

# Informatyka w Medycynie

## Projekt nr 2

Wykrywanie naczyń dna siatkówki oka

Ewelina Bamberska

- ❖ 136679
- ❖ Grupa L3

Weronika Jankowska

- ❖ 136720
- ❖ Grupa L3

## 1. Zastosowany język programowania oraz dodatkowe biblioteki

Analizę wykonano w języku programowania Python, w środowisku Jupyter Notebook.

Do przeprowadzenia analizy wykorzystano następujące biblioteki:

- Numpy,
- OpenCV,
- Scikit-image,
- Scikit-learn,
- Keras.

## 2. Opis zastosowanych metod

### 2.1. Przetwarzanie obrazów

Do przetworzenia obrazu wykorzystano biblioteki OpenCV oraz Scikit-image.

Kolejne kroki przetwarzania obrazu wykorzystane w algorytmie:

- Przetworzenie wstępne - zmniejszenie rozmiaru do rozmiaru 4 razy mniejszego, przekonwertowanie na szare kolory oraz zastosowanie rozmycia.
- Zastosowanie adaptywnej binaryzacji (thresholding) gaussowskiej.
- Usunięcie linii z soczewki oka poprzez wykorzystanie algorytmu wyszukiwania okręgów (o odpowiedniej średnicy) na obrazie HoughCircles. Usunięcie linii, które kryją się wokół znalezionych okręgów z pominięciem żył. Aby to zrobić, algorytm sprawdza jasność pikseli - naczynia krwionośne w soczewce są z reguły dużo ciemniejsze niż wszystkie elementy.
- Usunięcie tła oraz linii, która wynika z krawędzi pomiędzy tłem a właściwym obrazem, oraz usunięcie pozostałych pikseli poprzez zastosowanie dylatacji.
- Zastosowanie filtra frangi do wykrywania krawędzi.
- Zastosowanie dylatacji oraz erozji (closing), aby pozbyć się ewentualnych pustych pikseli wewnętrz naczyń krwionośnych.

Aby wyznaczyć cechy obrazu przez przetworzenie go specjalnie przygotowanymi do tego algorytmami, nie jest potrzebny zbiór oznaczonych danych, co może być przewagą tej metody nad metodą wykorzystującą machine learning. Przetwarzanie obrazu może być więc bardzo przydatne wszędzie tam, gdzie są dostępne jedynie mniejsze zbiory danych lub po to, aby przygotować obraz do dalszego przetwarzania przez machine learning.

Wadą takiego podejścia może być czas przetwarzania danych, który jest dużo dłuższy niż czas działania sieci neuronowej. Samo uczenie się sieci neuronowej jest bardzo długie, zależnie od tego, ile danych zostało wykorzystane, aby sieć stworzyć, jednak przewidywanie danych za jej pomocą jest krótsze, niż przetworzenie obrazu.

## 2.2. Uczenie maszynowe

Jako algorytm uczenia maszynowego zostały wykorzystane konwolucyjne sieci neuronowe (CNN). Jest to bardzo efektywna metoda przetwarzania obrazów, ponieważ opiera się na zasadzie wykorzystywania filtrów, które wykrywają cechy w obrazie. Filtry są przedstawione przykładowo w formie tablicy 3x3, a ich zadaniem jest kolejno przechodzić przez każdy piksel w tablicy, przetwarzać go, i dzięki temu wykrywać proste wzory, takie jak np. krawędzie. Wyniki filtrowania łączą się w coraz bardziej skomplikowane wzorce, a następnie wykrywają cechy.

Przedstawiony algorytm nie ma za zadania wykrywania bardziej skomplikowanych cech, a jedynie naczyń krwionośnych na obrazach w rozmiarach 5 pikseli na 5 pikseli oraz głębokości 3 (kolory R, G, B w zakresie 0 - 255). Sieć posiada jedną warstwę wejściową, która jest również warstwą CNN - wykorzystuje filtr 3x3, oraz jedną ukrytą CNN, z filtrem o takim samym rozmiarze. Warstwy jako funkcję aktywacji posiadają funkcję ReLU. Na sieć składa się również warstwa wyjściowa z dwoma neuronami (pierwszy neuron wskazuje 1 w przypadku, kiedy wykryto tło, a drugi 1 w przypadku, gdy wykryto naczynie krwionośne). Warstwa jako funkcję aktywacji przyjmuje funkcję softmax. Pomiędzy warstwą ukrytą oraz warstwą wyjściową znajduje się warstwa Flatten, która nie przeprowadza obliczeń, ale służy do przeniesienia wartości.

