

Dr hab. Bartosz Zieliński, prof. UJ
Wydział Matematyki i Informatyki
Uniwersytet Jagielloński
bartosz.zielinski@uj.edu.pl

Kraków, 21 grudnia 2025 r.

Recenzja rozprawy doktorskiej
mgra inż. Przemysława Dolaty
zatytułowanej:

Deep learning methods for computer vision-based industrial inspection applications

1. Problem badawczy i jego znaczenie

W swojej pracy doktorant podejmuje zagadnienia związane z zastosowaniem głębokich sieci neuronowych do klasyfikacji i segmentacji obrazów przemysłowych. Opracowuje autorskie metody oraz przeprowadza ich szczegółową ewaluację na podstawie eksperymentów komputerowych. Większość zaproponowanych rozwiązań ma charakter nowatorski i stanowi przykład udanego wykorzystania metod uczenia maszynowego.

2. Wkład autora

Celem pracy było rozwiązanie szeregu zadań przemysłowych, począwszy od klasyfikacji defektów ziarna, poprzez automatyczną ocenę napawania laserowego, pomiar minimalnej średnicy ziemniaka, aż po monitorowanie wzrostu mączniaka. Zakres oraz zawartość rozprawy oceniam pozytywnie. Bardzo dobrze zdefiniowano również trzy pytania badawcze oraz podpytania. Opis poszczególnych publikacji jest bardzo staranny, a jego lektura sprawiła mi dużą przyjemność. Wyjątkiem jest pierwsza połowa sekcji 4.1, która jest nieco rozmyta i cięższa w odbiorze, co może wynikać z trudności w połączeniu artykułu P4 z artykułami P5 i P6. Poza tym jednak treść była bardzo jasna, co zapewne w dużej mierze zawdzięczamy promotorowi, któremu chciałbym za to podziękować.

Szczegółowa analiza prac doktoranta pozwoliła zdefiniować następujące wnioski:

- **P1:** Doktorant stosuje sieci syjamskie do klasyfikacji obrazów ziarna z różnych rzutów kamery i testuje różne metody fuzji cech tych sieci. Metodologia ta została wprowadzona już kilka lat wcześniej¹, ale pracę oceniam pozytywnie z uwagi na ciekawe zastosowanie, przygotowanie własnej bazy danych, poprawną metodologię oraz wykorzystanie nowatorskich jak na tamten czas (zwłaszcza w Polsce) głębokich

¹ Yi et al. Deep Metric Learning for Person Re-identification. ICPR 2014

WPŁYNEŁO
22-12-2025

RDN-111/245/2025

2 14

sieci neuronowych. Istotnym uchybieniem w sekcji z wynikami jest brak informacji o wariancji dokładności, co uniemożliwia ocenę stabilności uzyskanych rezultatów.

- **P2:** Doktorant stosuje dwie osobne sieci, dedykowane odpowiednio górnej i dolnej stronie ziarna (ang. dorsal and ventral sides of a grain), oraz dodaje sieć klasyfikującą orientację ziarna. Praca jest poprawna, ale mniej wnikliwa niż P1 (brakuje np. ablacji) i również nie zawiera informacji o wariancji dokładności.
- **P3:** Doktorant proponuje trening U-Neta, podczas którego losowe warstwy są maskowane. Po treningu usuwane są kolejne warstwy odpowiadające największej rozdzielczości, dopóki nie powoduje to znacznego spadku dokładności metody. Dzięki temu uzyskiwana jest metoda, pięciokrotnie szybsza, stosowana do automatycznej oceny napawania laserowego (ang. laser cladding). Sam pomysł oceniam jako ciekawy, jednak szczegółowa lektura artykułu budzi wiele pytań, szczególnie z uwagi na brak porównania proponowanej metody z prostszymi alternatywami (np. segmentacją obrazów o mniejszej rozdzielczości przy użyciu U-Neta z mniejszą liczbą warstw).
- **P4:** Doktorant wprowadza metodę automatycznej oceny minimalnej średnicy (ang. minimal diameter) ziemniaka, opartą na segmentacji instancji za pomocą metody Mask R-CNN oraz informacji o kącie wykonania zdjęcia. Z uwagi na brak informacji o faktycznej wielkości ziemniaka doktorant stworzył syntetyczny zbiór danych, który został wykorzystany do trenowania sieci regresyjnej. Uważam to za ciekawy pomysł.
- **P5:** Doktorant przedstawia bardzo zaawansowaną metodę monitorowania wzrostu mączniaka (ang. mealworm growth monitoring). W celu ograniczenia liczby etykiet potrzebnych do wytrenowania modelu wprowadza on metodę sztucznego generowania obrazów, która generuje realistyczne obrazy, uwzględniające nawet cień rzucany przez mączniaka. Pozytywnie oceniam porównanie dokładności kilku bardziej aktualnych metod segmentacji instancyjnej (tj. Shape Priors i AISFormer).
- **P6:** Doktorant przedstawia alternatywę dla metody DETR, w której dla każdego zapytania obiektowego (ang. object query) nie uzyskujemy współrzędnych pola ograniczającego (ang. bounding box), lecz trzy punkty odpowiadające linii środkowej (ang. centreline). Jest to najciekawsza koncepcyjnie idea rozprawy, jednak pod względem metryk mAP i AP działa porównywalnie do standardowego DETR, co jest nieco rozczarowujące.

