

Sprawozdanie z Projektu „Dno oka”

1. Skład grupy

- Adam Detmer 155976
- Jakub Buler 155987

2. Zastosowany język programowania i biblioteki

- **Język:** Python
- **Środowisko:** Google Colab / Jupyter Notebook
- **Biblioteki:**
 - numpy, opencv-python, scikit-image – Przetwarzanie obrazu
 - scikit-learn – metryki klasyfikacji, modele ML
 - matplotlib, pandas – wizualizacja i raportowanie
 - torch, torchvision – Sieci neuronowe

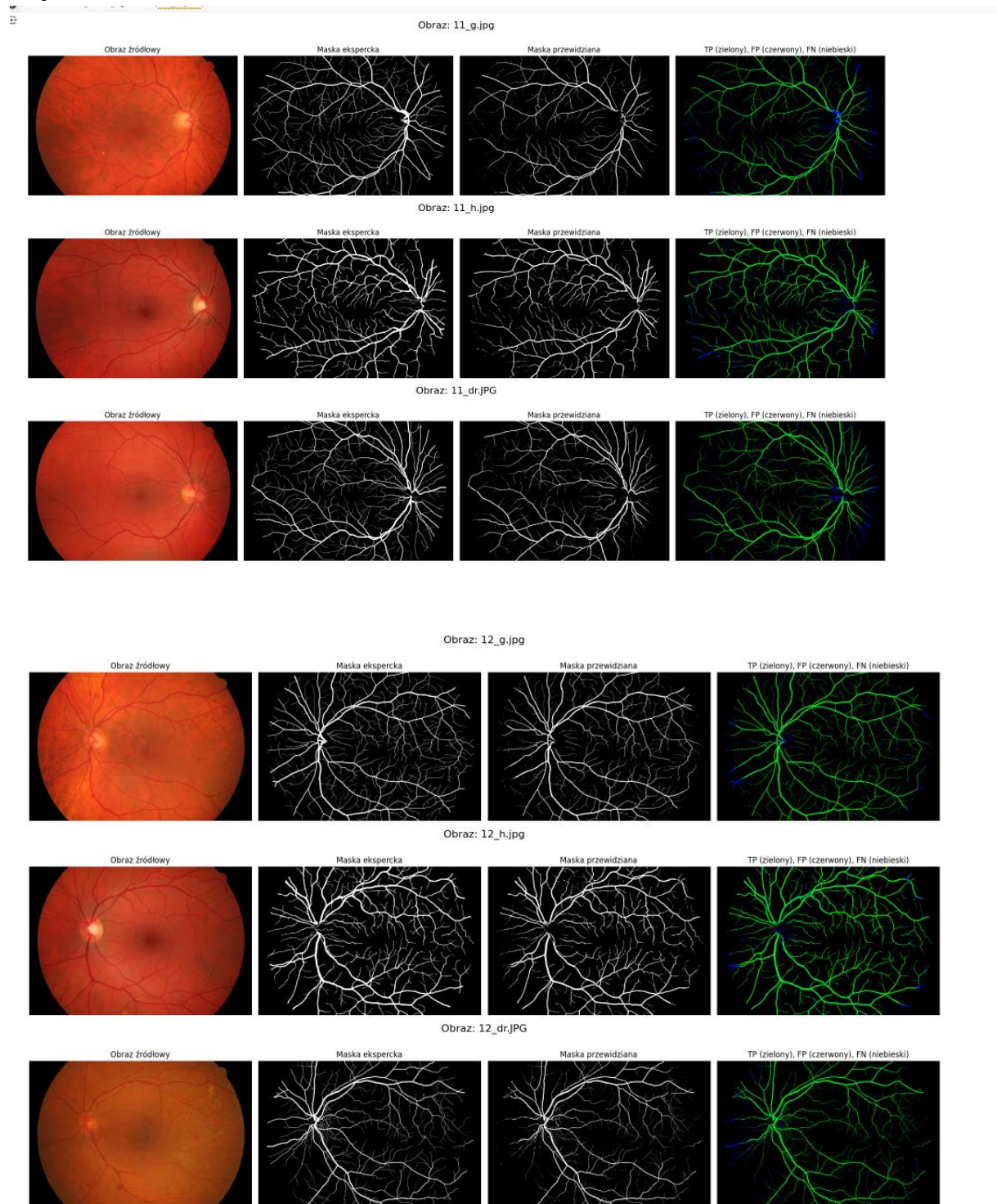
3. Opis zastosowanych metod

3.1 Przetwarzanie obrazów

- **Kanał zielony:** ekstrakcja informacji strukturalnej (naczynia najlepiej widoczne).
- **Denoising (FastNIMeansDenoising):** redukcja szumów bez utraty szczegółów.
- **Equalizacja histogramu:** wzmocnienie kontrastu naczyń.
- **Filtr Frangiego:** wykrywanie naczyń.
- **Progowanie:** ręczne progowanie odpowiednio dobrane eksperymentalnie.
- **Postprocessing:** usunięcie drobnych obiektów, zamknięcie morfologiczne.
- **Maskowanie FOV:** ograniczenie segmentacji do wnętrza siatkówki.

Uzasadnienie: Do analizy naczyń krwionośnych zastosowaliśmy klasyczne metody przetwarzania obrazu, skupiając się na poprawie kontrastu i redukcji szumu. Wybrano kanał zielony, ponieważ najlepiej uwidacznia naczynia na tle siatkówki. Filtr FastNIMeansDenoising usunął szum bez rozmywania cienkich struktur, a equalizacja histogramu zwiększyła kontrast naczyń względem tła. Zastosowana dylacja pozwoliła wzmocnić słabo widoczne naczynia przed dalszą segmentacją. Tak przygotowane obrazy były lepiej przystosowane do działania filtru Frangiego i progowania.

