

prof. dr hab. inż. Rafał Scherer
Katedra Sztucznej Inteligencji
Wydział Informatyki i Sztucznej Inteligencji
Politechnika Częstochowska
al. Armii Krajowej 36
42-200 Częstochowa

Częstochowa, dn. 15 maja 2025 r.

Recenzja

rozprawy doktorskiej mgr inż. Justyny Stypułkowskiej, pt.: Analiza obrazów wielospektralnych i RGB z zastosowaniem głębokich sieci neuronowych dla potrzeb automatycznej oceny parametrów kondycji kukurydzy przez robota polowego.

Niniejszą recenzję opracowano zgodnie z uchwałą nr 811/2024 Rady naukowej dyscypliny informatyka techniczna i telekomunikacja PW. Promotorem jest prof. dr hab. inż. Przemysław Rokita.

1. Charakterystyka tematu, celu i tezy badawczej rozprawy

Tematyka rozprawy wpisuje się w kluczowe wyzwania współczesnego rolnictwa precyzyjnego. W dobie rosnącej populacji, zmian klimatycznych i presji na zrównoważone gospodarowanie zasobami, opracowanie narzędzi umożliwiających precyzyjny monitoring upraw ma fundamentalne znaczenie dla bezpieczeństwa żywnościowego i ochrony środowiska. Głównym celem pracy jest stworzenie systemu opartego na sztucznej inteligencji, który integruje analizę obrazów wielospektralnych i RGB z autonomicznym działaniem robota polowego. Autorce przyświeca idea zastąpienia tradycyjnych, czasochłonnych metod oceny kondycji roślin (np. manualnej inspekcji pól) przez zautomatyzowane rozwiązania oparte na danych. Badania koncentrują się na trzech kluczowych parametrach agronomicznych: identyfikacji faz rozwojowych w skali BBCH, klasyfikacji poziomów nawodnienia oraz wykrywaniu porażenia wybranymi patogenami. Temat nabiera dodatkowej wagi w świetle wyzwań związanych z adaptacją rolnictwa do zmian klimatu. Opracowane rozwiązania nie tylko przyczyniają się do zwiększają efektywność produkcji, ale także do redukcji śladu węglowego poprzez precyzyjne zarządzanie agrochemikaliami.

2. Zawartość rozprawy

Recenzowana praca mgr inż. Justyny Stypułkowskiej składa się z dziewięciu rozdziałów, wykazu oznaczeń i skrótów, zmiennych, spisu rysunków, tabel, bibliografii oraz załączników. Dokument liczy 214 stron.

Pierwszy rozdział jest krótkim wprowadzeniem do tematyki, omówieniem projektów z którymi związane były prace wykonane w ramach rozprawy doktorskiej; zawiera również cele pracy: analiza faz rozwoju kukurydzy (w skali BBCH) przez zastosowanie głębokich sieci neuronowych do automatycznej detekcji i klasyfikacji na podstawie obrazów RGB i wielospektralnych, wybór najlepszego algorytmu i typu obrazowania; ocena poziomu nawodnienia kukurydzy za pomocą sieci neuronowych w analizie obrazów RGB i wielospektralnych w celu detekcji i klasyfikacji stanu nawodnienia; wykrywanie chorób kukurydzy przez automatyczną detekcję i klasyfikację porażenia patogenami na podstawie obrazów RGB z użyciem najskuteczniejszego algorytmu; stworzenie autorskich zbiorów danych składających się z oznaczonych obrazów (RGB i wielospektralnych) podzielonych na zestawy treningowe, walidacyjne i testowe, pochodzących z pola testowego. Rozdział kończy omówienie zawartości rozprawy.

