

Kierunek: **Informatyczne Systemy Automatyki (ISA)**
Specjalność: **Inteligentne Systemy Przemysłu 4.0 (IPS)**

PRACA DYPLOMOWA
MAGISTERSKA

(Projekt przejściowy)

Rozpoznawanie chorób oczu za pomocą
głębokich sieci neuronowych

inż. Dominik Goleniewski
inż. Aleksandra Rozmus

Opiekun pracy
Mgr inż. Teodor Niżyński

Słowa kluczowe: choroby oczu, sieci neuronowe, uczenie maszynowe

Spis treści

1	Wstęp	1
1.1	Problem badawczy	1
1.2	Cel pracy	1
1.2.1	Rozpatrywane choroby	1
1.3	Struktura pracy	3
2	Przegląd literaturowy	5
3	Zbiór danych uczących	7
3.1	Dobór danych	7
3.2	Rozkład liczości klas	8
4	Praca badawcza	9
4.1	Wykorzystane technologie i środowisko pracy	9
4.2	Architektura	9
4.3	Badanie ilości epok	9
4.4	Badanie doboru funkcji aktywacji i algorytmu uczenia	13
4.4.1	ReLu oraz Adam	14
4.4.2	ReLu oraz SGD	15
4.4.3	Tangens hiperboliczny oraz Adam	17
4.4.4	Tangens hiperboliczny oraz SGD	18
5	Wnioski i plany na przyszłość	21
5.1	Wnioski	21
5.2	Plany na przyszłość	21

1. Wstęp

W pracy formułuje się cele o charakterze badawczym wymagające doboru i zastosowania metod badawczych, wykorzystując wiedzę teoretyczną oraz naukową. Wskazane jest przedstawienie, co nowego jest zaproponowane w pracy oraz podanie ograniczeń i słabych/mocnych stron opracowanego rozwiązania (jeżeli dotyczy). Rozdział wprowadzający powinien służyć czytelnikowi do zrozumienia celu pracy.

1.1. Problem badawczy

Choroby oczu stwarzają znaczące wyzwania zdrowotne na globalną skalę. W dzisiejszych czasach identyfikacja niektórych schorzeń oczu przy użyciu tradycyjnych metod może być czasochłonna i skomplikowana, zwłaszcza jeśli chodzi o rzadkie przypadki. Wczesna diagnoza tych chorób może istotnie przyspieszyć proces leczenia, co w rezultacie może zapobiec utracie wzroku.

W miarę postępu technologicznego i rozwoju sztucznej inteligencji, naukowcy i przedsiębiorstwa coraz częściej szukają sposobów wykorzystania AI, zwłaszcza w dziedzinie medycyny. W obszarze opieki zdrowotnej, gdzie lekarze mają do czynienia z różnymi przypadkami chorych, sztuczna inteligencja oferuje potencjał poprawy jakości i dostępności leczenia. W okulistyce, na przykład, można wykorzystać uczenie maszynowe do analizy zdjęć dna oka, co pomaga w diagnozowaniu chorób. Głębokie sieci neuronowe, zwłaszcza sieci konwolucyjne, stają się kluczowym narzędziem w rozpoznawaniu obrazów.

1.2. Cel pracy

Celem pracy jest stworzenie konwolucyjnej sieci neuronowej do rozpoznania poszczególnych chorób oczu oraz przeprowadzenie badań wpływu zmiany parametrów sieci na jej skuteczność.

1.2.1. Rozpatrywane choroby

Jaskra

Jednym z powszechnie znanych schorzeń oczu jest jaskra, charakteryzująca się uszkodzeniem nerwu wzrokowego, silnie związana z podwyższonym ciśnieniem wewnątrzgałkowym. Według Światowej Organizacji Zdrowia (WHO) jaskra zajmuje drugie miejsce wśród przyczyn ślepoty na świecie, zaraz za zaćmą. Ważne jest, aby zauważyć, że najczęściej dotyka osoby po 40. roku życia. Objawy jaskry obejmują zmniejszone pole widzenia, zwłaszcza na obwodzie, oraz trudności z widzeniem nocnym. Podobnie jak w przypadku wielu chorób oczu, wczesne wykrycie jaskry zwiększa szanse na utrzymanie zdolności wzrokowej.

Zaćma

Zaćma stanowi główną przyczynę ślepoty na świecie, spowodowaną zmętnieniem soczewki, co ostatecznie prowadzi do utraty widzenia z powodu braku odpowiedniej ilości światła. Głównym czynnikiem ryzyka zaćmy jest starzenie się, gdy elastyczność i przejrzystość soczewki zanikają. Palenie tytoniu oraz nadmierna ekspozycja na promienie UV również przyspieszają jej rozwój. Nie istnieje idealna metoda zapobiegania starzeniu się soczewki, ale pewne środki ostrożności, takie jak noszenie okularów przeciwsłonecznych i unikanie palenia, mogą zmniejszyć ryzyko zachorowania. Chirurgiczne leczenie zaćmy polega na wymianie mętnej soczewki na sztuczną.

Nadciśnienie

Nadciśnienie wewnątrzgałkowe, a dokładniej nadciśnienie płynu wodnistego w oku jest ściśle związane z jaskrą, często stanowiącą jego wynik. Badanie, które mierzy ciśnienie w oku nazywa się tonometria, a odpowiedni zakres dla zdrowego człowieka określa się w milimetrach rtęci i wynosi między 10mm Hg, a 21mm Hg. Warto również nadmienić, że dzienne wahania ciśnienia w oku nie powinny wynosić więcej niż 5mm Hg. W leczeniu nadciśnienia wewnątrzgałkowego kluczową rolę odgrywa związek z jaskrą. Metody te obejmują stosowanie kropli do oczu, a w niektórych przypadkach problem można rozwiązać również drogą operacyjną.

Cukrzyca

Cukrzyca to choroba metaboliczna charakteryzująca się podwyższonym poziomem glukozy we krwi. W kontekście oczu może prowadzić do retinopatii cukrzycowej, powodującej uszkodzenie naczyń krwionośnych w siatkówce. Zaawansowana retinopatia może prowadzić do niszczenia istniejących naczyń, a także do powstawania nowych, co może skutkować krwawieniami i bliznami, w rezultacie prowadząc do utraty wzroku.

Patologiczna krótkowzroczność

Patologiczna krótkowzroczność występuje, gdy przekracza ona bardzo wysoki stopień, np. -6 dioptrii. W tych przypadkach gałka oczna jest nadmiernie wydłużona w osi strzałkowej, co skutkuje patologicznymi zmianami strukturalnymi narządów wzroku, takimi jak rogówka, siatkówka i naczynia krwionośne na gałce ocznej. Takie deformacje zwiększają ryzyko wielu chorób oczu. Leczenie obejmuje zazwyczaj noszenie okularów korekcyjnych lub specjalnych soczewek kontaktowych, takich jak soczewki twardochrzętne czy skleralne. Regularne wizyty kontrolne u okulisty są również ważne, częstsze niż w przypadku zdrowych osób. Istnieją także bardziej inwazyjne metody, takie jak przeszczep rogówki czy operacja refrakcyjna.

