

**Recenzja rozprawy doktorskiej mgr Pauliny Tomaszewskiej w związku
z postępowaniem w sprawie nadania stopnia doktora w dziedzinie nauk
inżynieryjno-technicznych w dyscyplinie informatyka techniczna
i telekomunikacja**

1. Uwagi wstępne i podstawy sporządzenia recenzji

Niniejszą recenzję sporządzono w odpowiedzi na pismo Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej z dnia 18 maja 2026 r., informujące o powołaniu mnie na recenzentkę rozprawy doktorskiej mgr Pauliny Tomaszewskiej uchwałą podjętą na posiedzeniu Rady w dniu 21 kwietnia 2026 r.

Przedmiotem recenzji jest rozprawa pt. „Role of spatial and temporal context in classification using deep neural networks”, przedłożona w postępowaniu w sprawie nadania mgr Paulinie Tomaszewskiej stopnia doktora w obszarze nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja. Rozprawa doktorska dotyczy istotnego działu informatyki, jakim są wyjaśnialne metody klasyfikacji oparte na głębokich sieciach neuronowych. Rozprawa została wykonana pod kierunkiem prof. dr hab. Przemysława Biecka.

Recenzja przedstawia szczegółową ocenę rozprawy w odniesieniu do wymagań stawianych rozprawom doktorskim.

2. Tematyka i cele rozprawy

Tematyka rozprawy jest ulokowana w bardzo szybko rozwijającym się obszarze połączenia metod wyjaśnialnej sztucznej inteligencji, głębokich sieci neuronowych oraz systemów wspierających podejmowanie decyzji. Rozprawa dotyczy znaczenia kontekstu przestrzennego i czasowego w procesie podejmowania decyzji przez klasyfikatory oparte na głębokich sieciach neuronowych. Szczególną uwagę poświęcono ciągom danych charakteryzującym się zmiennością w czasie oraz obrazom o zróżnicowanej zawartości.

Uzasadnienie podjęcia tematu Autorka przedstawiła w Rozdziale 2.1 (*Model evaluation*) oraz 2.2 (*Context*), gdzie doprecyzowała motywację do podjęcia badań nad wskazaną tematyką oraz wskazała ograniczenia popularnych technik wyjaśnialnej sztucznej inteligencji.

Cele i tezy dysertacji zostały określone przez Autorkę na stronach 14–16. w Rozdziale 1.1 (*Research hypothesis*) Autorka sformułowała następującą główną hipotezę badawczą:

Informacja kontekstowa odgrywa ważną rolę w procesie podejmowania decyzji przez klasyfikatory oparte na głębokim uczeniu.

Weryfikację głównej hipotezy badawczej uporządkowano za pomocą siedmiu pytań badawczych, obejmujących zależności czasowe oraz przestrzenne w różnych skalach.

Stwierdzam, że sformułowane hipotezy i pytania badawcze są odpowiednio sprecyzowane, trafnie odzwierciedlają zakres badań oraz w pełni pokrywają przedstawioną tematykę pracy doktorskiej.

3. Najważniejsze osiągnięcia rozprawy

Głównym osiągnięciem rozprawy jest opracowanie 8 rozwiązań odpowiadających na postawioną hipotezę badawczą i umożliwiających badanie roli kontekstu przestrzennego oraz czasowego w procesie podejmowania decyzji przez głębokie sieci neuronowe.

Zaproponowane podejścia różnią się skalą (lokalna, globalna), typem zależności (czasowe, przestrzenne) oraz stopniem zaawansowania rozwiązania.

Do najważniejszych osiągnięć rozprawy zaliczam:

1. Metodę **LIMES** (ang. *Lightweight Conditional Model Extrapolation for Streaming Data under Class-Prior Shift*), wykorzystującą kontekst czasowy do adaptacji modelu do zmian prawdopodobieństw a priori klas w danych strumieniowych.
2. Metodę **Deep Spatial Context**, obejmującą trzy miary służące analizie kontekstu przestrzennego.
3. Architekturę **Duel-IQ** uwzględniającą relacje przestrzenne.
4. Metodę oceny roli lokalnego kontekstu w obrazach histopatologicznych.
5. Protokół oceny **ReasonSMIL**, służący badaniu zdolności modeli uczenia wieloinstancyjnego (ang. *Multiple Instance Learning*, MIL) do uwzględniania kompozycyjnej struktury obrazów wraz z zestawem syntetycznych zbiorów danych **SyntheticSMIL**.

