

# Projekt: Klasyfikacja zmian skórnych - melanoma vs. benign

---

**Autor:** Zofia Pietrucha **Data:** Maj 2025

**Przedmiot:** Informatyka Praktyczna - Sztuczna Inteligencja

**Prowadzący:** mgr Grzegorz Madejski, mgr Maciej Stankiewicz

## 1. Opis projektu

### 1.1 Cel projektu

Celem projektu było stworzenie systemu klasyfikacyjnego do automatycznego rozpoznawania znamion skórnych jako łagodne (benign) lub złośliwe (melanoma). Projekt realizowałam w formie **50% badania AI + 50% tworzenie aplikacji**, zgodnie z wytycznymi prowadzącego.

### 1.2 Dataset

Wykorzystałam dataset **Skin Moles: Benign vs. Malignant (Melanoma) - ISIC 2019** zawierający obrazy znamion skórnych w dwóch klasach:

- **Benign** (łagodne znamiona)
- **Melanoma** (czerniak)

Link: [https://www.kaggle.com/datasets/adisongoh/skin-moles-benign-vs-malignant-melanoma-isic19/data?select=label\\_MEL](https://www.kaggle.com/datasets/adisongoh/skin-moles-benign-vs-malignant-melanoma-isic19/data?select=label_MEL)

### 1.3 Podział danych

Dane podzieliłam w najpierw w proporcji **70% treningowe / 30% testowe**. Jednak później zmieniłam na podział **80% treningowe / 20% testowe** z zachowaniem stratyfikacji klas, dla większej liczby obrazów treningowych.

## 2. Badane modele i techniki

Przebadalałam łącznie **9 różnych podejścia** do klasyfikacji, od tradycyjnych metod uczenia maszynowego po zaawansowane sieci neuronowe i modele hybrydowe.

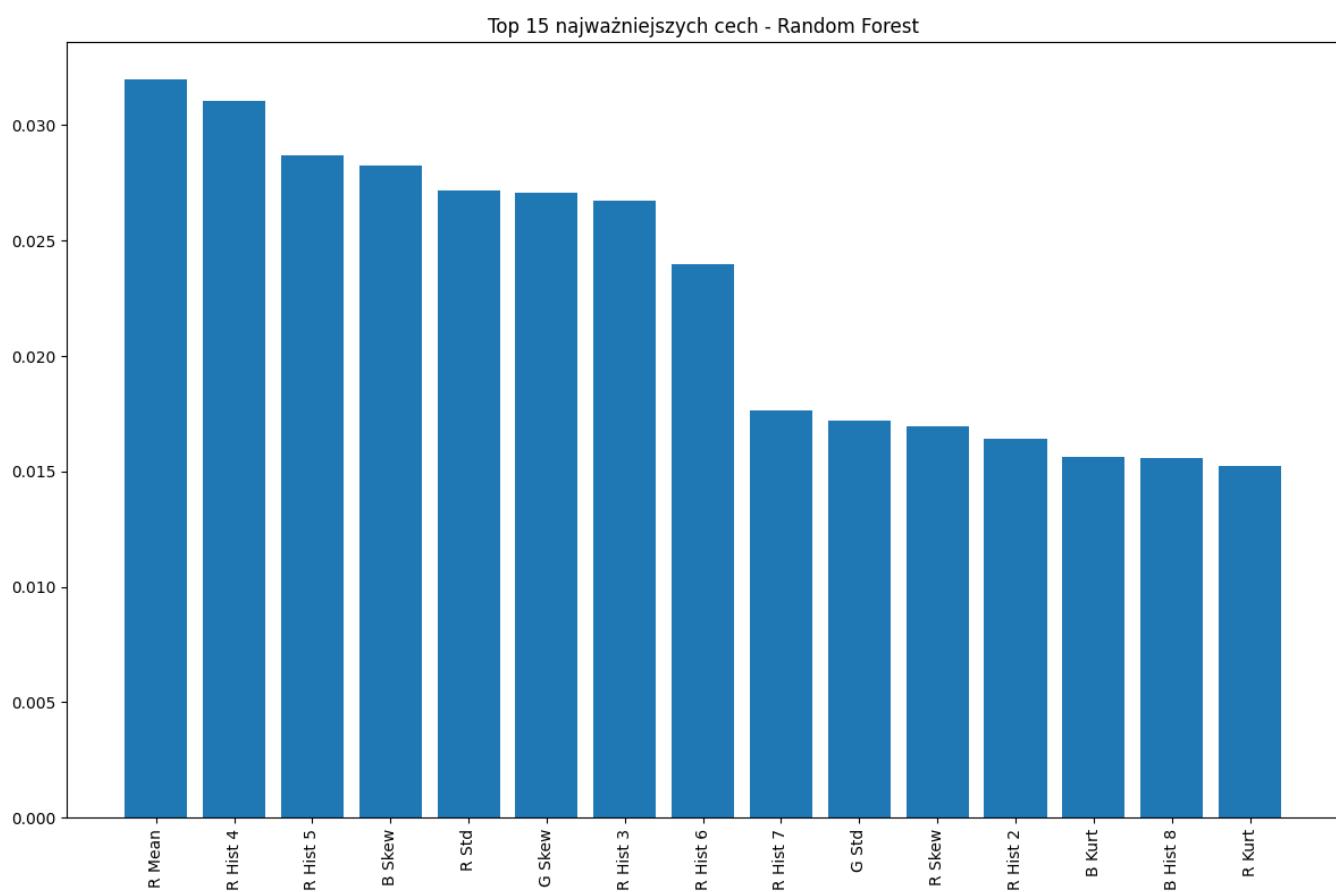
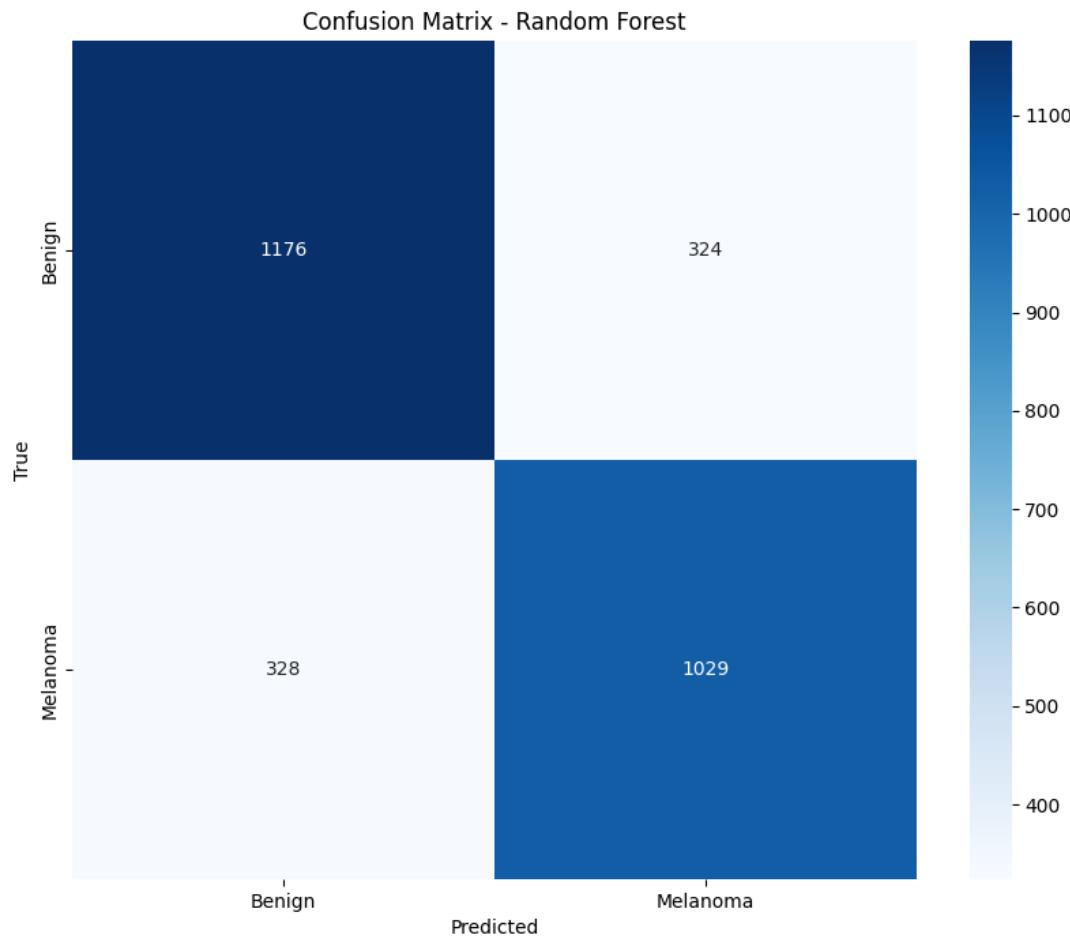
### 2.1 Random Forest (podstawowy)

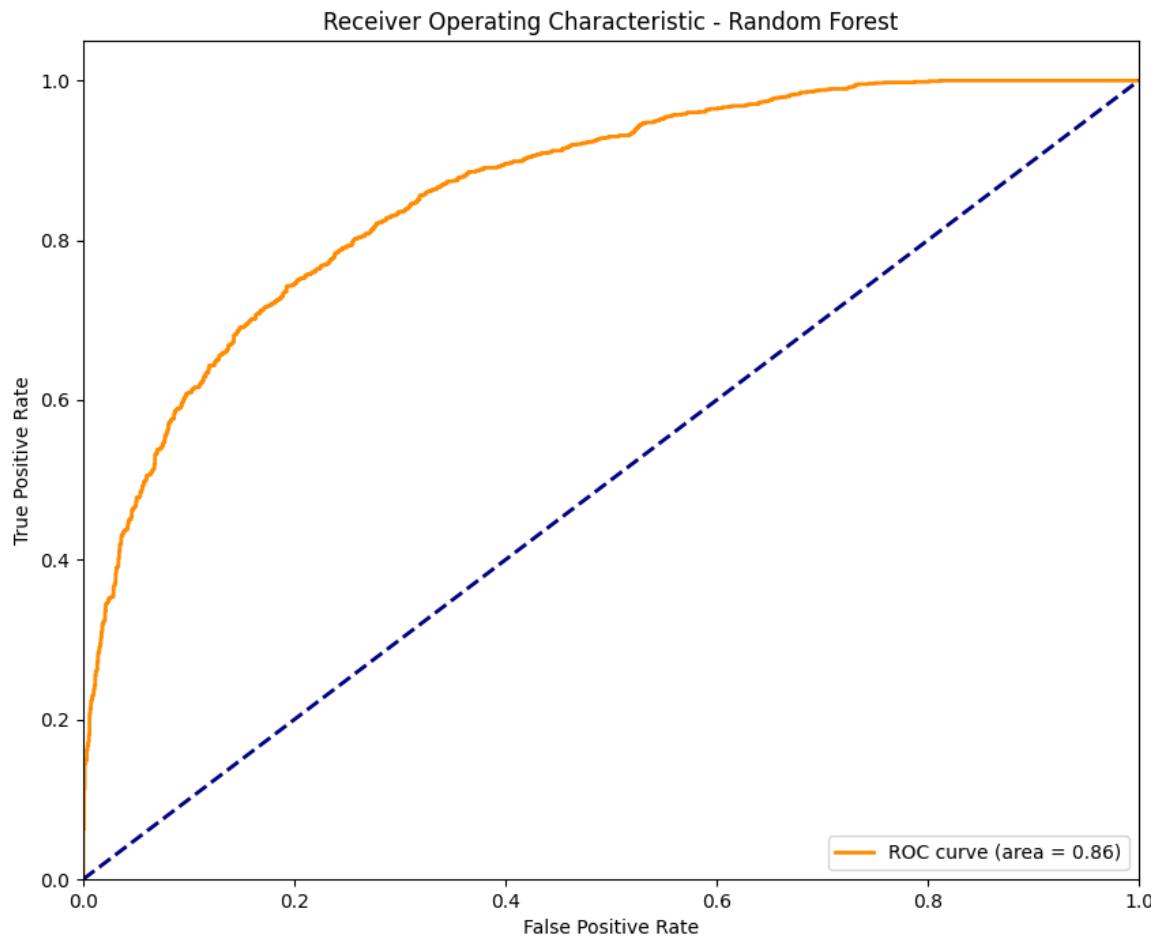
#### **Architektura i parametry:**

- Ekstrakcja podstawowych cech kolorów (RGB)
- Histogramy kolorów (20 przedziałów na kanał)
- Statystyki (średnia, odchylenie standardowe, skewness, kurtosis)
- Random Forest: 200 drzew, max\_depth=20
- Łącznie ~**72 cechy**

#### **Wyniki:**

- **Dokładność:** 75.9%
- **ROC AUC:** 0.86





**Analiza:** Podstawowy Random Forest osiągnął zadowalające wyniki, z najważniejszymi cechami związanymi ze średnimi wartościami kanałów kolorów. Histogramy kolorów również okazały się istotne dla klasyfikacji.

