

Opracowanie i porównanie skuteczności trzech modeli sieci neuronowych (bazowego, pośredniego i zoptymalizowanego) w zadaniu rozpoznawania kart do gry na podstawie obrazów cyfrowych. Projekt ma na celu zbadanie wpływu różnych architektur sieci neuronowych i technik uczenia głębokiego na dokładność klasyfikacji kart, a także identyfikację czynników wpływających na poprawę lub pogorszenie jakości predykcji modeli.

Dane

W projekcie wykorzystano publicznie dostępny zbiór danych "Cards Image Dataset-Classification" dostępny na [platformie Kaggle](#).

Zbiór danych zawiera obrazy kart do gry o wysokiej jakości, idealne do projektów z zakresu rozpoznawania obrazów. Każdy obraz ma wymiary 224 x 224 x 3 i jest zapisany w formacie JPG. Obrazy zostały przycięte tak, aby przedstawiały pojedyncze karty, zajmujące ponad 50% powierzchni obrazu.

Zbiór danych jest podzielony na trzy części:

- **Treningowa:** 7094 obrazy,
- **Testowa:** 530 obrazów,
- **Walidacyjna:** 530 obrazów.

Dane są zorganizowane w 53 podkatalogach, gdzie każdy katalog reprezentuje jeden z 53 rodzajów kart. Dodatkowo do zbioru dołączono plik CSV, który ułatwia ładowanie i zarządzanie danymi.

Modele

Model bazowy

Warstwa wejściowa:

- Input o wymiarach (224, 224, 3) - przyjmuje obrazy RGB o rozmiarze 224x224 pikseli
- Rescaling($1./255$) - normalizacja wartości pikseli do zakresu [0,1]

Warstwy ukryte:

- Conv2D z 128 filtrami o rozmiarze (3,3) z funkcją aktywacji ReLU
- MaxPooling2D z oknem (2,2) do redukcji wymiarowości
- Flatten() - spłaszczenie danych do jednowymiarowego wektora

Warstwa wyjściowa:

- Dense z 53 neuronami (odpowiadającymi klasom kart) z funkcją aktywacji softmax

Total params: 2,474,613

Trainable params: 2,474,613

Jest to najprostszy model w projekcie, wykorzystujący podstawowe warstwy konwolucyjne i gęste. Architektura jest stosunkowo płytka, zawierająca tylko jedną warstwę konwolucyjną, co może ograniczać zdolność modelu do wyodrębniania złożonych cech z obrazów kart.

Model pośredni

Warstwa wejściowa:

- Input o wymiarach (224, 224, 3) - przyjmuje obrazy RGB
- Rescaling($1./255$) - normalizacja wartości pikseli

Warstwy konwolucyjne:

- Pierwsza warstwa Conv2D: 32 filtry o rozmiarze (3,3) z aktywacją ReLU
- MaxPooling2D o rozmiarze (2,2)
- Druga warstwa Conv2D: 64 filtry o rozmiarze (3,3) z aktywacją ReLU
- MaxPooling2D o rozmiarze (2,2)
- Dropout(0.3)

Warstwy klasyfikacyjne:

- Flatten() - spłaszczenie danych
- Dense(64) z funkcją aktywacji ReLU
- Dense(13) z funkcją aktywacji softmax

Total params: 1,320,181

Trainable params: 1,320,181

Model ten jest bardziej zaawansowany od bazowego, wykorzystując dwie warstwy konwolucyjne, regularyzację poprzez dropout oraz dwie warstwy gęste o różnych rozmiarach.

Model zoptymalizowany

Warstwa wejściowa:

- **Input(shape=(224, 224, 3))**: Przyjmuje obrazy RGB o wymiarach 224x224.
- **Rescaling(1./255)**: Normalizuje wartości pikseli do zakresu [0, 1].