Rozdział drugi omawia sieci neuronowe w zastosowaniach wizji komputerowej. Omówiono neurony, sieci typu nieliniowy MLP, oraz zasadę uczenia sieci z danych oraz dokładnie działanie sieci splotowych (CNN). Omówiono również popularne architektury jak odmiany RsNet oraz Transformer (HTC, Mask2Former, ConvNeXt, EfficientNet, Vision Transformer, Swin Transformer). Dalej omówiono specyfikę przetwarzania obrazów dwuwymiarowych składających się z wielu kanałów, szczególnie wielospektralnych oraz podstawowe sposoby oceny jakości działania systemów uczących się. Jest to istotny przegląd istniejących rozwiązań w dziedzinie monitoringu upraw, w którym autorka nie tylko prezentuje aktualny stan wiedzy, ale także, co szczególnie cenne, identyfikuje luki badawcze, uzasadniając tym samym potrzebę i nowatorski charakter własnych badań. Wnikliwa analiza literaturowa obejmuje zarówno metody oparte na tradycyjnym przetwarzaniu obrazów, jak i najnowsze rozwiązania wykorzystujące głębokie uczenie.

Rozdział 3 omawia dane użyte w badaniach oraz proces ich przygotowania i oznaczania. Autorka przedstawia kompleksową metodykę przygotowania danych, implementacji modeli oraz procesów trenowania i walidacji, co stanowi solidną podstawę dla uzyskanych wyników. Przedstawiono również architektury sieci neuronowych zastosowane do detekcji i klasyfikacji obiektów, metody ich trenowania oraz opracowany klasyfikator głoszący, który poprawia wyniki analizy. Rozdział rozpoczyna się od opisu metodyki przygotowania danych, w tym pozyskiwania i oznaczania obrazów RGB oraz wielospektralnych z kamer MicaSense RedEdge MX DUAL i MapIR Survey3 Red+Green+NIR. Dane zostały zebrane z pól testowych, a następnie poddane procesowi oznaczania, który obejmował ręczną adnotację faz rozwojowych kukurydzy w skali BBCH, poziomów nawodnienia oraz oznak porażenia patogenami. Wykorzystano narzędzie Label Studio, co zapewniło spójność i dokładność oznaczeń. Dane zostały podzielone na zbiory treningowe, walidacyjne i testowe, co umożliwiło rzetelną ocenę modeli. Autorka ogólnie opisuje implementację zaawansowanych architektur głębokich sieci neuronowych, w tym HTC (HTC_x101, HTC_r101, HTC_r50), Mask2Former, ConvNeXt Small, ResNet50, EfficientNet-B3, Vision Transformer i Swin Transformer. Każdy z modeli został dostosowany do specyfiki zadania, takiego jak detekcja faz rozwojowych, klasyfikacja

poziomów nawodnienia czy identyfikacja patogenów. Wykorzystano frameworki m.in. PyTorch i MMDetection, co zapewniło efektywność obliczeniową i skalowalność rozwiązań. W rozdziale szczegółowo omówiono proces trenowania modeli, uwzględniając techniczne aspekty, takie jak optymalizacja hiperparametrów (np. współczynnik uczenia, liczba epok) oraz zastosowanie technik augmentacji danych. Autorka zastosowała metodę walidacji *hold-out*, która pozwoliła na ocenę skuteczności modeli na niezależnym zbiorze testowym. Monitorowano metryki takie jak *bbox mAP* i *segm mAP*, co umożliwiło precyzyjną ocenę jakości detekcji i segmentacji. W końcowej części rozdziału opisano proces detekcji i klasyfikacji obiektów, w tym zastosowanie klasyfikatora głosującego dla danych wielospektralnych. Model ten agregował wyniki z różnych kanałów spektralnych, zwiększając dokładność predykcji. Autorka podkreśliła również znaczenie interpretacji wyników i analizy błędów, co pozwoliło na identyfikację obszarów wymagających poprawy.

