

Częstochowa, dn. 16 marca 2021 r.

Dr hab. inż. Marcin Korytkowski, prof. Politechniki Częstochowskiej
Katedra Inteligentnych Systemów Informatycznych
Wydział Inżynierii Mechanicznej i Informatyki
Politechnika Częstochowska
al. Armii Krajowej 36
42-200 Częstochowa

Recenzja

rozprawy doktorskiej mgr inż. Dariusza Sychela pt.: Uczenie kaskady klasyfikatorów w celu redukcji oczekiwanej liczby wyznaczonych cech w zadaniach detekcji obiektów na obrazach

Promotor: dr hab. inż. Przemysław Klęsk, prof. ZUT

Niniejszą recenzję opracowano na zlecenie Prorektora ds. Nauki, Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie z dnia 26.01.2021 r.

1. Charakterystyka tematu, celu i tezy badawczej rozprawy

Wyszukiwanie obiektów na obrazach jest niezwykle ważnym obszarem zagadnień związanych z przetwarzaniem obrazów i uczeniem maszynowym. Zastosowanie wyników badań z tej tematyki jest niezwykle szerokie od diagnostyki medycznej aż po zwykłe rozpoznawanie i klasyfikowanie obiektów. Na podstawie analizy przeprowadzonych badań symulacyjnych, w recenzowanej pracy, doktorant skupił się na potwierdzeniu osiągniętych rezultatów w głównej mierze na przykładach związanych wyszukiwaniem twarzy, choć przedstawiono także symulacje z zakresu rozpoznawania syntetycznych liter.

Cel i tezę pracy sformułowano w następujący sposób:

Cel pracy: Opracowanie algorytmu uczenia kaskady klasyfikatorów redukującego oczekiwaną liczbę ekstrahowanych cech w porównaniu do algorytmu Violi i Jonesa przy zachowaniu końcowych wymagań dla czułości i specyficzności.

Teza: Poluzowanie wymagań cząstkowych z wykorzystaniem aktualizowanych średnich geometrycznych oraz technik przeszukiwania drzewa uczenia umożliwia redukcję wartości oczekiwanej liczby cech ekstrahowanych przez kaskadę klasyfikatorów.

2. Zawartość rozprawy

Recenzowana praca mgr inż. Dariusza Sychela składa się z 5 rozdziałów zasadniczych, streszczenia w języku polskim i angielskim, spisów tablic, rysunków oraz bibliografii. Dokument liczy 86 stron.

Pierwszy rozdział jest wprowadzeniem do tematyki, zawiera cel, tezę oraz strukturę pracy.

Rozdział drugi zawiera opis wraz z interpretacją trzech geometrycznych przykładów:

- a) „Kostka w rogu”
- b) „Kostka 3D”
- c) „Pułapka z szachownicą”

pokazujących, w jaki sposób uczenie kaskady z odgórnie nałożonymi wymaganiami na każdy etap w postaci miary częstości fałszywych alarmów (False Acceptance Rate (FAR)) oraz czułości może prowadzić do nieoptymalności tej metody. Autor wskazał tutaj przykład zbyt wczesnego przerwania budowy kaskady w danym etapie, co skutkować może wzrostem liczby cech w kolejnych krokach. Drugim wnioskiem autora jest stwierdzenie, że istnieje możliwość przerwania budowy kaskady w danym etapie pomimo niespełnienia wymaganej minimalnej czułości (d_{min}) oraz wartości maksymalnej FAR a założone wyniki końcowe nadal będą możliwe do osiągnięcia. Te dwa wnioski autora są głównym motywem do podjęcie dalszych badań.

W rozdziale trzecim zaprezentowano zestaw autorskich metod opracowanych na bazie wniosków przedstawionych w rozdziale drugim. Mają one na celu wprowadzenie nowych (poluzowanych) kryteriów optymalizujących proces budowy kaskady na każdym etapie. Rezultatem tych działań ma być opracowanie technik budujących klasyfikatory, w taki sposób by znacząco redukować wartość oczekiwaną liczby ekstrahowanych cech. Jednocześnie wprowadzenie tych nowych warunków nie wpływa na postawione wymagania końcowe dla całej kaskady w postaci wartości FAR i czułości. Autor wprowadza w tym celu dwa twierdzenia i przeprowadza ich proces dowodowy. W efekcie opracowane zostają trzy nowe metody bazujące na poluzowanych kryteriach, przy czym pierwsze dwie uczące kaskadę z poluzowanymi wymaganiami na etap (UGM oraz UGM-G) opisane pseudokodem (*TrainRelaxedCascade*) oraz jedną bazującą na przeszukiwaniu drzew decyzyjnych w dwóch wariantach: z przycinaniem dokładnym (*TrainTreeCascade*) i przycinaniem przybliżonym (*TrainTreeCascadeApprox*). Do każdej z tych metod autor przedstawił pseudokod oraz przedstawił ich dokładny opis.