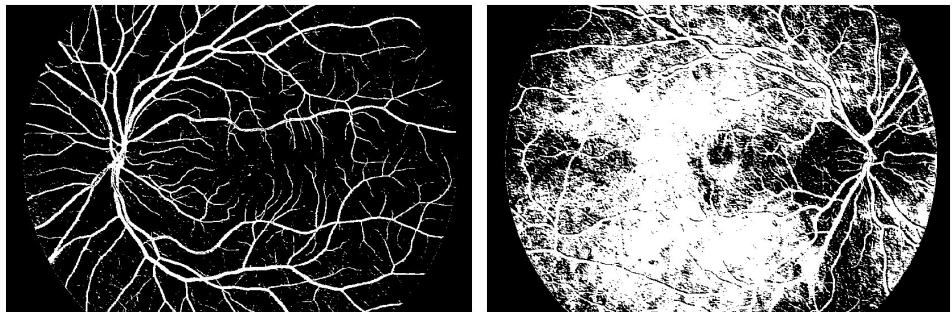
Przed przeprowadzeniem podziału obrazu na 5x5 pikseli, obraz został dodatkowo zmniejszony, tak jak w przypadku algorytmu cyfrowego przetwarzania.

Warstwa wejściowa składa się z 75 neuronów, a ukryta z 45 neuronów. Wartości te, jak i również współczynnik *Epoch* (determinuje liczbę przejść machine learningu przez cały zestaw danych) zostały wyznaczone z pomocą wstępnej oceny *10-fold cross validation*. Dla takich danych zarówno metryki - *accuracy* oraz *loss*, a także efekt wyjściowy były najlepsze.

Średnie accuracy	0,9627767941
Średnie sensitivity	0,9699631527
Średnie specificity	0,8742962184
Średnie negative predictive value	0,6848361773
Średnie precision	0,9898629513

Rys 1. Wyniki *10-fold cross validation*

Ta sama sieć neuronowa dla pewnych obrazów może działać bardzo dobrze, z kolei dla innych może sobie zupełnie nie radzić. Dlatego kluczowe było dobranie takich parametrów, aby w każdym przypadku sieć radziła sobie dobrze z wykrywaniem naczyń. Z tego powodu nie tylko średnie metryki są ważne, ale również najlepsze / najgorsze przypadki.

