



dr hab. inż. Piotr A. Kowalski, prof. AGH
Katedra Informatyki Stosowanej i Fizyki Komputerowej,
Wydział Fizyki i Informatyki Stosowanej,
Akademia Górniczo-Hutnicza w Krakowie,
al. Mickiewicza 30, 30-059 Kraków
email: pkowal@agh.edu.pl

oraz



Centrum Informatycznych Metod Analizy Danych
Instytut Badań Systemowych
Polskiej Akademii Nauk

Kraków, 1.07.2026

RECENZJA

rozprawy doktorskiej Pani mgr Pauliny Anny Tomaszewskiej
pt. „Role of spatial and temporal context in classification using deep neural networks”

1. Charakterystyka rozprawy

Przedłożona do recenzji rozprawa doktorska mgr Pauliny Anny Tomaszewskiej, zatytułowana „Role of spatial and temporal context in classification using deep neural networks”, została przygotowana w języku angielskim. Praca dotyczy roli informacji kontekstowej — przede wszystkim przestrzennej, a w mniejszym zakresie także temporalnej — w procesie klasyfikacji realizowanej z użyciem głębokich sieci neuronowych. Jest to temat aktualny, ważny poznawczo i praktycznie, mieszczący się na styku uczenia głębokiego, wyjaśnialnej sztucznej inteligencji, cyfrowej histopatologii oraz analizy modeli działających na danych złożonych.

Rozprawa ma przejrzystą strukturę. Otwiera ją wprowadzenie, w którym Autorka formułuje hipotezę badawczą oraz zestaw pytań badawczych. Następnie przedstawiono tło teoretyczne obejmujące zagadnienia ewaluacji modeli, wyjaśnianej sztucznej inteligencji oraz różnych rozumień kontekstu w modelach głębokich. Zasadnicza część rozprawy została zbudowana wokół ośmiu prac naukowych oznaczonych jako P1–P8, uzupełnionych rozdziałem wprowadzającym do cyfrowej patologii, rozdziałem podsumowującym oraz częścią dyskusyjną poświęconą dalszym kierunkom badań. Bibliografia jest obszerna i obejmuje zarówno prace klasyczne, jak i bardzo aktualne pozycje z obszaru XAI, modeli głębokich, histopatologii cyfrowej, modeli Multiple Instance Learning (MIL) oraz interpretowalności modeli.

Na szczególne podkreślenie zasługuje fakt, że rozprawa nie jest jedynie mechanicznym zestawieniem publikacji. Poszczególne rozdziały są poprzedzone komentarzami, które osadzają publikacje w szerszej narracji badawczej. Dzięki temu praca ma wyraźną oś przewodnią: pytanie o to, czy i w jaki sposób modele głębokie wykorzystują kontekst, a także jak można to badać, mierzyć i interpretować.

Oceniam rozprawę bardzo wysoko. Jest to praca dojrzała, ambitna, nowoczesna i dobrze osadzona w aktualnych problemach badawczych. Jej wartość polega nie tylko na zaproponowaniu konkretnych algorytmów, lecz także na zmianie perspektywy: od punktowego wykrywania fałszywych informacji do rozumienia szerszych, powtarzalnych struktur narracyjnych. Taka zmiana jest ważna, ponieważ dezinformacja rzadko funkcjonuje jako pojedyncze, odizolowane zdanie; częściej przyjmuje postać powracających opowieści, które adaptują się do nowych kontekstów, tematów i odbiorców.

2. Teza i cel pracy

Centralna hipoteza badawcza rozprawy została sformułowana następująco: informacja kontekstowa odgrywa istotną rolę w procesie podejmowania decyzji przez klasyfikatory oparte na głębokich sieciach neuronowych. Autorka analizuje tę hipotezę w kilku wymiarach: typu kontekstu, jego skali, obszaru zastosowania oraz etapu pracy z modelem — od projektowania modeli po ich wyjaśnianie i audytowanie.

Uważam, że tak sformułowana hipoteza jest trafna i dobrze dobrana do zakresu pracy. Jest wystarczająco ogólna, aby połączyć różne wątki badawcze, a zarazem wystarczająco konkretna, aby można ją było operacjonalizować poprzez zaproponowane pytania badawcze, eksperymenty, metody i studia przypadków.