Zwyrodnienie plamki żółtej związane z wiekiem

Zwyrodnienie plamki żółtej związane z wiekiem to choroba dotycząca obszaru siatkówki, odpowiedzialnego za centralne widzenie. Jak wskazuje sama nazwa, głównym czynnikiem wpływającym na jej wystąpienie jest wiek, zorientowany przede wszystkim na osoby po 50. roku życia. Dodatkowo, ekspozycja na promienie UV i palenie tytoniu, podobnie jak w innych chorobach oczu, odgrywają rolę. W zakresie leczenia najczęściej stosuje się terapie suplementacyjne oraz farmaceutyczne.

1.3. Struktura pracy

Pierwszy rozdział stanowi wstęp teoretyczny pozwalający czytelnikowi na zapoznanie się z badanym problemem, celem pracy oraz jej strukturą. Drugi rozdział stanowi przegląd literaturowy obecnych rozwiązań. Trzeci rozdział przedstawia zbiór danych użyty do wytrenowania stworzonego modelu. Czwarty rozdział zawiera przeprowadzone badania na stworzonym modelu. Piąty rozdział stanowi podsumowanie zrealizowanych prac, wnioski oraz możliwe plany na przyszłość.

2. Przegląd literaturowy

Sztuczna inteligencja i uczenie maszynowe (UM) stają się coraz bardziej istotne we wszystkich sferach życia, w tym szczególnie w dziedzinie medycyny. Z uwagi na ogromny potencjał poprawy jakości i dostępności opieki zdrowotnej, na uczelniach technicznych rozwijają się specjalności związane z wykorzystaniem informatyki i uczenia maszynowego w obszarze medycyny. Dodatkowo, coraz więcej naukowców skoncentrowanych jest na optymalizacji systemów opieki zdrowotnej, diagnozowaniu chorób oraz opracowywaniu innowacyjnych metod leczenia, co zostało wspomniane w artykule [6].

Jednym z kluczowych atutów dziedziny medycyny jest obfitość danych medycznych zgromadzonych w szpitalach i klinikach dotyczących każdego pacjenta. Ten obfity zasób informacji otwiera znakomite możliwości dla nauk takich jak sztuczna inteligencja i uczenie maszynowe, które z natury swojej potrzebują znacznych ilości danych. Ośrodkom medycznym łatwo jest pozyskiwać te dane w stosunkowo krótkim czasie, co zostało zobrazowane w artykule [3]. Przykładowo, w bardzo krótkim okresie udało się stworzyć obszerny zbiór danych dotyczący pacjentów z COVID-19.

Choroby oczu mogą prowadzić do nieodwracalnego pogorszenia wzroku, jeśli nie zostaną odpowiednio leczone w porę. Opracowano różne techniki obrazowania, aby pomóc w wykrywaniu chorób oczu, w tym powszechnie stosowaną fotografię siatkówki oka w kolorze. Badania mające na celu poprawę diagnozowania stanowią zainteresowanie inżynierów i stale prowadzone są w tym kierunku badania. Badanie chorób siatkówki oka jest popularnym tematem, ze względu na łatwą dostępność zbioru danych w internecie. Przykładowe badanie zostało przedstawione w artykule [4] i stanowiło inspirację do niniejszej pracy. W niniejszej pracy odniesiono się do metody opartej na destylacji wiedzy, zaproponowanej w artykule, celem poprawy automatycznej klasyfikacji chorób oczu opartej na obrazach. Przedstawiona technika, wykorzystująca dwie głębokie sieci neuronowe, umożliwia skuteczniejsze rozpoznawanie chorób oczu we wczesnych etapach, nawet w sytuacjach, gdzie symptomy są trudne do zauważenia. Eksperymenty potwierdzają, że zastosowanie tej metody znacząco zwiększa skuteczność diagnozy, eliminując jednocześnie konieczność korzystania z cech klinicznych.

Bardziej zaawansowany system został przedstawiony w pracy [1]. Jest to zaawansowany system klasyfikacji głębokiego uczenia Fundus-DeepNet, zdolny do precyzyjnej identyfikacji wielu chorób oczu poprzez integrację reprezentacji cech z par obrazów siatkówki oka. Wykorzystuje on m.in. bloki takie jak High-Resolution Network (HRNet) i blok uwagi (Attention Block), a także Dyskryminacyjną Maszynę Boltzmanna (DRBM) do generowania rozkładu prawdopodobieństwa identyfikującego osiem różnych chorób oczu. Eksperymenty na zestawie danych OIA-ODIR potwierdzają, że System Fundus-DeepNet osiąga wybitne wyniki, oferując obiecujące rozwiązanie dla wczesnej diagnozy i leczenia w dziedzinie okulistyki.

3. Zbiór danych uczących

3.1. Dobór danych

Aby przeprowadzić eksperymenty skorzystano z powszechnie znanej platformy kaggle.com, która oferuje obszerną kolekcję danych. Jest to jedna z najczęściej używanych stron przez profesjonalistów zajmujących się sztuczną inteligencją oraz uczeniem maszynowym. Po dokładnej analizie dostępnych danych udało się znaleźć zestaw danych dotyczący Rozpoznawania Chorób Oka. Zbiór ten obejmuje ponad 6000 zdjęć zebranych przez firmę Shanggong Medical Technology Co., Ltd z różnych szpitali w Chinach. Zdjęcia te przedstawiają tło oka i zostały wykonane przy użyciu aparatów różnych marek, takich jak Canon, Zeiss czy Kowa. Pacjenci, których dane są udostępnione, reprezentują różne płcie i wiek, w zakresie od 1 do 91 lat. Ponadto, ich oczy różnią się także pod względem koloru.

Portal z danymi oferuje zarówno oryginalne, jak i częściowo przetworzone zdjęcia. Warto zaznaczyć, że oryginalne zdjęcia zostały zarejestrowane przy użyciu sprzętu o wysokiej rozdzielczości, co skutkowało dużym zużyciem pamięci systemowej. Aby dostosować obrazy do procesu uczenia sieci neuronowych, konieczne było unifikowanie rozdzielczości. Z tego powodu twórca zbioru danych zaproponował już wstępnie przetworzone obrazy o rozmiarze $512 \times 512 \times 3$, gdzie 3 oznacza trzy kanały barwne. W celu efektywnego uczenia sieci zdecydowano się dodatkowo zmniejszyć rozdzielczość do $224 \times 224 \times 3$ - jest to powszechnie stosowany rozmiar w wielu znanych architekturach sieci neuronowych.