Wyniki Przetwarzania



| | accuracy | sensitivity | specificity | harmonic_mean | arithmetic_mean | TP | TN | FP | FN | image |
|---|----------|-------------|-------------|---------------|-----------------|--------|---------|----|--------|-----------|
| 0 | 0.9789 | 0.7241 | 1.0 | 0.8400 | 0.8621 | 453077 | 7559651 | 0 | 172616 | 11_g.jpg |
| 1 | 0.9823 | 0.8105 | 1.0 | 0.8954 | 0.9053 | 620346 | 7419989 | 0 | 145009 | 11_h.jpg |
| 2 | 0.9823 | 0.7908 | 1.0 | 0.8832 | 0.8954 | 547969 | 7492407 | 0 | 144968 | 11_dr.JPG |
| 3 | 0.9872 | 0.8445 | 1.0 | 0.9157 | 0.9223 | 571271 | 7508920 | 0 | 105153 | 12_g.jpg |
| 4 | 0.9845 | 0.8527 | 1.0 | 0.9205 | 0.9264 | 732763 | 7326049 | 0 | 126532 | 12_h.jpg |
| 5 | 0.9886 | 0.8325 | 1.0 | 0.9086 | 0.9162 | 463446 | 7628620 | 0 | 93278 | 12_dr.JPG |

(Błąd wliczeniu TN i specificity wynikał z pastowania wartości w masce na zły typ po braniu obszaru z fova)

3.2 Uczenie maszynowe *Decision Tree Classifier*

W podejściu opartym na uczeniu maszynowym segmentujemy obrazy siatkówki na małe fragmenty (patches) o stałym rozmiarze, np. 15×15 pikseli. Dla każdego takiego fragmentu obliczamy zestaw cech opisujących jego strukturę i teksturę – są to momenty Hu (odporne na rotacje i skalowanie) oraz wariancje jasności w kanałach RGB. Cechy te reprezentują właściwości wizualne naczyń krwionośnych i stanowią wejście dla klasyfikatora.

