

Klasyfikacja samochodów ze względu na typ nadwozia

Aleksander Lemiesz 244887

08.06.2022

Streszczenie

Artykuł opisuje problem klasyfikacji typu nadwozia samochodu i po krótko opisuje do czego może zostać wykorzystane rozwiązanie. W celu opracowania rozwiązania została wytrenowana sieć neuronowa bazująca na sieci AlexNet i własnym klasyfikatorze, klasyfikująca obrazy samochodów z bazy danych StanfordCars. Metoda oraz wyniki porównywane są z artykułem analizującym podobny problem. Wyniki osiągnięte dla zbioru z augmentacją i bez były odpowiednio 76% i 79% odpowiednio. Lepszy z wyników jest drugim najlepszym wynikiem w porównaniu z wynikami innego autora. Dodatkowo wynik ten został uzyskany dla modelu, który klasyfikuje więcej typów nadwozia niż w porównywanym artykule. Uzyskane wyniki mogą być spowodowane niebalansowaniem w zbiorze danych.

1. Wprowadzenie

Artykuł skupia się na problemie rozpoznawania typu samochodu ze względu na typ nadwozia. Może to zostać wykorzystane na stronie sprzedającej samochody używane do filtrowania ogłoszeń. Algorytm rozpoznaje nadwozia takie jak kabriolet, coupe, van, minivan, sedan itp. W tym celu wykorzystana została sieć neuronowa, trenowana na podstawie zbioru obrazów StanfordCars[2]. Na podstawie artykułu Bhanu Yerry, w którym nadwozia również były klasyfikowane. Następnie zaproponowana została nowa sieć służąca do klasyfikacji typów aut oraz przedstawione zostały wyniki. Zbiór danych StanfordCars jest również używany do klasyfikowania zdjęć samochodów do konkretnych modeli, czyli całych etykiet z tego zbioru[3] oraz klasyfikacji po marce produkującej dane auto.

2. Opis podobnych prac

Podobny problem został opisany w artykule *"Car Image Classification Using Features Extracted from Pre-trained Neural Networks"*[1] autorstwa Bhanu Yerry. Sugeruje on podobne użycie algorytmu rozpoznającego typy nadwozia samochodów. Do ekstrakcji cech użyte zostały pre-trenowane na bazie obrazów ImageNet modele z platformy Keras. Tymi pre-trenowanymi modelami były VGG 16, VGG 19, ResNet50V2, Inception V3, MobileNetV2. Autor pracy również trenował klasyfikator na zbiorze danych StanfordCars i przeprowadził operację oversampling tak aby poprawić dane. Na zbiór ten składa się ponad 16 tys. obrazów przypisanych do 196 klas. Autor pracy skonsolidował te klasy do pięciu klas nadwozia Kabriolet/Coupe, Sedan, SUV, Truck, Van/Minivan. Do porównania wyników autor skorzystał z dokładności klasyfikacji.

3. Opis metody

Aby przeprowadzić podobną klasyfikację pobrany został zbiór danych StanfordCars w, którym etykieta obrazka wygląda następująco, Make/Model/Year, przykładowo „Audi A5 Coupe 2012” zostały zgrupowane po przedostatnim członie do etykiety, przykładowo „Coupe”. W odróżnieniu od wcześniej wspomnianego artykułu zaproponowana sieć klasyfikuje samochody do ośmiu następujących klas Cab,

Convertible, Coupe, Hatchback, Minivan, Sedan, SUV, Van. Oryginalnie autorzy klasyfikują Convertible oraz Coupe i Minivan oraz Van jako jedną klasę oraz nie rozpoznają Hatchback'ów. Powodować to może dodatkowe trudności w rozpoznaniu nowych klas jednak zwiększa potencjał algorytmu, ponieważ może zostać użyty do filtrowania samochodów po większej ilości typów. Do ekstrakcji cech używana jest gotowa sieć neuronowa AlexNet z biblioteki torchvision z własnym klasyfikatorem. Klasyfikator ten składa się z warstw dropout, liniowej oraz funkcji aktywacji ReLU następujących po sobie dwukrotnie i kończącej się jedną warstwą liniową. Cała sieć została trenowana łącznie z AlexNetem. Uczenie odbywało się przez 40 epok. Funkcją straty jest Cross Entropy Loss a optimizerem SGD.