Rysunek 3.1: Przykładowe zdjęcie oka ze zbioru - zdrowe oko prawe.

3.2. Rozkład licznosci klas

Jednym z trudności wynikających z tego zbioru danych jest jego niezrównoważenie. Problem ten polega na tym, że niektóre klasy posiadają znacznie więcej danych niż pozostałe. To z kolei ma istotny wpływ na proces projektowania i uczenia modelu, który często wykazuje tendencję do błędnych predykcji na korzyść klas o większej liczbie przykładów. Ze względu na znaczną przewagę ilości zdjęć ze zdrowymi oczami zdecydowano się na redukcję liczności tej klas. Zdecydowano się w zbiorze pozostawić 400 losowych zdjęć zdrowego lewego oka oraz 400 losowych zdjęć zdrowego prawego oka.

4. Praca badawcza

4.1. Wykorzystane technologie i środowisko pracy

Model sieci neuronowej postanowiono stworzyć przy pomocy języka Python i dostępnych w nim bibliotek, takich jak **Keras**, **Tensorflow**, **Scikit-learn** oraz **Matplotlib** [2]. Badania zostały przeprowadzone przy użyciu platformy **Google Colaboratory**, gdzie dostępna jest bezpłatna infrastruktura obliczeniowa do uruchamiania kodu. Wykorzystanie GPU na tej platformie znacząco przyspieszyło proces uczenia sieci w porównaniu z lokalnymi zasobami sprzętowymi.

4.2. Architektura

Architektura stworzonego modelu zawierała: warstwy konwolucyjne i warstwy łączące, występujące na zmianę, warstwę spłaszczania, warstwy gęste oraz warstwy porzucania, także występujące naprzemiennie. Każda z warstw konwolucyjnych poza pierwszą posiadała filtry o rozmiarze 3×3 , natomiast w pierwszej warstwie miały one rozmiar 7×7 . Każda kolejna warstwa konwolucyjna podwajała licznosc występujących w niej filtrów względem poprzedniej. Liczba neuronów wyjściowych gęstych warstw wynosiła odpowiednio 128 oraz 64, wraz z oczywiście warstwą końcową posiadającą 8 neuronów wyjściowych, ze względu na przeprowadzoną klasyfikację 8 klas. Do pierwszych badań zdecydowano przyjąć się poniższe wartości hiperparametrów:

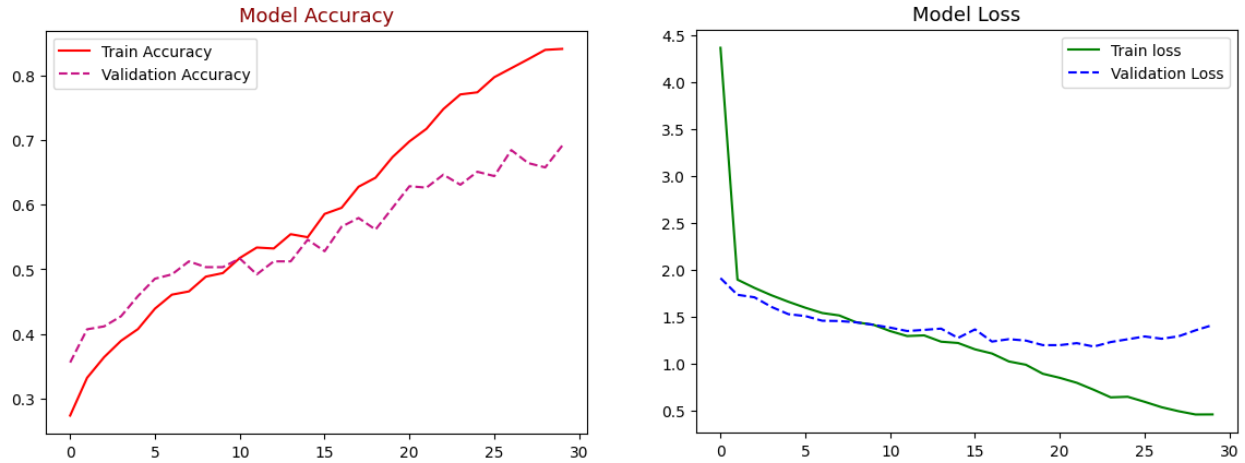
- ilość filtrów w pierwszej warstwie konwolucyjnej - 8
- wielkość wsadu - 32
- funkcje aktywacji w warstwach konwolucyjnych - RELU
- wartość współczynnika porzucania - 40%
- algorytm uczenia - Adam

W trakcie przeprowadzonych badań parametry te ulegały zmianie.

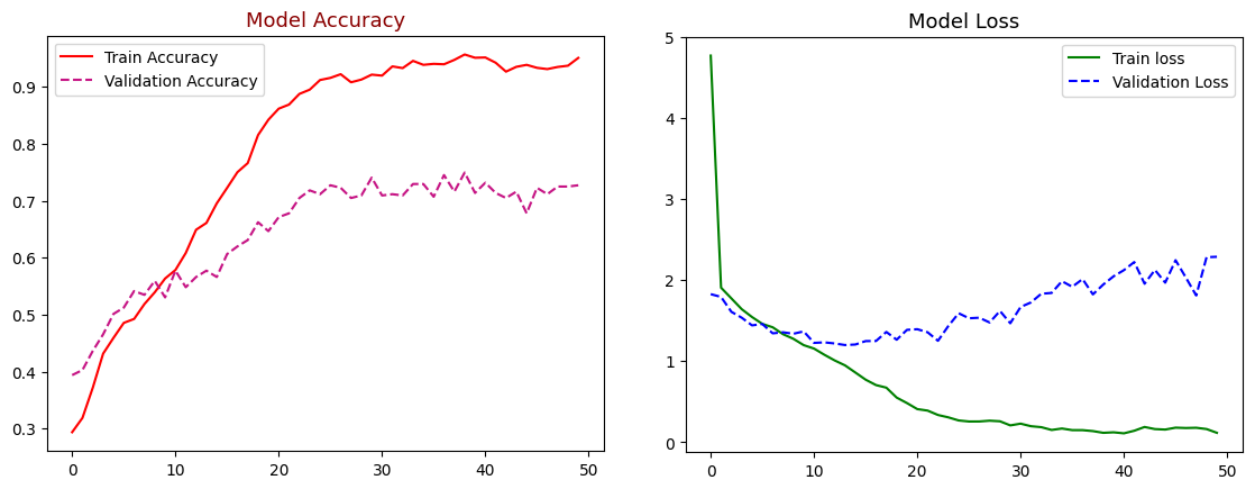
4.3. Badanie ilości epok

W kontekście sieci neuronowych, termin "epoka" odnosi się do etapu uczenia modelu, w którym ten jest jednorazowo trenowany na całym zbiorze treningowym. Po zakończeniu jednej epoki następuje ponowne przetworzenie wszystkich danych w celu kontynuacji treningu sieci. Optymalna liczba epok jest istotna, ponieważ zbyt duża może prowadzić do przetrenowania