## 2.2 Random Forest (ulepszony)

### Wprowadzone ulepszenia :

- **Cechy ABCDE dla melanomii:** asymetria, cechy brzegów, różnorodność kolorów
- **Dodatkowe przestrzenie kolorów:** HSV (odcień, nasycenie, jasność), LAB (jasność, osie kolorów niezależne od urządzenia)
- **Zaawansowane cechy tekstury:** LBP (lokalne wzorce binarne), GLCM (macierz współwystępowania pikseli), filtry Gabora (wykrywanie wzorców w różnych orientacjach)
- **Standaryzacja cech** (StandardScaler - normalizacja do średniej 0 i odch. std. 1)
- **Optymalizacja hiperparametrów** (GridSearchCV - automatyczne przeszukiwanie najlepszych parametrów)
- **Cross-validation** (5-fold - podział na 5 części dla rzetelnej oceny)
- Łącznie ~200+ cech

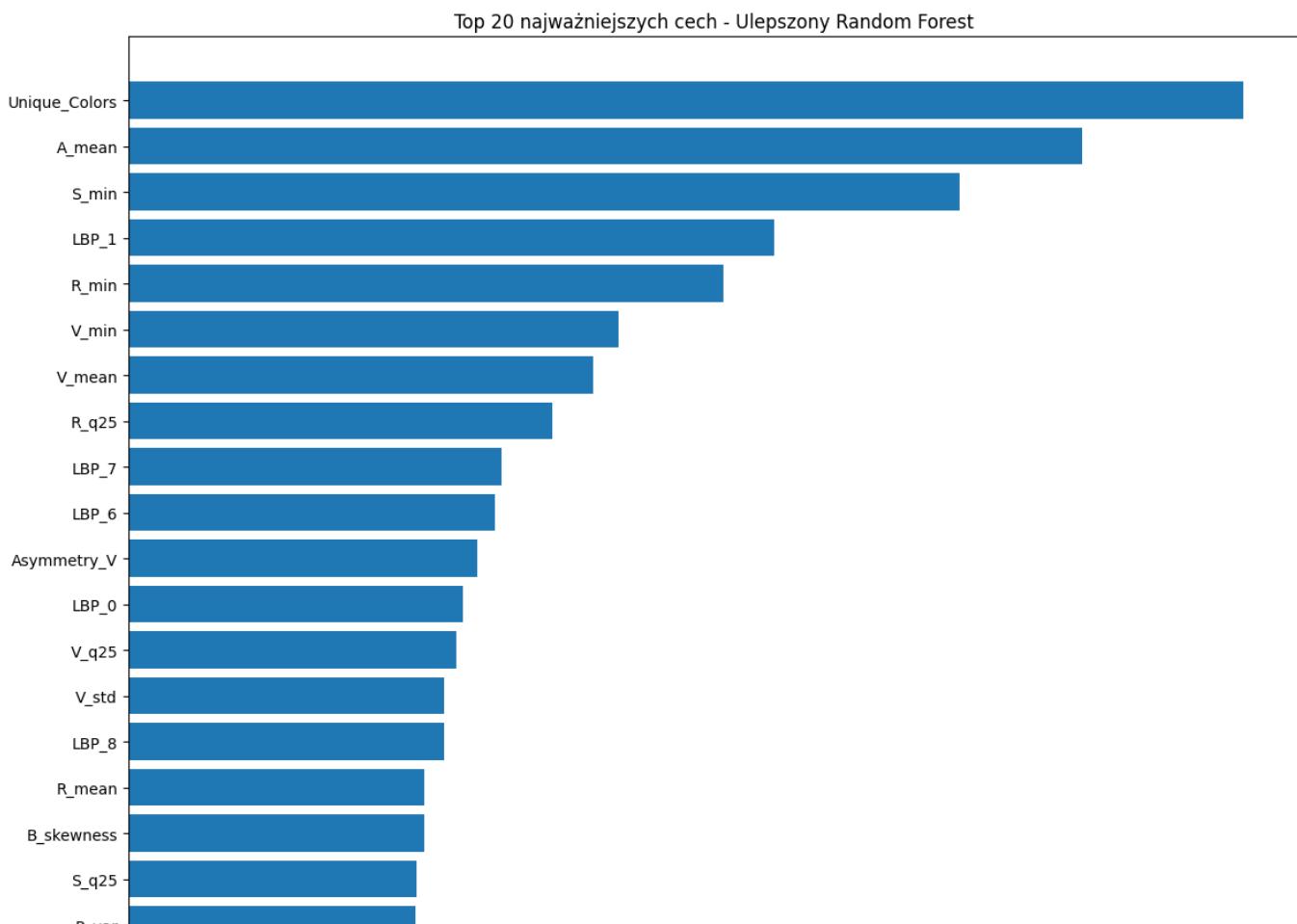
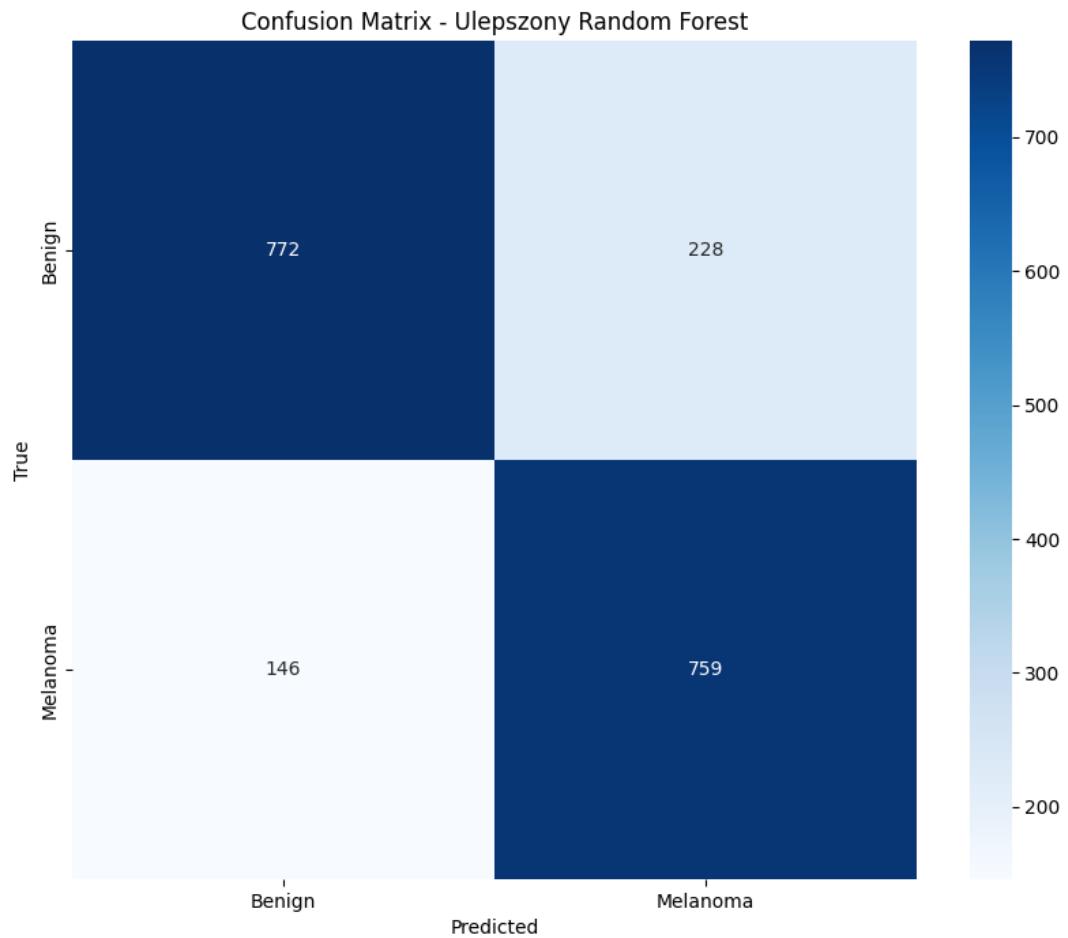
### Parametry po optymalizacji:

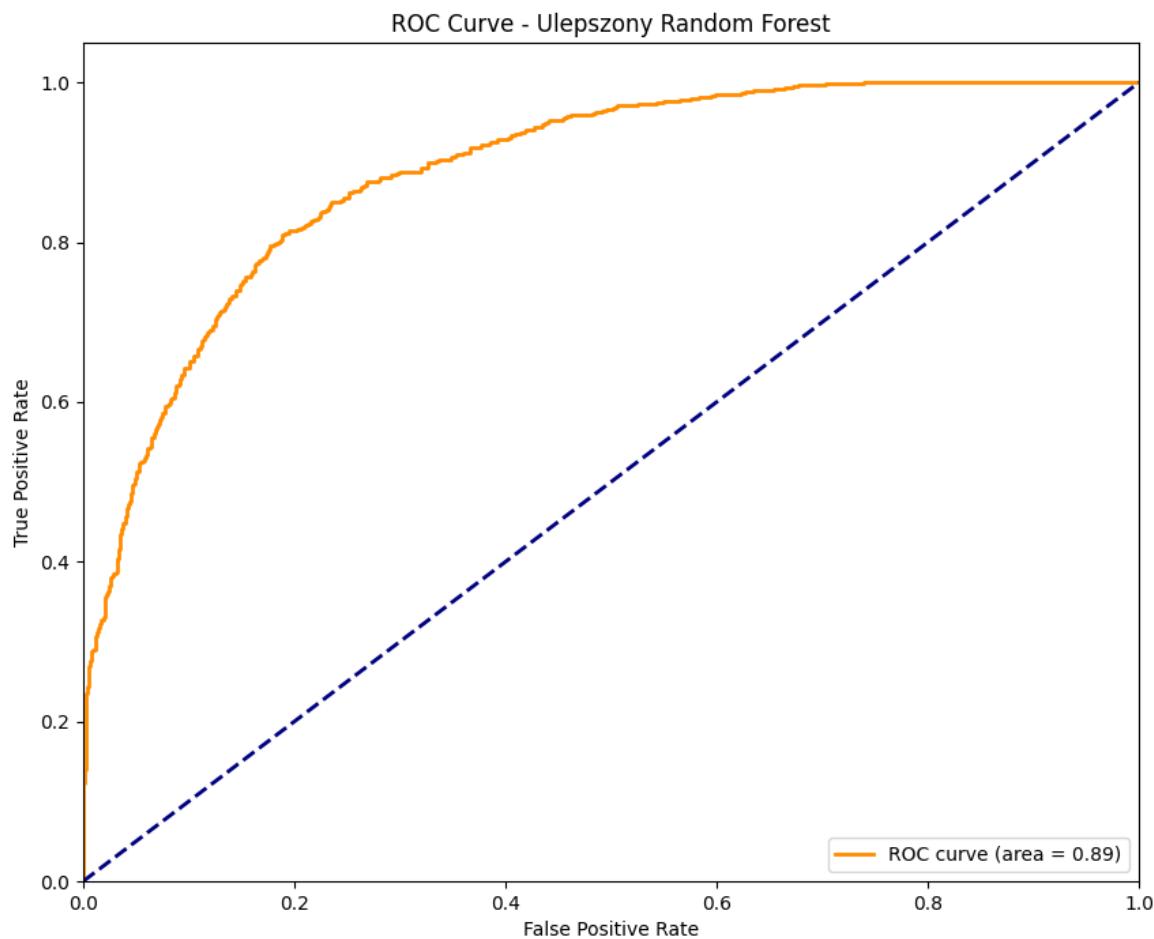
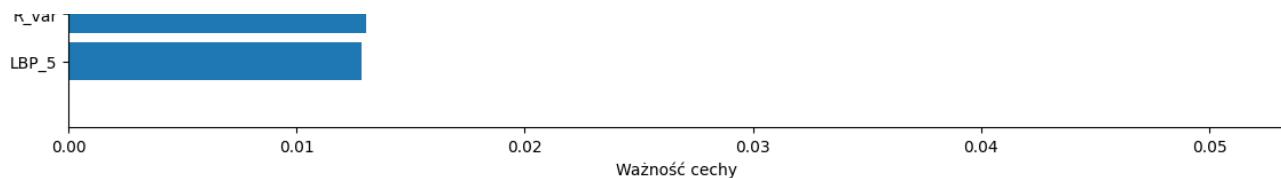
- n\_estimators=300, max\_depth=20
- class\_weight='balanced'

- Funkcja celu: F1-score

### Wyniki:

- **Dokładność:** 83.8%
- **ROC AUC:** 0.89
- **CV F1 Score:**  $0.78 \pm 0.04$





**Analiza:** Znacząca poprawa wyników dzięki cechom specyficznych dla melanom. Najważniejsze okazały się cechy związane z różnorodnością kolorów i asymetrią znamion.

