

Cel laboratorium

Celem laboratorium było zbudowanie własnej architektury sieci przy pomocy pakietów TensorFlow i Keras, a następnie przeprowadzenie eksperymentów badających jakość klasyfikacji dla wykorzystując obrazy z bazy danych AT&T ORL Face Database.

Własna architektura sieci

Poniżej zaproponowana własna architektura sieci splotowej. Najlepiej pokazać kod, który mówi sam za siebie:

```
def create_model(input_shape: tuple) -> Sequential:
    model = tf.keras.models.Sequential([
        Convolution2D(6, (3, 3), activation='relu', input_shape=input_shape,
padding='same'),
        BatchNormalization(),
        MaxPooling2D(pool_size=(2, 2)),
        Dropout(0.25),

        Convolution2D(12, (3, 3), activation='relu', padding='same'),
        MaxPooling2D(pool_size=(2, 2)),
        Dropout(0.25),

        Flatten(),
        Dense(128, activation='relu'),
        Dropout(0.1),
        Dense(50, activation='softmax'),
    ])
    model.compile(loss='sparse_categorical_crossentropy',
optimizer=SGD(momentum=0.9), metrics=['accuracy'])

    return model
```

Przebieg laboratorium

Zadaniem było przy pomocy własnej architektury sieci wykonanie 3 eksperymentów wykorzystując wspomnianą powyżej bazę danych. Wyniki należało przedstawić w formie tabel lub wykresów.

W eksperymentach liczba epok do uczenia modelu została ustawiona na 100 epok. Przy liczbie 10 epok jakość uczenia zazwyczaj nie przekraczała 50%. Przy 100 epokach uczenia jakość uczenia wynosiła w większości 100%, dlatego nie była ona brana pod uwagę przy badaniach.

Podział danych na uczące i testowe był wykonywany przy zachowaniu proporcji 50% / 50%, a random state został ustawiony na 0, aby móc porównać między sobą wyniki z poszczególnych eksperymentów w końcowym podsumowaniu.

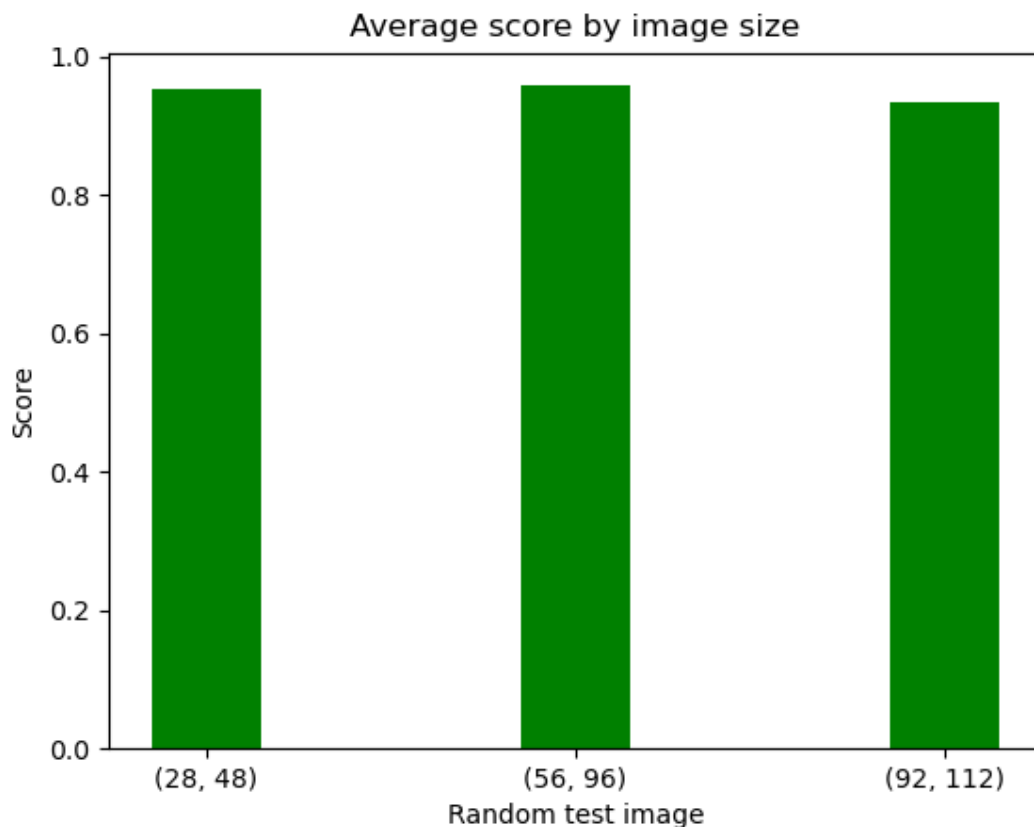
Wpływ rozdzielczości obrazu wejściowego na jakość klasyfikacji

Po wczytaniu obrazów, były one przeskalowywane do następujących rozdzielczości:

- (28, 48),
- (56, 96),
- (92, 112).