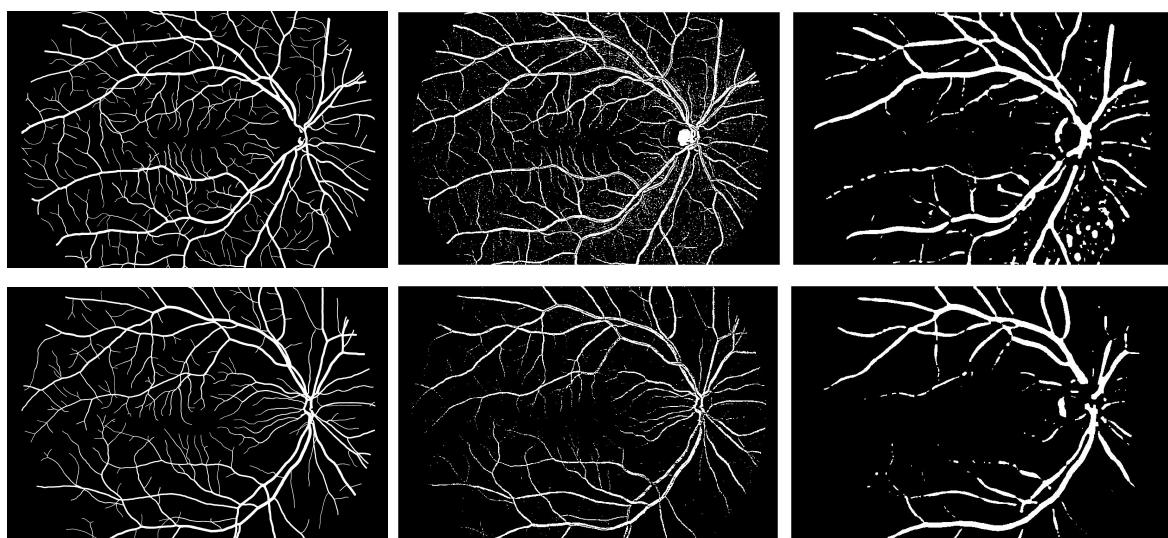


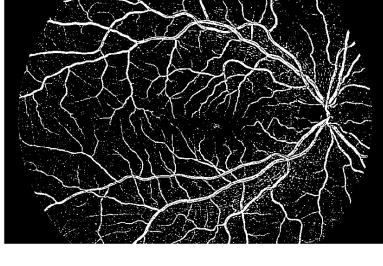
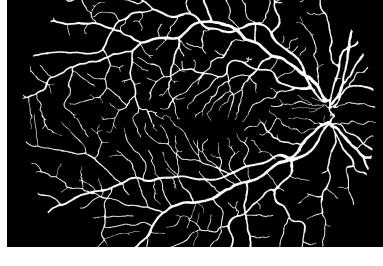
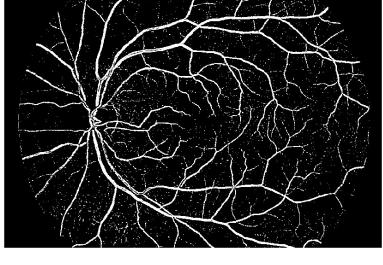
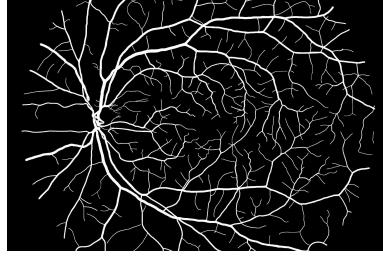
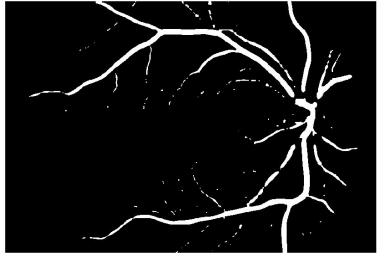
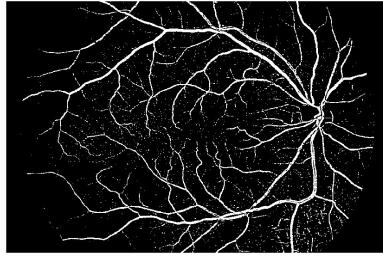
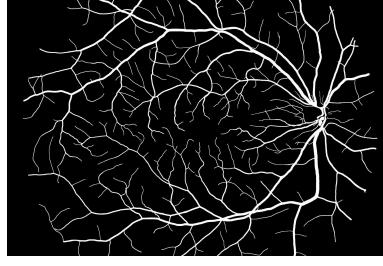
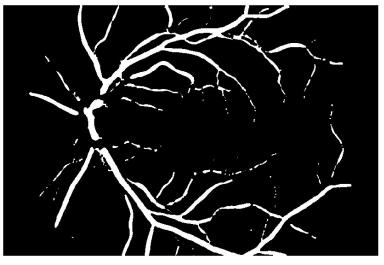
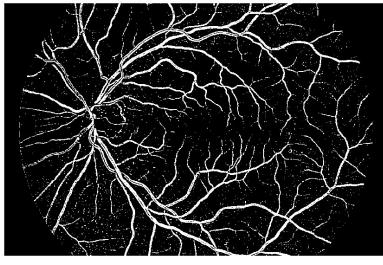
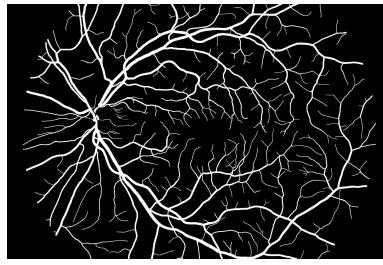
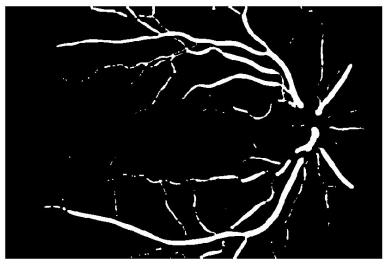
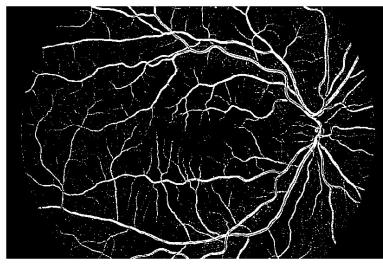
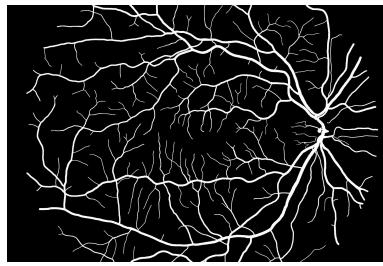
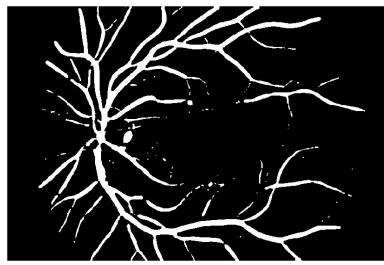
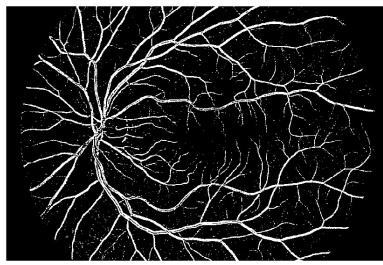
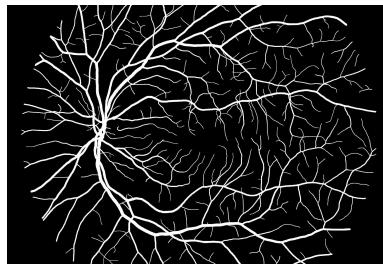
Rys. 2. Wyniki dla 25 neuronów na pierwszej warstwie i 50 na drugiej - dla niektórych przypadków sieć radzi sobie świetnie, dla innych - możemy zaobserwować przeuczenie.