## 2.3 CNN (podstawowy)

### Architektura:

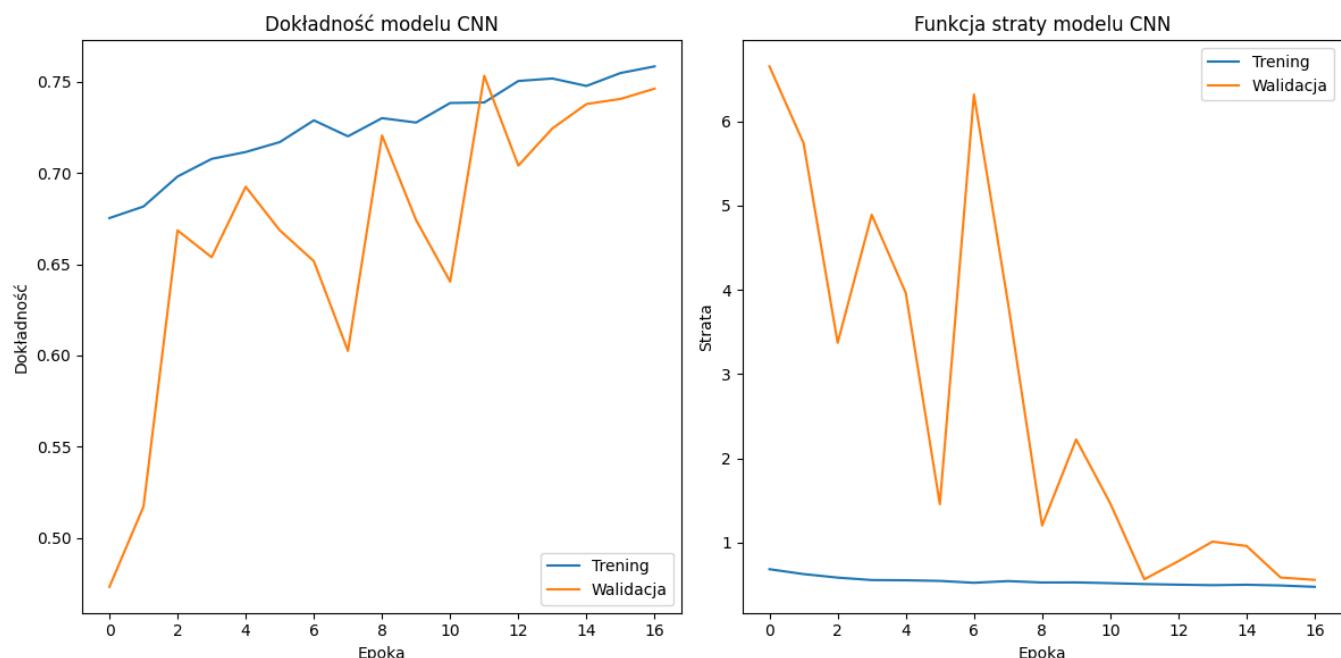
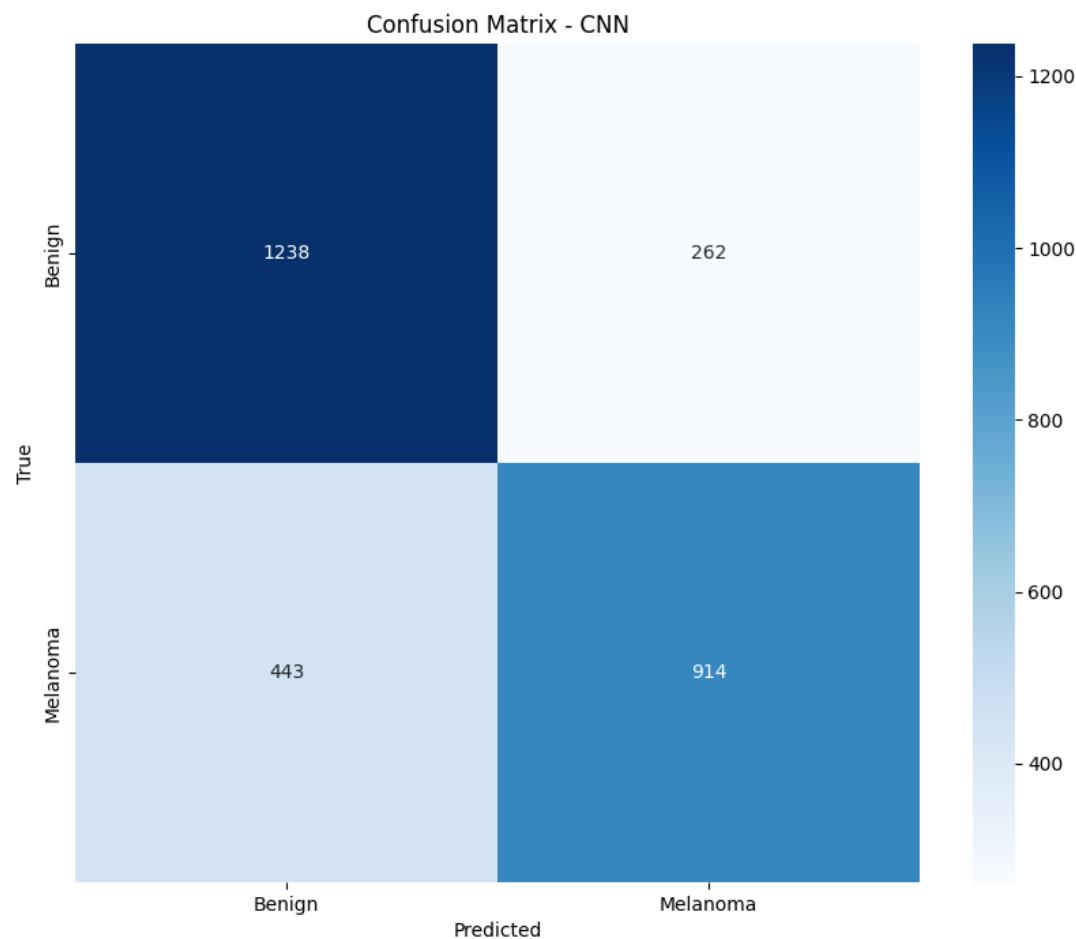
- 3 bloki konwolucyjne (32, 64, 128 filtrów)
- BatchNormalization + MaxPooling + Dropout
- Dwie warstwy Dense (512, 1)
- Input: 224x224x3

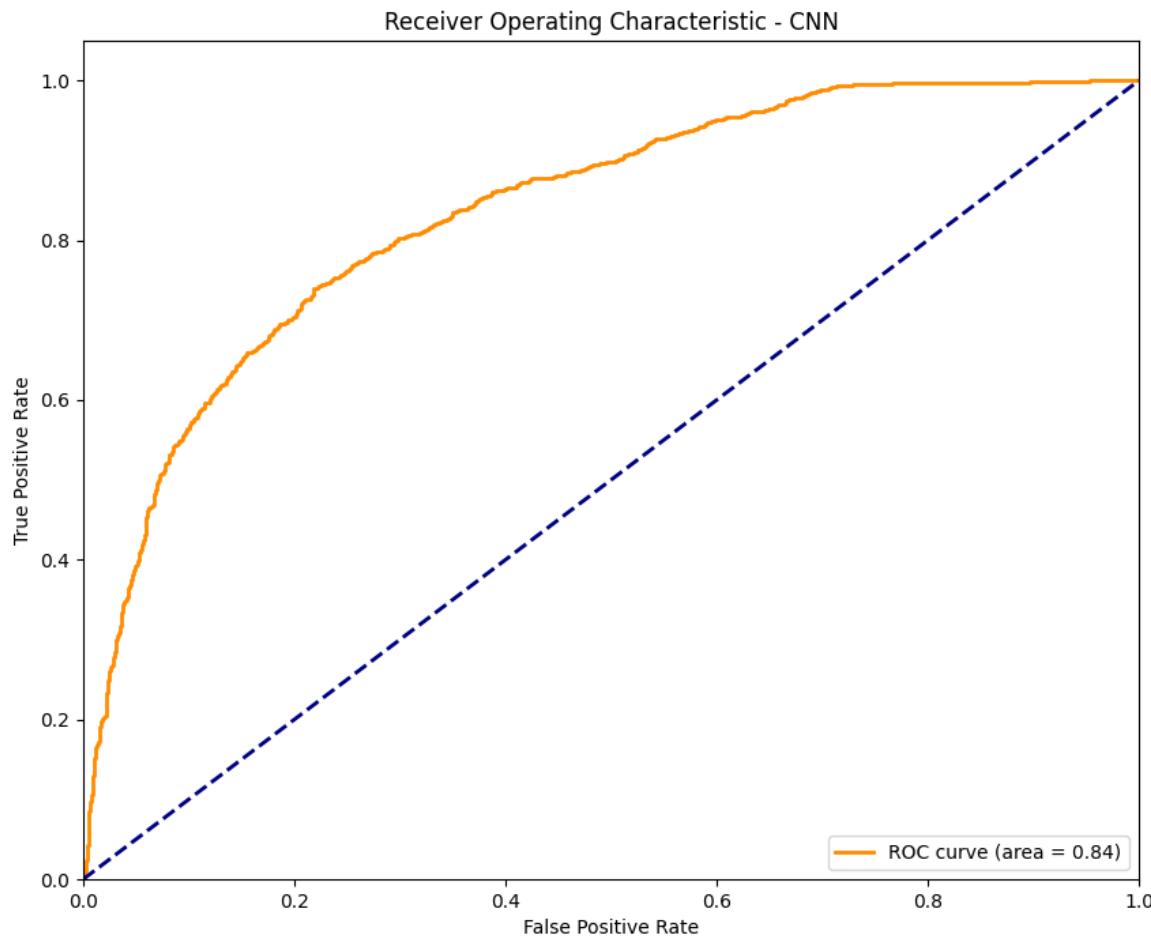
### Parametry treningu:

- Optimizer: Adam ( $\text{lr}=0.0001$ )
- Loss: binary\_crossentropy
- Epochs: 20, batch\_size=32
- Augmentacja: rotacja, odbicia, zoom

## Wyniki:

- **Dokładność:** 71.9%
- **ROC AUC:** 0.84
- **Czułość na melanomę:** 67.4%





**Analiza:** Podstawowy CNN wykazał stabilne uczenie się bez oznak przeuczenia - zarówno dokładność treningowa jak i walidacyjna rosły w podobnym tempie. Głównym problemem była niska czułość na melanomę, prawdopodobnie spowodowana niezbalansowaniem klas i brakiem odpowiedniego ważenia podczas treningu.

## 2.4 CNN (ulepszony)

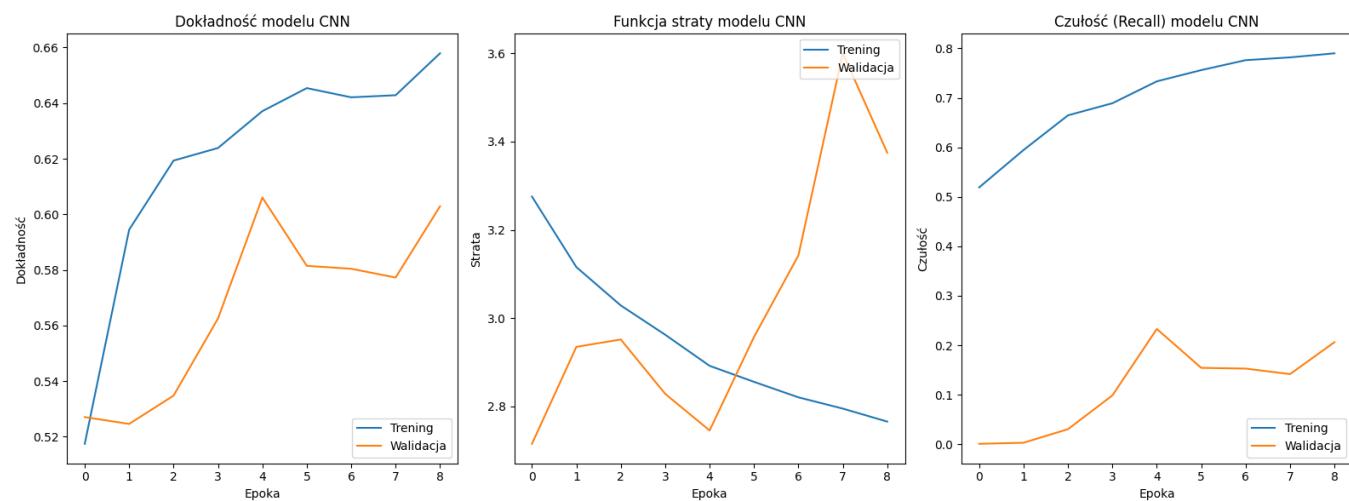
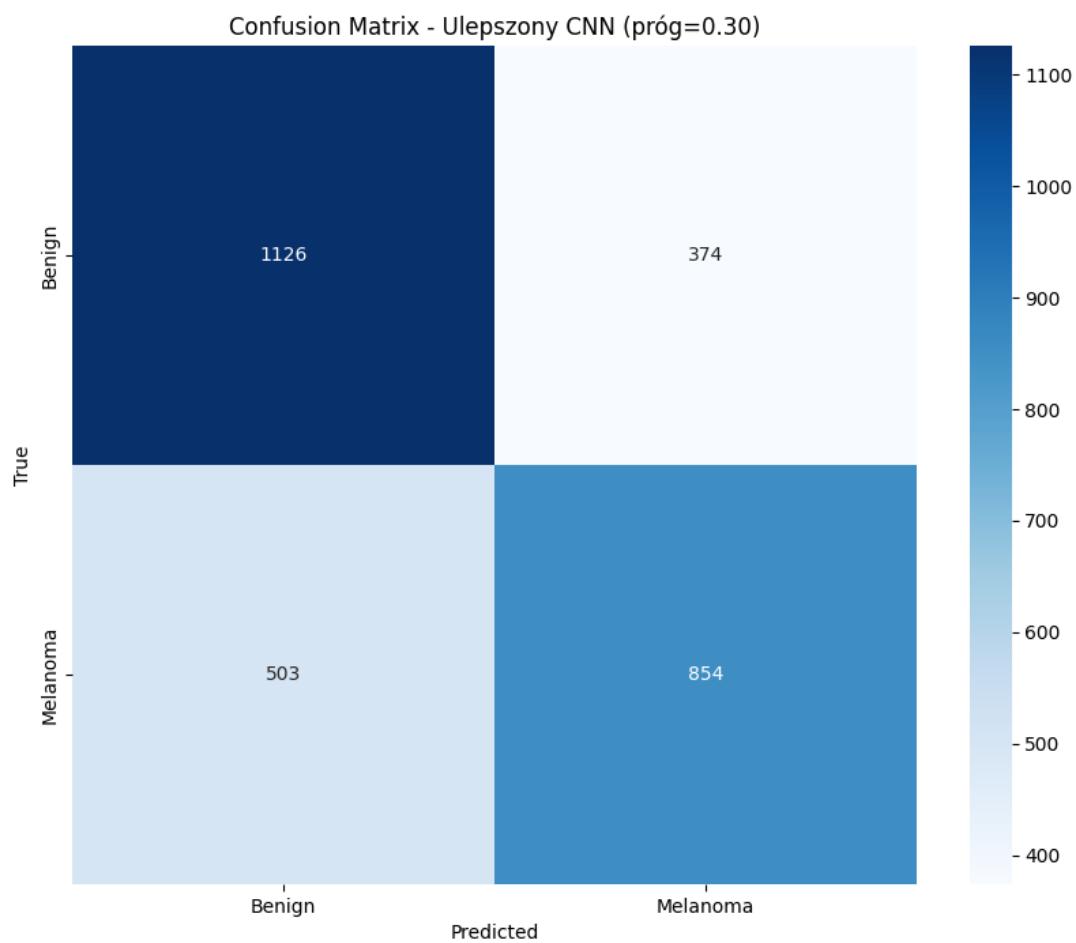
### Wprowadzone ulepszenia:

