

Raport 3 - MSI

Tomasz Sewastynowicz

Uniwersytet im. Adama Mickiewicza w Poznaniu

10.12.2023 r.

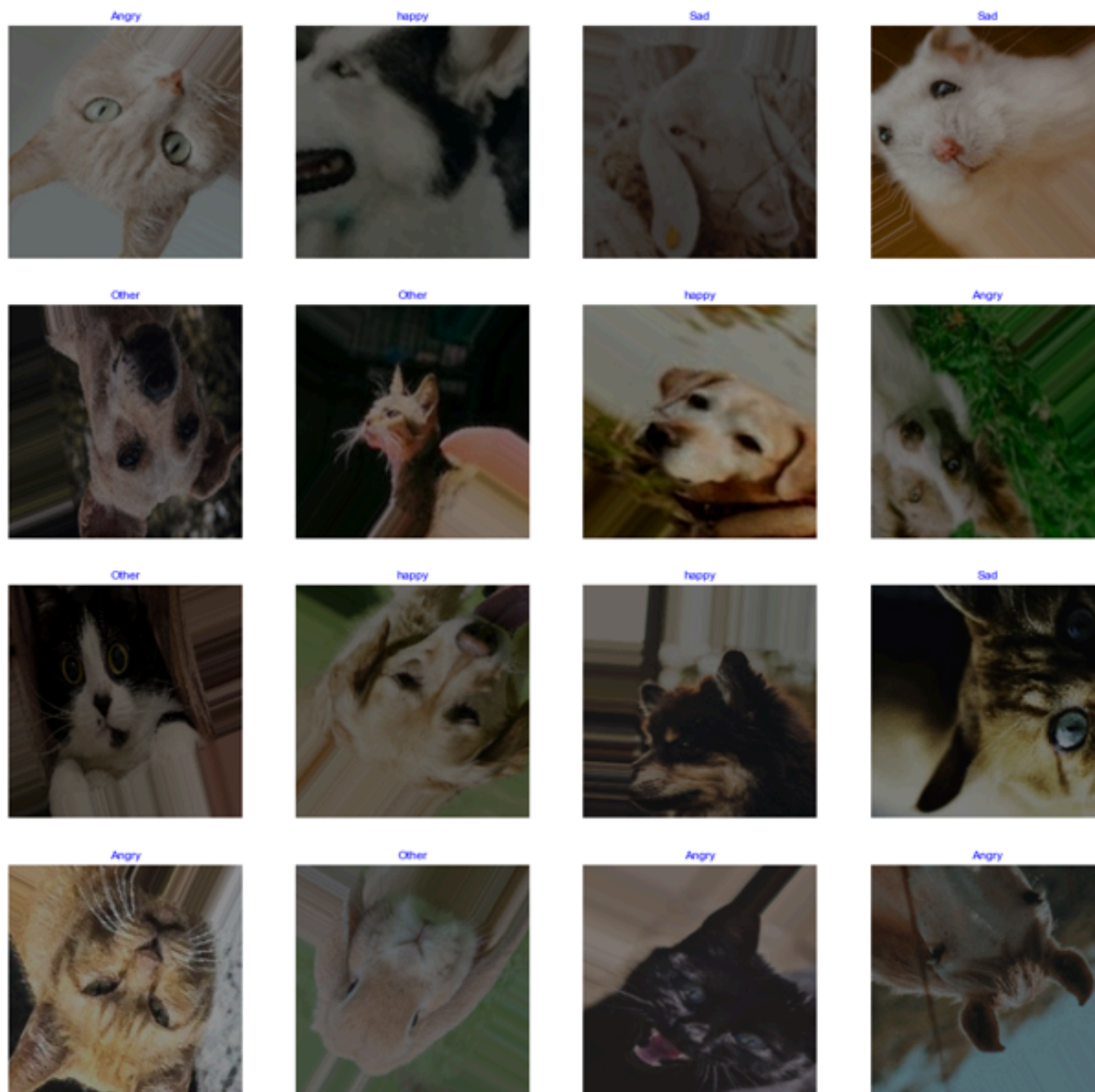
1. Cel pracy

Raport skupia się na zastosowaniu sieci konwolucyjnych w celu identyfikowania poprawnych emocji u różnych zwierząt. Zbiór zawiera różne zdjęcia twarzy psów, kotów, królików, chomików, owiec, koni i ptaków. Obrazy zawierają różnorodne emocje, jakie mogą odczuwać te zwierzęta i zawierają 4 typy stanów: radość, smutek, złość, oraz kategoria “inne”, zawierająca pozostałe, bardziej neutralne wyrazy twarzy u zwierząt.