W rozdziale 4 autorka prezentuje wyniki badań dotyczących detekcji i klasyfikacji faz rozwojowych kukurydzy w skali BBCH, wykorzystując obrazy RGB oraz wielospektralne z kamery MicaSense RedEdge MX Dual. Kluczowym elementem jest porównanie skuteczności obu typów obrazowania oraz ocena różnych architektur głębokich sieci neuronowych. Autorka przedstawia kompleksowy schemat opracowanego systemu, obejmujący: przetwarzanie wstępne (kalibracja i korekcja obrazów wielospektralnych, normalizacja danych RGB), implementację modeli (zastosowanie architektur HTC, Mask2Former i innych do detekcji obiektów oraz segmentacji instancji) oraz klasyfikację faz BBCH (przypisanie roślin do odpowiednich klas rozwojowych na podstawie analizy liści). Obrazy wielospektralne, dzięki dodatkowym kanałom (np. NIR, bliska podczerwień), zapewniły wyższą dokładność w zadaniu detekcji faz rozwojowych niż obrazy RGB. Najlepsze wyniki osiągnięto dla modeli HTC_x101 i Mask2Former, które skutecznie radziły sobie z różnymi warunkami oświetleniowymi i nakładaniem się roślin. Wykorzystanie danych wielospektralnych pozwoliło na lepszą segmentację liści, co przełożyło się na dokładniejsze określenie fazy BBCH. Modele oparte na RGB również osiągnęły dobre rezultaty, ale były bardziej podatne na zakłócenia związane z cieniami i zmiennymi warunkami świetlnymi. Wykazano, że obrazy RGB są wystarczające do podstawowej klasyfikacji, ale w przypadku wymagających warunków ich skuteczność spada. Z powodu problemów z identycznością obrazów poszczególnych kanałów multispektralnych opisanych w rozdziale 3, autorka zastosowała klasyfikator głosujący, który agregował wyniki z różnych kanałów spektralnych. Dzięki temu osiągnięto wyższą precyzję w porównaniu do pojedynczych modeli oraz lepsze uogólnienie na różne warunki środowiskowe. Analiza krzywych uczenia wykazała, że modele oparte na danych wielospektralnych szybciej osiągały zbieżność i miały niższy błąd walidacyjny. Obrazy wielospektralne zapewniły znacznie lepszą dokładność niż RGB w detekcji faz rozwojowych kukurydzy.

Rozdział piąty pracy doktorskiej kontynuuje temat detekcji faz rozwojowych kukurydzy w skali BBCH, skupiając się tym razem na analizie obrazów pozyskanych kamerą wielospektralną MapIR Survey3 (Red+Green+NIR) oraz porównaniu ich skuteczności z tradycyjnym obrazowaniem RGB. Autorka prezentuje spójną metodologię badawczą, w której szczególną

uwagę poświęca różnicom w efektywności obu typów obrazowania oraz optymalizacji procesu klasyfikacji. Kluczowym elementem rozdziału jest szczegółowa analiza wyników uzyskanych dla różnych architektur sieci neuronowych. Podobnie jak w poprzednim rozdziale, najlepsze rezultaty osiągnięto dla modeli HTC i Mask2Former, jednak w tym przypadku autorka zwraca szczególną uwagę na specyficzne zalety kamery MapIR Survey3. Jej zdolność do rejestracji kanału bliskiej podczerwieni (NIR) okazała się szczególnie cenna w identyfikacji subtelnych różnic w stanie fizjologicznym roślin, co przełożyło się na wyższą dokładność klasyfikacji w porównaniu do obrazów RGB. Co istotne, rozdział zawiera pogłębioną dyskusję na temat wpływu poszczególnych kanałów spektralnych na skuteczność detekcji, co stanowi wartościowy wkład w rozwój metod analizy obrazów w rolnictwie precyzyjnym. Autorka przedstawia również interesujące porównanie krzywych uczenia dla różnych konfiguracji modeli, wskazując na szybszą zbieżność i niższe wartości błędów w przypadku wykorzystania danych wielospektralnych. Szczegółowa analiza metryk jakościowych (mAP, precision, recall) potwierdza przewagę obrazowania wielospektralnego, zwłaszcza w trudniejszych warunkach, takich jak gęste ulistnienie czy nakładanie się roślin. Rozdział zawiera także cenne uwagi praktyczne dotyczące implementacji systemu w rzeczywistych warunkach polowych, co świadczy o dbałości autorki o aspekty aplikacyjne prowadzonych badań.

Rozdział szósty pracy doktorskiej wprowadza nowy, istotny wątek badawczy dotyczący automatycznej oceny stanu nawodnienia kukurydzy z wykorzystaniem obrazowania RGB i wielospektralnego. Autorka w sposób systematyczny prezentuje kompleksowe podejście do problemu, rozpoczynając od szczegółowego opisu przyjętej metodyki, przez analizę wyników, aż po praktyczne wnioski implementacyjne. Kluczowym osiągnięciem tego rozdziału jest opracowanie i walidacja trójklasowego systemu klasyfikacji (małe, średnie i duże nawodnienie), który wykorzystuje zarówno wizualne cechy morfologiczne roślin (widoczne w obrazach RGB), jak i charakterystyczne sygnatury spektralne (rejestrowane przez kamerę MapIR Survey3). Szczególnie wartościowe są przedstawione przez autorkę wyniki porównawcze, które jednoznacznie wskazują na wyższą skuteczność obrazowania wielospektralnego.