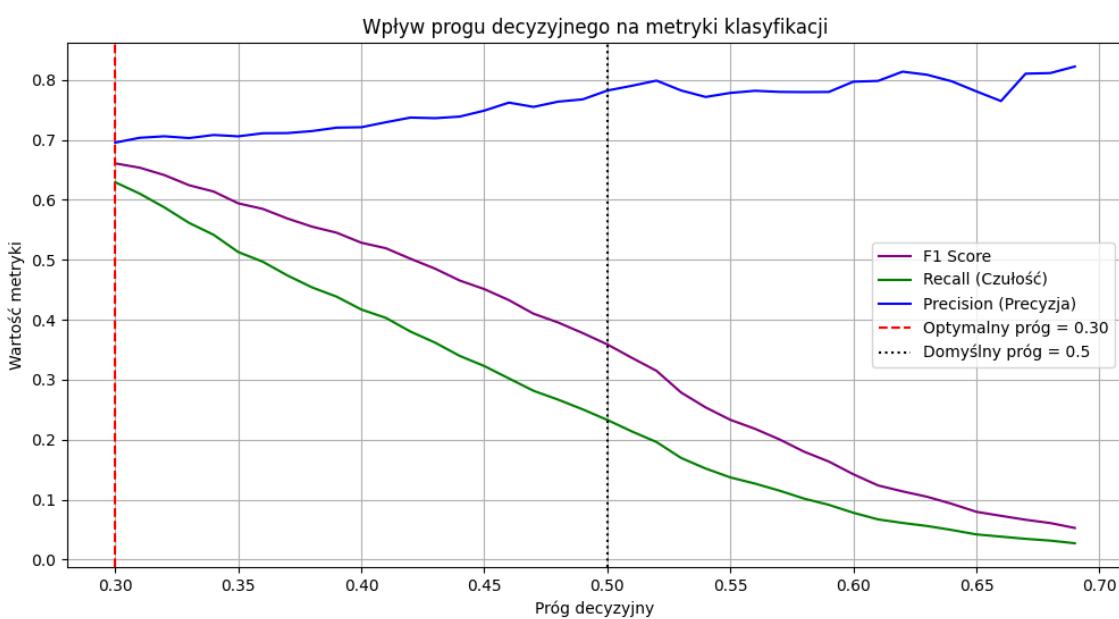
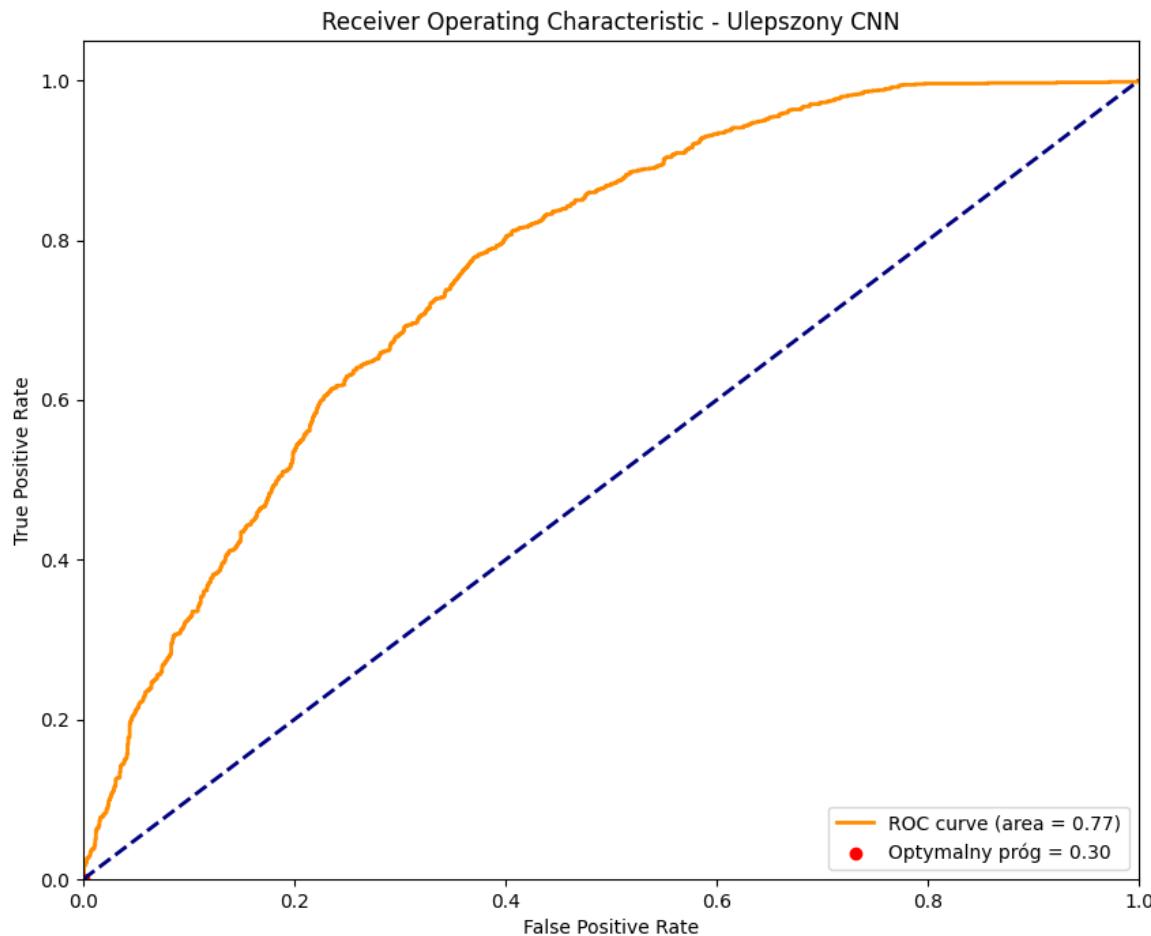
- **Ważenie klas:** melanoma otrzymała 1.5x większą wagę
- **Rozszerzona augmentacja:** brightness, shear, więcej rotacji
- **Regularyzacja L2 (kara za duże wagi - zapobiega przeuczeniu)** we wszystkich warstwach
- **Label smoothing (0.1)**
- **Cykliczna stopa uczenia**
- **Monitorowanie recall** zamiast accuracy
- Dodatkowy blok konwolucyjny (256 filtrów)

### Wyniki:

- **Dokładność:** 69.7%
- **ROC AUC:** 0.77
- **Optymalny próg:** 0.30

- Czułość na melanomę: 62.9%





**Analiza:** CNN ulepszony paradoksalnie gorzej wykrywał melanomę niż podstawowy - miał niższą czułość (62.9% vs 67.4%). Agresywne ważenie klas i niski próg (0.30) nie przyniosły oczekiwanej poprawy w wykrywaniu melanom.

## 2.5 CNN v3 (architektura rezydualna)

## Kluczowe innowacje:

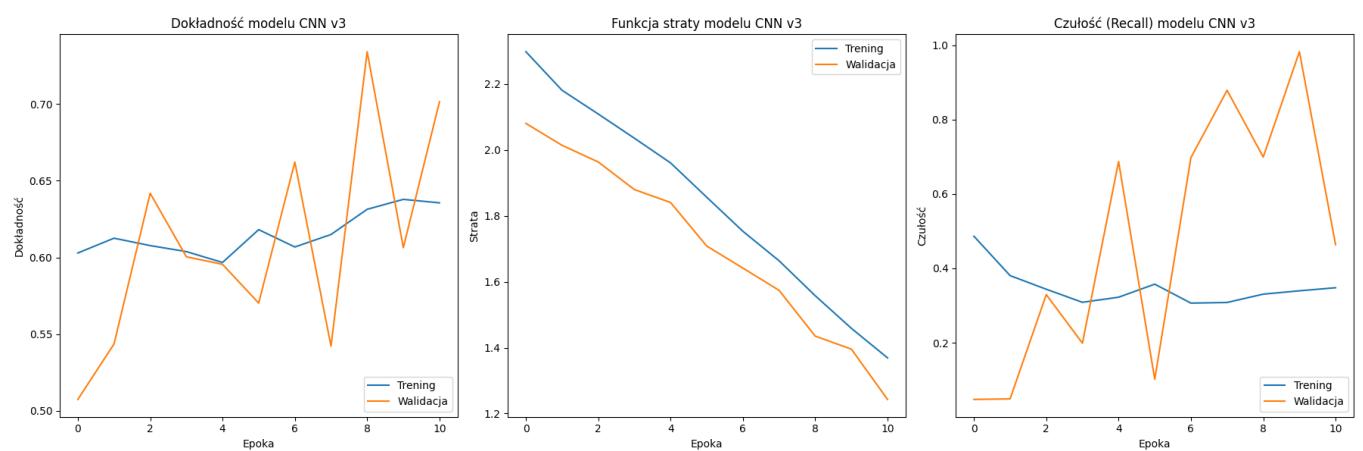
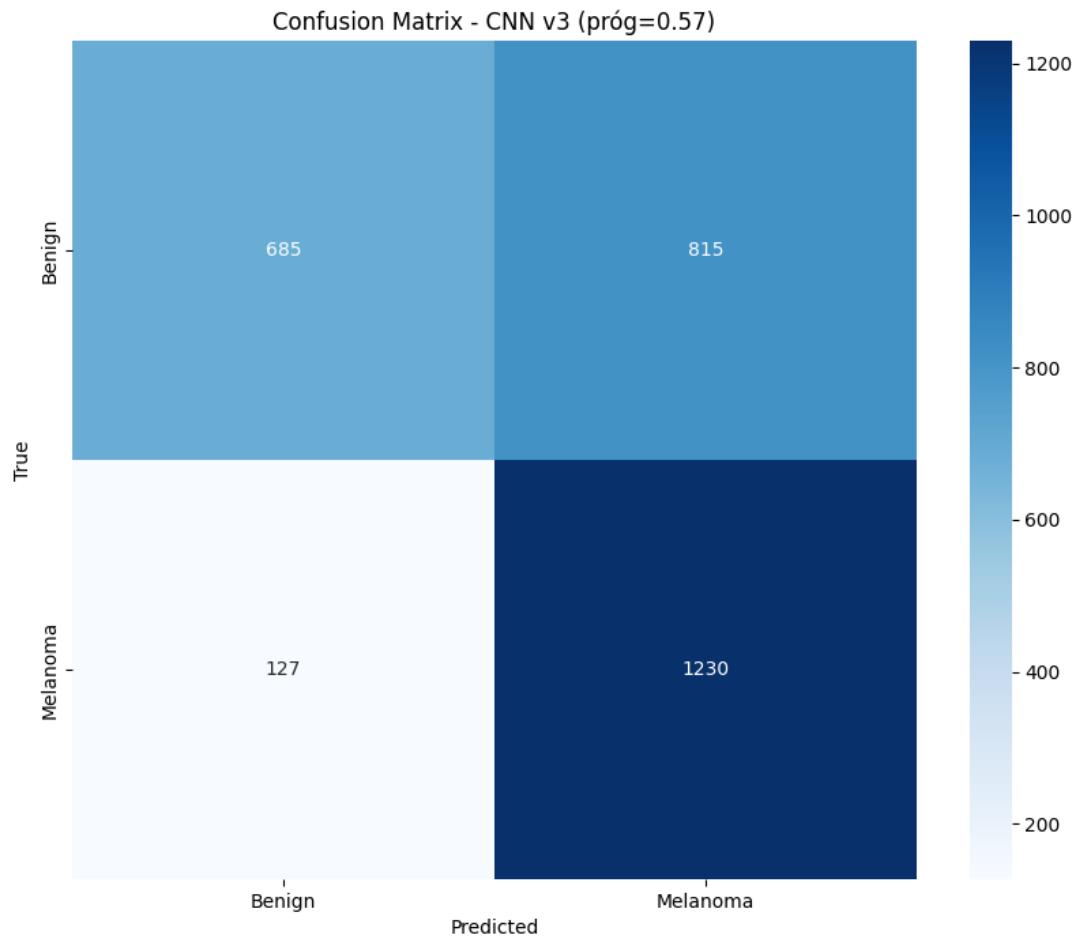
- **Architektura rezydualna** (ResNet-like) z shortcut connections (bezpośrednie połączenia omijające warstwy - umożliwiają uczenie głębszych sieci)
- **Focal Loss** zamiast binary crossentropy (skupia się na trudnych przypadkach i rzadkich klasach)
- **Znacznie silniejsze ważenie melanomów** (3x)
- **Zaawansowana augmentacja** z channel\_shift\_range (losowa zmiana intensywności kanałów RGB - symuluje różne warunki oświetlenia)
- **GlobalAveragePooling** zamiast Flatten (uśrednia każdy kanał do jednej wartości zamiast spłaszczenia wszystkich pikseli - mniej parametrów, lepsza generalizacja)
- 30 epok z większą cierpliwością

