Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej Nr 60

XXVIII Seminarium

ZASTOSOWANIE KOMPUTERÓW W NAUCE I TECHNICE 2018

Oddział Gdański PTETiS

doi: 10.32016/1.60.07

DIAGNOZOWANIE STANU RETINOPATII CUKRZYCOWEJ PRZY POMOCY GŁĘBOKICH SIECI NEURONOWYCH

Bartłomiej JARZEMBIŃSKI[1], Arkadiusz KWASIGROCH[2], Michał GROCHOWSKI[3]

1. Politechnika Gdańska, Wydział Elektrotechniki i Automatyki
   e-mail: bjarzembinski@yahoo.pl
2. Politechnika Gdańska, Wydział Elektrotechniki i Automatyki
   e-mail: arkadiusz.kwasigroch@pg.edu.pl
3. Politechnika Gdańska, Wydział Elektrotechniki i Automatyki
   e-mail: michal.grochowski@pg.edu.pl

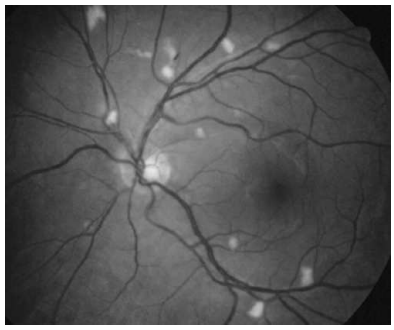
Streszczenie: W referacie opisano problem wykrywania oraz klasyfikacji stanu retinopatii cukrzycowej ze zdjęć dna oka przy pomocy głębokich sieci neuronowych. Retinopatia cukrzycowa jest chorobą oczu często występującą u osób z cukrzycą. Nieleczona prowadzi do uszkodzenia wzroku, a nawet ślepoty. W pracy badawczej opracowano system wykrywania retinopatii cukrzycowej na podstawie zdjęć dna oka. Opracowana sieć neuronowa przypisuje stan choroby w 5 stopniowej skali – od braku choroby do najbardziej zaawansowanego stanu choroby. Zaproponowano specjalny system kodowania klas w celu uchwycenia wielkości różnicy pomiędzy rzeczywistymi a predykowanymi stanami choroby. Uzyskano wysokie wyniki klasyfikacji na zbiorze testowym. W celu oceny skuteczności działania systemu wykorzystano miary statystyczne takie jak ważona Kappa i dokładność.

Słowa kluczowe: głębokie uczenie, sieci neuronowe, retinopatia cukrzycowa.

1. WSTĘP

Retinopatia cukrzycowa jest chorobą oczu związaną z występowaniem u pacjenta cukrzycy. Nawet do 45% pacjentów zdiagnozowanych z cukrzycą posiada pewien stan retinopatii cukrzycowej. Światowa Organizacja Cukrzycy przewiduje wzrost zachorowań na cukrzycę w następnych latach, efektem tego będzie wzrost występowania retinopatii cukrzycowej w populacji. Retinopatia cukrzycowa prowadzi do uszkodzenia wzroku, a nawet ślepoty. Obecnie jest jedną z najczęstszych przyczyn utraty wzroku.

Obecnie choroba diagnozowana jest na podstawie analizy zdjęć dna oka (rys 1.). Jest to czasochłonny proces wykonywany przez lekarza okulistę. Choroba stanowi poważny problem w rejonach gdzie dostęp do specjalistów jest utrudniony. Wczesne wykrycie choroby znacznie zmniejsza ryzyko utraty wzroku. Wymienione czynniki motywują do badań nad opracowaniem komputerowego systemu wspomagania analizy zdjęć dna oka.



Rys. 1. Zdjęcie dna oka

Proponowanym narzędziem do analizy zdjęć dna oka są głębokie konwolucyjne sieci neuronowe (CNN). Ich dynamiczny rozwój w ostatnich latach spowodowany był pojawieniem się dużych zbiorów danych oraz znacznym wzrostem mocy obliczeniowej. Metody głębokiego uczenia wywodzą się z klasycznych sieci neuronowych. Współczesne konwolucyjne sieci neuronowe, które używane są do rozwiązywania różnorodnych zadań, posiadają wiele warstw oraz miliony parametrów.