Rozdział siódmy pracy doktorskiej koncentruje się na kluczowym zagadnieniu diagnostyki fitosanitarnej, tj. automatycznej identyfikacji objawów porażenia kukurydzy przez wybrane patogeny. Autorka w sposób metodyczny prezentuje kompleksowe rozwiązanie oparte wyłącznie na analizie obrazów RGB, co stanowi istotne novum w kontekście dotychczasowych badań wykorzystujących głównie dane spektralne. Centralnym elementem rozdziału jest opracowanie i walidacja systemu detekcji głównych patogenów oraz śladów żerowania szkodnika. Autorka szczegółowo opisuje proces tworzenia specjalistycznego zbioru danych, uwzględniającego zróżnicowane stadia rozwoju objawów chorobowych, co miało kluczowe znaczenie dla skuteczności późniejszej klasyfikacji. W rozdziale szczególnie wartościowa jest pogłębiona analiza charakterystycznych cech wizualnych wykorzystywanych przez algorytmy do identyfikacji patogenów. Autorka przedstawia innowacyjne podejście polegające na integracji systemu diagnostycznego z platformą robota polowego, ze szczególnym

uwzględnieniem optymalizacji parametrów akwizycji obrazów w warunkach polowych. Praktyczne aspekty implementacji, w tym analiza wpływu warunków oświetleniowych na skuteczność detekcji oraz propozycje algorytmów kompensacji zmiennego natężenia światła, stanowią szczególnie wartościowy element pracy.

Rozdział ósmy stanowi praktyczne zwieńczenie badań przedstawionych w pracy, koncentrując się na implementacyjnych aspektach opracowanego systemu do automatycznej oceny kondycji kukurydzy. Autorka prezentuje dwa kluczowe elementy systemu: interaktywną aplikację opartą na bibliotece Gradio oraz mechanizm eksportu wyników do plików CSV, które razem tworzą kompletną platformę analityczną dla użytkowników końcowych. Aplikacja wykorzystująca bibliotekę Gradio została zaprojektowana jako intuicyjne narzędzie umożliwiające przetwarzanie obrazów w czasie rzeczywistym. Autorka szczegółowo opisuje jej modułową architekturę, która integruje wcześniej wytrenowane modele do detekcji faz rozwojowych, poziomów nawodnienia i patogenów. Interfejs użytkownika został zoptymalizowany pod kątem praktycznego zastosowania w rolnictwie precyzyjnym, oferując możliwość przesyłania pojedynczych obrazów lub całych sekwencji z pól uprawnych. W aplikacji zaimplementowano także funkcje wizualizacji wyników, w tym nakładanie masek segmentacyjnych i podświetlanie obszarów zainteresowania, co znacznie ułatwia interpretację wyników przez agronomów. Drugim istotnym komponentem systemu jest generowanie plików CSV zawierających szczegółowe parametry przetwarzania oraz wyniki analizy. Mechanizm ten umożliwia łatwą integrację z istniejącymi systemami zarządzania uprawami oraz dalszą analizę statystyczną zgromadzonych danych. W rozdziale podkreślono także znaczenie standaryzacji formatu wyjściowego, co jest szczególnie istotne w kontekście przyszłej skalowalności rozwiązania.

Rozdział 9 podsumowuje rozprawę oraz proponuje dalsze możliwości rozbudowy zaproponowanych metod. Dalej następuje spis ilustracji i tabel, bibliografia oraz załączniki.