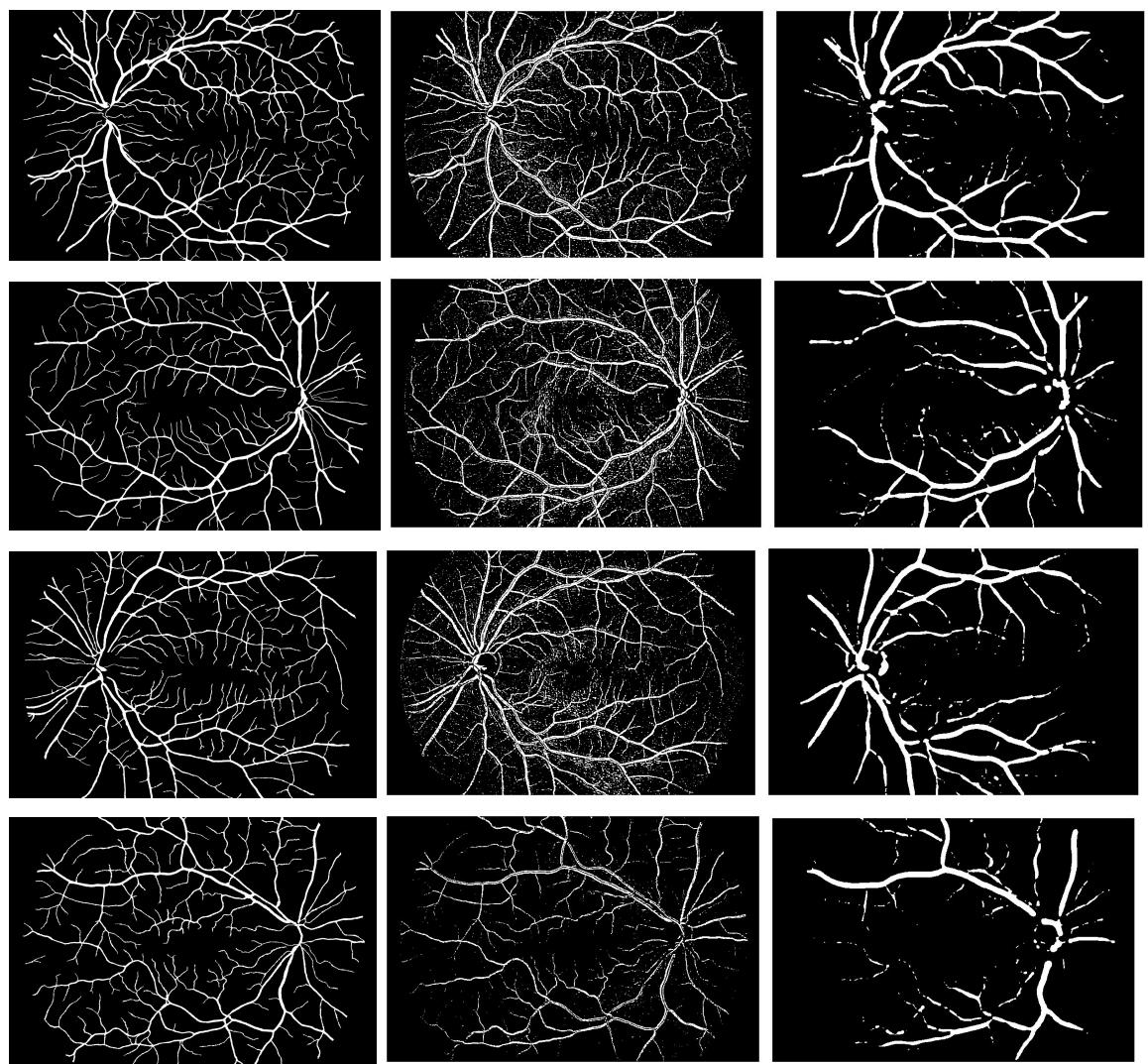
Z powodu specyficznych cech sieci CNN, nie zostały wyznaczone żadne dodatkowe cechy obrazu. W ten sposób sieć może po kolei przefiltrowywać każdy piksel i dzięki temu wykrywać wzorce. Zupełnie inaczej działają np. drzewa binarne, czy też inne klasyfikatory, dla których wyznaczenie momentów Hu, saturacji pikseli lub innych cech może być bardzo pomocne, lub nawet konieczne do działania.