Po ich normalizacji model był uczony na przeskalowanych obrazach. Następnie dla 500 iteracji wybierany był losowo obraz testowy i na nim była liczona dokładność. Po zakończeniu zebrane dokładności były uśredniane dla danej rozdzielczości obrazu.

Wyniki zebrano w postaci histogramu:



Pomimo małych przeskalowanych obrazów, które dla człowieka są nieczytelne i dają mu żadnej informacji, model jest w stanie wyekstrahować z niego cechy, na bazie których jest w stanie się nauczyć i poprawnie klasyfikować nowe obrazy.

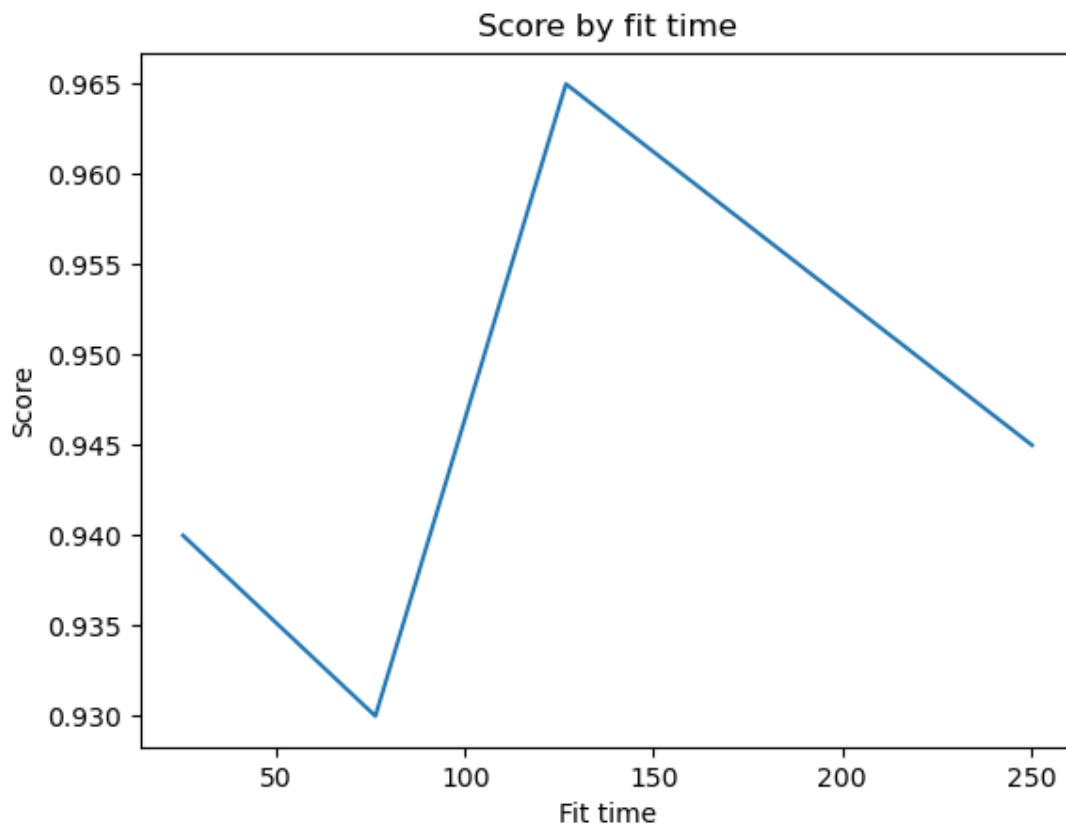
Co ciekawe najniższa jakość klasyfikacji jest dla oryginalnego rozmiaru obrazu, jednak może być to zależne od losowości stanów w modelu. Pomimo tego, wciąż obrazy rozpoznawano się w wysoką dokładnością.

Zaobserwowano w trakcie uczenia, że im mniejszą rozdzielczość ma obraz, tym szybciej klasyfikator się uczy.

Wpływ czasu uczenia do jakości klasyfikacji

Uczenie modelu było przeprowadzane przy następującej ilości epok uczenia: 100, 300, 500, 1000.