3. Ocena rozprawy

W ramach rozprawy doktorskiej Pani mgr inż. Justyna Stypułkowska zaproponowała oryginalne rozwiązanie, które łączy nowatorskie rozwiązania uczenia maszynowego z aplikacjami dla rolnictwa precyzyjnego, oferując narzędzia do monitoringu upraw w skali pojedynczych roślin. Tematyka pracy jest bardzo aktualna i potrzebna, oryginalny dorobek autorki polega na:

- stworzeniu systemu automatyzującego ocenę kondycji kukurydzy, integrującego detekcję faz rozwojowych (skala BBCH), poziomów nawodnienia i porażenia patogenami z wykorzystaniem robota polowego,
- udowodnieniu przewagi obrazowania wielospektralnego nad RGB,
- zaproponowanie klasyfikatora głoszącego dla danych wielospektralnych co poprawiło stabilność predykcji poprzez agregację wyników z różnych kanałów spektralnych,

- detekcji patogenów wyłącznie na bazie obrazów RGB, pomijając kosztowne dane spektralne,
- stworzeniu praktycznej implementacji jako aplikacji i systemu eksportu wyników (CSV) gotowych do wdrożenia w gospodarstwach,
- stworzenie autorskich zbioru danych faz BBCH i poziomów nawodnienia, wypełniając lukę badawczą,
- walidację w warunkach zbliżonych do rzeczywistych przez testy na polatkach doświadczalnych z użyciem prototypu robota polowego.

Rozprawa doktorska uwidacznia wysoką ogólną wiedzę teoretyczną i praktyczną oraz umiejętność prowadzenia pracy naukowej mgr inż. Justyny Stypułkowskiej. Opracowała wprowadzenie do tematyki i przegląd literatury na temat architektur sieci neuronowych związanych z tematyką pracy. Rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego. Zaproponowane metody mają duże znaczenie dla nauk technicznych oraz przemysłu, zarówno teoretyczne, jak i aplikacyjne.

Niezależnie od mojej dobrej oceny pracy, nasunęły mi się następujące pytania i uwagi:

- Podstawowe sieci typu MLP i CNN omówione są bardzo dokładnie, a wszystkie złożone architektury używane w badaniach bardzo pobieżnie.
- Brak jest szczegółowej dyskusji na temat geograficznej i sezonowej zmienności danych treningowych (np. czy uwzględniono różne regiony klimatyczne lub ekstremalne warunki pogodowe), co może ograniczać generalizację modeli w rzeczywistych warunkach polowych. Niewystarczająco opisano metody radzenia sobie z nie zrównoważonymi klasami (np. rzadkie fazy BBCH lub rzadkie patogeny), co mogło wpłynąć na wyniki metryk.
- Brak jest omówienia problemów z oznaczaniem obrazów RGB i wielospektralnych; kto dokonywał takiej adnotacji?
- Brak jest podania dokładnych meta parametrów użytych sieci neuronowych oraz szczegółów jak obrazy wielospektralne były przetwarzane przez warstwy splotowe.
- Część zawartości, np. Tabela 4.1 (wklejona jako obraz) oraz np. Tabela 4.4 pochodzą z pracy Stypułkowska, J., & Rokita, P. (2025). Classification of Maize Growth Stages Using Deep Neural Networks with Voting Classifier. Machine Graphics and Vision, 3-4):29–53, 2024, bez jej cytowania.
- Zazwyczaj zamiast pisać np. „algorytm HTC”, mówimy o modelu czy architekturze HTC.
- Do oznaczenia wymiaru czasami stosowany jest znak x, a czasami ×.

4. Wnioski końcowe recenzji

Podsumowując recenzję stwierdzam, że rozprawa doktorska „Analiza obrazów wielospektralnych i RGB z zastosowaniem głębokich sieci neuronowych dla potrzeb automatycznej oceny parametrów kondycji kukurydzy przez robota polowego” prezentuje

oryginalne rezultaty stanowiące rozwiązanie problemu naukowego oraz wkład w rozwój dyscypliny informatyka techniczna i telekomunikacja. Pani Justyna Stypułkowska wykazała się umiejętnością samodzielnej pracy badawczej, znajomością literatury światowej i wiedzą w zakresie wizji komputerowej. Rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydatki. Recenzowana praca spełnia wymagania Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2022 r. poz. 574 z późn. zm.). w dyscyplinie naukowej informatyka techniczna i telekomunikacja. Wnoszę o jej przyjęcie i dopuszczenie do dalszych etapów postępowania doktorskiego.