### 3. Wizualizacja wyników

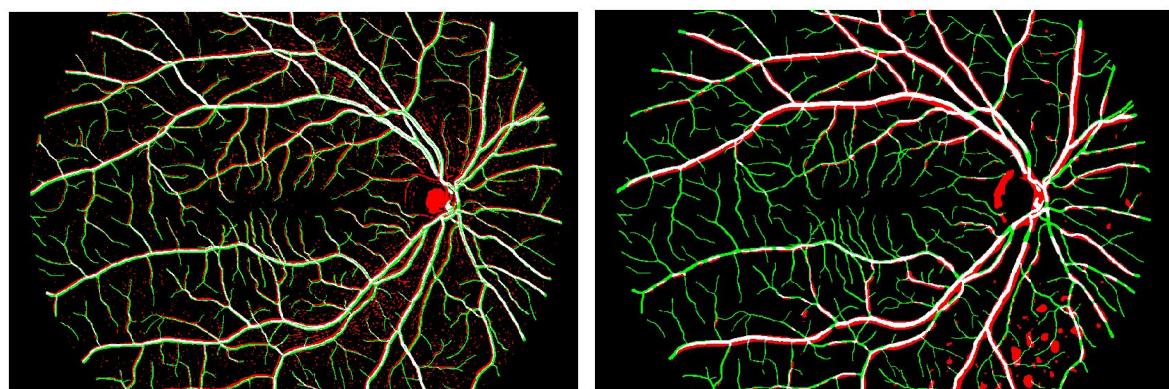
Do przetestowania algorytmów została wykorzystana baza obrazów HRF. Sieć uczy się na drugim, trzecim oraz czwartym z nich, a pozostałe zostały wykorzystane do zaprezentowania wyników.