3. Streszczenie rozprawy

Rozprawa dotyczy jednego z fundamentalnych problemów współczesnego uczenia głębokiego: modele osiągają coraz lepsze wyniki, lecz nadal często nie wiemy, na czym dokładnie opierają swoje decyzje. Autorka wskazuje, że dotychczasowe metody wyjaśniania sztucznej inteligencji dla obrazów koncentrują się głównie na pytaniu „gdzie” znajdują się istotne fragmenty wejścia albo „co” przedstawiają wskazane obszary. Tymczasem w wielu rzeczywistych zastosowaniach kluczowe jest także pytanie, „jak” elementy obrazu są względem siebie położone. Jest to bardzo ważna intuicja badawcza, ponieważ relacje przestrzenne — odległość, sąsiedztwo, orientacja, kompozycja czy wzajemne położenie struktur — mogą być równie ważne jak obecność pojedynczych obiektów.

W pierwszym nurcie pracy Autorka analizuje kontekst temporalny. Zaproponowana metoda LIMES wykorzystuje informację o historycznych rozkładach klas jako kontekst temporalny wspierający adaptację modelu działającego w warunkach przesunięcia rozkładu klas. W drugim, dominującym nurcie rozprawy, Autorka koncentruje się na kontekście przestrzennym. Najpierw wskazuje ograniczenia popularnych metod XAI, które nie są przystosowane do wyjaśniania decyzji zależnych od relacji przestrzennych. Następnie proponuje i bada rozwiązania pozwalające

analizować rolę kontekstu w cyfrowej histopatologii, w modelach Multiple Instance Learning (MIL) oraz w zadaniach abstrakcyjnego rozumowania wizualnego.

Szczególnie istotny jest blok prac dotyczących obrazów Whole Slide Images (WSI). Autorka bada, czy modele MIL potrafią korzystać z kompozycji obrazu i zależności sąsiedztwa, proponuje zbiory SyntheticSMIL, protokół ReasonSMIL oraz metodę Deep Spatial Context. Przeprowadza również analizę porównawczą modeli trenowanych na całych obrazach tkanek oraz na podzbiorach łątek, w tym regionach zainteresowania wskazanych przez lekarza. Wyniki te prowadzą do ważnej praktycznej obserwacji: ręczny wybór obszarów zainteresowania (ROI) nie musi poprawiać jakości modeli w porównaniu z wykorzystaniem pełnego kontekstu dostępnego w WSI.

4. 4. Najważniejsze osiągnięcia rozprawy

Za najważniejsze osiągnięcia rozprawy uważam:

- Wyraźne postawienie problemu przestrzennego XAI. Autorka przekonująco pokazuje, że klasyczne mapy istotności nie wystarczają, gdy decyzja modelu zależy od relacji pomiędzy obiektami lub regionami obrazu.
- Zaproponowanie zmiany perspektywy z „where” i „what” na „how”. Jest to bardzo wartościowy wkład koncepcyjny: Autorka nie tylko wskazuje lukę, ale proponuje język opisu problemu.
- Opracowanie zbiorów SyntheticSMIL oraz protokołu ReasonSMIL. Są to narzędzia pozwalające badać, czy modele MIL rozumieją kompozycyjność obrazu i czy podobnie zachowują się przy różnych podziałach danych.
- Zaproponowanie metody Deep Spatial Context. Metoda ta wnosi ilościowe, matematycznie umocowane podejście do badania kontekstu przestrzennego w modelach uwagowych, co jest szczególnie cenne w obszarze XAI, gdzie wiele metod ma charakter głównie wizualny lub jakościowy.
- Analiza lokalnego kontekstu w klasyfikacji łątek histopatologicznych. Zaproponowane podejście oparte na ograniczaniu kontekstu poprzez maskowanie pozwala badać stabilność decyzji modelu przy zmniejszaniu dostępnej informacji o otoczeniu tkanki.
- Wartościowe wyniki praktyczne dla cyfrowej histopatologii. W szczególności interesująca jest obserwacja, że modele trenowane na pełnych WSI mogą korzystać z różnorodności i pełniejszego kontekstu w sposób trudny do zastąpienia przez ręcznie wskazane ROI.
- Projekt Duel-IQ. Autorka pokazuje, że idea porównywania odpowiedzi kandydackich w zadaniach wizualnych testów IQ może być skuteczna i że rozumowanie relacyjne może być wspierane przez odpowiednią architekturę procesu decyzyjnego.