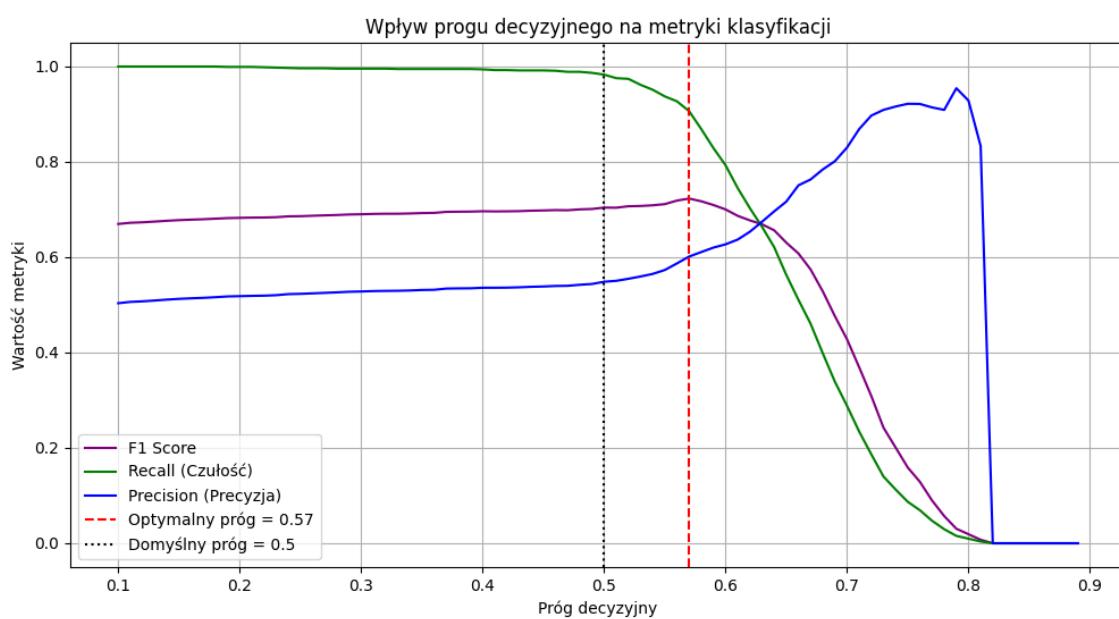
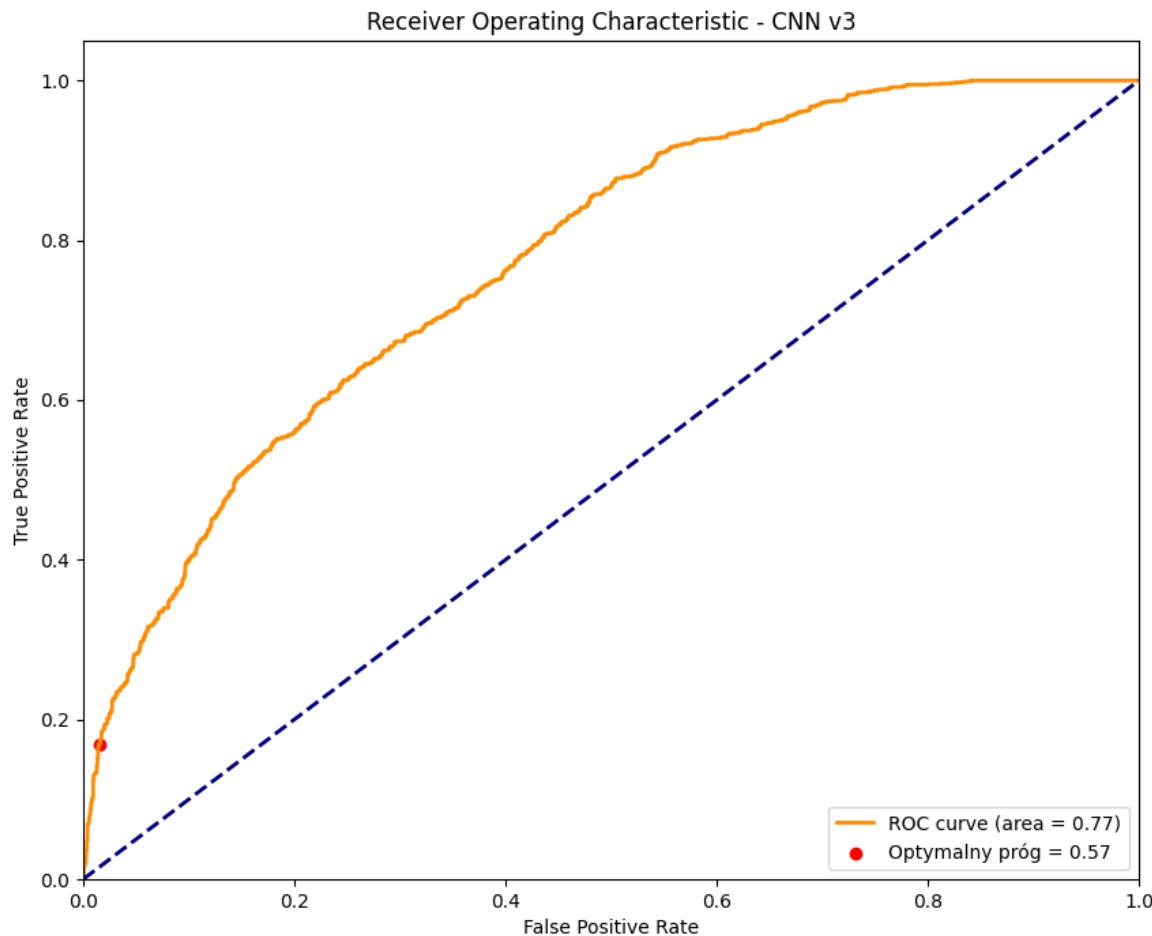
## Parametry Focal Loss:

- gamma=2.0 (współczynnik skupienia na trudnych przypadkach - wyższe wartości bardziej ignorują łatwe przykłady)
- alpha=0.25 (waga dla klasy pozytywnej melanoma - balansuje niezrównoważone klasy)

## Wyniki:

- **Dokładność:** 66.3%
- **ROC AUC:** 0.77
- **Optymalny próg:** 0.57
- **Recall dla melanomów:** 90.6%





**Analiza:** Model osiągnął bardzo wysoką czułość na melanomę kosztem ogólnej dokładności. Focal Loss skutecznie skupił się na trudnych przypadkach.

## 2.6 CNN v4 (architektura hybrydowa)

## Zaawansowane techniki:

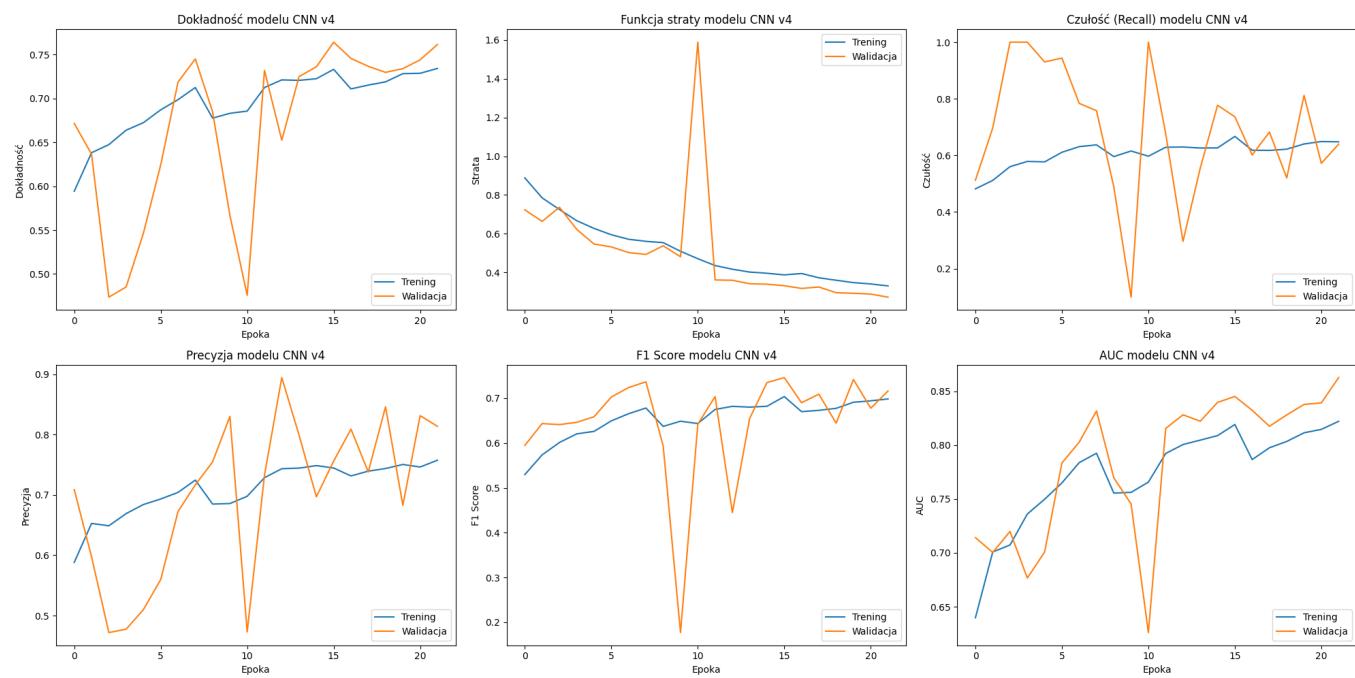
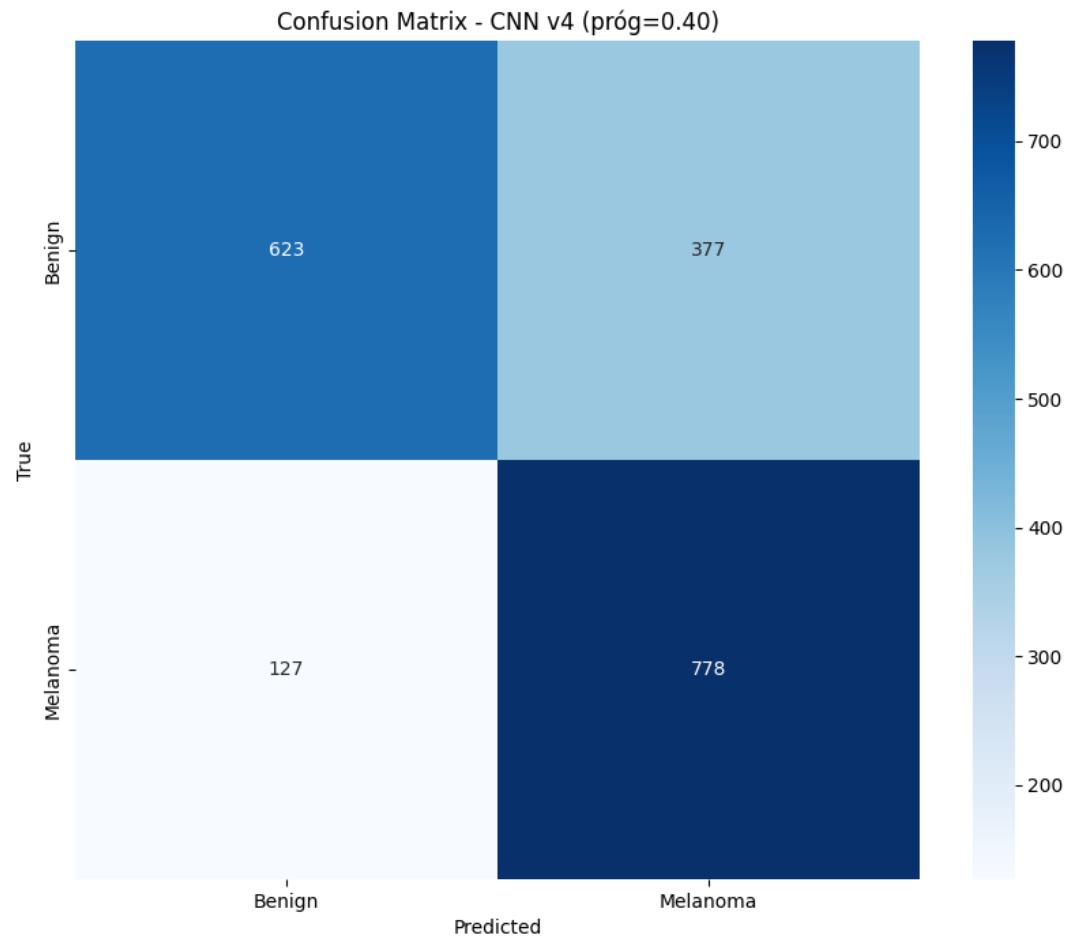
- **Bloki Squeeze-and-Excitation** (mechanizm uwagi który uczy się ważności różnych kanałów - wzmacnia istotne cechy) dla adaptacyjnej uwagi
- **Separowalne konwolucje** (SeparableConv2D - dzieli konwolucję na dwie operacje: depthwise i pointwise - mniej parametrów, szybsze obliczenia)
- **Kombinowana funkcja straty:** Focal Loss + Dice Loss (łączy skupienie na trudnych przypadkach z miarą podobieństwa segmentów - lepsze dla niezbalansowanych danych)
- **Spatial Dropout** zamiast zwykłego Dropout (wyłącza całe kanały zamiast pojedynczych neuronów - lepsza regularyzacja dla CNN)
- **Cosine Annealing LR Scheduler** z warm restarts (cykliczne zmniejszanie i resetowanie stopy uczenia - pomaga uniknąć lokalnych minimów)
- **Zrównoważone ważenie:** melanoma 2.5x (zmniejszone z 3x)
- 40 epok