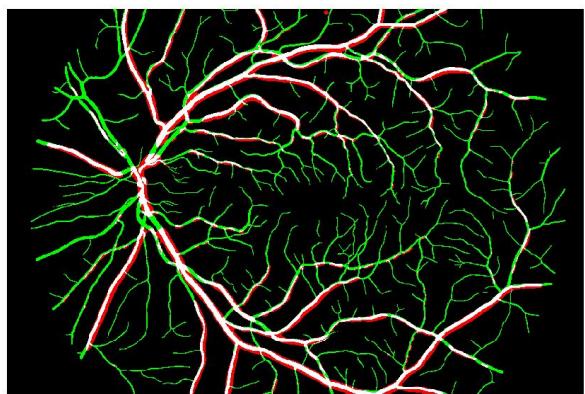
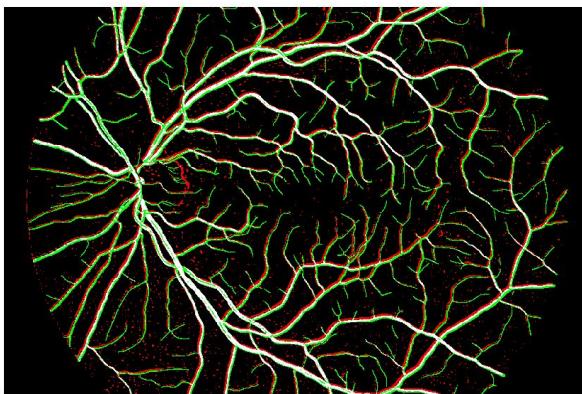
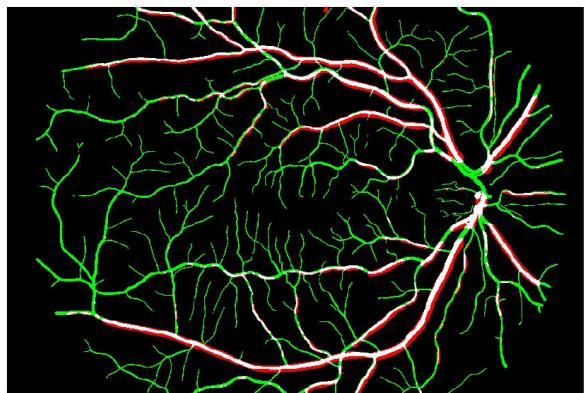
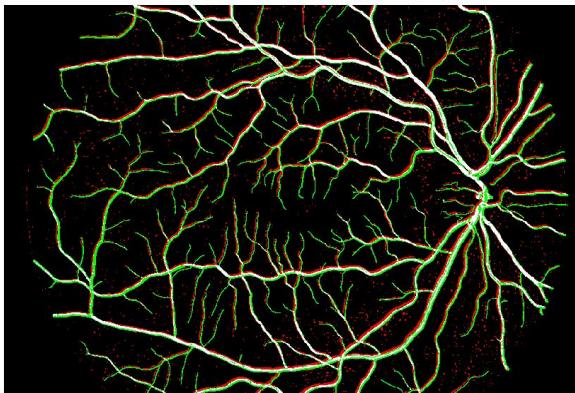
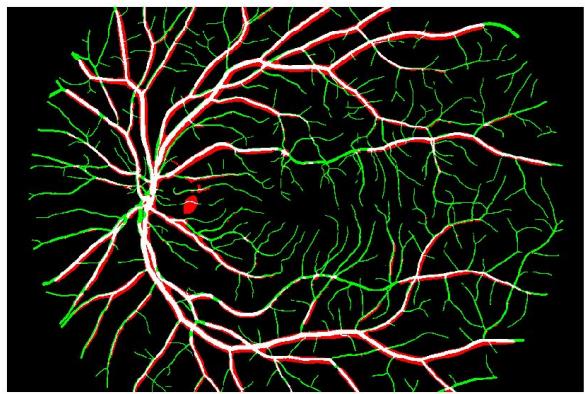
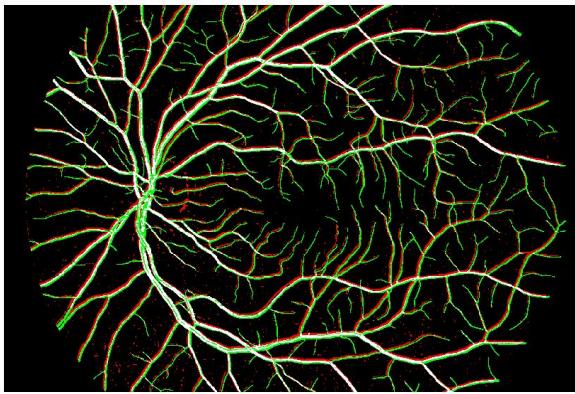
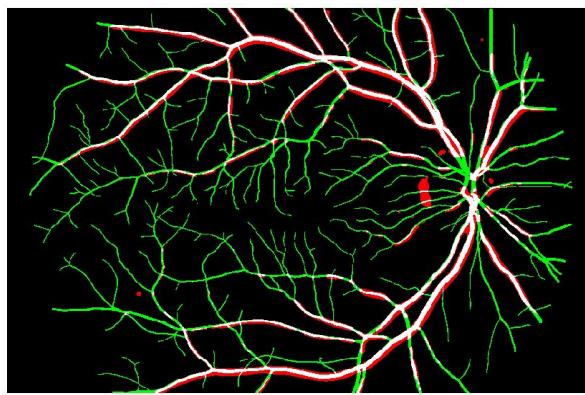
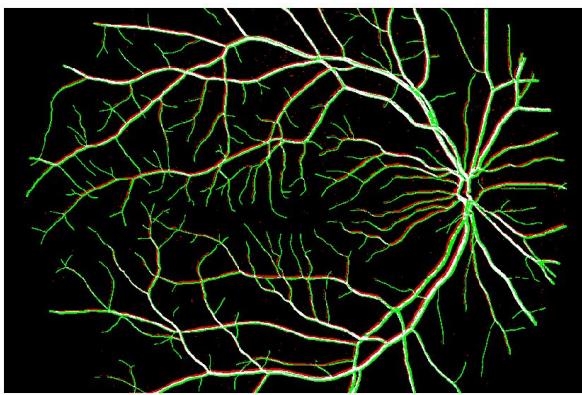