Warstwy konwolucyjne i poolingowe:

- **Conv2D(32, (3, 3), kernel_initializer='he_normal', use_bias=False, padding='same')**: 32 filtry o rozmiarze (3,3) z inicjalizacją He Normal
- **Activation('relu')**: Funkcja aktywacji ReLU.
- **Conv2D(32, (3, 3), kernel_initializer='he_normal', use_bias=False, padding='same')**: Kolejna warstwa z 32 filtrami o identycznych parametrach.
- **Activation('relu')**: Funkcja aktywacji ReLU.
- **MaxPooling2D((4, 4))**: Pooling z oknem (4,4).
- **SpatialDropout2D(0.3)**: Dropout przestrzenny z prawdopodobieństwem 30%

- **Conv2D(64, (3, 3), kernel_initializer='he_normal', use_bias=False, padding='same')**: 64 filtry z inicjalizacją He Normal.
- **Activation('relu')**
- **Conv2D(64, (3, 3), kernel_initializer='he_normal', use_bias=False, padding='same')**
- **Activation('relu')**
- **MaxPooling2D((2, 2))**: Pooling z oknem (2,2).
- **SpatialDropout2D(0.3)**

- **Conv2D(128, (3, 3), kernel_initializer='he_normal', use_bias=False, padding='same')**
- **BatchNormalization()**: Normalizacja wsadowa.

- **Activation('elu')**: Funkcja aktywacji ELU.
- Powtórzone trzykrotnie dla warstw **Conv2D(128)** z identycznymi parametrami.
- **MaxPooling2D((2, 2))**
- **SpatialDropout2D(0.3)**
- **Conv2D(256, (3, 3), kernel_initializer='he_normal', use_bias=False, padding='same')**
- **BatchNormalization()**.
- **Activation('elu')**
- Powtórzone czterokrotnie dla warstw **Conv2D(256)** z identycznymi parametrami.
- **MaxPooling2D((2, 2))**
- **SpatialDropout2D(0.3)**

Warstwy klasyfikacyjne:

1. **Flatten()**: Spłaszczenie map cech do jednowymiarowego wektora.
2. **Pierwsza warstwa Dense**:
 - **Dense(256, kernel_initializer='he_normal', use_bias=False)**: 256 neuronów, inicjalizacja He Normal
 - **BatchNormalization()**
 - **Activation('elu')**
 - **Dropout(0.2)**: Dropout z prawdopodobieństwem 20%.
3. **Druga warstwa Dense**:
 - **Dense(128, kernel_initializer='he_normal', use_bias=False)**: 128 neuronów, inicjalizacja He Normal
 - **BatchNormalization()**
 - **Activation('elu')**
 - **Dropout(0.2)**
4. **Warstwa wyjściowa**:
 - **Dense(53, activation='softmax')**: 53 neuronów odpowiadających klasom kart, z funkcją aktywacji softmax (klasyfikacja wieloklasowa).

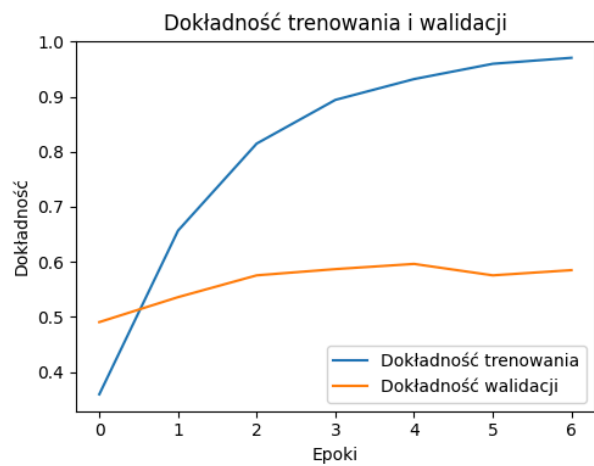
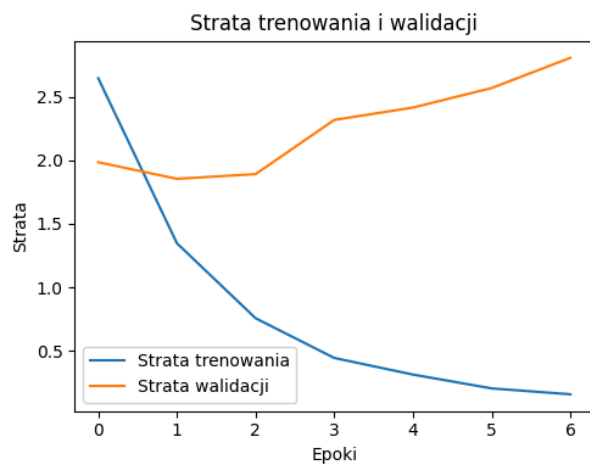
Total params: 5,756,437