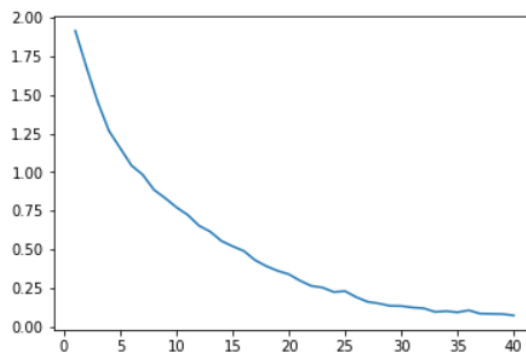


Fig. 3. 1. Wykres funkcji straty podczas treningu dla 40 epok.

Dodatkowo dla porównania została przeprowadzona augmentacja danych według augmentation policy CIFAR10, IMAGENET oraz SVHN. Dla każdej z polityk zostało wygenerowane tyle obrazów aby zbilansować zbiór danych. Wcześniej wspomniany model oparty o AlexNet został wytrenowany przez pięć epok z takimi samymi parametrami i jedyną różnicą była augmentacja danych.

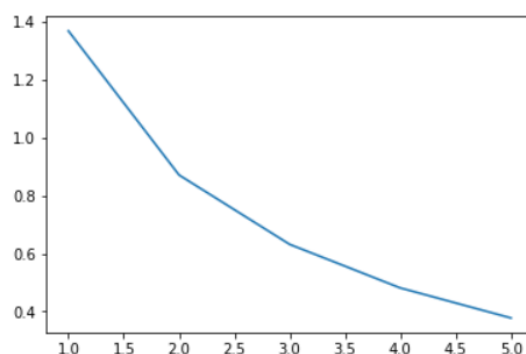


Fig. 3. 2. Wykres funkcji straty podczas treningu dla 5 epok dla augmentowanego zbioru danych.

Aby wyniki klasyfikacji algorytmu mogły zostać porównane z wynikami przedstawionymi we wcześniej wspomnianym artykule została również wytrenowana sieć klasyfikująca na takie same typy nadwozia. Dodatkowo aby poprawić jej skuteczność trening odbywał się na augmentowanym i zbalansowanym zbiorze danych. Nauka trwała 24 epoki i zakończyła się ze względu na wzrost wartości funkcji straty.

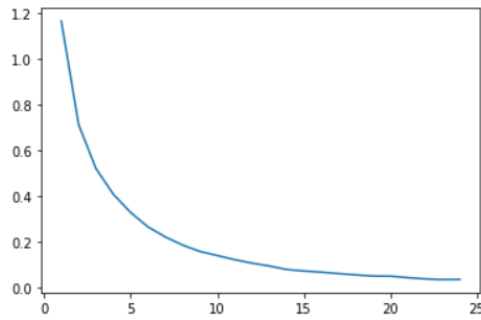


Fig. 3. 3. Wykres funkcji straty podczas treningu dla 24 epok dla augmentowanego i zbilansowanego zbioru danych.

Ostatecznie aby zweryfikować jak zachowa się sieć nauczona klasyfikacji większej ilości klas ze zbilansowanym augmentowanym zbiorem danych kolejny eksperyment został przeprowadzony. Nauka trwała 34 epoki i zakończyła się ze względu na wzrost wartości funkcji straty.

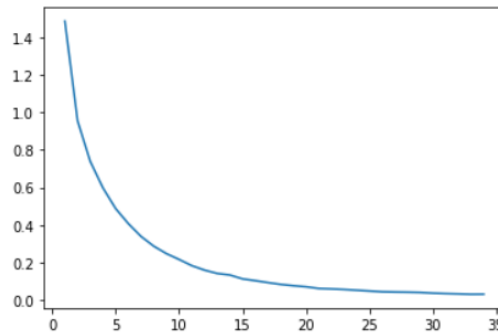


Fig. 3. 4. Wykres funkcji straty podczas treningu dla 34 epok dla augmentowanego i zbilansowanego zbioru danych rozpoznającego 9 klas.

4. Wyniki

Algorytm bez augmentacji osiąga skuteczność 79% dla zbioru testowego oraz 100% dla zbioru treningowego. Najgorzej radzi sobie z rozpoznawaniem Hatchbacków a najlepiej z SUV'ami. Jest to najprawdopodobniej spowodowane tym że gorzej nauczone klasy są gorzej reprezentowane w zbiorze danych.