W ramach pracy został wykorzystany zbiór danych "Pet's Facial Expression Image Dataset", z platformy Kaggle (<https://www.kaggle.com/datasets/anshtanwar/pets-facial-expression-dataset>). Zbioru tego użyto do trenowania i testowania sieci konwolucyjnych i stworzenia modeli, który będą jak najpoprawniej klasyfikowały grafiki.

2. Zbiór danych

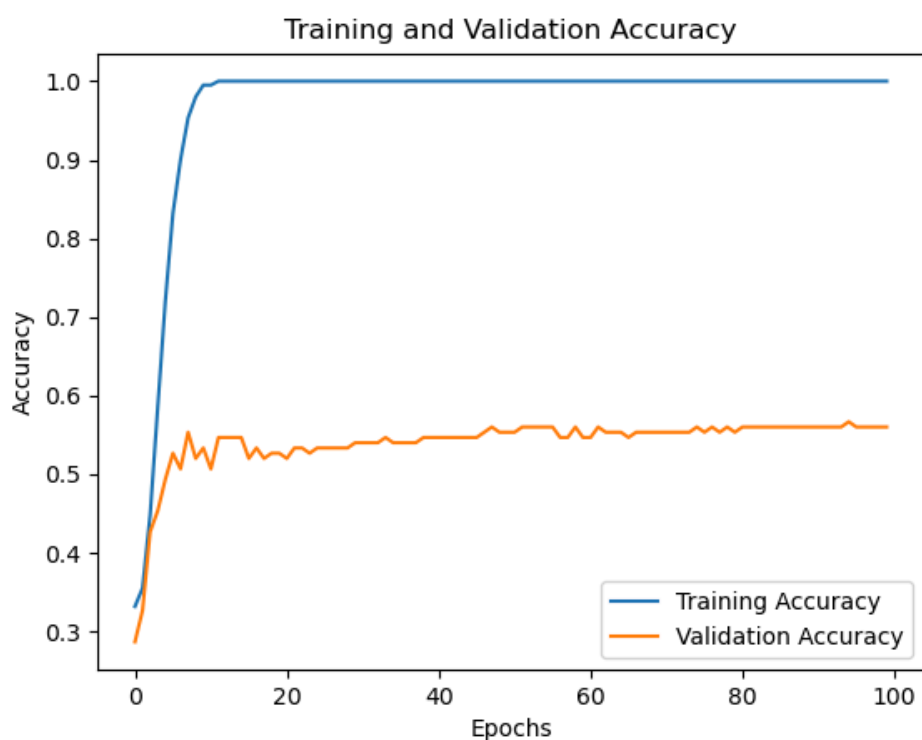
Zbiór ten dzieli łącznie 1000 zdjęć zwierząt na 4 kategorie: Angry, Sad, Happy, Other. Na **rysunku 1.** są widoczne przykładowe obrazki z tego zbioru z etykietami o danej emocji, które dane zwierzę okazuje.



Rysunek 1. Grafiki przedstawiające różne zwierzęta i emocje.