Kluczową część dorobku rozprawy stanowią wyniki badań empirycznych wskazujące na znaczenie kontekstu oraz różnorodności fragmentów obrazów dla jakości predykcji i zdolności modeli do uogólniania. Uzyskane wyniki wskazują również, że uwzględnienie kontekstu może poprawiać jakość klasyfikacji. Autorka wykazała ponadto różnice między modelami CLAM (ang. *Clustering-constrained Attention Multiple Instance Learning*) oraz TransMIL (ang. *Transformer-based Multiple Instance Learning*) pod względem wykorzystania informacji przestrzennej i stabilności mechanizmu uwagi.

Autorka dołożyła starań, aby potwierdzić praktyczną skuteczność swoich rozwiązań. Rozprawa obejmuje różne obszary zastosowań, w tym analizę obrazów histopatologicznych, testów inteligencji oraz wpisów w serwisie Twitter. Szczególnie wysoko oceniam proponowane rozwiązania dotyczące zastosowań w obszarze histopatologii, gdzie analiza kontekstu przestrzennego wspiera wyjaśnianie predykcji i ocenę ich zgodności z wiedzą medyczną. Aspekt praktycznej użyteczności wyników badań jest wyraźnie uwidoczniiony i ma istotne znaczenie.

4. Ocena struktury rozprawy

Rozprawa napisana jest w języku angielskim. Podstawę rozprawy stanowi osiem prac oznaczonych jako P1–P8, zestawionych w Rozdziale 1.3 i powiązanych z pytaniami badawczymi.

Rozprawa składa się z trzynastu rozdziałów, których układ jest logiczny i spójny z przyjętym zakresem badań. Rozdziały rozprawy odpowiadające artykułom naukowym rozpoczynają się wprowadzeniem umieszczającym daną pracę na tle postawionych pytań badawczych.

Rozdział 1 przedstawia główną hipotezę i pytania badawcze, wykaz prac wchodzących w skład rozprawy oraz jej strukturę.

Rozdział 2 zawiera podstawy oceny modeli, w tym zagadnienia wykrywania pojęć w sieciach neuronowych oraz przegląd literatury.

Część poświęconą autorskim rozwiązaniom rozpoczyna Rozdział 3, w którym przedstawiono metodę LIMES, wykorzystującą kontekst czasowy do adaptacji modelu w warunkach zmian prawdopodobieństw a priori klas (ang. *class-prior shift*).

W Rozdziale 4 Autorka omawia ograniczenia popularnych technik wyjaśnialnej sztucznej inteligencji w zadaniach, w których poprawna klasyfikacja zależy od relacji przestrzennych między obiektami. Rozdział ten pełni istotną funkcję argumentacyjną, uzasadniając potrzebę opracowania metod przedstawionych w dalszej części rozprawy.

Rozdział 5 prezentuje algorytm Duel-IQ, przeznaczony do rozwiązywania testów inteligencji.

Rozdział 6 wprowadza modele głębokiego uczenia stosowane w analizie obrazów histopatologicznych, w szczególności modele uczenia wieloinstancyjnego (ang. *Multiple Instance Learning*, MIL), przygotowując podstawy dla kolejnych analiz. Ich uporządkowanie według skali i sposobu badania kontekstu oceniam jako trafne.

W Rozdziale 7 przedstawiono protokół oparty na przesłanianiu części obrazu, służący ocenie znaczenia lokalnego otoczenia analizowanego fragmentu tkanki.

Rozdział 8 wprowadza syntetyczne zbiory danych SyntheticSMIL i protokół ReasonSMIL, umożliwiające badanie zdolności modeli MIL do uwzględniania kompozycyjnej struktury obrazów w kontrolowanych warunkach. Następnie, w Rozdziale 9, rozważania przeniesiono na rzeczywiste obrazy histopatologiczne, analizując możliwość odtworzenia struktury tkanki na podstawie podobieństwa cech fragmentów obrazu.

Rozdział 10 przedstawia metodę Deep Spatial Context oraz trzy miary służące badaniu zachowania informacji przestrzennej w cechach, wielkościach docelowych i resztach związanych z modelami MIL. Analizy obejmują różne podzbiory danych oraz warianty i rodziny modeli. Rozdział 11 uzupełnia tę część rozprawy o porównanie wykorzystania pełnych obrazów preparatów (ang. *Whole Slide Images*, WSI) z wykorzystaniem podzbiorów ich fragmentów, wybieranych losowo, wskazywanych przez lekarza lub wyznaczanych na podstawie wartości mechanizmu uwagi. Uwzględnienie miar różnorodności oraz oceny lekarskiej łączy perspektywę metodologiczną z praktycznym wymiarem prowadzonych badań.