Trainable params: 5,752,853

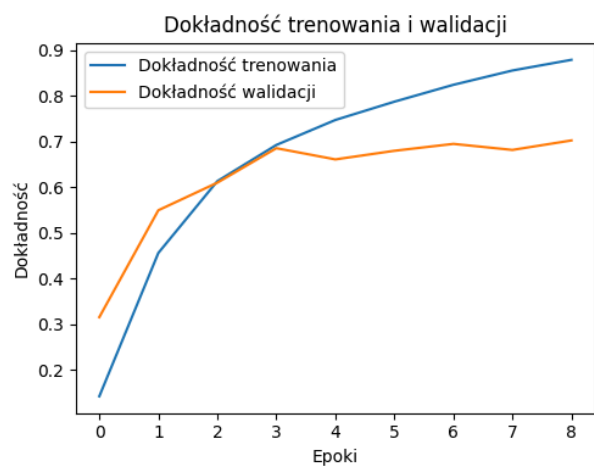
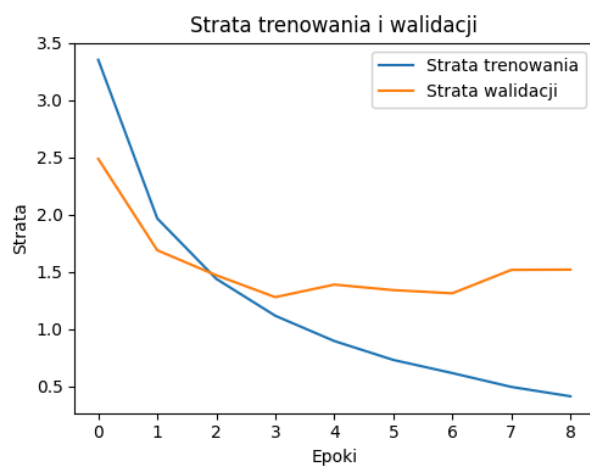
Jest to jeden z najbardziej złożonych modeli wykorzystanych w projekcie dający najlepsze wyniki. Dzięki wielu warstwom konwolucyjnym o zwiększającej się liczbie filtrów, model jest w stanie wydobywać hierarchiczne cechy z obrazów, co pozwala na klasyfikację 53 klas kart.

Wykresy przebiegu uczenia

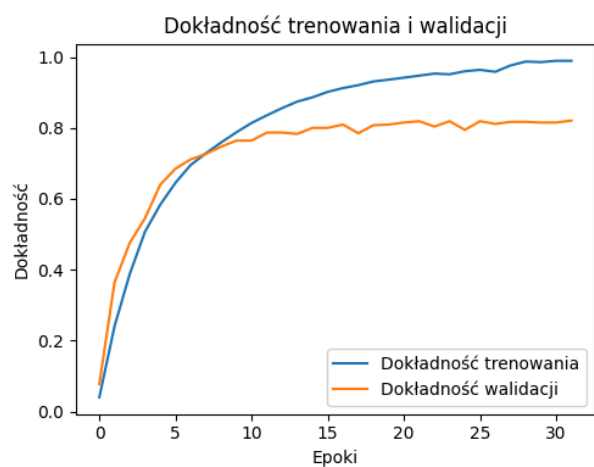
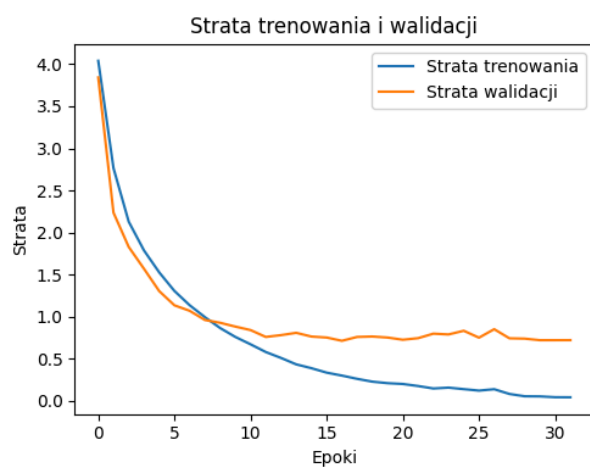
Model bazowy



Model pośredni



Model optymalny



Analiza modeli

Model bazowy:

- Wykazuje oznaki przeuczenia - dokładność trenowania osiąga około 95%, podczas gdy dokładność walidacji zatrzymuje się na poziomie 60%
- Strata walidacji rośnie wraz z kolejnymi epokami, co potwierdza problem z przeuczeniem

Model pośredni:

- Pokazuje lepszą generalizację niż model bazowy
- Dokładność walidacji osiąga około 70%
- Mniejsza różnica między stratą trenowania a stratą walidacji wskazuje na lepszą równowagę modelu

Model optymalny:

- Prezentuje najlepszą charakterystykę uczenia
- Dokładność trenowania i walidacji zbiegają się do podobnych wartości (około 80-90%)
- Strata trenowania i walidacji stabilizują się na niskim poziomie
- Wykazuje najmniejsze oznaki przeuczenia spośród wszystkich modeli

Wyniki F1-Score

Class	Bazowy F1-Score	Pośredni F1-Score	Optymalny F1-Score
Ace of Clubs	0.73	0.84	0.89
Ace of Diamonds	0.67	0.71	0.80
Ace of Hearts	0.78	0.86	0.95
Ace of Spades	0.74	0.61	0.82
Eight of Clubs	0.36	0.63	0.74
Eight of Diamonds	0.43	0.47	0.82
Eight of Hearts	0.36	0.59	0.86
Eight of Spades	0.67	0.53	0.86
Five of Clubs	0.75	0.78	0.84
Five of Diamonds	0.71	0.80	0.95
Five of Hearts	0.67	0.70	0.86
Five of Spades	0.59	0.62	0.70
Four of Clubs	0.62	0.75	0.78
Four of Diamonds	0.44	0.78	0.89
Four of Hearts	0.46	0.64	0.63
Four of Spades	0.64	0.73	0.89
Jack of Clubs	0.24	0.53	0.82
Jack of Diamonds	0.35	0.56	0.84
Jack of Hearts	0.53	0.50	0.74
Jack of Spades	0.36	0.35	0.84
Joker	0.80	0.90	0.95
King of Clubs	0.82	0.67	0.82
King of Diamonds	0.57	0.44	0.80
King of Hearts	0.54	0.27	0.67
King of Spades	0.25	0.32	0.76
Nine of Clubs	0.50	0.75	0.80
Nine of Diamonds	0.57	0.57	0.60
Nine of Hearts	0.75	0.63	0.90
Nine of Spades	0.78	0.90	0.91
Queen of Clubs	0.55	0.61	0.87
Queen of Diamonds	0.33	0.26	0.76
Queen of Hearts	0.18	0.47	0.84
Queen of Spades	0.89	0.71	0.86
Seven of Clubs	0.44	0.60	0.84
Seven of Diamonds	0.54	0.61	0.87
Seven of Hearts	0.31	0.50	0.76
Seven of Spades	0.78	0.76	0.91
Six of Clubs	0.55	0.76	0.86
Six of Diamonds	0.70	0.83	0.87
Six of Hearts	0.45	0.74	0.63
Six of Spades	0.47	0.70	0.82
Ten of Clubs	0.44	0.63	0.70
Ten of Diamonds	0.40	0.47	0.90
Ten of Hearts	0.74	0.91	0.95
Ten of Spades	0.67	0.70	0.76
Three of Clubs	0.71	0.78	0.71
Three of Diamonds	0.74	0.86	0.91
Three of Hearts	0.78	0.90	1.00
Three of Spades	0.75	0.80	0.91
Two of Clubs	0.29	0.78	0.82
Two of Diamonds	0.42	0.57	0.89
Two of Hearts	0.60	0.67	0.80
Two of Spades	0.40	0.67	0.78

Metric	P Bazowy	R Bazowy	F1 Bazowy	P Pośred.	R Pośred.	F1 Pośred.	P Optymal	R Optymal	F1 Optymal
Accuracy	-	-	0.55	-	-	0.66	-	-	0.83
Macro Avg	0.64	0.55	0.56	0.68	0.66	0.65	0.84	0.83	0.83
Weighted Avg	0.64	0.55	0.56	0.68	0.66	0.65	0.84	0.83	0.83

Analiza wyników F1-Score

Porównanie modeli:

- Model bazowy osiąga średnią dokładność i średni F1-Score
- Model pośredni uzyskuje średnią dokładność i średni F1-Score
- Model optymalny pokazuje najlepsze wyniki ze średnią dokładnością i średnim F1-Score