Wyniki uzyskiwane w trakcie pracy nad rozprawą zostały opublikowane w trzech przyzwoitych czasopismach z listy filadelfijskiej, a także przedstawione na konferencjach ACIIDS (CORE B) i FedCSIS (multi-konferencja). Brak publikacji w wiodących czasopismach i na prestiżowych konferencjach stanowi największy mankament pracy, obok dużego rozrzutu czasowego publikacji (pierwszą i ostatnią pracę dzieli aż osiem lat).

3. Poprawność

Praca stanowi przewodnik po wspomnianych publikacjach autora, opublikowanych w trzech czasopismach i na dwóch konferencjach. Po wprowadzeniu w sekcji 1 następuje opis trzech zagadnień, kolejno w sekcjach 2, 3 i 4. Każda z sekcji zawiera podsekcję opisującą

poszczególne publikacje, napisaną w sposób niezwykle przejrzysty i przemyślany. Pracę wieńczy zwięzłe podsumowanie, po którym zamieszczono bibliografię oraz pełne teksty prac.

Lektura rozprawy doprowadziła do sformułowania poniższych pytań:

1. Czy metody zaproponowane w P1, P2, P3 i P4 są obecnie stosowane w przemyśle?
2. Wszystkie warstwy U-Neta w P3 mają taką samą liczbę kanałów, co zaburza podstawową zasadę głębokich sieci konwolucyjnych, która zakłada zwiększanie liczby kanałów w miarę zmniejszania rozdzielczości w kolejnych warstwach. Czy nie lepiej sprawdziłaby się sieć U-Net z mniejszą liczbą warstw, ale zachowująca tę zasadę? Na obronie prosiłbym o przedstawienie przekonujących wyników, które pozwolą zweryfikować tę hipotezę.
3. Czy próbował Pan testować więcej punktów dla linii środkowej (np. 4 lub 5) w P6?
4. Dlaczego wyniki w P6 dla linii środkowej są niewiele lepsze od wyników dla pól ograniczających? Jakie widzi Pan kolejne kroki, aby wykorzystać potencjał nowego rozwiązania?
5. Czy Rysunek 9 w P4 był wygenerowany dla przykładów z bazy syntetycznej? Jeżeli tak, to faktyczna dokładność metody nie jest chyba znana.
6. Dlaczego P5 została opublikowana w IEEE Access, a P6 w ACIIDS? Wydaje mi się, że obie publikacje mają większy potencjał publikacyjny. Czy chodzi o niewystarczające wyniki na standardowych benchmarkach?
7. Czy obrazy mączniaka wygenerowane w P5 są zawsze tak realistyczne? Jeżeli nie, czy może Pan przedstawić kilka mniej udanych obrazów?
8. Jak dodatkowa sieć klasyfikująca górną i dolną stronę ziarna w P2 wpływa na możliwość wykorzystania proponowanej techniki w czasie rzeczywistym?
9. Proszę dokładnie opisać proces anotacji 500 danych treningowych i walidacyjnych do treningu sieci klasyfikującej górną i dolną stronę ziarna w P2. Proszę także opisać sposób postępowania w przypadku bocznego ułożenia ziarna (ang. sidewise grain orientation), prezentowanego na rysunku 3 w P2.

4. Wiedza kandydata

Na podstawie lektury uważam, że doktorant posiada ugruntowaną wiedzę z zakresu informatyki technicznej i telekomunikacji, w szczególności w obszarze metod klasyfikacji i segmentacji opartych na głębokich sieciach neuronowych. Doktorant poprawnie posługuje się zaawansowanym aparatem matematycznym, a także potrafi zaplanować i przeprowadzić eksperyment komputerowy w celu oceny jakości proponowanych metod.

Przegląd literatury dotyczącej zagadnień przedstawionych w pracach doktoranta pozwala stwierdzić, że doktorant posiada aktualną wiedzę w zakresie tematyki rozprawy, a także potrafi dokonać krytycznego przeglądu źródeł w celu wskazania interesujących kierunków badań.

5. Podsumowanie

Recenzowana dysertacja przedstawia rozwiązanie ważnych i oryginalnych problemów, stanowiąc przykład udanego wykorzystania metod uczenia maszynowego w obrazach przemysłowych.

Uwagi przedstawione w punkcie 3 recenzji mają jedynie charakter dyskusyjny i w żaden sposób nie deprecjonują osiągniętych przez doktoranta rezultatów ani nie wpływają na pozytywne wrażenie o przedłożonej rozprawie.

Dlatego, biorąc pod uwagę powyższe, stwierdzam, że rozprawa mgra inż. Przemysława Dolaty pt. „Deep learning methods for computer vision-based industrial inspection applications” spełnia wymagania ustawy Prawo o szkolnictwie wyższym i nauce (Dz. U. 2024, poz. 1571). W związku z tym wnoszę o jej przyjęcie i dopuszczenie mgra inż. Przemysława Dolaty do publicznej obrony.

Z poważaniem,