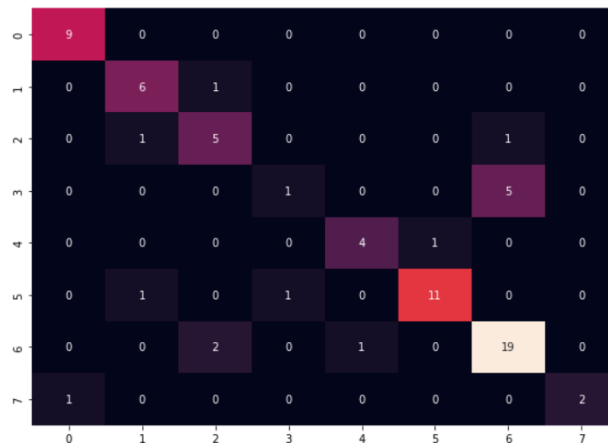


Fig 4. 1. Macierz pomyłek dla podstawowego zbioru danych.

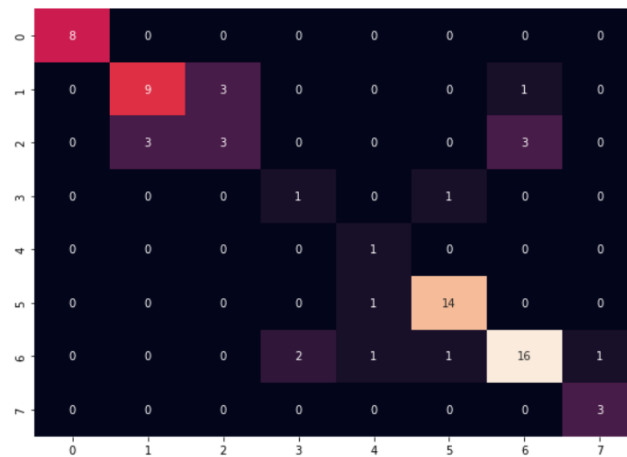


Fig 4. 2. Macierz pomyłek dla augmentowanego zbioru danych.

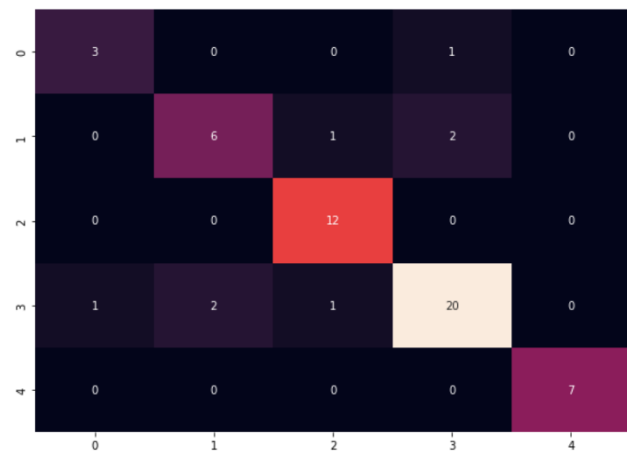


Fig 4. 3. Macierz pomyłek dla augmentowanego i zbalansowanego zbioru danych.

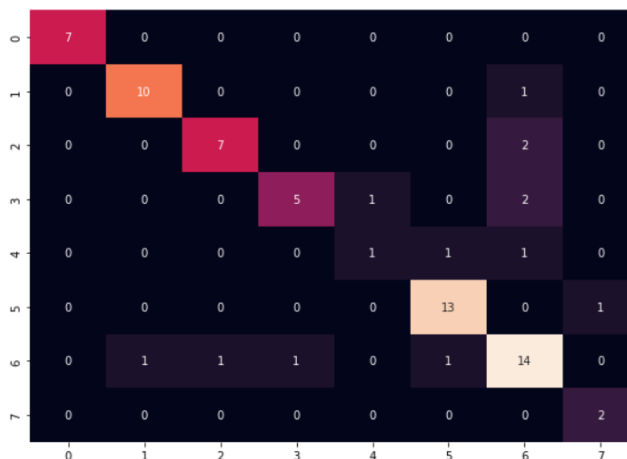


Fig 4. 4. Macierz pomyłek dla augmentowanego i zbalansowanego zbioru danych dla 9 typów klas.

```
'Cab': 0, 'Convertible': 1, 'Coupe': 2, 'Hatchback': 3, 'Minivan': 4, 'SUV': 5, 'Sedan': 6, 'Van': 7
```

```
{'Cab': 0, 'ConvertibleCoupe': 1, 'SUV': 2, 'Sedan': 3, 'VanMinivan': 4}
```

Fig 4. 5. Słowniki etykiet i ich indeksów.