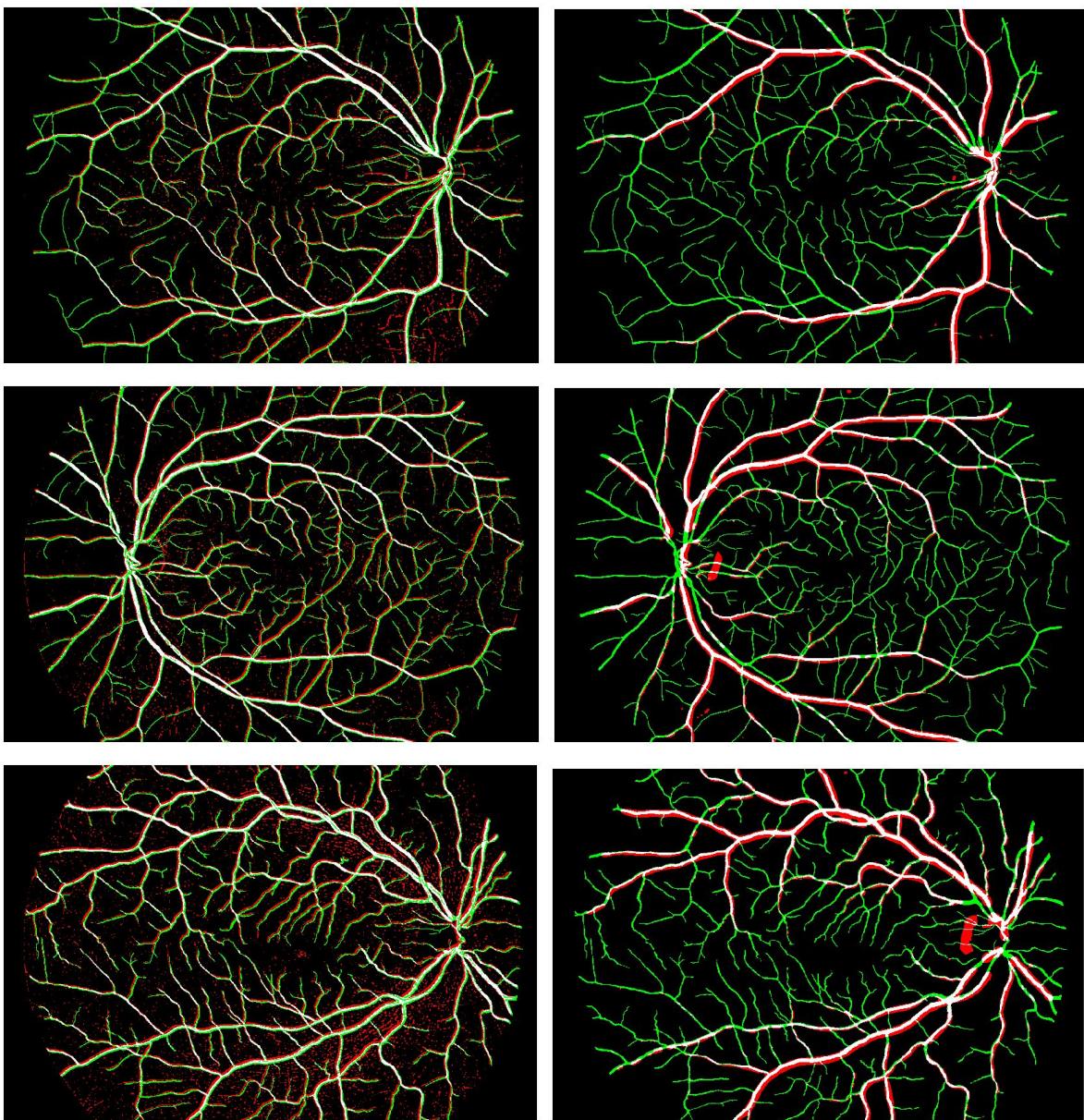




Rys 3. Kolejne obrazy z bazy HRF. Od lewej - ground truth, wynik działania machine learningu, cyfrowe przetwarzanie obrazów

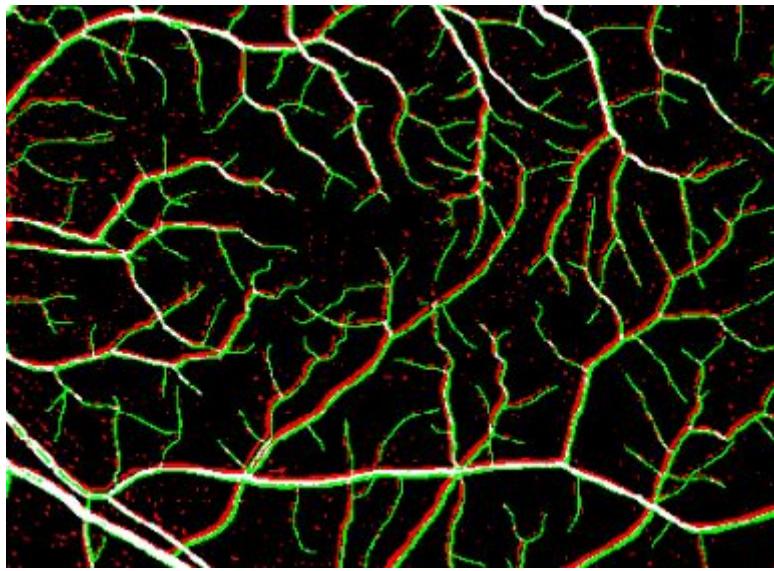






Rys 4. Nałożenie wynikowych obrazów na ground truth. Po stronie lewej - machine learning, po prawej - cyfrowe przetwarzanie obrazu. Kolorem czerwonym zaznaczono błędnie sklasyfikowane piksele jako naczynia, kolorem zielonym - błędnie sklasyfikowane tło.

Na rysunku nr 4 można zauważyć, że sieć neuronowa w przeważającej większości miejsc dobrze sklasyfikowała kształt naczyń krwionośnych, ale nie do końca dobrze sobie poradziła z umieszczeniem naczyń w przestrzeni, zwłaszcza tych najdrobniejszych, które często są przesunięte o kilka pikseli. Należy zauważać, że wszystkie piksele, które zostały w taki sposób przemieszczone, będą mimo to liczone jako niepoprawnie sklasyfikowane w analizie wyników. Aby wyeliminować problem, należałoby dalej pracować nad parametrami sieci, tak, aby filtry działały jak najdokładniej.