- Dbłość o reprodukowalność. W rozprawie odnotowano udostępnienie kodu dla kilku prac składających się na rozprawę, co istotnie zwiększa wartość naukową i praktyczną przedstawionych wyników.

5. Ocena rozprawy

a. Uwagi krytyczno-polemiczne:

Moja ocena rozprawy jest zdecydowanie pozytywna. Praca jest ambitna, wielowątkowa, dobrze osadzona w aktualnym stanie badań i zawiera zarówno wkład koncepcyjny, jak i metodyczny oraz eksperymentalny. Poniższe uwagi traktuję jako punkty do dyskusji, które mogą być inspirujące podczas obrony lub w dalszych pracach Autorki.

1. Rozprawa operuje szerokim rozumieniem kontekstu: przestrzennym, temporalnym, lokalnym, semiglobalnym, globalnym, a także odnoszonym do różnych domen. Jest to siła pracy, ale jednocześnie rodzi to pytanie, czy możliwe byłoby jeszcze bardziej formalne, wspólne ujęcie kontekstu niezależne od typu danych.
2. Wątek temporalny, reprezentowany głównie przez metodę LIMES, jest interesujący, ale zdecydowanie mniejszy objętościowo niż część poświęcona kontekstowi przestrzennemu. Warto podczas obrony doprecyzować, czy Autorka widzi oba typy kontekstu jako równorzędne elementy tej samej teorii, czy raczej jako dwa przykłady szerszego zjawiska.
3. Wyniki dotyczące WSI i modeli MIL są bardzo cenne, ale naturalnym pytaniem jest to, w jakim stopniu obserwacje te przenoszą się na inne typy obrazów medycznych, inne nowotwory, inne zadania kliniczne oraz inne sposoby przygotowania łątek.
4. Autorka świadomie analizuje attention scores, ale w literaturze trwa dyskusja, czy jest to zawsze dobrym nośnikiem wyjaśnienia. Warto rozważyć, czy Deep Spatial Context można w przyszłości łączyć z dodatkowymi testami kontrfaktycznymi lub perturbacyjnymi.
5. Bardzo interesujące byłoby pogłębienie analizy wrażliwości wyników na dobór liczby sąsiadów, sposób konstruowania macierzy wag przestrzennych, poziom powiększenia obrazu oraz typ ekstraktora cech.
6. Zbiory syntetyczne są bardzo użyteczne, ponieważ dają kontrolę nad relacjami przestrzennymi. Jednocześnie pozostaje pytanie o granice transferu wniosków z danych syntetycznych na złożone, heterogeniczne obrazy histopatologiczne.
7. Obserwacja wskazująca, że ręczny wybór regionów zainteresowania nie poprawia jakości modelu, jest bardzo ważna z praktycznego punktu widzenia. Warto jednak doprecyzować, czy dotyczy ona przede wszystkim analizowanego zadania i protokołu, czy może wskazuje na szerszą zasadę projektowania modeli w cyfrowej patologii.
8. Rozprawa dobrze wpisuje się w nurt Red XAI, czyli wyjaśnień dla twórców i audytorów modeli. Ciekawym dalszym krokiem byłoby zbadanie, które z proponowanych miar i wizualizacji mogą

być realnie użyteczne dla patologów lub lekarzy podejmujących decyzje kliniczne. Uwaga ta stanowi bardziej pewnego rodzaju motywację do dalszych działań niż krytyczną uwagę w niniejszej recenzji.

W szczególności z racji na obowiązki recenzenta, muszę w tym miejscu wskazać jeszcze kilka uwag związanych z pojedynczymi niedociągnięciami stylistycznymi. Doktorant również nie ustrzegł się nielicznych mankamentów natury technicznej, takich jak błędy interpunkcyjne, czy też w pracy występują nieliczne usterki edytorskie oraz literówki.

Poniższe uwagi mają charakter porządkujący i nie podważają wartości pracy.