Analiza klas:

Najwyższe wyniki F1-Score dla modelu optymalnego osiągnięto dla kart:

- Three of Hearts (1.00)
- Ace of Hearts (0.95)
- Joker (0.95)
- Five of Diamonds (0.95)
- Ten of Hearts (0.95)

Najniższe wyniki F1-Score dla modelu optymalnego zanotowano dla:

- King of Spades (0.70)
- Ten of Clubs (0.70)
- King of Hearts (0.67)
- Four of Hearts (0.63)
- Six of Hearts (0.63)

Model bazowy

Confusion Matrix

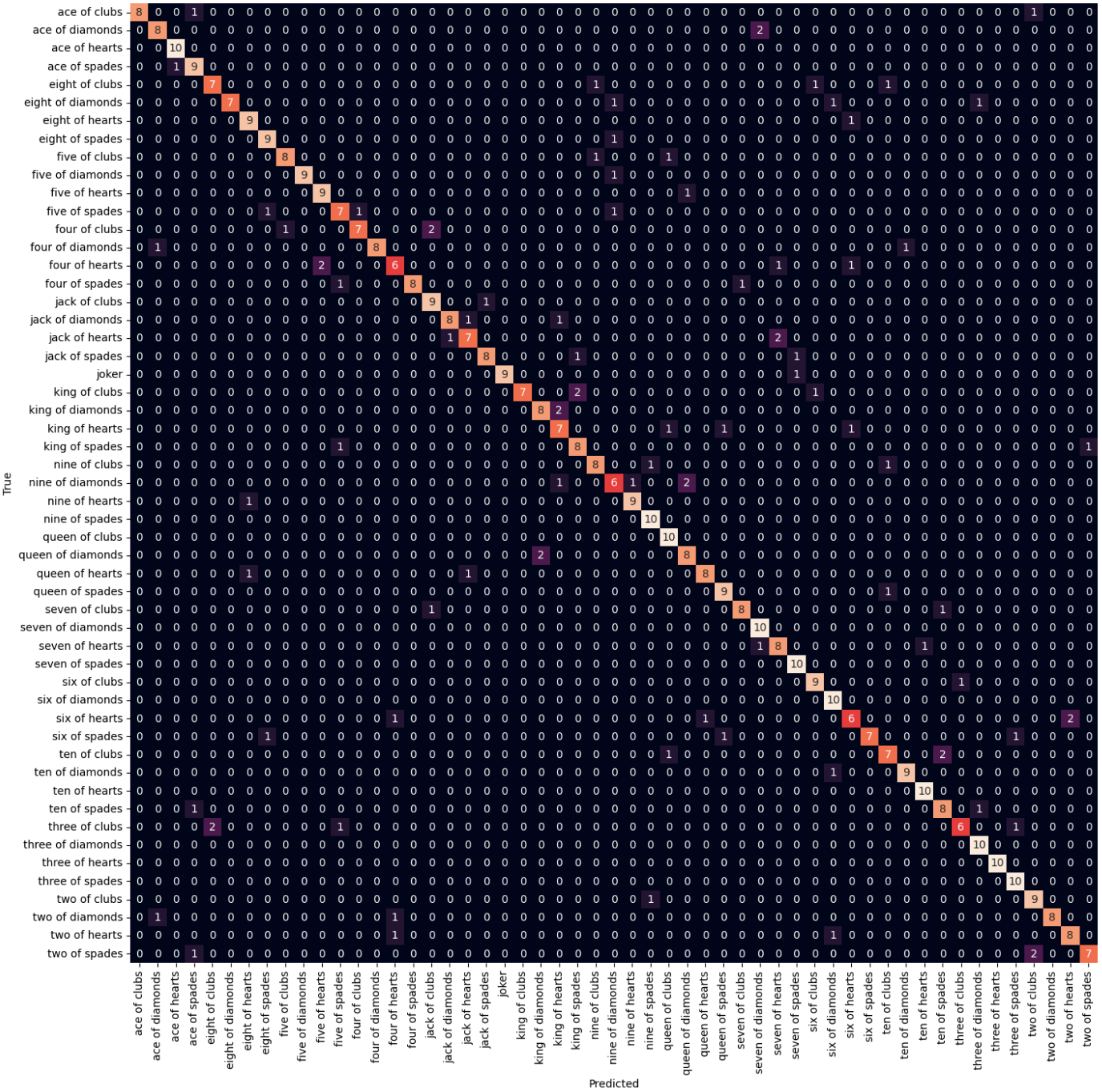
[illegible]