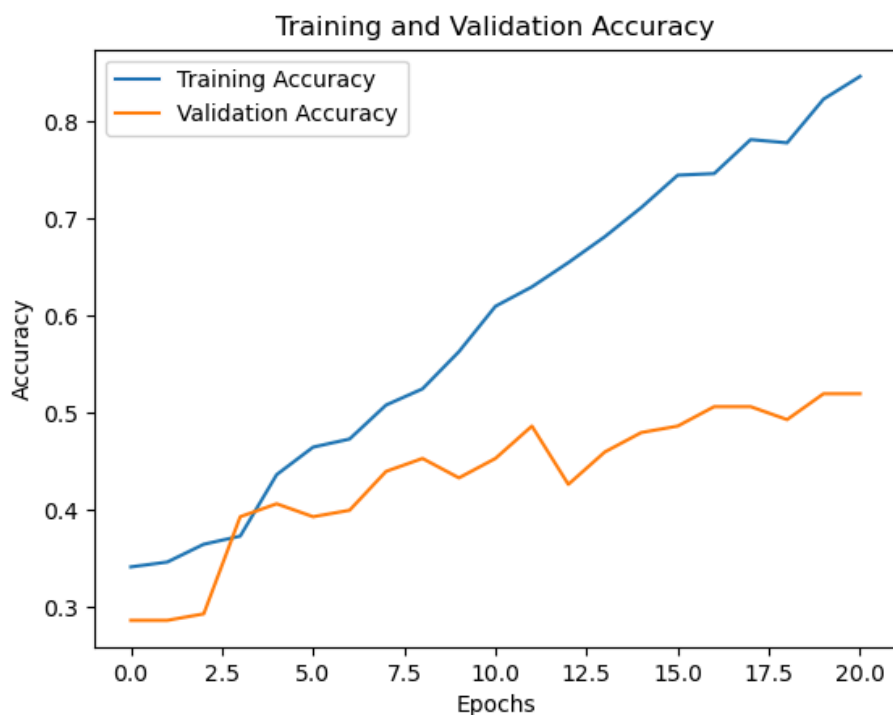
3. Modele sieci konwolucyjnych

Łącznie zostały stworzone 4 modele, których dokładność na zbiorze walidacyjnym wynosiła od 48% do 70%, oraz został użyty gotowy model DenseNet (**rysunek 6.**), który osiągnął dokładność na poziomie 63.3%. Modele różniły się między sobą pod względem liczby warstw i ilości technik zapobiegania przeuczeniu, oraz wartościami tych technik. Najprostszy model (**rysunek 2.**) nie zawierający żadnych technik zapobiegania przeuczenia osiągnął accuracy na zbiorze testowym na poziomie 100% i 56% na zbiorze walidacyjnym.



Rysunek 2. Accuracy pierwszego modelu - model przeuczony.

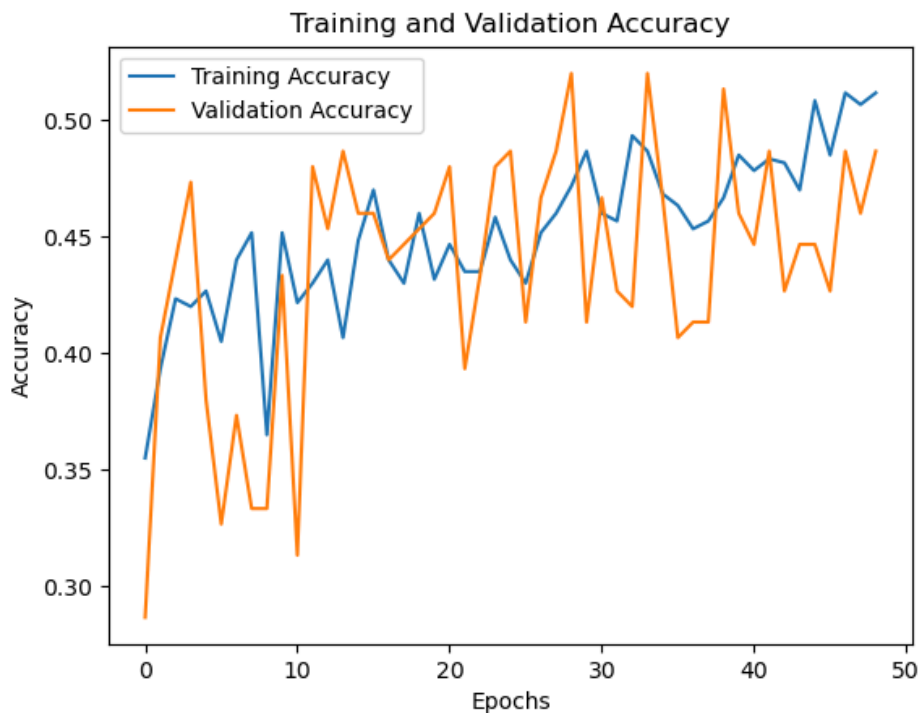
W następnym modelu (**rysunek 3.**) zastosowano kilka technik zapobiegania przeuczenia. Dropout został użyty na poziomie 0.25, regularyzacja L2 na poziomie 0.001, oraz Early Stopping na 5 epoch. Mimo to, model nie osiągnął lepszego, a gorszy wynik na zbiorze walidacyjnym od poprzedniego, a zastosowane techniki nie zapobiegły przeuczeniu.