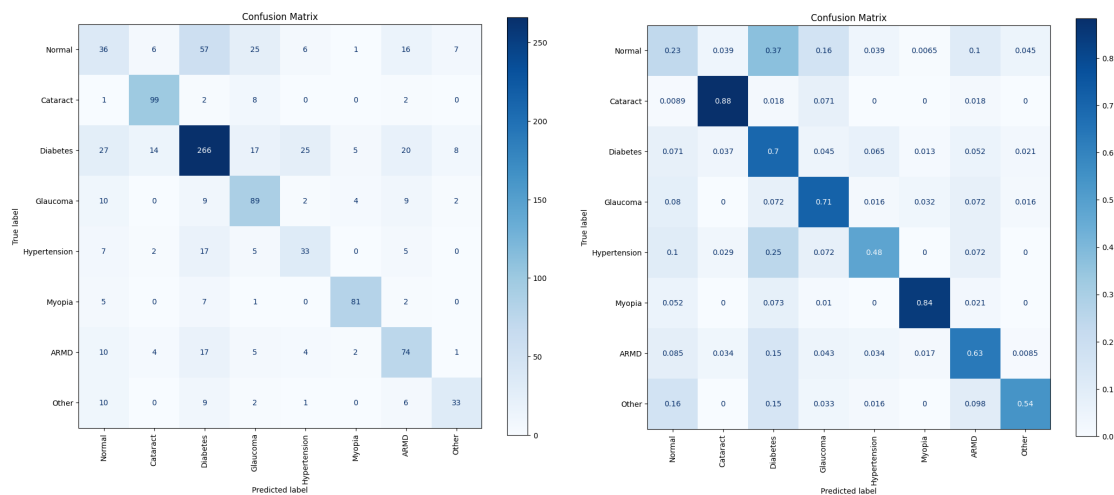
modelu, co oznacza utratę zdolności do ogólnego uogólniania, natomiast zbyt mała może skutkować niedostatecznym przeszkoleniem modelu i obniżoną jakością wyników. W ramach pierwszego badania, zdecydowano się sprawdzić rezultaty osiągane przez sieć dla dwóch różnych ilości epok: 30 i 50.



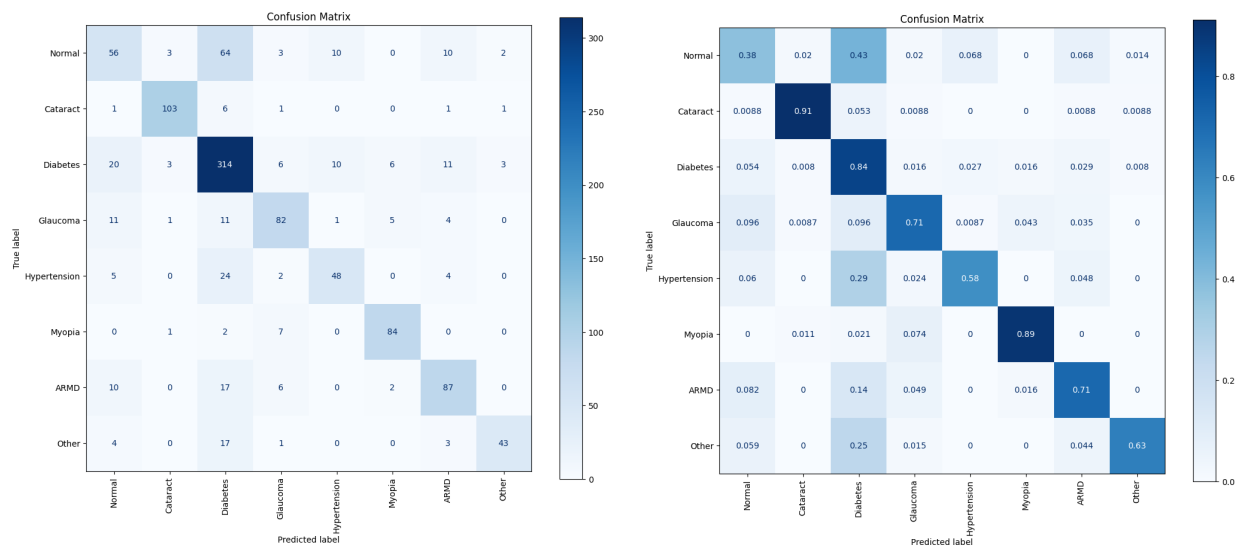
Rysunek 4.1: Wykresy dokładności i straty dla 30 epok



Rysunek 4.2: Wykresy dokładności i straty dla 50 epok



Rysunek 4.3: Macierz pomyłek dla 30 epok



Rysunek 4.4: Macierz pomyłek dla 50 epok

Listing 4.1: Wyniki uzyskane dla 30 epok

	precision	recall	f1-score	support
Normal	0.34	0.23	0.28	154
Cataract	0.79	0.88	0.84	112
Diabetes	0.69	0.70	0.69	382
Glaucoma	0.59	0.71	0.64	125
Hypertension	0.46	0.48	0.47	69
Myopia	0.87	0.84	0.86	96
ARMD	0.55	0.63	0.59	117
Other	0.65	0.54	0.59	61
accuracy			0.64	1116
macro avg	0.62	0.63	0.62	1116
weighted avg	0.63	0.64	0.63	1116

Suma f1 score: 1.9261832077540657

Srednia f1 score: 0.2407729009692582

Listing 4.2: Wyniki uzyskane dla 50 epok

	precision	recall	f1-score	support
Normal	0.52	0.38	0.44	148
Cataract	0.93	0.91	0.92	113
Diabetes	0.69	0.84	0.76	373
Glaucoma	0.76	0.71	0.74	115
Hypertension	0.70	0.58	0.63	83
Myopia	0.87	0.89	0.88	94
ARMD	0.72	0.71	0.72	122
Other	0.88	0.63	0.74	68
accuracy			0.73	1116
macro avg	0.76	0.71	0.73	1116
weighted avg	0.73	0.73	0.73	1116

Suma f1 score: 5.817949757372596

Srednia f1 score: 0.7272437196715745

4.4. Badanie doboru funkcji aktywacji i algorytmu uczenia

W przypadku sieci neuronowych konwolucyjnych (CNN) kluczową rolę odgrywa starannie dobrany algorytm uczenia. Jego głównym zadaniem jest dostosowywanie wag sieci neuronowej w trakcie procesu uczenia, mając na celu poprawę jakości modelu i minimalizację funkcji kosztu. Istotną cechą efektywnego algorytmu uczenia jest zdolność do unikania utknięcia w minimach lokalnych oraz utrzymanie stabilności modelu. Dodatkowo, różne algorytmy mają zdolność do szybszego uczenia się niż inne.