Model pośredni

Confusion Matrix

[illegible]

Model optymalny



Analiza macierzy pomyłek dla poszczególnych modeli

Model bazowy charakteryzuje się dużym rozproszeniem predykcji. Widoczne są liczne błędne klasyfikacje, szczególnie w przypadku kart o podobnych cechach. Model często myli karty tego samego koloru lub wartości. Szczególnie problematyczne jest rozpoznawanie figur (króli, dam i waletów), gdzie występuje znacząca liczba błędnych klasyfikacji.

Model pośredni wykazuje zauważalną poprawę w stosunku do modelu bazowego. Wartości na głównej przekątnej macierzy są wyższe, co wskazuje na lepszą skuteczność klasyfikacji. Nadal występują pewne błędy w rozpoznawaniu kart o podobnych cechach, ale ich liczba jest znacząco mniejsza. Widoczna jest poprawa w klasyfikacji figur, choć nadal zdarzają się pomyłki między kartami tego samego koloru.

Model optymalny prezentuje najlepszą skuteczność klasyfikacji spośród wszystkich trzech modeli. Macierz błędów pokazuje wyraźną dominację wartości na głównej przekątnej, co świadczy o wysokiej precyzji klasyfikacji. Błędne predykcje są znacznie rzadsze, a najgorszy przypadek to 4 błędne klasyfikacje na 10 przykładów.

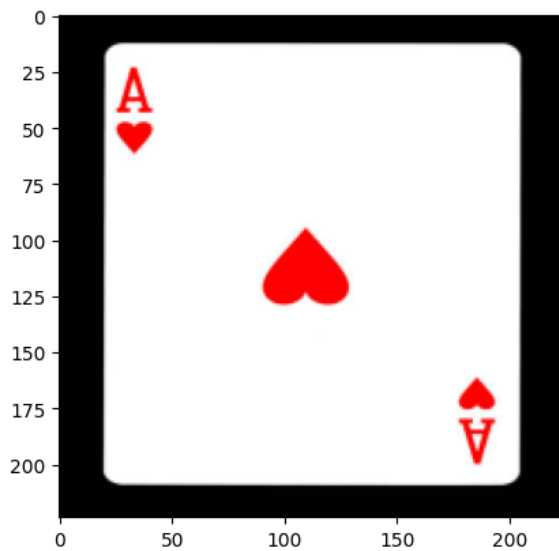
Podsumowanie

Wypracowany model optymalny osiąga najlepsze wyniki pod względem generalizacji i stabilności uczenia. Wprowadzone w nim modyfikacje skutecznie przeciwdziałają przeuczeniu, które było widoczne w modelu bazowym. Dłuższy czas treningu (30 epok) pozwolił na lepszą stabilizację parametrów modelu.

Średni F1-Score i dokładności na poziomie 0.83 świadczy o dość wysokiej skuteczności modelu w zadaniu rozpoznawania kart, jednak ze względu na ograniczoną liczbę danych testowych i walidacyjnych (po 10 obrazów dla każdej klasy), rzeczywiste wyniki modelu mogą się znacząco różnić od uzyskanych.

Macierze błędów dla wszystkich trzech modeli mimo ograniczonej ilości danych pokazują progresywną poprawę skuteczności klasyfikacji.

Wykorzystanie najlepszego modelu do rozpoznania przykładowej karty



1/1 ————— 1s 835ms/step

	Prediction	In %
ace of hearts	0.210318	21.0
four of spades	0.141025	14.1
four of clubs	0.106215	10.6
ten of diamonds	0.093023	9.3
two of hearts	0.058674	5.9
four of hearts	0.042337	4.2
king of spades	0.041617	4.2
ace of spades	0.039988	4.0
two of diamonds	0.036283	3.6
nine of diamonds	0.029797	3.0
four of diamonds	0.024819	2.5
queen of spades	0.018896	1.9
five of spades	0.017962	1.8
six of diamonds	0.013450	1.3
nine of hearts	0.012031	1.2
ten of hearts	0.010699	1.1

W tym przypadku model prawidłowo dopasował klasę "ace of hearts" z prawdopodobieństwem 21%