Rysunek 3. Accuracy drugiego modelu - model przeuczony.

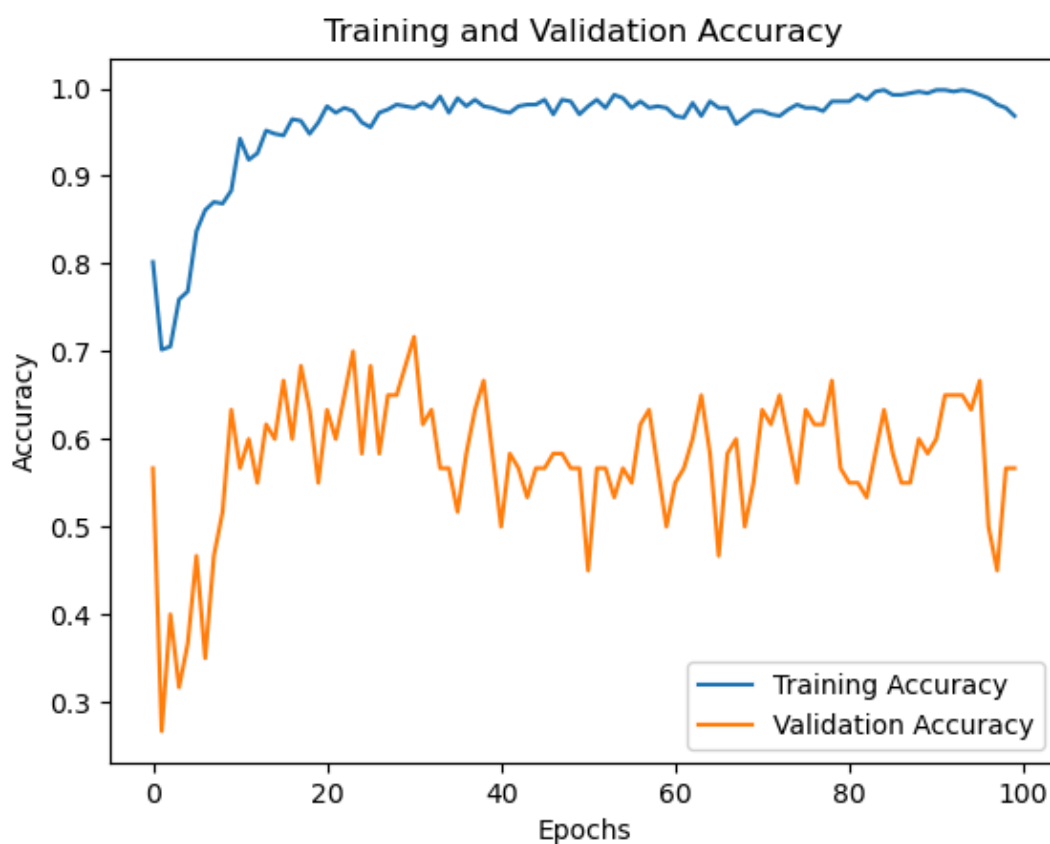
W 3 modelu, widocznym na **rysunku 4.** po zwiększeniu dropoutu do 0.3, regularyzacji L2 na 0.01, użyciu SeparableConv2D zamiast Conv2D, udało się uniknąć przeuczenia, ale dokładność zarówno dla zbioru treningowego, jak i walidacyjnego drastycznie spadła w porównaniu z poprzednimi modelami. Techniki zapobiegania przeuczenia zamiast zmniejszyć

dokładność tylko na zbiorze testowym, zmniejszyły również na zbiorze walidacyjnym, przez co zbiór osiągnął gorsze wyniki od modeli przeuczonych.