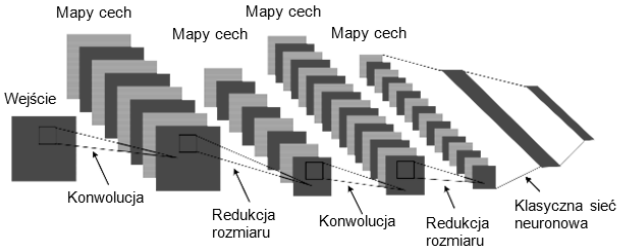
Prawie każdy system klasyfikacji obrazu można podzielić na 2 moduły – moduł ekstrakcji cech oraz moduł klasyfikacji. Moduł ekstrakcji cech wydobywa ze zdjęcia najważniejsze cechy, które następnie przekazywane są na wybrany klasyfikator (np. algorytm regresji liniowej, SVM, sieci neuronowe). Przekazanie obrazu bezpośrednio na klasyfikator nie przynosi dobrych efektów ze względu na duży rozmiar wektorów wejściowych np. małe zdjęcie 200x200 pikseli zawiera aż 40 tys. pikseli. Z tego powodu istotna jest ekstrakcja cech, która zmniejsza ilość cech (wejść do sieci) podawanych na klasyfikator, dodatkowo cechy te znacznie lepiej opisują obraz. Klasyczne systemy klasyfikacji obrazu wymagają ręcznego przygotowania modułu ekstrakcji cech. Należy wybrać odpowiednie algorytmy oraz ich parametry. Zaletą głębokich sieci neuronowych oraz przyczyną ich sukcesu była możliwość automatycznej ekstrakcji cech. Z tego powodu głębokie sieci neuronowe są w stanie przetwarzać surowe obrazy bez konieczności ich intensywnego przetwarzania. Ponadto moduł ekstrakcji cech oraz klasyfikacji są jednym systemem, który uczony jest podczas procesu optymalizacji sieci

neuronowej. W klasycznych rozwiązaniach oba moduły są zwykle oddzielnym zbiorem algorytmów.

W publikacji zaprezentowano automatyczny, wspomagany komputerowo system wykrywania oraz przypisywania stopnia retinopatii cukrzycowej na podstawie zdjęć dna oka. Do uczenia systemu wykorzystano bazę zdjęć dna oka składająca się z ponad 80 tys. obrazów. Działanie systemu sprawdzono na dwóch problemach: wykrywanie retinopatii cukrzycowej (problem binarny) oraz przypisywanie zaawansowania retinopatii cukrzycowej (problem wieloklasowy).

2. GŁĘBOKIE KONWOLUCYJNE SIECI NEURONOWE

Głębokie konwolucyjne sieci neuronowe składają się z wielu warstw. Najczęściej wykorzystywanymi warstwami są: warstwa konwolucyjna, warstwa aktywacji oraz warstwa redukująca rozmiar. Warstwy te budują moduł ekstrakcji cech. Moduł klasyfikacji zbudowany jest natomiast z klasycznych warstw sieci neuronowej.



Rys. 2. Konwolucyjna sieć neuronowa [1]

Warstwy konwolucyjne przeprowadzają wielokrotną operację filtrowania obrazu wejściowego, tworząc tzw. mapy cech. Następnie mapy cech poddawane są funkcji aktywacji, która wprowadza nieliniowość w sieci neuronowej, zwiększając możliwość reprezentacji sieci. Najczęściej wykorzystywaną funkcją aktywacji jest funkcja ReLU (Rectified Linear Unit) opisana wzorem $f(x)=\max(x,0)$. Funkcja sprowadza wszystkie wartości pikseli mniejszych od zera do zera. W literaturze, warstwa aktywacji najczęściej traktowana jest jako część składowa warstwy konwolucyjnej i nie uwzględnia się jej w tabelach opisujących architekturę. Pomiędzy niektórymi warstwami konwolucyjnymi wprowadzana jest warstwa redukująca rozmiar obrazu. Redukcja rozmiaru obrazu prowadzi do lepszych możliwości generalizacji sieci neuronowej, ponadto znacząco redukuje czas uczenia sieci neuronowej.