Rozdziały 12 i 13 zawierają odpowiednio podsumowanie i analizę dalszych kierunków badań.

Kolejność rozdziałów jest logiczna i prowadzi od oceny wpływu lokalnego otoczenia, przez kontrolowane eksperymenty na danych syntetycznych, do analizy złożonych zależności przestrzennych i doboru danych wejściowych w rzeczywistych zastosowaniach.

5. Szczegółowa ocena rozprawy doktorskiej

Przedstawiona rozprawa stanowi istotny wkład w rozwój dyscypliny informatyka techniczna i telekomunikacje, w szczególności w zakresie badania roli kontekstu przestrzennego i czasowego w klasyfikacji z wykorzystaniem głębokich sieci neuronowych. Zaproponowane metody umożliwiają analizę sposobu wykorzystania informacji kontekstowej przez modele oraz ocenę jej wpływu na podejmowane decyzje. Szczególne znaczenie mają zastosowania w analizie obrazów histopatologicznych, w których wyjaśnialność predykcji i możliwość weryfikacji ich podstaw przez lekarza są istotne dla wiarygodności systemów wspomagających diagnostykę. Opracowane rozwiązania mają zatem wartość zarówno naukową, jak i aplikacyjną, wspierając rozwój metod oceny i wyjaśniania działania modeli z potencjałem zastosowania w medycynie.

Prace P1–P8 obejmują artykuły konferencyjne, pracę warsztatową (P4), pracę przedstawioną w ramach konsorcjum doktoranckiego (P7) oraz manuskrypt oznaczony w rozprawie jako zgłoszony na konferencję MICCAI 2026 (ang. *International Conference on Medical Image Computing and Computer Assisted Intervention*) (P8). Praca P2 ma charakter pracy stanowiskowej (ang. *position paper*). Na uwagę zasługuje zróżnicowany charakter prac odpowiadający przyjętej konstrukcji rozprawy, łączącej rozważania koncepcyjne, proponowane rozszerzenia metodologiczne i badania empiryczne. Paulina Tomaszewska jest pierwszą autorką wszystkich wymienionych prac, a w przypadku P7 jedyną autorką.

Pomimo licznych zalet rozprawy i oryginalności proponowanych rozwiązań niektóre zagadnienia wymagają doprecyzowania lub szerszej dyskusji. Poniżej przedstawiam uwagi i pytania dotyczące przyjętych definicji, założeń metodologicznych oraz możliwości uogólnienia uzyskanych wyników. Ich celem jest pogłębienie dyskusji nad zakresem stosowalności opracowanych metod, ich ograniczeniami oraz kierunkami dalszego rozwoju.

1. Pojęcie kontekstu wymaga doprecyzowania. Definicje przedstawione w Rozdziale 2.2 (*Context*) w niewystarczającym stopniu nawiązują do istniejącej literatury. Ze względu na kluczową rolę pojęcia kontekstu w rozprawie zasadne byłoby szersze omówienie jego dotychczasowych ujęć oraz wskazanie, które definicje zaczerpnięto z wcześniejszych prac, a które stanowią ich modyfikację lub autorską propozycję. Warto również uzasadnić wybór przyjętych definicji i omówić poczynione założenia. W szczególności warto doprecyzować, jak odległą historię obejmuje kontekst czasowy, w jaki sposób ustala się jej zakres oraz jak przyjęte ujęcie odnosi się do klasycznych metod opisu i modelowania zależności w analizie szeregów czasowych. Dyskusję warto uzupełnić zarówno o wybrane podstawowe metody analizy szeregów czasowych, takie jak modele wektorowej autoregresji (ang. *Vector Autoregression*, VAR), uwzględniające zależności między wieloma zmiennymi i ich opóźnionymi wartościami, jak i o metody uczenia maszynowego w trybie online (ang. *online learning*).
2. W odniesieniu do kontekstu przestrzennego zasadne byłoby szersze omówienie aktualnego stanu badań nad metodami uwzględniania autokorelacji przestrzennej oraz zależności między obserwacjami. Pozwoliłoby to precyzyjniej określić, w jakim zakresie proponowane podejścia rozwijają istniejące metody modelowania zależności przestrzennych.