1. Wśród słów kluczowych dwukrotnie pojawia się „kontekst przestrzenny”. Najprawdopodobniej jedno z wystąpień powinno brzmieć „kontekst temporalny” — str. 8, słowa kluczowe.
2. W sformułowaniu „to facilitate other researchers reproduction” brakuje poprawnej konstrukcji dopełniaczowej. Można było to ująć: „to facilitate other researchers’ reproduction of the results” albo „to facilitate reproduction by other researchers” — str. 17, akapit po liście publikacji P1–P8.
3. Występuje niespójność numeracji publikacji P6 i P7. Na liście publikacji P6 oznacza pracę Deep Spatial Context, a P7 pracę Information Injection..., natomiast w spisie treści rozdział 9 opisano jako P6: Information Injection..., a rozdział 10 jako P7: Deep Spatial Context... .
4. W sformułowaniu „masks annotated by human” występuje błąd językowy. Poprawna forma to np. „masks annotated by humans” albo „human-annotated masks” — str. 28.
5. W podpisie pod rysunkiem 4.9 pojawia się literówka „dspites dataset”. — str. 68, podpis pod rysunkiem 4.9.
6. Zdanie „One of examples is situation...” wymaga korekty językowej. Poprawna wersja mogłaby brzmieć np. „One example is a situation...” albo „One of the examples is a situation...” — str. 185.
7. W sformułowaniu „opposed to scenario...” warto zastosować poprawniejszą konstrukcję „as opposed to a scenario...” — str. 186.

6. Podsumowanie

Rozprawa doktorska mgr Pauliny Anny Tomaszewskiej jest pracą dojrzałą, ambitną i wartościową naukowo. Autorka podjęła ważny problem badawczy, zaproponowała własne metody i protokoły ewaluacyjne, przeprowadziła szeroki zestaw eksperymentów oraz pokazała, że kontekst — szczególnie przestrzenny — może mieć istotne znaczenie zarówno dla działania modeli głębokich, jak i dla ich wyjaśniania.

Na uznanie zasługuje konsekwencja, z jaką Autorka prowadzi badania od rozpoznania luki w istniejących metodach XAI, przez propozycję nowej perspektywy badawczej, aż po konkretne narzędzia i wyniki eksperymentalne. Praca ma także wymiar praktyczny, szczególnie w części dotyczącej cyfrowej histopatologii i modeli MIL.

Wskazane powyżej uwagi mają charakter dyskusyjny i porządkujący. Stanowią raczej przyczynek do dalszego rozwoju bardzo interesującego kierunku badań niż zastrzeżenia wobec wartości rozprawy. Nie wpływają one na moją jednoznacznie pozytywną ocenę pracy.

Konkludując, uważam, że rozprawa doktorska Pani mgr Pauliny Anny Tomaszewskiej spełnia wymagania stawiane w odpowiednich przepisach rozprawom doktorskim i wobec tego stawiam wniosek o jej dopuszczenie do dalszych, przewidzianych Ustawą, etapów postępowania o nadanie stopnia doktora.

Wnoszę o wyróżnienie rozprawy doktorskiej mgr Pauliny Anny Tomaszewskiej pt. „Role of spatial and temporal context in classification using deep neural networks” ze względu na jej **wysoką wartość naukową, oryginalny i aktualny wkład w rozwój metod wyjaśnialnej sztucznej inteligencji oraz bardzo dojrzałe ujęcie roli kontekstu przestrzennego i temporalnego w działaniu głębokich sieci neuronowych.**

Rozprawa **wyróżnia się spójną koncepcją badawczą, znaczącym dorobkiem publikacyjnym, propozycją nowych metod i protokołów ewaluacyjnych, a także istotnym potencjałem praktycznym, szczególnie w obszarze cyfrowej histopatologii i analizy modeli Multiple Instance Learning.** W mojej ocenie **poziom merytoryczny, samodzielność badawcza Autorki oraz jakość uzyskanych wyników uzasadniają skierowanie do Rady Dyscypliny Politechniki Warszawskiej wniosku o wyróżnienie dysertacji.**

Signed by /
Podpisano przez:



Piotr Kowalski

Date / Data: 2026-
07-06 10:20

dr hab. inż. Piotr A. Kowalski, prof. AGH