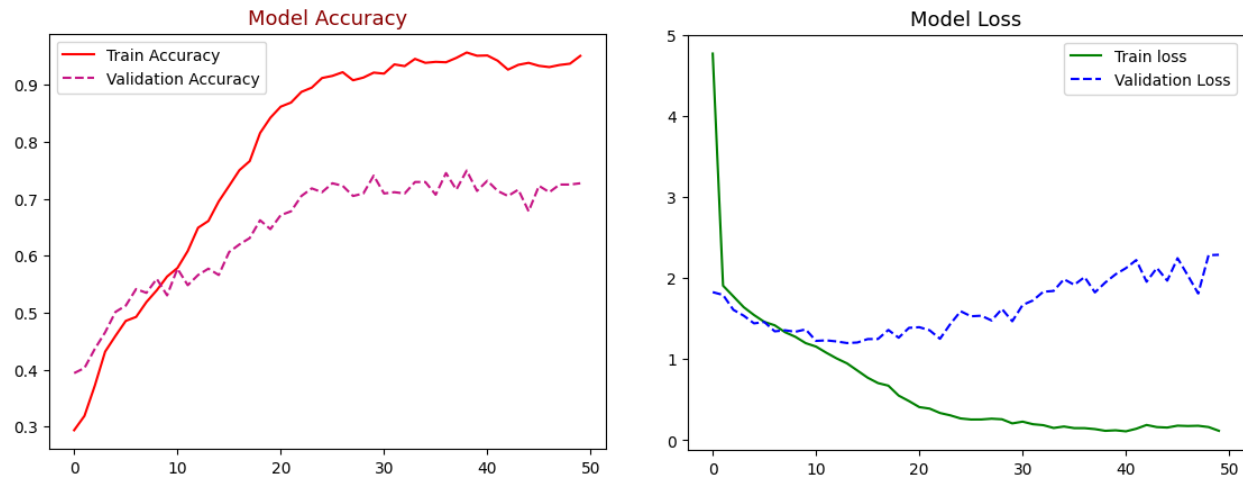
Prawdopodobnie najczęściej wykorzystywaną funkcją aktywacji w CNN jest funkcja ReLU (Rectified Linear Unit). Dużą jej zaletą jest to, że jest znacznie prostsza obliczeniowo od takich funkcji jak sigmoidalna, czy tangens hiperboliczny. Ponadto rozwiązuje problem zanikającego gradientu, gdyż w przypadku, gdy wartość wejściowa jest większa od 0 to gradient wynosi 1.

Algorytm Adam (Adaptive Moment Estimation) to prawdopodobnie najczęściej używany algorytm uczenia w sieciach neuronowych konwolucyjnych (CNN). Jego kluczowym elementem jest zastosowanie momentum, które polega na wykorzystaniu momentu gradientu w celu przyspieszenia procesu uczenia. W przypadku dużych zmian w modelu, czyli gdy gradienty są duże, algorytm Adam przyspiesza proces aktualizacji wag. Natomiast gdy zmiany nie są tak znaczące, wagi zmieniają się mniej dynamicznie [5].

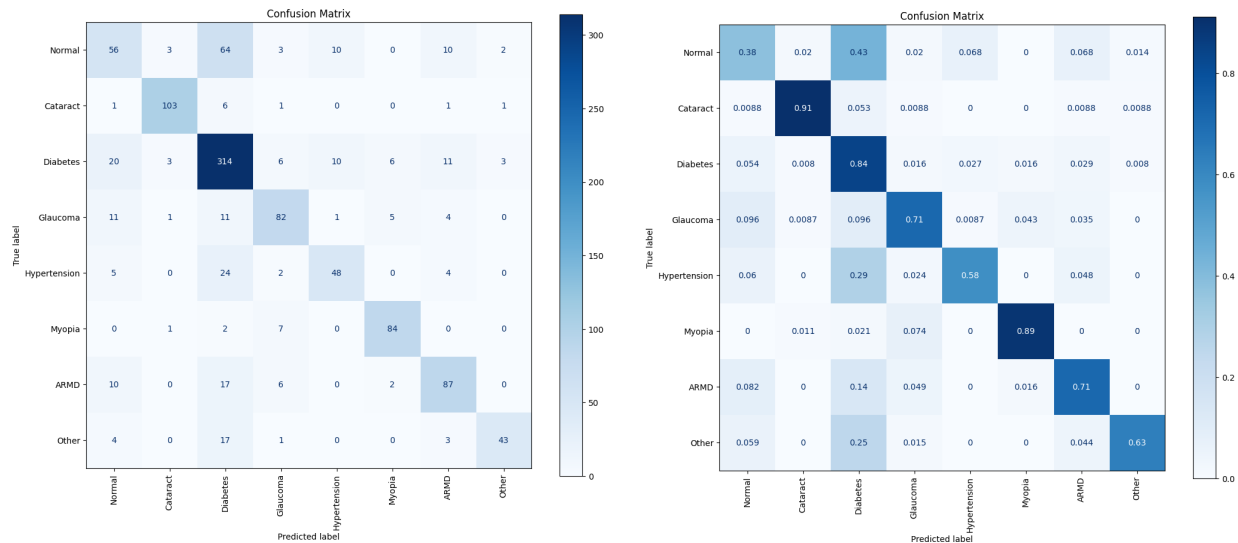
SGD (Stochastic Gradient Descent) to jeden z algorytmów, który był najwcześniej wykorzystywany w CNN. Polega on przede wszystkim na losowym wyborze próbki, następnie wykonywane jest obliczanie gradientu względem funkcji celu, po czym dokonywana jest aktualizacja wag. Ten proces jest następnie powtarzany. Wadą tego algorytmu w porównaniu do pozostałych jest występowanie większych oscylacji w trakcie uczenia oraz ryzyko utkania w lokalnych minimach.

Funkcja tangens hiperboliczny (\tanh) przypomina funkcję sigmoidalną, lecz różni się zakresem, ponieważ przyjmuje wartości w przedziale od -1 do 1, w przeciwieństwie do sigmoidy, która ma zakres od 0 do 1. Jest to także jedna z powszechnie używanych funkcji aktywacji. Charakteryzuje się ona symetrią względem 0, co jest szczególnie korzystne, gdy wartości są skoncentrowane wokół zera.