Kolejną warstwą często stosowaną w głębokich sieciach neuronowych jest warstwa dropout. Warstwa znacznie poprawia możliwości generalizacji sieci neuronowej. Podczas treningu warstwa usuwa losowo część z neuronów na czas kilku przykładów uczących, następnie aktualizuje wagi pozostałych neuronów. Po aktualizacji wag neurony są przywracane i proces się powtarza. Warstwy dropout działają tylko podczas uczenia sieci neuronowej, w fazie testowej natomiast wszystkie neurony są przywracane, a ich wagi przeskalowane.

3. ZBIÓR DANYCH

Podczas opracowywania systemu skorzystano z reprezentatywnej bazy danych udostępnionej przez organizację EyePACS [2]. Zbiór składa się z około 80 tys. zdjęć wysokiej rozdzielczości, podzielonych na 5 klas odpowiadającym różnemu stopniowi zaawansowania choroby. Każde ze zdjęć ocenione zostało przez kwalifikowanego okulistę. Dostarczone zdjęcia wykonane zostały różnym sprzętem oraz w różnych warunkach. Ponadto część zdjęć zawiera artefakty kamery, które przypominają symptomy pierwszego stopnia retinopatii cukrzycowej. Kolejnym problemem jest silne niezbilansowanie klas, które znacznie utrudnia uczenie głębokich sieci neuronowych. Na przykład: pierwsze klasa zawiera aż 60 tys. zdjęć, natomiast piąta klasa reprezentująca najwyższy stopień choroby zawiera około 2 tys. zdjęć.

4. ARCHITEKTURA

W celu realizacji zadania klasyfikacji zaimplementowano sieć VGG16 [3], która była wcześniej wytrenowana na zbiorze Imagenet. Struktura architektury przedstawiona jest w tabeli 1.

Tabela 1. Struktura sieci neuronowej

VGG-D
Warstwa wejściowa (224x224x3)
Konwolucja (3x3, 64) x2
MaxPooling
Konwolucja (3x3, 128) x2
Maxpooling
Konwolucja (3x3, 256) x3
MaxPooling
Konwolucja (3x3, 512) x3
MaxPooling
Konwolucja (3x3, 512) x3
MaxPooling
4048 KN Dropout x2
Wyjście (4 neurony sigmoidalne)

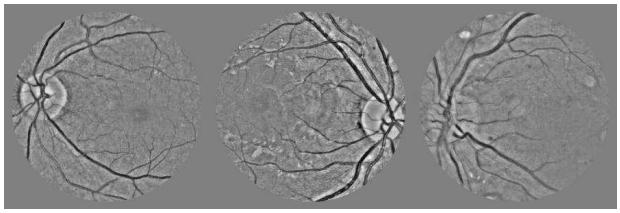
Struktura przyjmuje na wejście kolorowe zdjęcia. Wszystkie użyte filtry konwolucyjne posiadają rozmiar 3x3. Ilość filtrów w warstwach wynosi od 64 do 512. Warstwa MaxPooling zmniejsza rozmiar map cech. Ostatnie warstwy głębokiej sieci neuronowej składają się z klasycznych warstw w których skład wchodzą klasyczne neurony (KN). Cztery neurony sigmoidalne zostały wykorzystane w celu zaproponowanego zakodowania 5 klas.

5. PROCES UCZENIA

5.1. Wstępne przetwarzanie danych

Rozmiar zdjęć został przeskalowany tak aby dopasować go do założonego rozmiaru wejścia sieci. Dodatkowo zdjęcia zostały tak skadrowane, aby widoczne na nich oczy miały równy promień na wszystkich zdjęciach. Następnie przeprowadzono normalizacje danych w celu osiągnięcia średniej zbioru równą zero oraz odchylenia standardowego równego jeden. Kolejnym krokiem było odjęcie lokalnego koloru średniego (rys. 3). Operacje te umożliwiły wyszczególnienie istotnych informacji w zdjęciu, które są najbardziej przydatne w procesie klasyfikacji.

MOST WIEDZY Pobrano z mostwiedzy.pl



Rys. 3. Zdjęcia po odjęciu lokalnego koloru

W celu zwiększenia możliwości klasyfikacji sieci neuronowych przeprowadzono rozszerzanie ilości danych polegające na odbiciach zdjęć, przybliżaniu, rozciąganiu itp.