Rys 5. Nałożenie wynikowych obrazów na ground truth w przybliżeniu. Wynik działania sieci neuronowej.

#### 4. Analiza wyników

Do analizy wyników zastosowano następujące miary:

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)}$$

$$Sensitivity = \frac{TP}{(TP + FN)}$$

$$Specificity = \frac{TN}{(TN + FP)}$$

$$Precision = \frac{TP}{(TP + FP)}$$

$$Negative\ Predictive\ Value = \frac{TN}{(TN + FN)}$$

Przetwarzanie obrazu	Accuracy	Sensitivity	Specificity	Negative predictive value	Precision	Średnia NPV oraz Precision
Średnie	0,907	0,556	0,939	0,958	0,451	0,705
Minimalne	0,894	0,506	0,931	0,944	0,361	0,652
Maksymalne	0,919	0,599	0,946	0,972	0,541	0,757
Odchylenie standardowe,	0,008	0,026	0,004	0,010	0,061	0,035

Rys 5. Wyniki pomiarów dla przetwarzania obrazów

CNN	Accuracy	Sensitivity	Specificity	Negative predictive value	Precision	Średnia NPV oraz Precision
Średnie	0,948	0,972	0,766	0,750	0,971	0,860
Minimalne	0,930	0,960	0,609	0,575	0,943	0,759
Maksymalne	0,959	0,979	0,911	0,835	0,993	0,914
Odchylenie standardowe,	0,010	0,006	0,093	0,075	0,017	0,046
Poprawa średnich względem cyfrowego przetwarzania obrazu	1,05	1,75	0,82	0,78	2,15	1,22

Rys 6. Wyniki pomiarów dla CNN

## 5. Wnioski

Miara *accuracy* mówi o poprawne sklasyfikowanych jako pozytywne pikselach, biorąc pod uwagę wszystkie piksele. W obu algorytmach średnio wynosi ona ponad 90% (odpowiednio 95% oraz 90,7% dla sieci neuronowych oraz cyfrowego przetwarzania). W tym przypadku nie jest to miarodajna metryka, ponieważ mamy do czynienia z niezrównoważonym rozkładem klas.

Lepszymi miarami są *sensitivity*, które mówi o tym, jaki procent naczyń krwionośnych został pozytywnie sklasyfikowany (odpowiednio 97% oraz 56%, co jest znaczącą różnicą), oraz *specificity* - które definiuje procent poprawnie sklasyfikowanego tła (77% oraz 94%).

Dla danych niezrównoważonych bardzo dobrą metryką jest *precision*. Metryka ta jest proporcją poprawnie sklasyfikowanych wartości klasy pozytywnej. Nie bierze ona pod uwagę klasy negatywnej, dlatego dla badanego przypadku, gdzie naczyń krwionośnych jest bardzo mało w porównaniu do tła, jest to najlepsza metryka badająca skuteczność (spośród wybranych). Można zauważyć, że w badanych algorytmach CNN osiągnęło *precision* ponad dwa razy wyższe niż cyfrowe przetwarzanie obrazu (97% i 45%).

Odwrotną miarą jest *negative predictive value*, które bierze pod uwagę tylko negatywną klasę, a więc mówi o tym, jaki procent tła zostało dobrze sklasyfikowane. W przedstawionych algorytmach wynosi ona 75% i 95% odpowiednio dla CNN i przetwarzania obrazu.

Takie wyniki można przewidzieć patrząc na wynikowe obrazy - przetwarzanie obrazu bardzo dużą część naczyń krwionośnych pomija, a więc zarówno *sensitivity* jak i *precision* będzie dla tej metody niskie, a *specificity* oraz *NPV* - wysokie. Ostateczne

porównanie otrzymamy patrząc również na średnią arytmetyczną *NPV* oraz *precision*, które jest zauważalnie większe dla CNN (86% i 70%).

Jak dla tak małej ilości danych wejściowych (3 obrazy), sieć radzi sobie bardzo dobrze. Obrazy wyjściowe są bardzo zbliżony do oryginalnej maski, zdecydowanie bardziej, niż w przypadku cyfrowego przetwarzania obrazów. Większość błędów sieci neuronowej jest związana z przesunięciem kształtów o kilka pikseli. Metoda przetwarzania obrazów zdecydowanie nie jest w stanie wykryć drobnych naczyń krwionośnych, które po prostu je pomija.

Najlepsze efekty dawałyby więc kombinacja obu metod, w których wykorzystanie cyfrowego przetwarzania mogłoby pomóc w usunięciu przesunięcia.