Maska ekspercka (ground truth) jest dzielona na te same fragmenty i wykorzystywana do oznaczenia każdego segmentu jako naczyniowy (1) lub tło (0), zależnie od wartości środkowego piksela. Z uwagi na dużą nierównowagę klas (dużo więcej tła niż naczyń), stosujemy *undersampling*, by zrównoważyć liczbę przykładów obu klas w zbiorze treningowym.

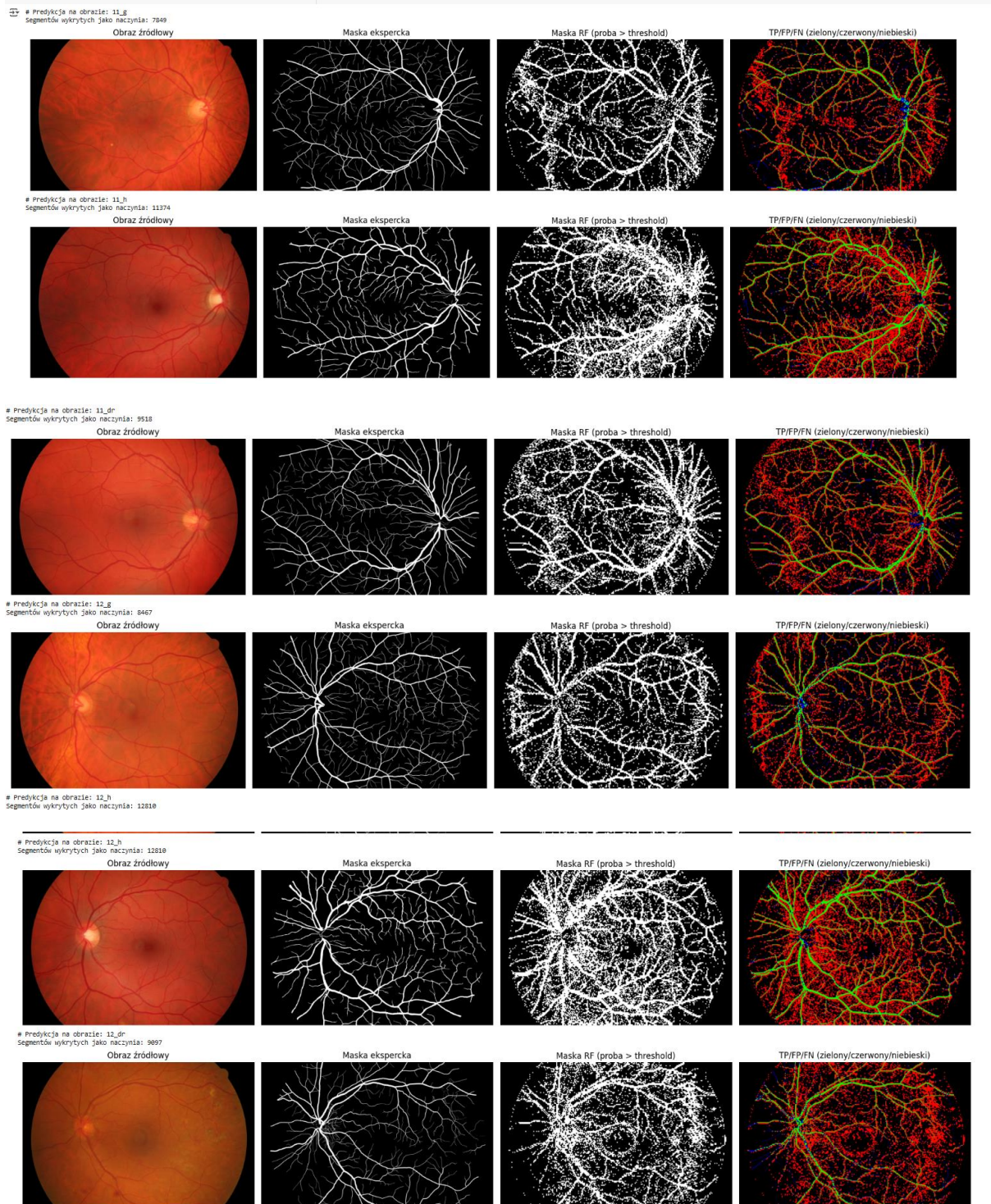
Do klasyfikacji zastosowano model drzewa decyzyjnego `DecisionTreeClassifier` z biblioteki `scikit-learn`. W celu kontroli złożoności modelu i uniknięcia przeuczenia przyjęto następujące parametry:

