniki, zwłaszcza w przypadkach ograniczonych danych treningowych lub zasobów obliczeniowych.