4.4.1. ReLu oraz Adam



Rysunek 4.5: Wykresy dokładności i uzyskane dla funkcji aktywacji ReLu i algorytmu uczenia Adam



Rysunek 4.6: Macierz pomyłek uzyskana dla funkcji aktywacji ReLu i algorytmu uczenia Adam

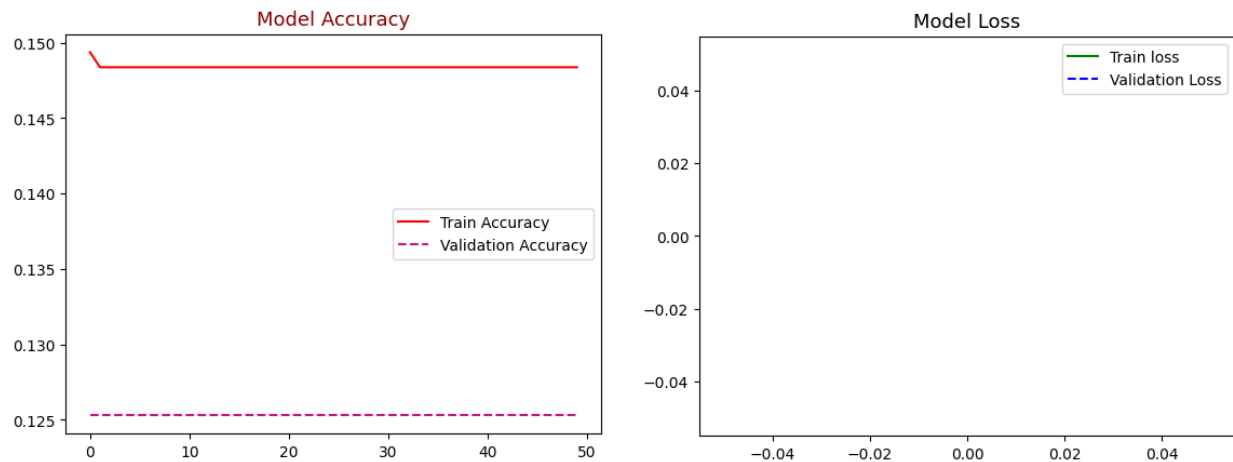
Listing 4.3: Wyniki uzyskane dla funkcji aktywacji ReLu i algorytmu uczenia Adam

	precision	recall	f1-score	support
Normal	0.52	0.38	0.44	148
Cataract	0.93	0.91	0.92	113
Diabetes	0.69	0.84	0.76	373
Glaucoma	0.76	0.71	0.74	115
Hypertension	0.70	0.58	0.63	83
Myopia	0.87	0.89	0.88	94
ARMD	0.72	0.71	0.72	122
Other	0.88	0.63	0.74	68
accuracy			0.73	1116
macro avg	0.76	0.71	0.73	1116
weighted avg	0.73	0.73	0.73	1116

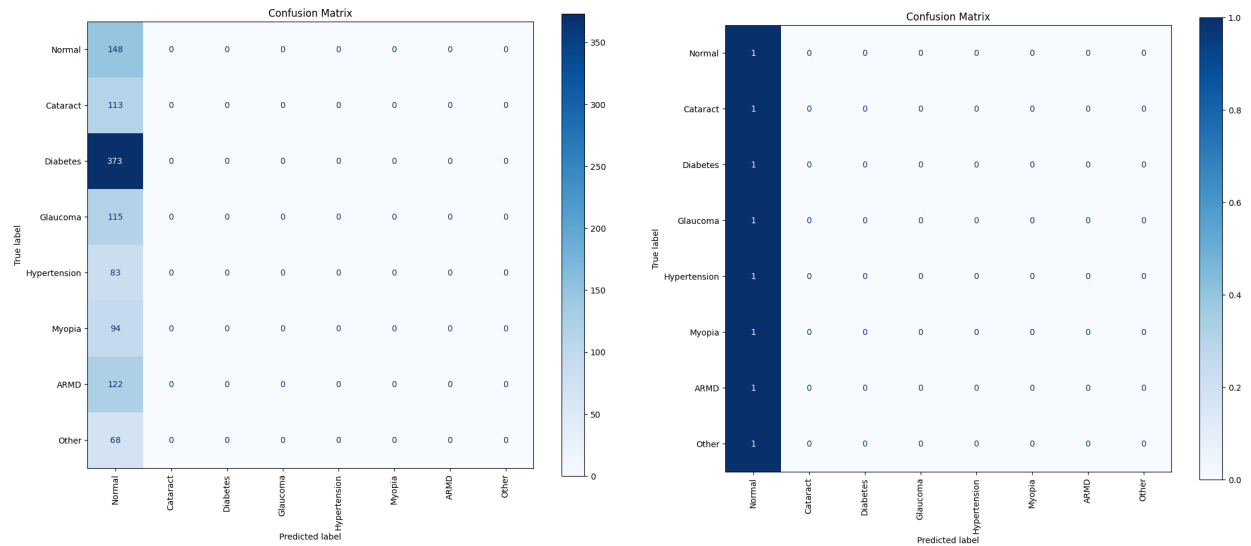
Suma f1 score: 5.817949757372596

Srednia f1 score: 0.7272437196715745

4.4.2. ReLu oraz SGD



Rysunek 4.7: Wykresy dokładności i uzyskane dla funkcji aktywacji ReLu i algorytmu uczenia SGD



Rysunek 4.8: Macierz pomyłek uzyskana dla funkcji aktywacji ReLu i algorytmu uczenia SGD

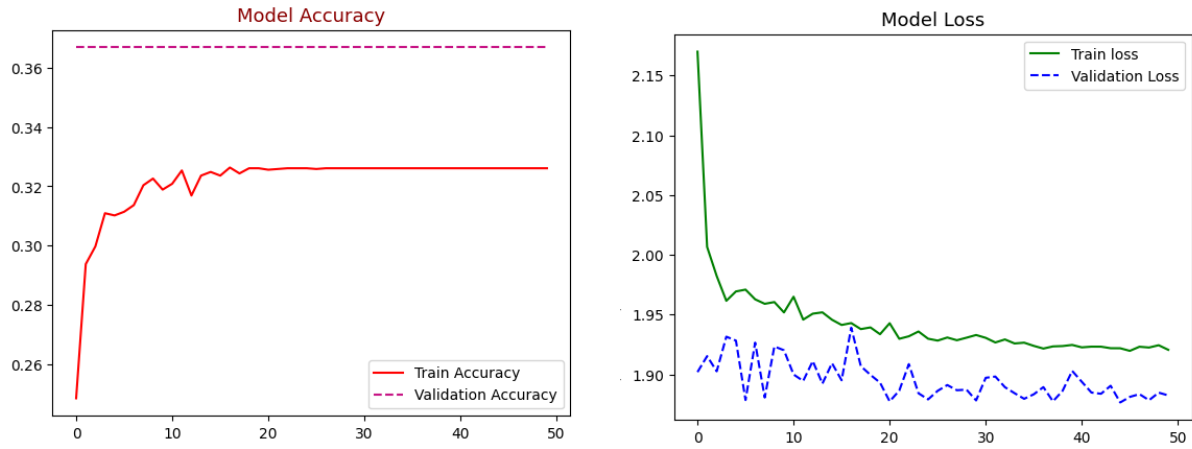
Listing 4.4: Wyniki uzyskane dla funkcji aktywacji ReLu i algorytmu uczenia SGD

	precision	recall	f1-score	support
Normal	0.13	1.00	0.23	148
Cataract	0.00	0.00	0.00	113
Diabetes	0.00	0.00	0.00	373
Glaucoma	0.00	0.00	0.00	115
Hypertension	0.00	0.00	0.00	83
Myopia	0.00	0.00	0.00	94
ARMD	0.00	0.00	0.00	122
Other	0.00	0.00	0.00	68
accuracy			0.13	1116
macro avg	0.02	0.12	0.03	1116
weighted avg	0.02	0.13	0.03	1116