5.2. Specjalne kodowanie klas

Zadanie klasyfikacji zdjęć dna oka jest zadaniem wieloklasowym. Nie jest to jednak klasyczna klasyfikacja wieloklasowa, ze względu na uwzględnianie różnicy pomiędzy rzeczywistym a predykowanym stanem choroby. Jest to szczególnie istotne w zastosowaniach medycznych, gdzie niewykrycie choroby może skutkować utratą zdrowia pacjenta. W celu uchwycenia tej informacji zastosowano specjalne kodowanie klas. Na przykład dla problemu 4-klasowego kolejne wektory kodujące wyglądają następująco: [0 0 0], [1 0 0], [1 1 0], [1 1 1]. Wykorzystanie takiego kodowania umożliwia określenie różnicy pomiędzy klasami, ponadto w łatwy sposób pozwala na włączenie tej informacji w algorytm optymalizacyjny. Ponadto zastosowano klasyczne kodowanie „1 z n" w celu badania efektywności zaproponowanego kodowania.

5.3. Uczenie sieci neuronowej

Głęboka sieć neuronowa uczona jest przy pomocy algorytmu SGD (Stochastic Gradient Descent), który minimalizuje funkcję celu wykorzystując informację o gradientach względem parametrów sieci neuronowej. Ze względu na wykorzystanie specjalnego kodowania, zmodyfikowana została funkcja celu. Funkcja celu dla pojedynczego przykładu uczącego jest sumą entropii krzyżowych pomiędzy odpowiadającymi sobie wartościami wektorów rzeczywistych oraz predykowanych.

5.4. Sprzęt i oprogramowanie

Ze względu na wysokie koszty obliczeniowe głębokich sieci neuronowych, uczenie sieci wymaga wydajnego sprzętu wyposażonego w jednostki GPU. W celu przeprowadzenia uczenia sieci neuronowej wykorzystano jednostkę komputerową wyposażoną w następujące komponenty: karta graficzna GeForce GTX 770 z pamięcią 2 GB, procesor Intel Core i7-5820K, pamięć RAM 64 GB.

System opracowano wykorzystując język Python. Skorzystano z bibliotek TensorFlow, która umożliwia budowanie grafów obliczeniowych oraz automatyczne różniczkowanie. Ponadto użyto biblioteki Keras, która pozwala na szybkie prototypowanie architektur sieci neuronowych.

6. WYNIKI

6.1. Wykorzystane miary statystyczne

W celu oceny jakości działania stworzonego systemu wykorzystano dwie miary statystyczne – dokładność oraz współczynnik Kappa. Dokładność jest stosunkiem ilości poprawnie sklasyfikowanych zdjęć do ilości wszystkich zdjęć. Współczynnik Kappa jest miarą zgodności pomiędzy dwiema ocenami. Jej wartość waha się od -1 do 1, gdzie 1 oznacza całkowitą zgodność ocen. Współczynnik dostarcza znacznie więcej informacji na temat systemu niż dokładność. Na przykład, w przypadku dwóch lekarzy: jeden z nich ocenia każdą chorobę myląc się o jeden stopień (np. brak choroby ocenia jako słaby objaw retinopatii cukrzycowej), a drugi lekarz myli się o znacznie wyższy stopień (np. najcięższy stopień retinopatii ocenia jako brak choroby). Obaj lekarze będą mieli zerową dokładność ponieważ każdy z nich się myli. Natomiast współczynnik kappy pierwszego lekarza będzie znacznie wyższy ze względu na to, że popełniał mniejsze błędy.