3. Doprecyzowania wymagają pojęcia skali lokalnej i globalnej ze szczególnym uwzględnieniem skali semiglobalnej. Ponadto w jaki sposób przypisuje się analizowane sąsiedztwo do danej skali i czy przypisanie to zależy od wielkości obrazu oraz rozmiaru jego fragmentów? Takie uzupełnienie pozwoliłoby lepiej umiejscowić rozważania na tle bieżącej literatury i precyzyjniej określić zakres stosowalności wprowadzonych pojęć.
4. Weryfikacja eksperymentalna metody LIMES została ograniczona do jednego zbioru danych, co utrudnia ocenę możliwości uogólnienia uzyskanych wyników. Ponadto, ocena metody LIMES opiera się na ograniczonym zestawie miar jakości klasyfikacji. Ze względu na analizowane zmiany prawdopodobieństw a priori klas zasadne byłoby uwzględnienie także precyzji (ang. *precision*) i czułości (ang. *recall*), raportowanych dla poszczególnych klas oraz w kolejnych przedziałach czasu. Warto rozszerzyć badania na zbiory różniące się liczbą klas, stopniem ich niezrównoważenia oraz dynamiką zmian prawdopodobieństw a priori klas. Interesujące byłoby również określenie ograniczeń metody w przypadku innych rodzajów zmian rozkładu danych (ang. *concept-drift*). Pozwoliłoby to lepiej określić warunki, w których wykorzystanie kontekstu czasowego przynosi korzyści. Rozszerzenie analiz oraz pogłębienie dyskusji w tym zakresie znacząco wzmocniłyby tę część rozprawy pod względem metodologicznym.
5. W odniesieniu do pytania RQ4 interesujące byłoby rozszerzenie dyskusji poza kontekst zawarty w pojedynczym obrazie. Jakie są możliwości i ograniczenia dostosowania proponowanych podejść do analizy danych multimodalnych, w których informacja kontekstowa pochodzi z różnych źródeł, na przykład obrazów histopatologicznych oraz zmieniających się w czasie danych klinicznych? Jakie zmiany w definicji kontekstu i sposobie oceny jego wpływu na predykcję byłyby wówczas konieczne?
6. W rozważaniach dotyczących pytania badawczego RQ6 Autorka odwołuje się do usuwania informacji nieistotnej z danych wejściowych. Warto doprecyzować, według jakich kryteriów informacja jest uznawana za nieistotną i kto dokonuje takiej oceny. Jakie podejścia do identyfikacji i usuwania informacji nieistotnej zastosowano w badaniach dotyczących RQ6, a jakie rozważano jako dalsze kierunki prac? w jaki sposób rozróżniono ocenę istotności informacji przez eksperta od jej przydatności dla zadania predykcyjnego?
7. W przypadku rozważań dotyczących pytania badawczego RQ7 proszę o szersze uzasadnienie doboru miar podobieństwa cech fragmentów obrazu oraz omówienie wpływu tego wyboru na możliwość odtworzenia struktury tkanki. Czy badano wrażliwość uzyskanych wyników na wybór miary podobieństwa? Warto również doprecyzować, w jakim stopniu podobieństwo reprezentacji wyznaczonych przez sieć odzwierciedla rzeczywiste relacje przestrzenne między fragmentami tkanki.
8. Opis metody przedstawionej w Rozdziale 7.4 wymaga większej formalizacji oraz wyraźniejszego oddzielenia założeń metodologicznych od szczegółów przeprowadzonego eksperymentu. Warto jednoznacznie zdefiniować sposób ograniczania kontekstu, parametry procedury oraz kryteria oceny zmian predykcji. Pozwoliłoby to lepiej określić ogólny charakter proponowanej metody i możliwości jej zastosowania poza analizowanym eksperymentem.

9. Opis metody Deep Spatial Context w rozdziale 10 wymaga uzupełnienia o ocenę spełnienia założeń zastosowanych modeli regresji. Usuwanie cech na podstawie współczynnika korelacji Pearsona ogranicza silne zależności liniowe między parami cech, lecz nie wyklucza współliniowości wynikającej z liniowych kombinacji większej liczby cech. Warto przedstawić diagnostykę liniowości, struktury wariancji błędów oraz autokorelacji reszt po dopasowaniu modeli oraz dokonać krytycznej analizy założeń. Czy przyrost współczynnika determinacji oceniano na danych wykorzystanych do dopasowania modeli? Warto doprecyzować również interpretację przyrostu współczynnika determinacji.
10. Wnioski końcowe (Rozdział 12) są trafne, jednak warto uzupełnić je o pogłębioną, krytyczną analizę uzyskanych wyników, bardziej szczegółowe omówienie ograniczeń proponowanych metod oraz jasno sformułowane rekomendacje dotyczące sposobów ograniczenia ich wpływu.