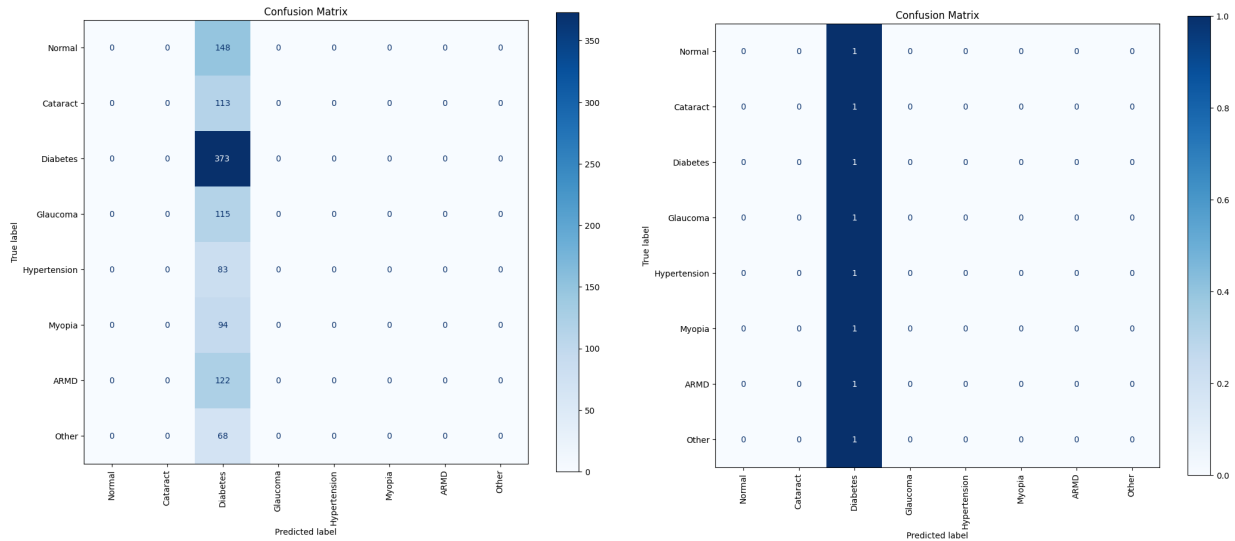
Suma f1 score: 0.23417721518987342

Srednia f1 score: 0.029272151898734177

4.4.3. Tangens hiperboliczny oraz Adam



Rysunek 4.9: Wykresy dokładności i uzyskane dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia Adam



Rysunek 4.10: Macierz pomyłek uzyskana dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia Adam

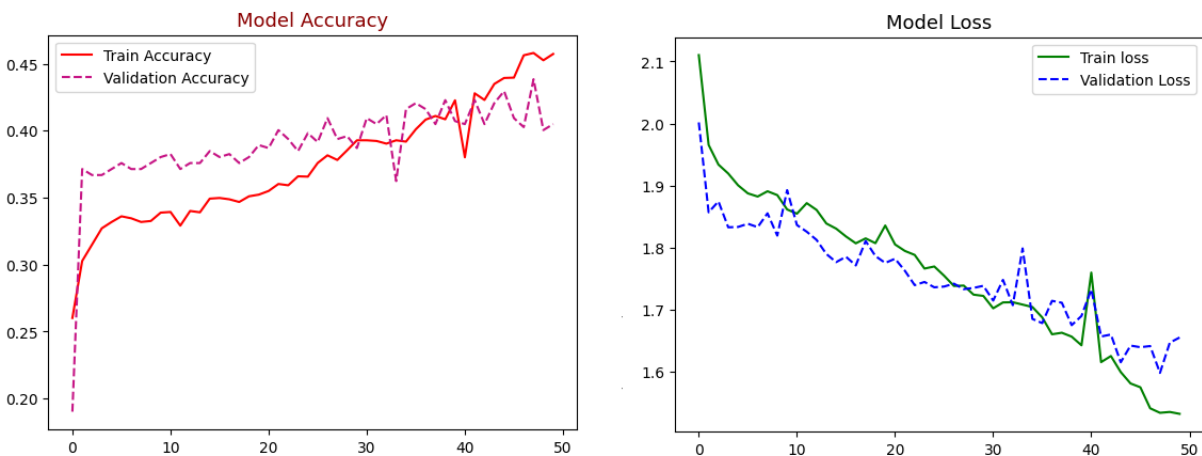
Listing 4.5: Wyniki uzyskane dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia Adam

	precision	recall	f1-score	support
Normal	0.00	0.00	0.00	148
Cataract	0.00	0.00	0.00	113
Diabetes	0.33	1.00	0.50	373
Glaucoma	0.00	0.00	0.00	115
Hypertension	0.00	0.00	0.00	83
Myopia	0.00	0.00	0.00	94
ARMD	0.00	0.00	0.00	122
Other	0.00	0.00	0.00	68
accuracy			0.33	1116
macro avg	0.04	0.12	0.06	1116
weighted avg	0.11	0.33	0.17	1116