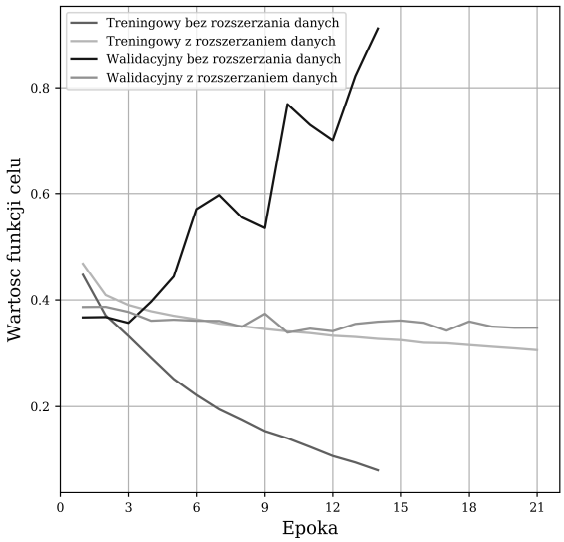
Dodatkowo wprowadzono zadanie klasyfikacji dwuklasowej, w której algorytm rozpoznawał brak choroby lub jej obecność. W celu oceny tego systemu wprowadzono dwie miary: czułość oraz swoistość. Czułość jest miarą poprawnie sklasyfikowanych przykładów pozytywnych (obecność choroby) do całkowitej ilości przykładów pozytywnych w zbiorze. Natomiast swoistość jest analogiczną miara, w której pod uwagę bierze się przykłady negatywne (brak choroby). Wysoka czułość jest pożądana wśród lekarzy jak i w systemach automatycznego wykrywania chorób, ze względu na to, że wiele niewykrytych chorób rozwija się komplikując przyszły proces leczenia.

6.2. Porównanie systemu kodowania klas

Sprawdzono działanie zaproponowanego kodowania klas z klasycznym podejściem „1 z n" - najczęściej wykorzystywanym w zadaniach wieloklasowych. Dzięki wprowadzeniu tej funkcjonalności uzyskano wzrost dokładności klasyfikacji z 43,3% do 50,8% oraz wzrost współczynnika Kappa z 0,691 do 0,776, w porównaniu do klasycznego kodowania „1 z n".

6.3. Wpływ algorytmów poszerzania danych

Przenalizowano wpływ rozszerzania danych na proces uczenia. Krzywe uczenia widoczne są na rys. 4.



Rys. 4. Wpływ rozszerzania danych na proces uczenia

Uczenie sieci bez rozszerzenia danych bardzo szybko prowadzi do przeuczenia sieci. Wprowadzenie rozszerzania danych znacznie pogorszyło wyniki uczenia zbioru treningowego, ze względu na to, że algorytmowi znacznie trudniej nauczyć się danych. Natomiast znacząco poprawiły się wyniki klasyfikacji zbioru walidacyjnego. Sieć nie przetrenowuje się, z epoki na epokę sieć zdobywa coraz lepsze możliwości generalizacji.

MOST WIEDZY Pobrano z mostwiedzy.pl

6.4. Wpływ uczenia transferowanego

Zastosowanie uczenia transferowanego znacznie poprawiło możliwości klasyfikacji głębokiej sieci neuronowej oraz przyspieszyło uczenie. Sieć zainicjalizowana losowymi wagami wymagała aż 128 epok uczących w celu osiągnięcia najwyższego wyniku klasyfikacji, czas treningu wynosił 5,5 doby. Zastosowanie uczenia transferowanego zmniejszyło liczbę potrzebnych epok do 34, a czas treningu do 1.5 doby. Wykorzystanie wcześniej zdobytej wiedzy znacząco redukuje wymagany czas treningu oraz ma pozytywny wpływ na możliwości klasyfikacji sieci.

6.5. Wyniki

Najlepsze wyniki uzyskano przy użyciu zmodyfikowanej przez nas sieci VGG. Zastosowano rozszerzanie danych, uczenie transferowane oraz specjalne kodowanie klas. Zaproponowany system uzyskał wartość współczynnika kwadratowej kappy równą 0,776 oraz dokładność równą 50,8%. Uzyskana dokładność jest całkiem wysoka ze względu na 5 klas – im większa ilość klas tym trudniej uzyskać odpowiednią dokładność. Ponadto zagadnienia medyczne charakteryzują się rozmytymi granicami pomiędzy poszczególnymi klasami, różni lekarze mogą różnie ocenić dany stan choroby. Macierz błędów popełnianych przez sieć widoczna jest w tabeli 2. Ocena 1 oznacza brak choroby, kolejne oceny są etapami choroby z rosnącym stopniem zaawansowania.