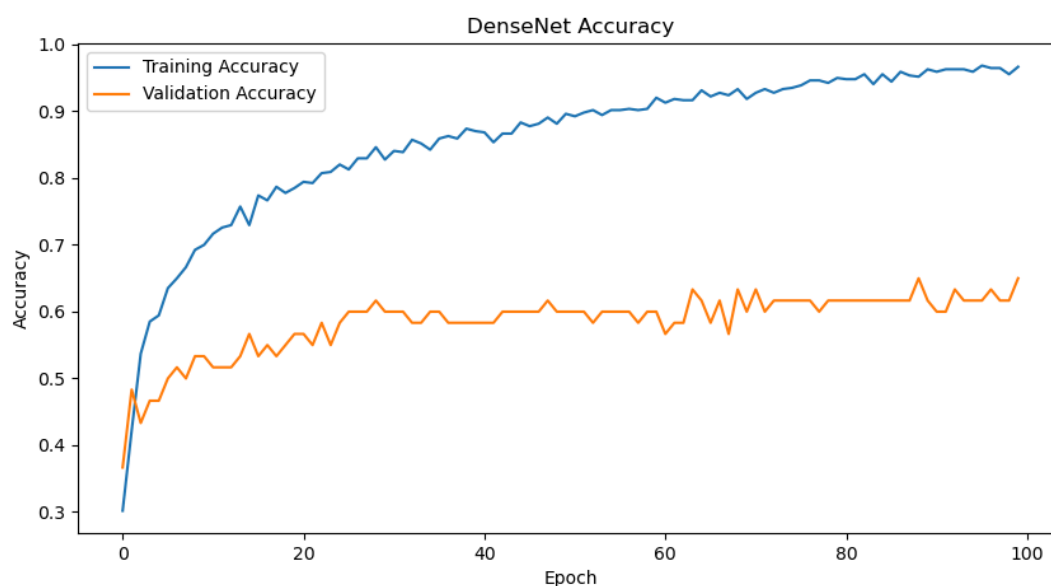
Rysunek 4. Accuracy trzeciego modelu - brak przeuczenia.

Model 4 przedstawiony na **rysunku 5**. zawiera te same techniki, ale znacznie większą ilość warstw. Dropout pozostał na poziomie 0.3, regularyzacja L2 na poziomie 0.001, zastosowano ponownie Conv2D zamiast SeparableConv2D, dzięki czemu udało się osiągnąć najlepszy dotychczasowy wynik dla zbioru walidacyjnego na poziomie blisko 70%.



Rysunek 5. Accuracy czwartego modelu.

Ostatni model widoczny na **rysunku 6.** został stworzony dzięki gotowej sieci DenseNet i osiągnął gorsze wyniki od 4 modelu osiągając wynik 63.3%. Model ten również przeuczył się na tej bazie danych. Porównując wyniki tego modelu do poprzednich można wyciągnąć wnioski, że zrobienie dobrego modelu, który się nie przeuczy jest trudne do osiągnięcia na tym zbiorze danych.



Rysunek 6. Accuracy piątego modelu - DenseNet.

4. Podsumowanie

Zbiór danych "Pet's Facial Expression Image Dataset" był niezwykle trudnym zbiorem do stworzenia nie przeuczonego modelu, który osiągnąłby dobry wynik na zbiorze walidacyjnym. Udało się to osiągnąć raz w modelu 3, ale dokładność modelu mocno spadła w porównaniu z przeuczonymi modelami. Może być to spowodowane małym zbiorem danych (1000 zdjęć), lub samą jakością zdjęć, lub też zbyt dużą różnorodnością zwierząt jak na taką ilość dostępnych zdjęć. Zbiór zawierający większą ilość grafik mógłby być bardziej odpowiednim zbiorem, który wpłynąłby pozytywnie na unikanie przeuczenia.