Rozprawa jest napisana bardzo starannie pod względem językowym. Można dostrzec jedynie pojedyncze redakcyjne niedoskonałości, np. odniesienie do sekcji 8.6 na stronie 42 jest niejasne. Należy zadbać o rozwinięcie skrótów przy pierwszym użyciu, np. w przypadku nazwy modelu CLAM pojawiającej się w pytaniu RQ6. Tabele i wykresy są w ogólności czytelne, lecz mogłyby zawierać wyraźniejsze oznaczenia kluczowych wyników. Podpis tabeli 8.2 niewystarczająco informuje o prezentowanych symbolach. Ryciny 10.2 i 10.11 mogłyby być bardziej wyraźne. Powyższe mankamenty stanowią drobne niedociągnięcia redakcyjne, nie są to istotne błędy lub zastrzeżenia merytoryczne.

W mojej opinii rozprawa doktorska zawiera obszerny, dobrze opracowany materiał badawczy, który oceniam pozytywnie pod względem formalnym i merytorycznym. Autorka umiejętnie połączyła opracowanie nowych metod analizy i wyjaśniania działania modeli z ich weryfikacją eksperymentalną, w szczególności w zastosowaniach do obrazów histopatologicznych.

Na uwagę zasługuje także obszerny i zróżnicowany dorobek publikacyjny Doktorantki. Oprócz ośmiu prac stanowiących podstawę rozprawy, wśród których znajdują się publikacje na wiodących konferencjach ICML (*International Conference on Machine Learning*), AISTATS (*International Conference on Artificial Intelligence and Statistics*), ICPR (*International Conference on Pattern Recognition*) i AIME (*International Conference on Artificial Intelligence in Medicine*), Autorka przedstawiła wykaz siedmiu dodatkowych prac, niewłączonych do rozprawy. Obejmują one między innymi artykuły w czasopismach *Pattern Recognition*, *Nutrients* oraz *Journal of Medical Internet Research*.

Tematyka publikacji dotyczy m.in. wykorzystania głębokiego uczenia w analizie obrazów medycznych, zastosowań w ochronie zdrowia, odtwarzalności badań. Dorobek ten świadczy o szerokich zainteresowaniach naukowych Autorki i jej umiejętności podejmowania współpracy interdyscyplinarnej oraz międzynarodowej.

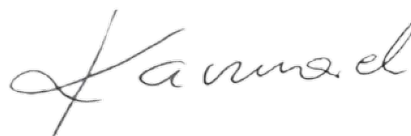
Dodatkowym atutem jest deklarowane udostępnienie kodu związanego z pracami P1, P3, P4, P5, P6 i P8 w repozytoriach GitHub, co ułatwi odtwarzanie wyników i wykorzystanie proponowanych rozwiązań w dalszych badaniach.

7. Konkluzja

Autorka rozprawy opracowała nowe algorytmy oraz komponenty większego systemu, które stanowią oryginalne rozwiązanie problemu naukowego dotyczącego roli kontekstu przestrzennego i czasowego w klasyfikacji z wykorzystaniem głębokich sieci neuronowych. Aby potwierdzić ich zasadność, przeprowadziła odpowiednie eksperymenty. Postawione cele badawcze zostały w pełni zrealizowane.

Oceniam wkład Autorki pozytywnie. Uzyskane wyniki są wartościowe i oryginalne, stanowią istotny wkład w rozwój dyscypliny informatyka techniczna i telekomunikacja, a w szczególności w rozwój wyjaśnialnych modeli klasyfikacyjnych. Bardzo istotne jest powiązanie prowadzonych badań z aktualnymi problemami i praktycznymi zastosowaniami.

Uważam, że rozprawa w pełni spełnia wymagania stawiane rozprawom doktorskim i wnoszę o dopuszczenie Pani mgr Pauliny Tomaszewskiej do publicznej obrony.



dr hab. Katarzyna Kaczmarek-Majer, prof. IBS PAN