Mimo że trenowany model bazujący na AlexNet nie osiąga najlepszego wyniku to nadal jest on drugi najlepszy a jest w stanie klasyfikować typy nadwozia na więcej kategorii niż w przypadku pozostałych, które zostały wytrenowane przez Bhanu Yerra. Mimo augmentacji danych model na nim się uczący nie osiągnął lepszych wyników, jednak model uczący się na zbalansowanym augmentowanym zbiorze danych podzielonym na taką samą liczbę klas jak w artykule wspomnianego autora osiąga najlepsze wyniki bo aż 86% skuteczności. Dla zbalansowanego zbioru danych dla 9 klas a więc większej ilości niż w porównywanym artykule skuteczność osiąga wynik 82%, a więc i tak osiąga lepsze wyniki niż najlepsza sieć ze wspomnianego artykułu mimo rozpoznawania większej ilości klas nadwozia.

Trenowany Model	Aug AlexNet	AlexNet	AlexNet Balance 9 Classes	AlexNet Balanced	VGG 16	VGG 19	ResNet50 V2	Inceptio nV3	MobileNetV2
Skuteczność	76%	79%	82%	86%	76%	75%	77%	81%	78%

Tabela 4. 1. Przedstawiająca wyniki skuteczności dla nowo wytrenowanej sieci oraz sieci wytrenowanych przez Bhanu Yerra.

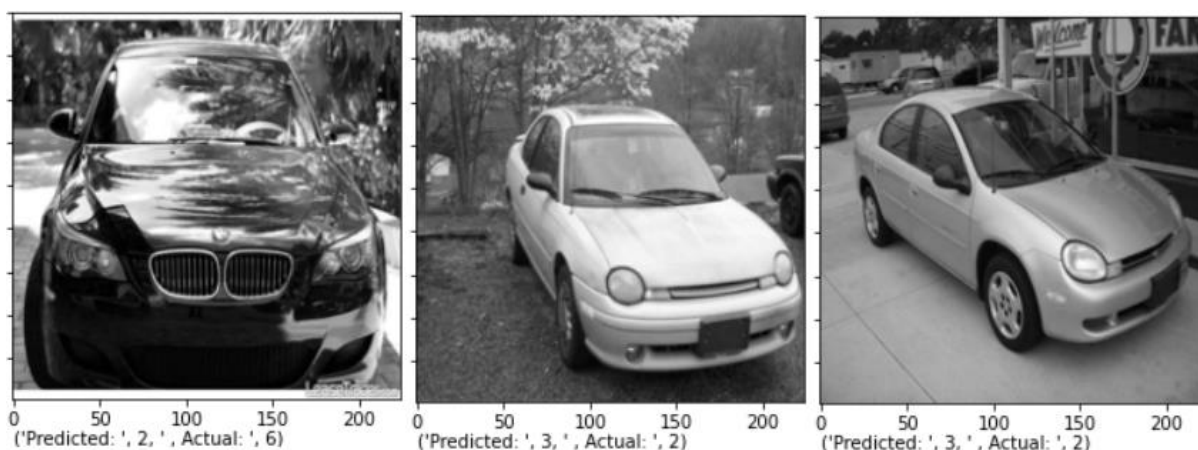


Fig 4.2. Przykłady złej klasyfikacji.

5. Podsumowanie

Można zatem powiedzieć, że wytrenowany model osiąga zadowalające wyniki szczególnie po porównaniu z tym jakie wyniki uzyskał Bhanu Yera. Nieprawidłowa klasyfikacja jest również często powiązana z kątem z jakiego przedstawiony jest pojazd. Na przykład jeśli auto jest przedstawione od przodu i widać tylko przód samochodu to nawet dla człowieka może być ciężko przewidzieć jakiego typu nadwozie posiada. Dodatkowo można powiedzieć że augmentacja i zbalansowanie danych pomaga uzyskać pożądane wyniki.

6. Literatura

1. "Car Image Classification Using Features Extracted from Pre-trained Neural Networks", Bhanu Yerra, <https://towardsdatascience.com/classifying-car-images-using-features-extracted-from-pre-trained-neural-networks-39692e445a14>
2. StanfordCars dataset https://ai.stanford.edu/~jkrause/cars/car_dataset.html
3. "Stanford Cars Classification using EfficientNet PyTorch", Sovit Ranjan Rath, <https://debuggercafe.com/stanford-cars-classification-using-efficientnet-pytorch/>