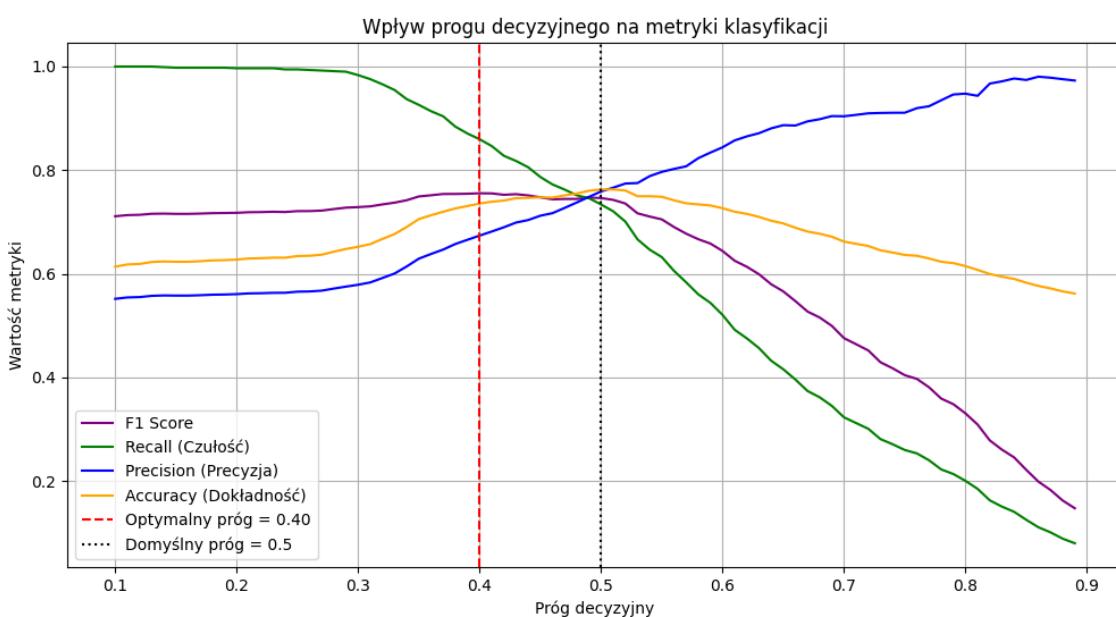
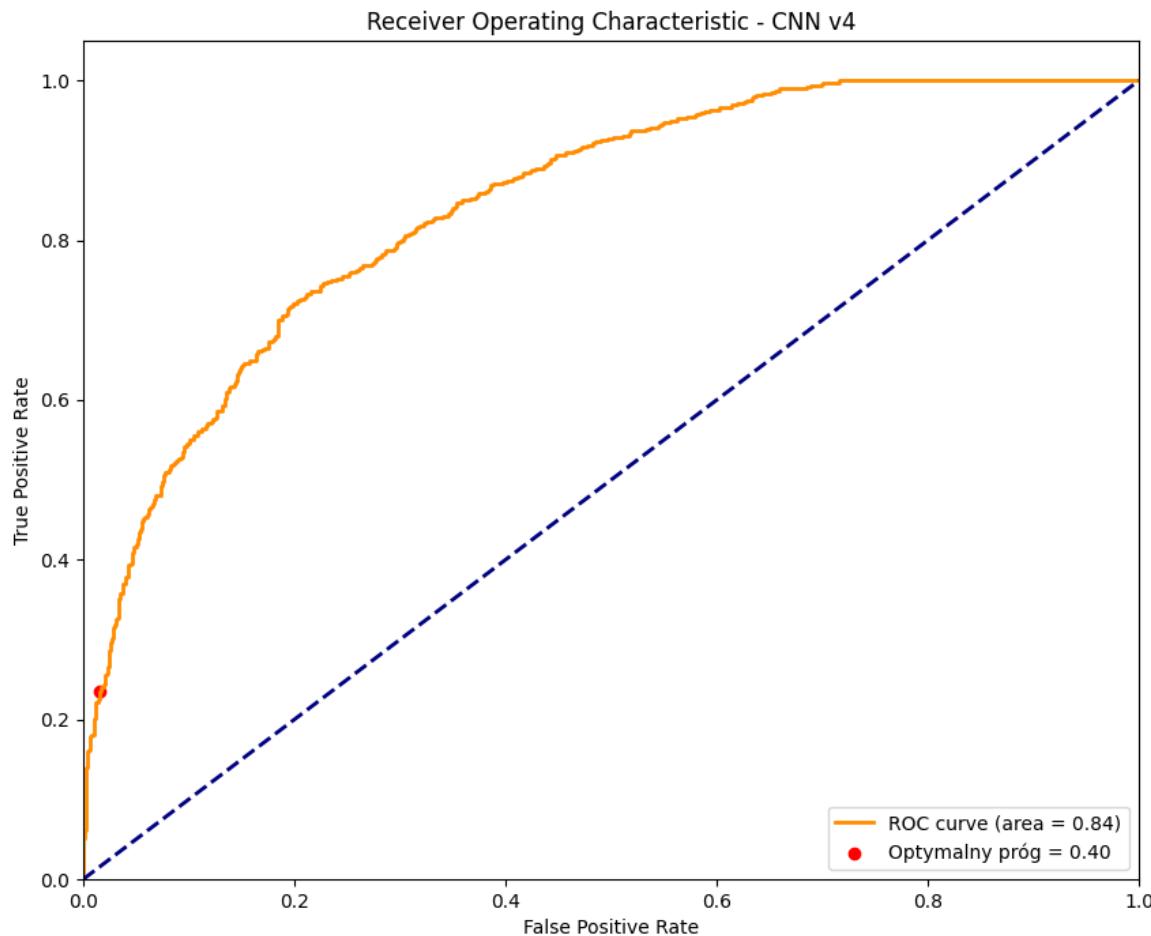
## Parametry funkcji straty:

- Focal Loss: gamma=2.0, alpha=0.25
- Dice Loss weight: 0.3

## Wyniki:

- **Dokładność:** 75.6%
- **ROC AUC:** 0.84
- **Optymalny próg:** 0.40
- **F1 Score:** 0.76





**Analiza:** Najlepszy kompromis między czułością a precyją wśród modeli CNN. Kombinowana funkcja straty i zaawansowane techniki regularizacji przyniosły stabilne wyniki.

## 2.7 Transfer Learning (MobileNetV2)

**Architektura:**

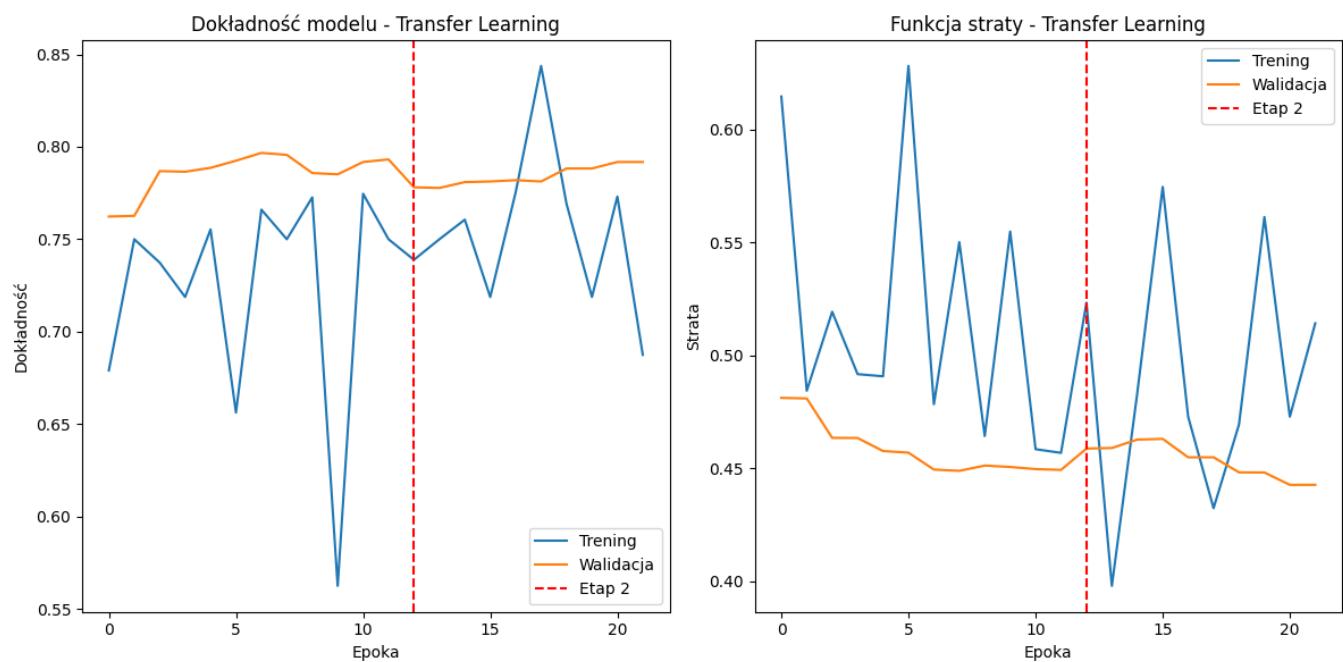
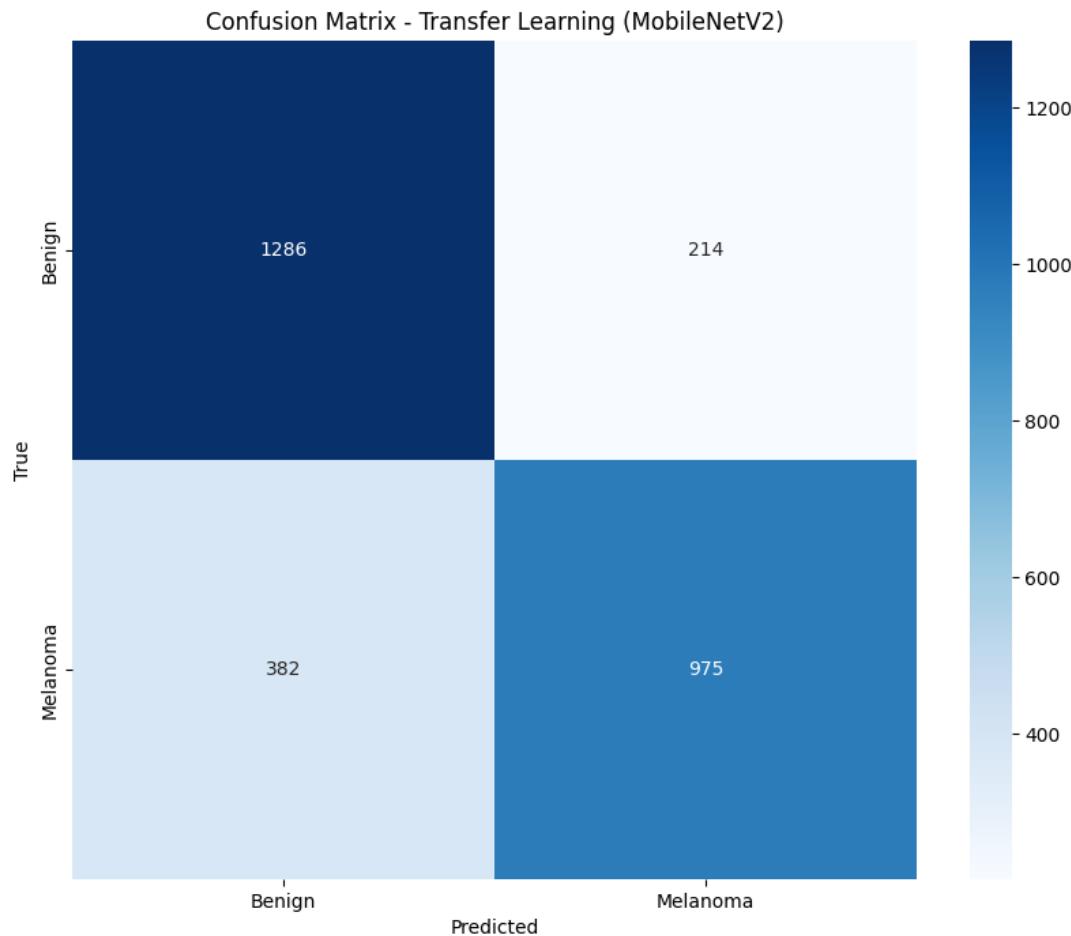
- **Bazowy model:** MobileNetV2 pre-trenowany na ImageNet
- **Dwuetapowe uczenie:**
  - Etap 1: Zamrożony backbone, trening tylko głowy klasyfikacyjnej
  - Etap 2: Fine-tuning ostatnich 20 warstw
- **Głowa klasyfikacyjna:** GlobalAveragePooling → Dense(256) → Dropout(0.5) → Dense(1)