Poniższy wykres przedstawia zależność jakości klasyfikacji od czasu uczenia (czas uczenia mierzony w sekundach, czego brakuje na wykresie):

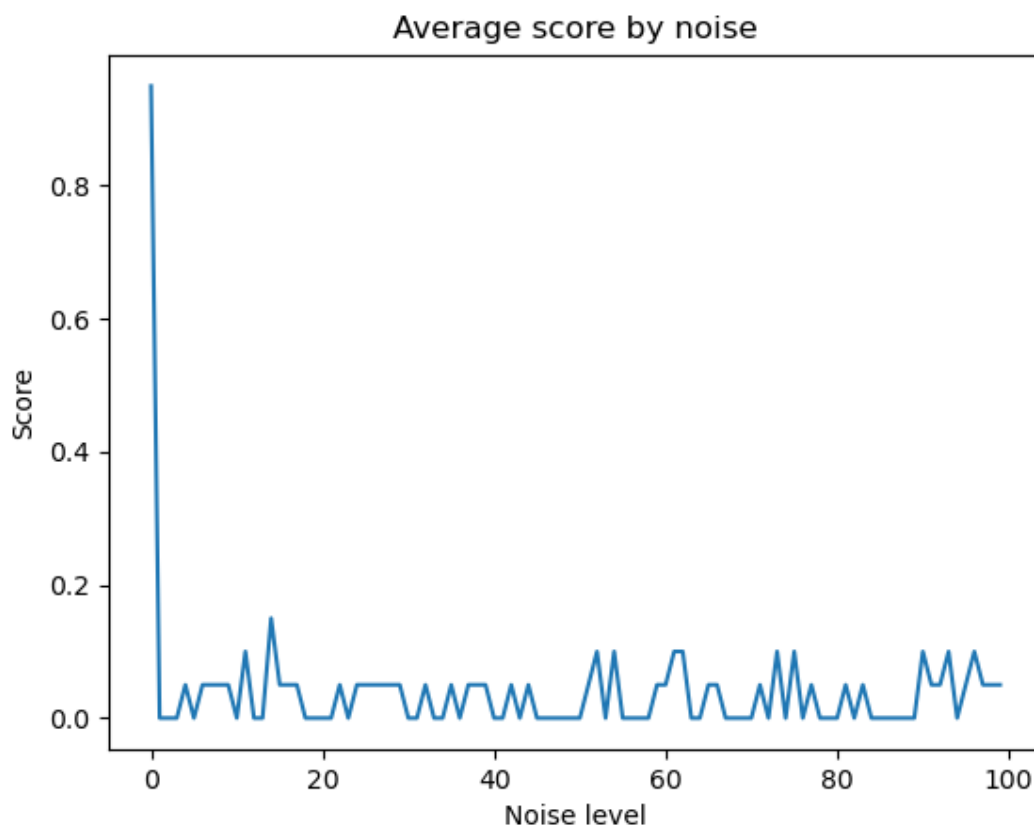


Wykres obrazuje, że odpowiednia liczba epok uczenia jest w stanie zwiększyć jakość klasyfikacji, nawet o 3.5 punktu procentowego. Jednak zbyt dużo epok uczenia pogarsza jakość klasyfikacji, co świadczy o przeuczeniu modelu.

Wpływ zaszumienia obrazu testowego na jakość klasyfikacji

Po nauczaniu modelu, wykonano pętlę w zakresie $<0; 100)$ z krokiem 1 jako poziom szumu. Następnie w 20 iteracjach na każdy poziom szumu losowo wybierano obraz, który zaszumiano. Na koniec wyniki jakości klasyfikacji dla obrazów zaszumionych uśredniano.

Poniższy wykres przedstawia uśrednioną dokładność klasyfikacji w zależności od poziomu szumu:

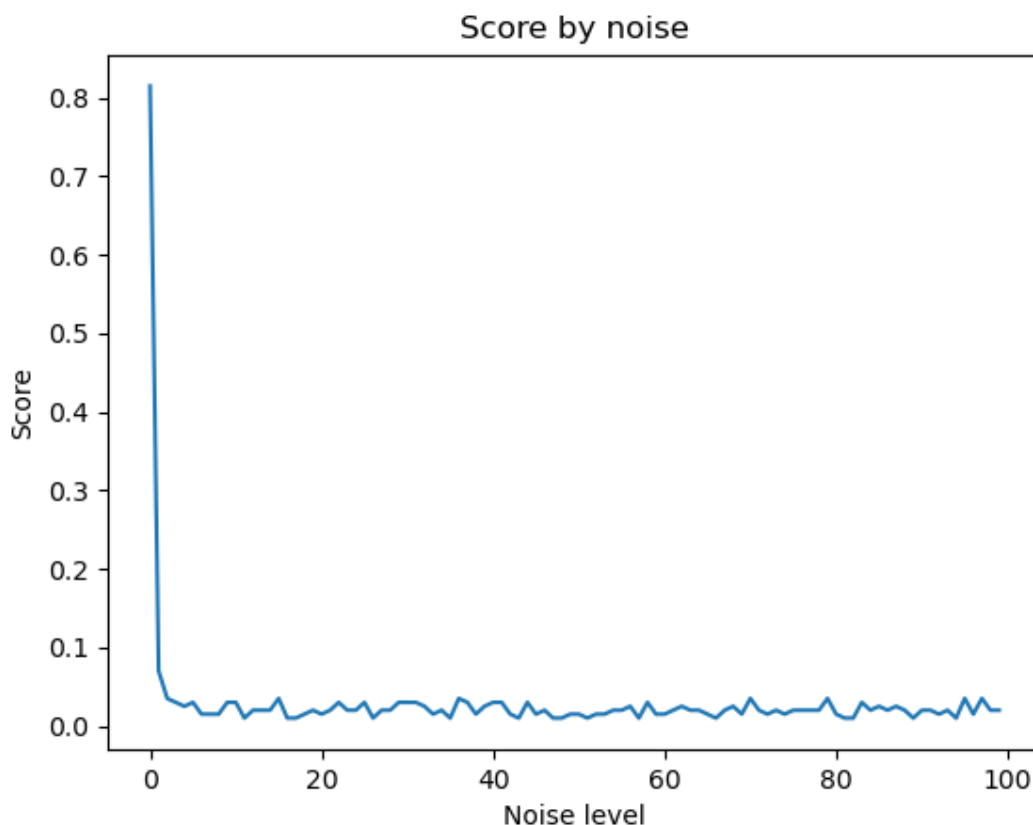


Uśredniona dokładność klasyfikacji jest bardzo niska. Brak szumu (pierwsza wysoka dokładność) pozwala jeszcze uzyskać wysoką dokładność klasyfikacji, podczas gdy ciągle jego zwiększanie bardzo już ją obniża. Zwiększanie poziomu szumu nie zmniejsza proporcjonalnie dokładności klasyfikacji, a spoglądając na powyższy wykres wygląda to na losowe.

Wpływ zaszumienia obrazów uczących na jakość klasyfikacji

Eksperyment odwrotny do poprzedniego. Przed uczeniem modelu, obrazy uczące były zaszumiane według aktualnego poziomu szumu. Następnie w 20 iteracjach losowany był 1 obraz testowy i liczona była dla niego dokładność klasyfikacji. Po wykonaniu tych iteracji dokładności klasyfikacji zebrane z losowych obrazów testowych były uśredniane.

Poniższy wykres przedstawia uśrednioną dokładność klasyfikacji w zależności od poziomu zaszumienia obrazów uczących:



Dokładność klasyfikacji jest bardzo niska, prawie wszystkie przypadki nie osiągają nawet 10%, średnio utrzymując się na 5%. Brak szumu standardowo pozwala uzyskać wysoką dokładność klasyfikacji, natomiast dodanie szumu do danych uczących powoduje znaczne pogorszenie dokładności klasyfikacji.

Również spoglądając na wykres, zwiększanie poziomu szumu nie zmniejsza proporcjonalnie dokładności klasyfikacji i tak jak w poprzednim eksperymencie, wpływ poziomu szumu zdaje się być losowy na dokładność klasyfikacji.

Wnioski

Zmniejszenie rozdzielczości obrazów wykorzystując odpowiednią metodę interpolacji daje równie dobre dokładności klasyfikacji jak obraz w oryginalnej rozdzielczości, a pozwala do tego zaoszczędzić dużo czasu na procesie uczenia.

Odpowiednia liczba epok uczenia jest w stanie zwiększyć otrzymywaną dokładność klasyfikacji, jednak zbyt duża ich liczba może przeuczyć model. Natomiast zbyt mała ich liczba nie jest w stanie poprawnie nauczyć modelu (tak jak wspomniałem, 10 epok to zdecydowanie za mało, 100 jest optymalnie i czasowo i wynikowo).

Zaszumienie jak widać stanowi spory problem dla zaproponowanego modelu i nie jest on na niego odporny. Zarówno w przypadku zaszumienia obrazów testowych jak i obrazów przygotowanych do procesu uczenia.

Gdyby bardziej się zapoznać z parametrami modeli, to można by tak zmodyfikować architekturę sieci, aby była odporna na zaszumienia.

Po wykonaniu eksperymentów i dobieraniu parametrów do zbudowania architektury sieci sprawdza się to co powiedział profesor na wykładzie: "Deep learning to jest bardziej art niż science".