Rozdział czwarty zatytułowany „Badania” zawiera opis przeprowadzonych symulacji komputerowych z wykorzystaniem metod wskazanych w poprzednim punkcie. W głównej mierze autor stosuje swoje metody do wykrywania twarzy na obrazach bazując na cechach Harra i HOG. W punkcie 4.3 autor pokazuje zastosowanie swoich metod do wykrywania litery „A” w ciągu treningowym wygenerowanym syntetycznie. W tym eksperymencie użyto cech obrotowo niezmienniczych bazujących na momentach Zernike’a. Na podstawie uzyskanych wyników można uznać, że zaproponowane przez autora modyfikacje w budowie kaskad klasyfikatorów mają pozytywny wpływ na wydajność algorytmów budujących zestawy klasyfikatorów.

Rozdział piąty zawiera podsumowanie całej rozprawy wraz z krótkim wskazaniem kierunku dalszych możliwych badań.

Ogólnie zasadnicze i oryginalne rezultaty pracy można podsumować następująco:

- wyczerpujący opis oraz nowe rezultaty naukowe dotyczące optymalizacji budowy kaskad klasyfikatorów,
- Opracowanie metod, które w procesie budowy kaskad, poza jakością działania, koncentrują się na optymalizacji wydajności kaskady. Zatem doktorant starał się w głównej mierze tak skonstruować swoje metody, by ich działanie prowadziło do redukcji wartości oczekiwanej liczby ekstrahowanych cech
- Wskazanie trzech przypadków geometrycznych uzasadniających powstawanie nieoptymalności w kaskadach budowanych z wykorzystaniem stałych wymagań na dany etap przy użyciu metody Violi i Jonesa.

Wymienione oryginalne metody przedstawione w rozdziałach 2,3,4 zostały opublikowane w artykułach naukowych w materiałach konferencyjnych oraz czasopismach z tzw. Listy Filadelfijskiej. Zaprezentowany materiał pokazuje, że Doktorant udowodnił tezę pracy.

3. Uwagi krytyczne i wskazówki dotyczące rozprawy

Praca jest napisana dobrze, jednak w czasie czytania nasunęły mi się następujące wątpliwości: W zaproponowanych przez autora rozwiązaniach poszczególne kaskady klasyfikatorów budowane są z wykorzystaniem metod z rodziny Boosting. W algorytmach tych każdy kolejny

klasyfikator budujący zespół otrzymuje ocenę proporcjonalną do jakości jego skuteczności (wartość z zakresu $(0,1)$) lub też, tak jak w przypadku metody Real Boost, „waga” ta zaszyta jest w odpowiedzi klasyfikatora. W procesie tworzenia w sposób automatyczny takiego zespołu może się okazać, że jeden z nich dominuje nad pozostałymi członkami zespołu (wartość przypisanej oceny jest większa od sumy wartości ocen pozostałych klasyfikatorów w zespole). W efekcie, istnienie takiego modułu mijają się całkowicie z celem- odpowiedź modułu jest zawsze odpowiedzią najlepszego klasyfikatora. Nie ma zatem sensu wykonywanie obliczeń dla wszystkich modułów. W treści rozprawy nie potrafiłem odnaleźć czy wprowadzono w tym zakresie stosowne ograniczenie.

Autor w badaniach symulacyjnych w Rozdziale 4 pokazujących zasadność zaproponowanych rozwiązań użył zestawów testujących, które w głównej mierze sam tworzył. Trzeba zauważyć, że wskazane liczby używanych w doświadczeniach próbek wskazują na brak ich zbalansowania. Niestety trudno mi jest porównać w tak zdefiniowanym przebiegu doświadczeń uzyskane wyniki z rezultatami dostępnymi w literaturze światowej, chociażby w zakresie uzyskanych wartości jakości i wydajności w problemach związanych z detekcją twarzy na obrazach.

Autor w swojej rozprawie wykonując badania symulacyjne ograniczył się do wyboru metod bazujących na budowie kaskad klasyfikatorów. Czy można porównać dokładność i wydajność zaproponowanych algorytmów z innymi np. wykorzystującymi sieci spłotowe, które obecnie są jednymi z najpopularniejszych metod znanych w literaturze?

Oczywiście powyższe uwagi mogą odnosić się do wielu prac naukowych, gdyż ze względu na ograniczony czas każdego przedsięwzięcia i ogromną liczbę publikowanych prac w nauce światowej, a tym samym nowych metod, nie jest możliwe wyczerpujące przetestowanie konkurencyjnych algorytmów czy zbiorów danych.

Patrząc pod kątem edytorskim można zauważyć kilka błędów, tzw. literówek. Należy także zwrócić uwagę na pisownię „nie” z imiesłowami i przymiotnikami (np. „niekończące się”). Oczywiście celem tak postawionych pytań i uwag nie jest jakiegokolwiek umniejszenie wartości naukowej rozważanej rozprawy doktorskiej.

4. Wnioski końcowe recenzji

Podsumowując recenzję stwierdzam, że Pan mgr inż. Dariusz Sychel w rozprawie doktorskiej „Uczenie kaskady klasyfikatorów w celu redukcji oczekiwanej liczby wyznaczonych cech w zadaniach detekcji obiektów na obrazach” zrealizował cel i zweryfikował tezę postawioną w pracy. Jednocześnie poprzez realizację licznych symulacji komputerowych autor poprał swoją tezę ich wynikami w praktyce.

Recenzowana praca spełnia wymagania ustawy o tytule i stopniach naukowych. Wnoszę o jej przyjęcie i dopuszczenie do publicznej obrony.