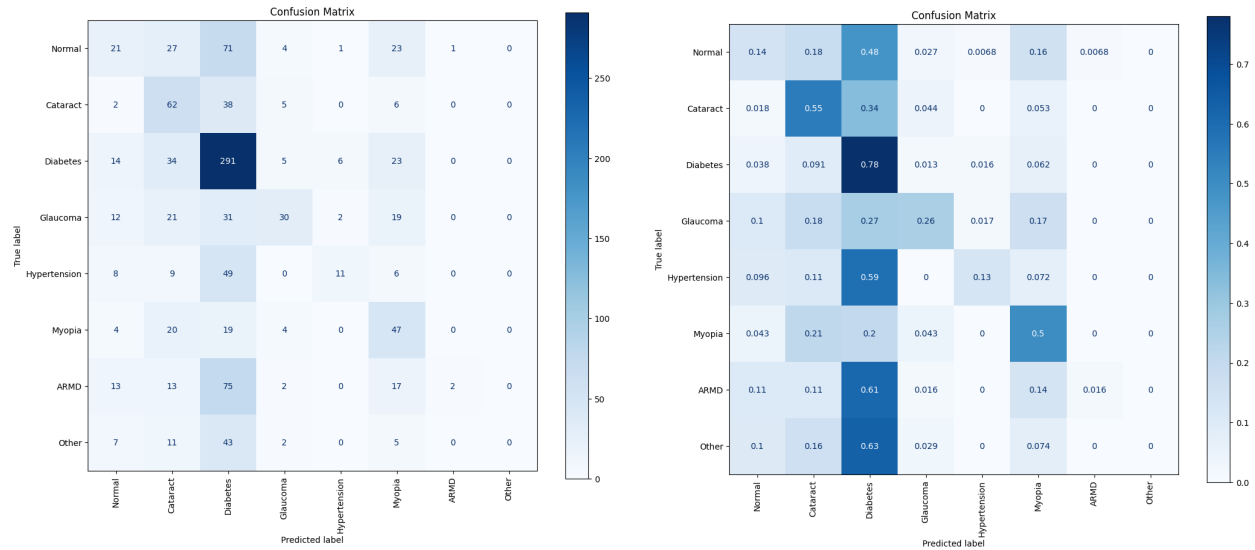
Suma f1 score: 0.5010073875083949

Srednia f1 score: 0.06262592343854936

4.4.4. Tangens hiperboliczny oraz SGD



Rysunek 4.11: Wykresy dokładności i uzyskane dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia SGD



Rysunek 4.12: Macierz pomyłek uzyskana dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia SGD

Listing 4.6: Wyniki uzyskane dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia SGD

	precision	recall	f1-score	support
Normal	0.26	0.14	0.18	148
Cataract	0.31	0.55	0.40	113
Diabetes	0.47	0.78	0.59	373
Glaucoma	0.58	0.26	0.36	115
Hypertension	0.55	0.13	0.21	83
Myopia	0.32	0.50	0.39	94
ARMD	0.67	0.02	0.03	122
Other	0.00	0.00	0.00	68
accuracy			0.42	1116
macro avg	0.40	0.30	0.27	1116
weighted avg	0.42	0.42	0.35	1116

Suma f1 score: 2.1678252382180294

Srednia f1 score: 0.2709781547772537

5. Wnioski i plany na przyszłość

5.1. Wnioski

- Wraz ze wzrostem ilości epok uzyskano lepsze wyniki.
- Stwierdzono, że funkcja aktywacji ReLU wykazuje lepsze wyniki pod względem szybkości uczenia w porównaniu do funkcji tanh. Model z ReLU może być bardziej skuteczny w rozpoznawaniu wzorców i dostosowywaniu wag, co przekłada się na efektywność uczenia.
- Funkcja tanh, charakteryzująca się symetrią względem zera i zakresem wartości od -1 do 1, może być korzystna w pewnych przypadkach, zwłaszcza gdy dane są scentrowane wokół zera. Jednak ReLU, pomimo braku symetrii, może być skuteczny w praktyce, szczególnie dla dużych zbiorów danych.
- Algorytm Adam wydaje się być bardziej zaawansowanym i często skutecznym wyborem dla wielu problemów uczenia maszynowego. Działa on z dynamicznie dostosowującym się współczynnikiem uczenia, co pozwala na bardziej efektywne dostosowanie modelu w trakcie uczenia. A SGD (Stochastic Gradient Descent) nadal może być użyteczny, zwłaszcza w prostszych zadaniach. Jednak Adam oferuje pewne ulepszenia, takie jak adaptacyjne dostosowywanie współczynnika uczenia, co sprawia, że może lepiej radzić sobie z różnymi rodzajami danych.

5.2. Plany na przyszłość

W kontekście planów na przyszłość, można dokonać badań odnośnie porównania większej ilości funkcji aktywacji (np. ELU, SELU) oraz odnośnie różnych algorytmów uczenia (np. SGD, RMSProp, Adamax). Można także skoncentrować się na temacie uczenia transferowego. Kolejnym z planów może być także korzystanie z większych zestawów danych, z większą ilością chorób oraz stworzyć inne sieci neuronowe.

Można także zrobić analizę porównawczą gotowych architektur sieci z bibliotek.

Bibliografia

- [1] S. Al-Fahdawi, A. S. Al-Waisy, D. Q. Zeebaree, R. Qahwaji, H. Natiq, M. A. Mohammed, J. Nedoma, R. Martinek, and M. Deveci. Fundus-deepnet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. *Information Fusion*, 102:102059, 2024.
- [2] F. Chollet. *Deep learning with Python*. Simon and Schuster, 2021.
- [3] I. Hawryluk, T. A. Mellan, H. Hoeltgebaum, S. Mishra, R. P. Schnekenberg, C. Whittaker, H. Zhu, A. Gandy, C. A. Donnelly, S. Flaxman, et al. Inference of covid-19 epidemiological distributions from brazilian hospital data. *Journal of the Royal Society Interface*, 17(172):20200596, 2020.
- [4] J. He, C. Li, J. Ye, Y. Qiao, and L. Gu. Self-speculation of clinical features based on knowledge distillation for accurate ocular disease classification. *Biomedical Signal Processing and Control*, 67:102491, 2021.
- [5] T. Kalaiselvi, S. Padmapriya, K. Somasundaram, and S. Praveenkumar. E-tanh: a novel activation function for image processing neural network models. *Neural Computing and Applications*, 34(19):16563–16575, 2022.
- [6] A. I. Poon and J. J. Sung. Opening the black box of ai-medicine. *Journal of Gastroenterology and Hepatology*, 36(3):581–584, 2021.

Spis rysunków

3.1	Przykładowe zdjęcie oka ze zbioru - zdrowe oko prawe.	7
4.1	Wykresy dokładności i straty dla 30 epok	10
4.2	Wykresy dokładności i straty dla 50 epok	10
4.3	Macierz pomyłek dla 30 epok	11
4.4	Macierz pomyłek dla 50 epok	11
4.5	Wykresy dokładności i uzyskane dla funkcji aktywacji ReLu i algorytmu uczenia Adam	14
4.6	Macierz pomyłek uzyskana dla funkcji aktywacji ReLu i algorytmu uczenia Adam	14
4.7	Wykresy dokładności i uzyskane dla funkcji aktywacji ReLu i algorytmu uczenia SGD	15
4.8	Macierz pomyłek uzyskana dla funkcji aktywacji ReLu i algorytmu uczenia SGD	16
4.9	Wykresy dokładności i uzyskane dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia Adam	17
4.10	Macierz pomyłek uzyskana dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia Adam	17
4.11	Wykresy dokładności i uzyskane dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia SGD	18
4.12	Macierz pomyłek uzyskana dla funkcji aktywacji Tangens hiperboliczny i algorytmu uczenia SGD	19