- `max_depth=30` – maksymalna głębokość drzewa została ograniczona do 30 poziomów, co pozwala modelowi uchwycić złożone zależności, ale zapobiega nadmiernemu dopasowaniu do szumu w danych;
- `min_samples_leaf=5` – minimalna liczba próbek w liściu ustalona na 5 sprawia, że drzewo nie tworzy zbyt małych i niestabilnych reguł;
- `random_state=42` – ustalono stały seed dla zapewnienia powtarzalności wyników;
- użycie domyślnego algorytmu podziału i funkcji entropii

Model uczony jest na przygotowanych cechach wyciągniętych z obrazu 01_dr.JPG, który nie pojawia się w zbiorze testowym.

W trakcie predykcji każdy fragment nowego obrazu jest klasyfikowany przez wytrenowany model. Jeśli prawdopodobieństwo klasy "naczynie" przekracza ustalony próg, to cały segment jest oznaczany jako naczyniowy w masce wynikowej. W ten sposób tworzymy binarną maskę wykrytych naczyń krwionośnych.

Wyniki Klasyfikatora



| | accuracy | sensitivity | specificity | harmonic_mean | arithmetic_mean | TP \ |
|---|----------|-------------|-------------|---------------|-----------------|--------|
| 0 | 0.814826 | 0.710321 | 0.823476 | 0.762724 | 0.766898 | 444443 |
| 1 | 0.750783 | 0.851661 | 0.740378 | 0.792129 | 0.796019 | 651823 |
| 2 | 0.782440 | 0.772744 | 0.783337 | 0.778004 | 0.778040 | 535463 |
| 3 | 0.802400 | 0.722965 | 0.809555 | 0.763813 | 0.766260 | 489031 |
| 4 | 0.723042 | 0.869954 | 0.705810 | 0.779332 | 0.787882 | 747547 |
| 5 | 0.777786 | 0.717668 | 0.782174 | 0.748533 | 0.749921 | 399543 |

| | TN | FP | FN | image |
|---|---------|---------|--------|-------|
| 0 | 6225188 | 1334463 | 181250 | 11_g |
| 1 | 5493593 | 1926396 | 113532 | 11_h |
| 2 | 5869078 | 1623329 | 157474 | 11_dr |
| 3 | 6078886 | 1430034 | 187393 | 12_g |
| 4 | 5170798 | 2155251 | 111748 | 12_h |
| 5 | 5966905 | 1661715 | 157181 | 12_dr |

Model drzewa decyzyjnego osiągnął dobre i stabilne wyniki — accuracy utrzymywało się w granicach 77–81%, a czułość (sensitivity) nawet powyżej 0.85, co oznacza skuteczne wykrywanie naczyń krwionośnych. Wyniki były zrównoważone, a harmoniczna średnia sięgała 0.79, co świadczy o dobrym kompromisie między wykrywaniem naczyń a unikaniem fałszywych alarmów.

Zastosowaliśmy drzewo decyzyjne, ponieważ jest szybkie, dobrze radzi sobie z nieliniowymi zależnościami, pozwala kontrolować złożoność modelu i jest łatwe do interpretacji. To sprawia, że dobrze sprawdza się w zadaniach klasyfikacji segmentów obrazu.