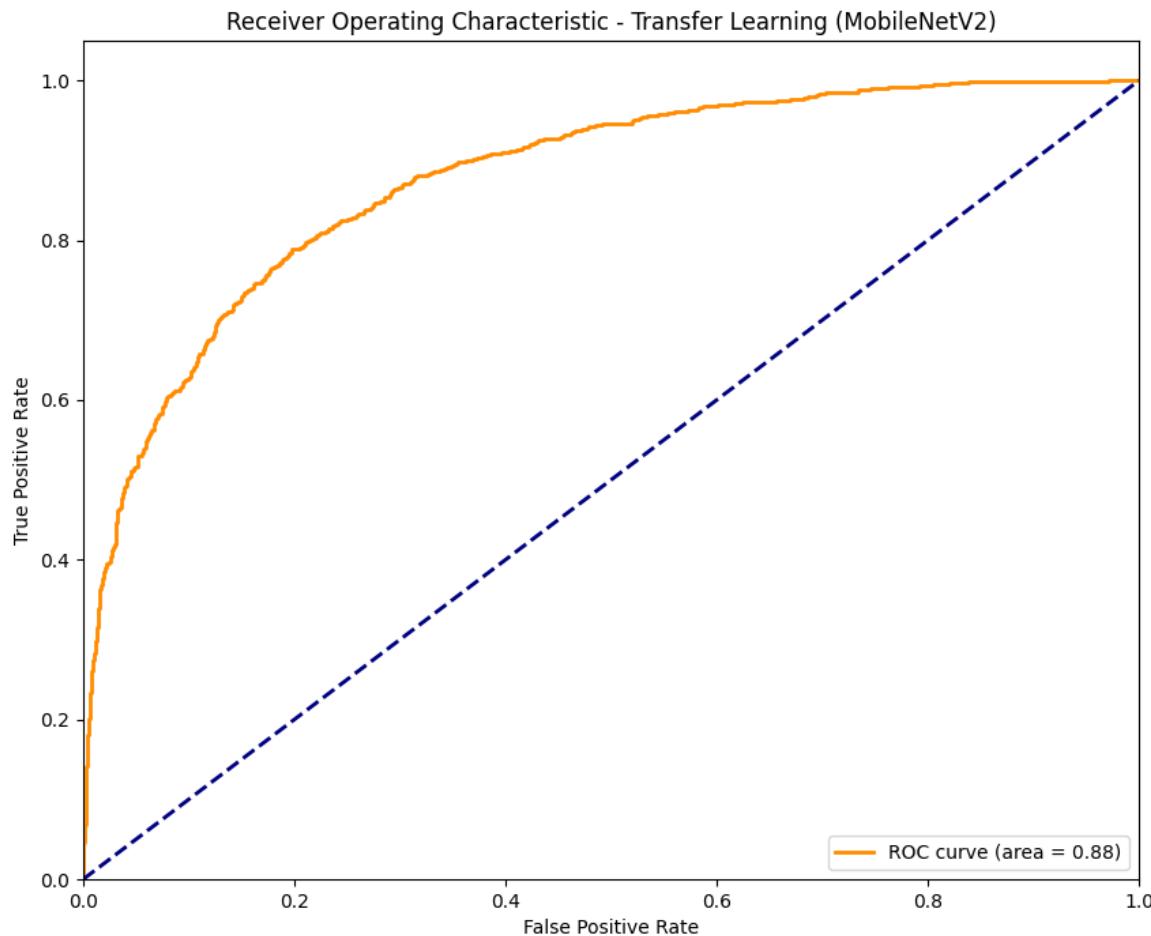
**Parametry treningu:**

- Etap 1: lr=0.0001 (wyższa stopa uczenia dla nowych warstw), 20 epok
- Etap 2: lr=0.00001 (10x niższa stopa dla delikatnego fine-tuningu pre-trenowanych warstw), 10 epok
- Input: 224x224x3

**Wyniki:**

- **Dokładność:** 78.5%
- **ROC AUC:** 0.88
- **Optymalny próg:** 0.49





**Analiza:** Transfer Learning z MobileNetV2 dał jedne z najlepszych wyników wśród pojedynczych modeli. Dwuetapowe uczenie pozwoliło na optymalne wykorzystanie pre-trenowanych cech.

## 2.8 XGBoost

### Ekstrakcja cech:

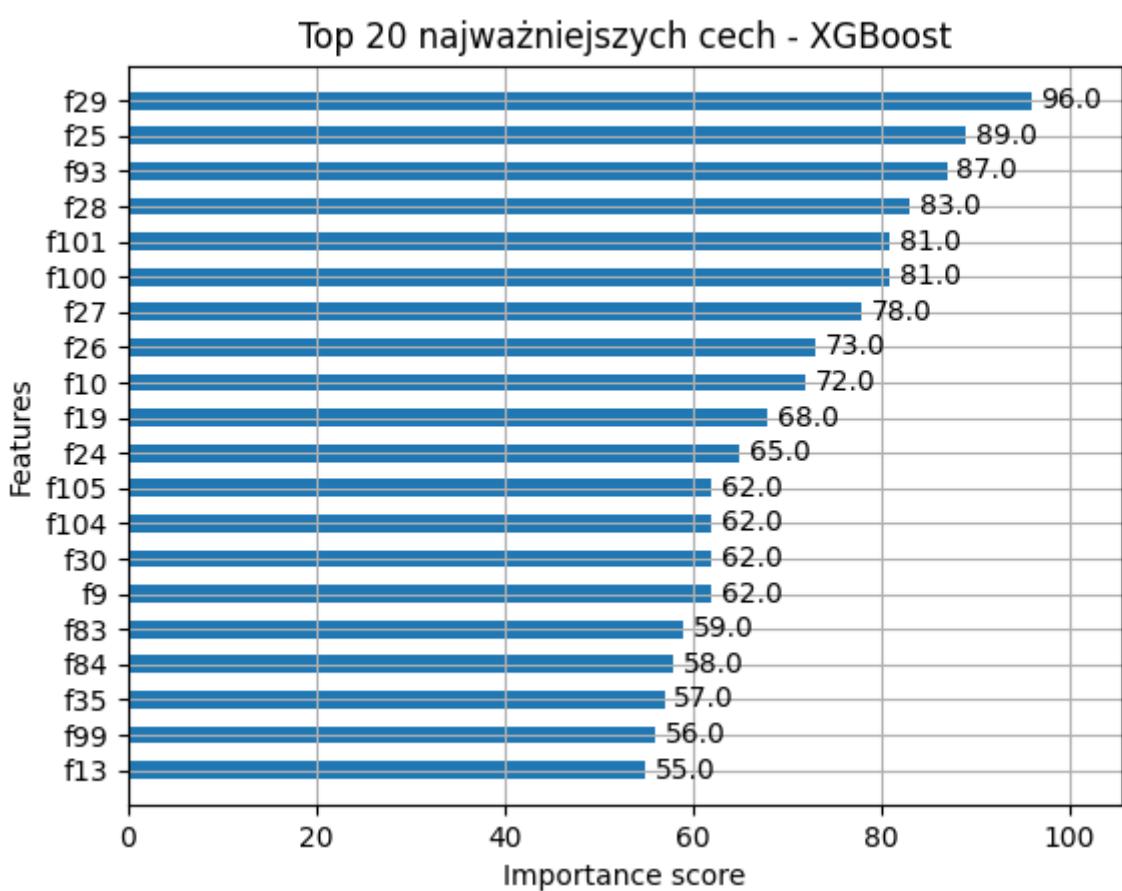
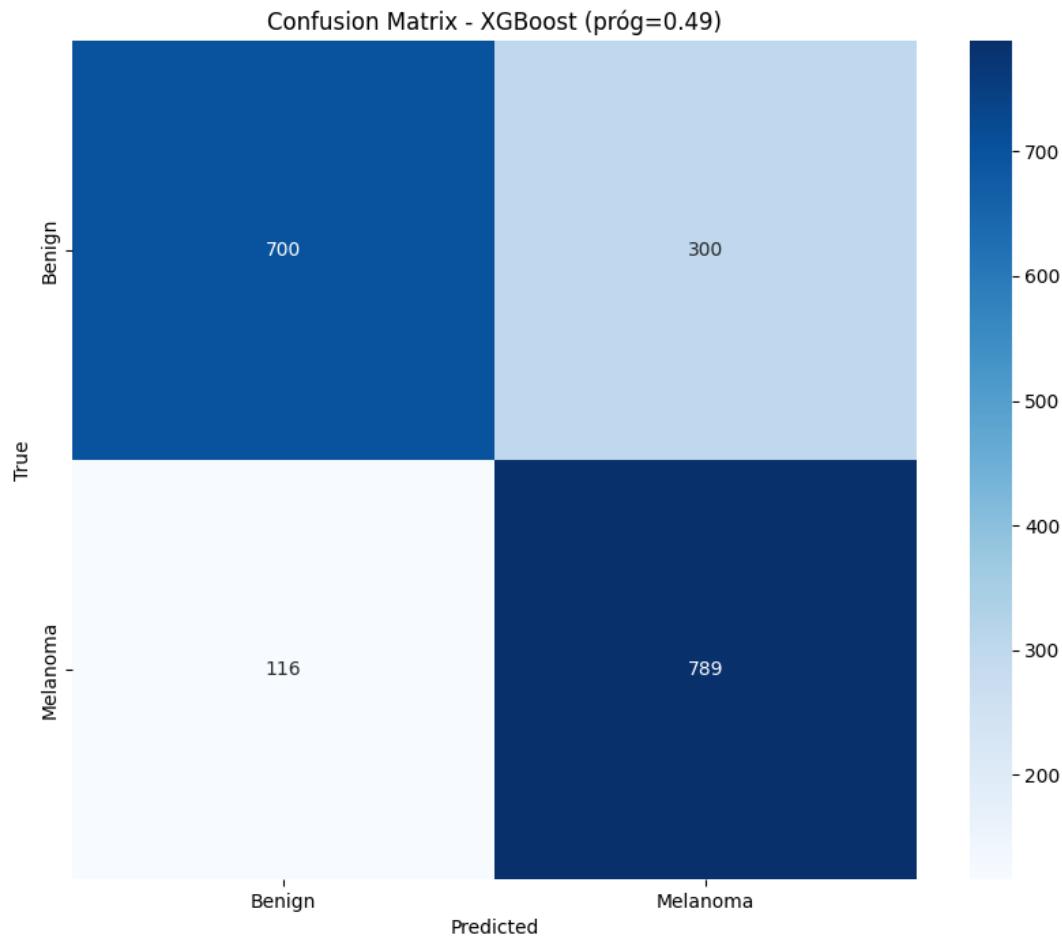
- **Przestrzenie kolorów:** RGB, HSV
- **Cechy asymetrii ABCD:** asymetria pozioma/pionowa
- **Cechy brzegów:** gęstość brzegów, ścisłość
- **Cechy tekstuury:** gradienty Sobela (wykrywanie krawędzi poziomych i pionowych), Laplacian (wykrywanie zmian intensywności - miara ostrości)
- **Cechy kształtu:** znormalizowane pole, obwód, okrągłość
- **Entropie kolorów** (miara różnorodności rozkładu kolorów - wyższa entropia = więcej różnych kolorów)
- Łącznie ~50 cech