Tabela 2. Macierz błędów

Kategorie	1 od CNN	2 od CNN	3 od CNN	4 od CNN	5 od CNN
1 od okulisty	101	83	12	2	2
2 od okulisty	64	111	20	4	1
3 od okulisty	16	64	46	59	15
4 od okulisty	2	14	23	141	20
5 od okulisty	2	7	13	69	109

Kolejnym krokiem było sprawdzenie w jaki sposób system klasyfikuje obrazy w zadaniu klasyfikacji 2 klasowej. Przyjęto dwie klasy: brak choroby oraz obecność choroby – w skład tej klasy wchodziły wszystkie przypadki z dowolnym etapem zaawansowania choroby. System taki uzyskał dokładność 81,7%. Czułość badanego systemu wynosiła 89,5%, a swoistość 50,5%. Wysoka czułość pożądana jest w systemach klasyfikacji chorób. Poprawne wykrycie retinopatii cukrzycowej zmniejsza ryzyko uszkodzenia wzroku i ślepoty.

7. PODSUMOWANIE

W referacie zaproponowano system automatycznej oceny stopnia retinopatii cukrzycowej na podstawie kolorowego zdjęcia dna oka. System oparto o głębokie, konwolucyjne sieci neuronowe. Udało się uzyskać wysoki wynik klasyfikacji oraz wysoką wartość współczynnika Kappa mierzącego zgodność pomiędzy ocenami lekarzy a ocenami sieci neuronowych. Warstwy konwolucyjne nabywają podczas treningu umiejętność ekstrakcji cech. Możliwość ekstrakcji cech przez głębokie sieci neuronowe pozwoliła na uczenie, a następnie klasyfikację obrazów o niskim stopniu wstępnego przetwarzania danych. Rozszerzanie danych oraz uczenie transferowane znacznie poprawiło jakość klasyfikacji głębokich sieci neuronowych.

Zaproponowany specjalny system kodowania klas umożliwił włączenie informacji o wielkości różnicy pomiędzy rzeczywistym a predykowanym stanem choroby. Użycie kodowania powiększyło współczynnik zgodności Kappa.

Uzyskane wyniki zachęcają do dalszych badań nad zastosowaniem sieci neuronowych do analizy obrazów medycznych. Opracowane aplikacje mogą służyć jako narzędzie wspomagające pracę lekarza, redukując czas potrzebny na manualną analizę obrazów.

Wzrost wydajności jednostek obliczeniowych oraz wzrost ilości danych w przyszłości będzie prowadziło do coraz lepszych wyników klasyfikacji.

8. UWAGI

Praca naukowa finansowana ze środków budżetowych na naukę w latach **2017-2021**, jako projekt badawczy w ramach programu „Diamentowy Grant".

9. BIBLIOGRAFIA

1. Y. LeCun, K. Kavukcuoglu, i C. Farabet, „Convolutional networks and applications in vision.", w *ISCAS*, 2010, t. 2010, s. 253–256.
2. „Diabetic Retinopathy Screening - EyePACS, LLC". [Online]. Dostępne na: http://www.eyepacs.com/. [Udostępniono: 01-paź-2017].
3. K. Simonyan i A. Zisserman, „Very deep convolutional networks for large-scale image recognition", *ArXiv Prepr. ArXiv14091556*, 2014.

CLASSIFICATION OF THE STAGE OF THE DISEASE BY DEEP NEURAL NETWORKS

In the paper we described computer aided detection system of diabetic retinopathy based on fundus photos of retina. Diabetic retinopathy is an eye disease associated with diabetes. Non-treated diabetic retinopathy leads to sight degeneration and even blindness. Early detection is crucial due to provide effective treatment. Currently, diabetic retinopathy detection is time consuming process, done manualy by medical specialist. The disease is dangerous issue in places where the availability of phisicians is limited. We employed the computer system that detect diabetic retinopathy and assess a stage of the disease based on retinal photo of fundus. We used one of the best image classification system – deep neural networks. Employed system assess the stage of the disease in 5 level scale – from absence of disease to the most severe stage of disease. We employed transfer learning and data augmentation to enhance classification result. Moreover we proposed special class coding system to catch the difference between real and predicted stage of disease. We tested employed system using different statistical measures like accuracy, sensitivity, specificity and Kappa score.

Keywords: deep learning, neural networks, diabetic retinopathy.

MOST WIEDZY Pobrano z mostwiedzy.pl