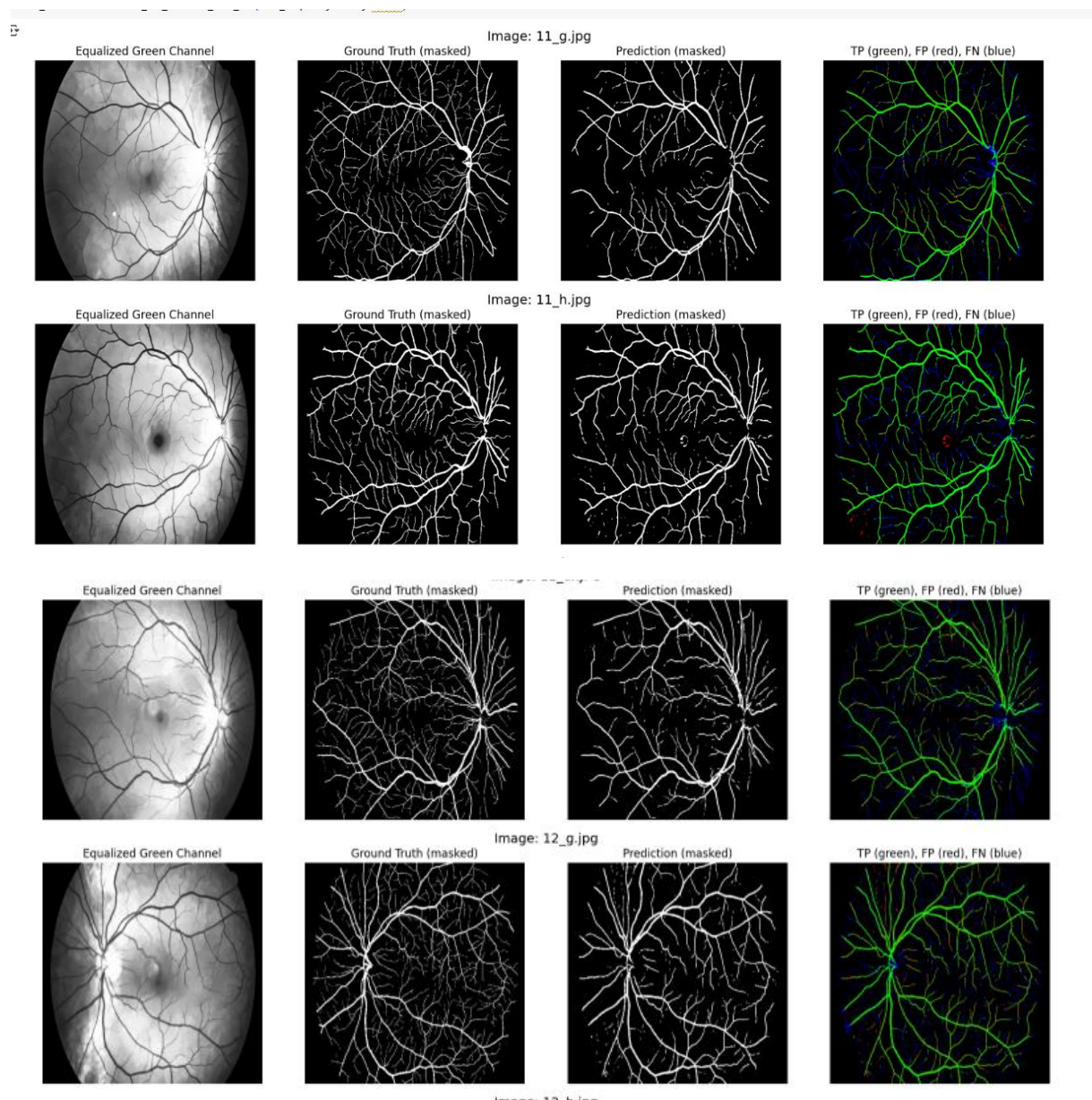
3.3 Sieć neuronowa

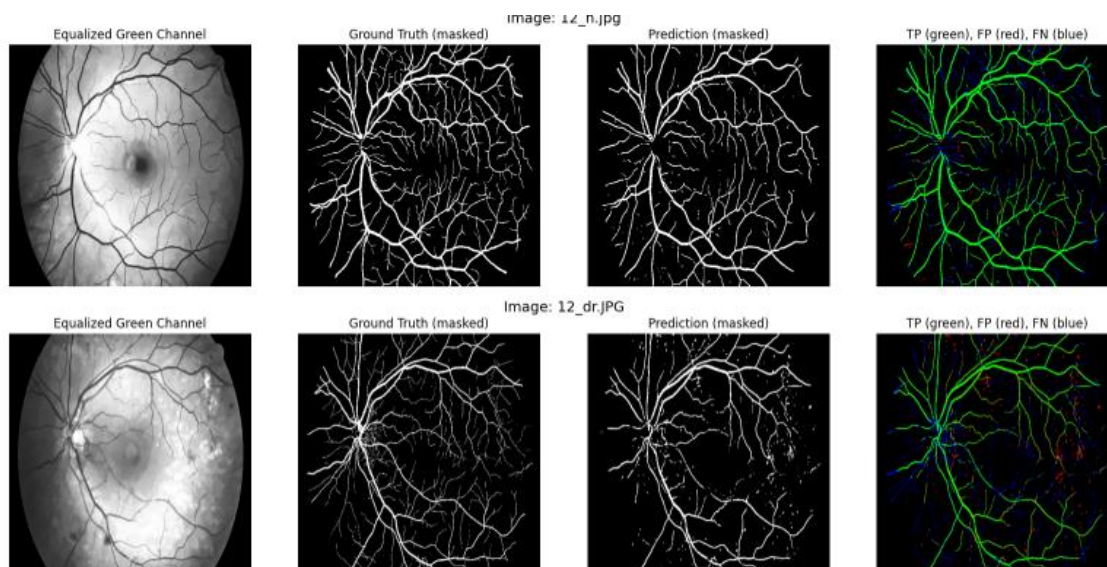
W celu dokładniejszego rozpoznawania drobnych struktur naczyń zastosowano segmentacyjną sieć neuronową U-Net, która sprawdza się w zadaniach segmentacji medycznej. Obrazy wejściowe były wcześniej przetwarzane – wykorzystywano tylko kanał zielony, który został poddany denoisingowi, dylacji i equalizacji histogramu, aby zwiększyć kontrast naczyń. Dla poprawy równowagi klas sieć uczona była na zbalansowanych fragmentach obrazu (patchach 128×128 px), zawierających jednocześnie naczynia i tło.

Maski eksperckie oraz obrazy były ograniczane przez FOV, dzięki czemu sieć nie uczyła się na czarnych obszarach spoza siatkówki. Model U-Net zawiera cztery poziomy enkodera i dekodera, z dodatkowymi warstwami konwolucyjnymi i transpozycyjnymi, oraz funkcję aktywacji sigmoid na wyjściu. Trenowano go przez 40 epok z użyciem funkcji kosztu będącej sumą **Binary Cross Entropy** i **Dice loss**, co pozwalało lepiej wyważyć wpływ zarówno obszarów tła, jak i naczyń. Zastosowany optymalizator to Adam z krokiem uczenia 1e-4.

To podejście pozwoliło sieci skutecznie uczyć kształtu i rozmieszczenia naczyń krwionośnych na podstawie lokalnego kontekstu i struktury obrazu.