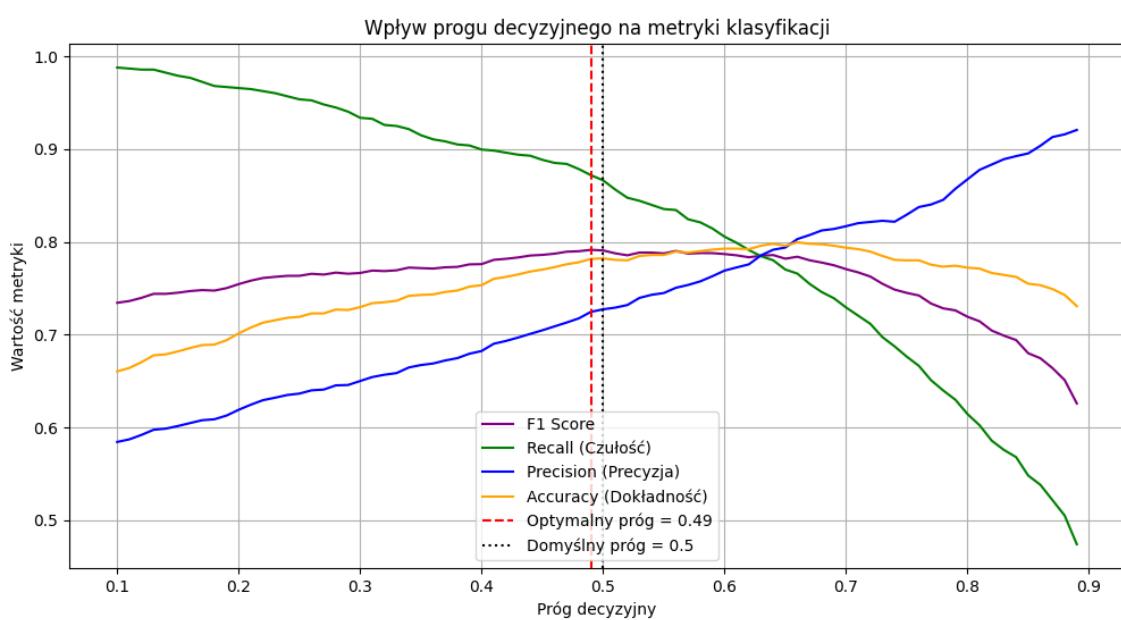
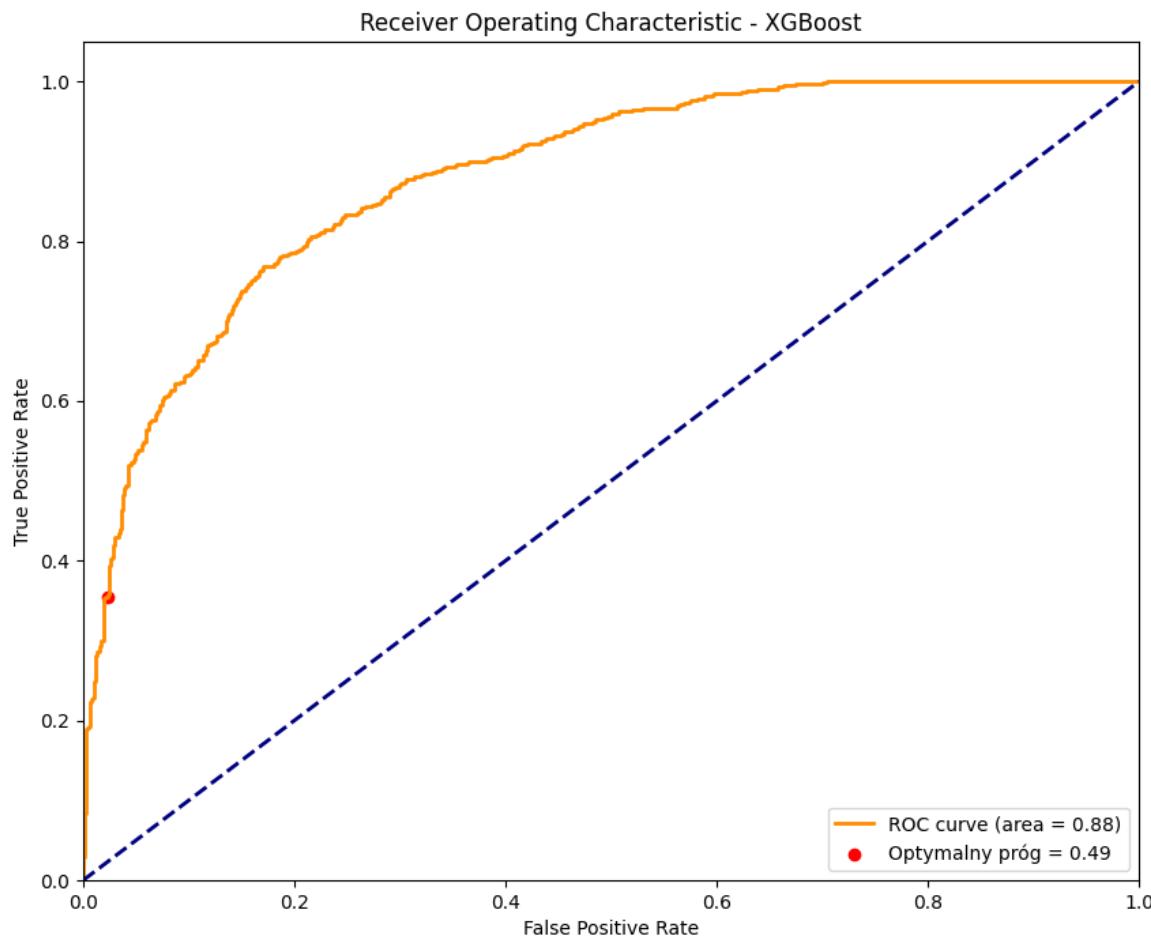
### Parametry modelu:

- n\_estimators=200, learning\_rate=0.1
- max\_depth=5, scale\_pos\_weight=2 (2x większa waga dla klasy melanoma - balansuje niezrównoważone dane)
- early\_stopping\_rounds=15

**Wyniki:**

- **Dokładność:** 74.3%
- **ROC AUC:** 0.88
- **Optymalny próg:** 0.49





**Analiza:** XGBoost osiągnął wysokie ROC AUC dzięki skutecznej ekstrakcji cech ręcznych. Najważniejsze okazały się cechy kolorów i gradientów.

## 2.9 Model Hybrydowy (Transfer Learning + XGBoost)

**Koncepcja:** Połączenie głębokich cech z MobileNetV2 z tradycyjnymi cechami obrazu w jednym modelu XGBoost.

#### Architektura:

- **Ekstraktor cech:** MobileNetV2 (1280 cech głębokich)
- **Cechy tradycyjne:** ~50 cech (RGB, HSV, tekstura, kształt)
- **Klasyfikator:** XGBoost
- **Łącznie:** ~1330 cech

#### Parametry XGBoost:

- n\_estimators=300, learning\_rate=0.05
- max\_depth=6, scale\_pos\_weight=2.5
- early\_stopping\_rounds=20

**Standaryzacja:** StandardScaler dla wszystkich cech

#### Wyniki:

- **Dokładność:** 80.1%
- **ROC AUC:** 0.903 ★ (najlepszy wynik)
- **Optymalny próg:** 0.51
- **F1 Score:** 0.79
- **CV F1 Score:** 0.78 ± 0.05

#### Analiza wkładu cech:

- **Cechy głębokie (MobileNetV2):** 91.2%
- **Cechy tradycyjne:** 8.8%

**WSTAW TUTAJ OBRAZY 9-13 z drugiej tury (Confusion Matrix, Feature contribution, ROC curve, Threshold analysis dla modelu hybrydowego)**

**Analiza:** Model hybrydowy osiągnął najlepsze wyniki ze wszystkich badanych podejść. Dominacja cech głębokich (91.2%) potwierdza siłę reprezentacji learned przez MobileNetV2, ale tradycyjne cechy nadal wnoszą wartościową informację.

### 3. Porównanie wszystkich modeli

Model	Dokładność	ROC AUC	F1 Score	% wykrywanie melanom	Optymalny próg
Random Forest (podstawowy)	77.2%	0.86	-	75.8%	0.5
Random Forest (ulepszony)	80.0%	0.89	0.78	83.9%	-
CNN (podstawowy)	75.3%	0.84	-	67.4%	0.5
CNN (ulepszony)	69.3%	0.77	-	62.9%	0.30

Model	Dokładność	ROC AUC	F1 Score	% wykrywanie melanomy	Optymalny próg
CNN v3	67.0%	0.77	0.72	90.6%	0.57
CNN v4	73.4%	0.84	0.76	85.9%	0.40
Transfer Learning	79.1%	0.88	-	71.8%	0.49
XGBoost	78.2%	0.88	-	87.2%	0.49
<b>Hybrydowy TL+XGB</b>	<b>80.1%</b>	<b>0.903</b>	<b>0.79</b>	<b>85.7%</b>	<b>0.51</b>

#### Legenda:

- Dokładność** - odsetek poprawnie sklasyfikowanych przypadków  $(TP + TN) / (TP + TN + FP + FN)$
- ROC AUC** - pole pod krzywą ROC, miara ogólnej zdolności klasyfikacyjnej modelu (0-1, wyższe = lepsze)
- F1 Score** - harmoniczna średnia precyzji i czułości, szczególnie ważna dla niebalansowanych danych
- % wykrywanie melanomy** - czułość/recall dla klasy melanoma  $(TP / (TP + FN))$  - odsetek poprawnie wykrytych czerniaków
- Optymalny próg** - próg decyzyjny maksymalizujący F1 Score dla klasy melanoma (gdy dostępny)

## 4. Kluczowe wnioski z badań

### 4.1 Najważniejsze odkrycia

- Model hybrydowy jako zwycięzca:** Połączenie głębokich cech z Transfer Learning z tradycyjnymi cechami obrazu dało najlepsze wyniki (ROC AUC = 0.903).
- Dominacja cech głębokich:** W modelu hybrydowym cechy z MobileNetV2 stanowiły 91.2% wkładu w predykcję, ale tradycyjne cechy (8.8%) nadal były istotne.
- Znaczenie optymalizacji progu:** Dla wszystkich modeli optymalizacja progu decyzyjnego znacząco poprawiła czułość na melanomę.
- Transfer Learning jako silna podstawa:** MobileNetV2 okazał się doskonałą podstawą zarówno jako samodzielny model, jak i jako ekstraktor cech.
- Cechy ABCD dla melanomy:** Tradycyjne cechy medyczne (asymetria, brzegi, kolory) znacząco poprawiły wyniki Random Forest.

### 4.2 Wyzwania napotkane

- Niebalansowanie klas:** Problem melanomii jako klasy mniejszościowej wymagał zastosowania ważenia klas i specjalistycznych funkcji straty.
- Trade-off czułość vs. precyzja:** Modele skupione na wysokiej czułości (CNN v3) osiągały to kosztem ogólnej dokładności.

### 4.3 Skuteczne techniki

- Focal Loss:** Bardzo skuteczna dla niebalansowanych danych

2. **Ważenie klas:** Kluczowe dla poprawy wykrywania melanomu
3. **Augmentacja danych:** Szczególnie rotacje i odbicia
4. **Standaryzacja cech:** Istotna dla modeli XGBoost
5. **Cross-validation:** Niezbędna dla rzetelnej oceny

## 5. Aplikacja webowa

### 5.1 Implementacja

Stworzyłem aplikację webową wykorzystującą **model hybrydowy (Transfer Learning + XGBoost)** - najlepszy z przebadanych modeli do klasyfikacji przesłanych zdjęć znamion.

#### Technologie:

- **Backend:** Flask (Python)
- **Frontend:** HTML, CSS, Bootstrap, JavaScript
- **Model:** Hybrydowy TL+XGBoost z optymalnym progiem 0.49

#### Funkcjonalności:

- Upload zdjęć znamion skórnnych
- Automatyczna ekstrakcja hybrydowych cech obrazu (MobileNetV2 + tradycyjne)
- Predykcja z procentową pewnością
- Wyświetlanie wyniku z odpowiednimi ostrzeżeniami

### 5.2 Mechanizm obliczania pewności

Aplikacja oblicza procent pewności w następujący sposób:

1. **Model generuje prawdopodobieństwo melanomu** (0-1) za pomocą `predict_proba()`
2. **Zastosowanie optymalnego progu 0.49:**
  - Jeśli prawdopodobieństwo  $> 0.49 \rightarrow$  klasyfikacja: "Melanoma"
  - Jeśli prawdopodobieństwo  $\leq 0.49 \rightarrow$  klasyfikacja: "Benign"
3. **Obliczanie pewności:**
  - **Dla melanomu:** pewność = prawdopodobieństwo  $\times 100\%$
  - **Dla benign:** pewność =  $(1 - \text{prawdopodobieństwo}) \times 100\%$

#### Przykład:

- Prawdopodobieństwo melanomu = 0.75  $\rightarrow$  Wynik: "Melanoma" z pewnością 75%
- Prawdopodobieństwo melanomu = 0.25  $\rightarrow$  Wynik: "Benign" z pewnością 75%

Ten mechanizm zapewnia, że procent pewności zawsze odzwierciedla siłę przekonania modelu o danej klasyfikacji.

### 5.3 Struktura plików

```
|__ app.py           # Główna aplikacja Flask
|__ templates/
|   __ index.html    # Interfejs użytkownika
```

```
└── static/
    └── css/
        └── styles.css      # Stylowanie
    └── js/
        └── main.js        # js
└── uploads/                      # Folder na przesłane obrazy
└── results/
    └── hybrid_tl_xgb/
        └── models/
            ├── xgboost_hybrid.pkl    # Wytrenowany model XGBoost
            ├── feature_extractor/    # Model MobileNetV2 (TensorFlow)
            ├── scaler.pkl           # StandardScaler dla cech
            └── model_config.pkl     # Konfiguracja modelu (próg, etc.)
```

## 6. Aspekt badawczy - wykroczenie poza materiał

### 6.1 Zaawansowane techniki wykorzystane

Wykroczyłam znacząco poza materiał z wykładu, implementując:

1. **Focal Loss:** Specjalistyczna funkcja straty dla niezbalansowanych danych
2. **Squeeze-and-Excitation blocks:** Mechanizm uwagi w CNN
3. **Architektura rezydualna:** Shortcut connections w sieciach
4. **Transfer Learning z fine-tuningiem:** Dwuetapowe uczenie
5. **Model hybrydowy:** Połączenie deep learning z traditional ML
6. **Separowalne konwolucje:** Efektywniejsze obliczeniowo
7. **Cosine Annealing:** Zaawansowane harmonogramowanie stopy uczenia
8. **Cross-validation z early stopping:** Rygorystyczna walidacja

## 7. Podsumowanie

Projekt zakończył się sukcesem, osiągając **ROC AUC = 0.903** dla najlepszego modelu hybrydowego. Kluczowym czynnikiem sukcesu było:

1. **Systematyczne podejście:** Testowanie od prostych do zaawansowanych metod
2. **Iteracyjne ulepszenia:** Każda wersja modelu wykorzystywała wnioski z poprzedniej
3. **Kombinowanie podejść:** Model hybrydowy połączył mocne strony różnych technik
4. **Optymalizacja dla problemu medycznego:** Skupienie na czułości wykrywania melanom

## 8. Uruchomieniennie

Projekt uruchamia się polecienniem `python app.py`, strona działa na `http://127.0.0.1:5000`