Wyniki przetwarzania





| | accuracy | sensitivity | specificity | harmonic_mean | arithmetic_mean | TP | TN | FP | FN | image |
|---|----------|-------------|-------------|---------------|-----------------|-------|--------|------|------|-----------|
| 0 | 0.9567 | 0.6879 | 0.9830 | 0.8094 | 0.8355 | 16086 | 234714 | 4047 | 7297 | 11_g.jpg |
| 1 | 0.9656 | 0.7658 | 0.9889 | 0.8632 | 0.8773 | 20909 | 232228 | 2613 | 6394 | 11_h.jpg |
| 2 | 0.9545 | 0.7014 | 0.9819 | 0.8183 | 0.8417 | 17938 | 232289 | 4281 | 7636 | 11_dr.JPG |
| 3 | 0.9515 | 0.7403 | 0.9740 | 0.8412 | 0.8571 | 18686 | 230741 | 6162 | 6555 | 12_g.jpg |
| 4 | 0.9635 | 0.7514 | 0.9913 | 0.8548 | 0.8713 | 22807 | 229764 | 2027 | 7546 | 12_h.jpg |
| 5 | 0.9534 | 0.6504 | 0.9797 | 0.7818 | 0.8151 | 13612 | 236322 | 4894 | 7316 | 12_dr.JPG |

4. Porównanie Wyników Drzewa i Sieci

Sieć neuronowa U-Net osiągnęła bardzo dobre wyniki — **accuracy na poziomie 95–96%**, przy **high specificity (do 99%)** oraz zauważalnej poprawie równowagi między czułością a precyzją.

Harmoniczna średnia (łącznie sensitivity i specificity) wynosiła od **0.78 do 0.86**, co oznacza skuteczne i zrównoważone wykrywanie naczyń.

W porównaniu z wcześniejszym klasyfikatorem drzewa decyzyjnego, U-Net:

- **zdecydowanie lepiej** radził sobie z odróżnianiem naczyń od tła (wyższa *specificity*),
- utrzymał **porównywalną lub lepszą czułość** (*sensitivity* do 0.76),
- i znacząco **zmniejszył liczbę fałszywych pozytywów (FP)** oraz fałszywych negatywów (FN).

Model uczył się na wycinkach obrazów, ale potrafił skutecznie generalizować wiedzę na całe obrazy testowe. U-Net lepiej odwzorowywał strukturę naczyń, zwłaszcza cienkich i lokalnie słabo widocznych, co klasyfikator oparty na patchach często pomijał. W praktyce oznacza to, że sieć neuronowa zapewnia bardziej spójne i dokładne segmentacje w zastosowaniach medycznych.

5. Podsumowanie

Porównanie średniej harmonicznej

| image_normalized | Drzewo | Klasyczna | UNet |
|------------------|----------|-----------|--------|
| 11_dr | 0.778004 | 0.8832 | 0.8183 |
| 11_g | 0.762724 | 0.84 | 0.8094 |
| 11_h | 0.792129 | 0.8954 | 0.8632 |
| 12_dr | 0.748533 | 0.9086 | 0.7818 |
| 12_g | 0.763813 | 0.9157 | 0.8412 |
| 12_h | 0.779332 | 0.9205 | 0.8548 |

Porównanie sensitivity

| image_normalized | Drzewo | Klasyczna | UNet |
|------------------|----------|-----------|--------|
| 11_dr | 0.772744 | 0.7908 | 0.7014 |
| 11_g | 0.710321 | 0.7241 | 0.6879 |
| 11_h | 0.851661 | 0.8105 | 0.7658 |
| 12_dr | 0.717668 | 0.8325 | 0.6504 |
| 12_g | 0.722965 | 0.8445 | 0.7403 |
| 12_h | 0.869954 | 0.8527 | 0.7514 |

Powyższe tabele przedstawiają porównanie wyników dla obrazów testowych. W zakresie czułości (sensitivity), metoda klasyczna dominuje w większości przypadków, osiągając najwyższe wartości dla 12_dr, 12_g i 11_dr, co sugeruje dobrą zdolność wykrywania prawdziwych naczyń krwionośnych. UNet ma czułość niższą niż klasyczna metoda, ale stabilną, podczas gdy drzewo decyzyjne pokazuje wysoką zmienność – od bardzo dobrej (11_h) do przeciętnej (12_dr). Ostatecznie klasyczna metoda oferuje najlepszy kompromis w zakresie skutecznego wykrywania naczyń.

Zdecydowanie przy większych zasobach ciekawe byłoby próba ulepszenia sieci neuronowej.

