

Częstochowa, dn. 7 marca 2025 r.

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

Recenzja

rozprawy doktorskiej mgra Jacka Kałużnego, pt.: Generowanie i walidacja syntetycznych zbiorów danych obrazów do trenowania modeli sztucznej inteligencji w wizji komputerowej.

Niniejszą recenzję opracowano zgodnie z uchwałą Rady naukowej dyscyplin matematyka i informatyka UAM z dnia 22/10/2024. Promotorem jest prof. UAM dr hab. Krzysztof Dyczkowski, a promotorem pomocniczym dr Wojciech Pałubicki.

1. Charakterystyka tematu, celu i tezy badawczej rozprawy

Generowanie danych syntetycznych oraz ich integracja z danymi rzeczywistymi odgrywa kluczową rolę w uczeniu maszynowym, ponieważ pozwala na przewyższenie ograniczeń związanych z dostępnością, jakością i różnorodnością danych, które są niezbędne do skutecznego trenowania modeli. W wielu zastosowaniach rzeczywiste dane mogą być trudne do pozyskania, niekompletne, obciążone błędami lub niezrównoważone, co negatywnie wpływa na wydajność i generalizację modeli. Dzięki generowaniu danych syntetycznych możliwe jest uzupełnienie brakujących informacji, zwiększenie różnorodności próbek, a także symulowanie rzadkich lub ekstremalnych scenariuszy, które są trudne do uchwycenia w rzeczywistych zbiorach danych. Integracja danych syntetycznych z rzeczywistymi pozwala na tworzenie bardziej reprezentatywnych i kompleksowych zestawów treningowych, co z kolei prowadzi do poprawy jakości modeli. Recenzowana praca stanowi wkład w generowanie złożonych danych w dziedzinie grafiki trójwymiarowej.

2. Zawartość rozprawy

Recenzowana praca mgra Jacka Kałużnego składa się z pięciu rozdziałów, spisu rysunków, tabel oraz bibliografii. Dokument liczy 96 stron.

Pierwszy rozdział jest krótkim wprowadzeniem do tematyki, zawiera trzy bardzo ogólne hipotezy: 1. Sztuczne sieci neuronowe mogą być wykorzystywane do zbierania, wyodrębniania

i analizy danych; 2. Modele proceduralne generujące zbiory danych syntetycznych są rozwiązaniem problemu małych zbiorów danych; 3. Integracja danych syntetycznych z rzeczywistymi zbiorami danych zwiększa odporność i uogólnialność modeli uczenia maszynowego, oraz omawia zawartość całej pracy.

Rozdział drugi prezentuje metodę odtworzenia geometrii z pojedynczego zdjęcia dwuwymiarowego. W tym celu zaproponowano modelowanie formy drzewa za pomocą struktury Radial Bounding Volume (RBV), składającej się z cylindrów, warstw i sektorów. Stworzono syntetyczny zestaw drzew za pomocą renderingu bazującego na fizyce z użyciem oświetlenia Blinn-Phong w silniku Unreal 4. Stworzony zbiór danych użyty został do uczenia sieci DeepLab-V3 segmentacji semantycznej. Zbiór poddano szerokiej augmentacji, takiej jak zmiana konfiguracji renderingu, zmiana pozycji kamer, oświetlenia, zmiany koloru, jasności i kontrastu, rozmycia Gaussa, odbicie i przycinanie obrazu. Do tworzenia struktur RBV użyto autorskiej sieci splotowej o czterech „głowach” (heads) tworzącej RBV w czterech rozdzielczościach, korzystając z osadzenia (embedding) obrazu drzewa. Z wektora osadzenia korzysta również „głowa” klasyfikująca obraz drzewa na gatunek. Wzrost drzewa odbywa się za pomocą dwóch algorytmów z literatury. Stworzono również dynamiczne modele drzew z wykorzystaniem prętów Cosserata. Przeprowadzono uczenie zaproponowanych modeli przy różnych metaparametrach oraz porównano z metodami z literatury. Podobieństwo zrekonstruowanych drzew do oryginałów zbadano za pomocą ankiet na platformie Mechanical Turk.

Rozdział 3 proponuje metodę przewidywania obszaru liścia wraz z autorskim zbiorem danych. Pierwszym etapem jest generowanie papieru milimetrowego za pomocą autorskiej funkcji z różnymi ustawieniami modułu cieniującego i rodzajami szumu. Następnie generowany jest model liścia za pomocą parametrycznej krzywej wielomianowej, tekstuowania za pomocą ruchu żółwia wraz elementem stochastycznym. Liście przechodzą cztery etapy renderowania i dodawane są dodatkowe elementy w celu zwiększenia realizmu. Obrazy syntetyczne zostały użyte do generowania fotorealistycznych liści za pomocą modelu ControlNet. Do tworzenia powierzchni liści użyta została sieć MobileNet V3.

Rozdział 4 omawia metodę tworzenia mapy głębokości sadzonek. Autor stosuje modele U-Net i Depth Anything. Modele uczone są na syntetycznym zbiorze danych wygenerowanym za pomocą systemu Lindenmayera w środowisku Unity. Modele neuronowe oryginalne i metodą autora zostały zwizualizowane i porównane z głębokością z lidar.

Rozdział 5 podsumowuje rozprawę. Dalej następuje bibliografia składająca się z 99 pozycji.

3. Ocena rozprawy

W ramach rozprawy doktorskiej Pan mgr Jacek Kałużny zaproponował oryginalne rozwiązanie związane z generowaniem danych syntetycznych. Tematyka pracy jest bardzo aktualna i potrzebna, oryginalny dorobek autora polega na stworzeniu rozbudowanych metod:

- tworzenia realistycznych modeli drzew na podstawie zdjęć dwuwymiarowych,
- generowania fotorealistycznych obrazów liści roślin,
- tworzenia mapy głębokości zdjęcia dwuwymiarowego.

Rozprawa doktorska uwidacznia wysoką ogólną wiedzę teoretyczną i praktyczną oraz umiejętność prowadzenia pracy naukowej mgra Jacka Kałużnego, który współpracował z bardzo dobrymi zespołami naukowymi oraz użył nowoczesnych modeli sieci neuronowych i metod grafiki komputerowej. Opracował wprowadzenie do tematyki i przegląd literatury na temat metod wizji komputerowej 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:

- Hipotezy stawiane w pracy są enigmatyczne i zakładają stworzenie metod ogólnych, a rozprawa porusza bardzo wąską tematykę grafiki komputerowej, a w niej drzew i liści. Hipoteza 1. *Sztuczne sieci neuronowe mogą być wykorzystywane do zbierania, wyodrębniania i analizy danych* jest obecnie oczywista i udowodniona w niezliczonej liczbie publikacji. Hipoteza 2. *Modele proceduralne generujące zbiory danych syntetycznych są rozwiązaniem problemu małych zbiorów danych* dotyczy tylko pewnych obszarów. Na przykład może nie mieć zastosowania w cyberbezpieczeństwie, medycynie, finansach itp. Hipoteza 3. *Integracja danych syntetycznych z rzeczywistymi zbiorami danych zwiększa odporność i uogólnialność modeli uczenia maszynowego* również dotyczy wybranych obszarów zastosowań uczenia maszynowego.
- Tytuł rozprawy, tak jak hipotezy, jest nadmiernie ambitny – rozprawa nie rozwiązuje w sposób uogólniony problemu małej ilości danych we wszystkich aspektach wizji komputerowej.
- Rozdział 2 jest dokładnym tłumaczeniem pracy Li, Bosheng, Jacek Kałużny, Jonathan Klein, Dominik L. Michels, Wojtek Pałubicki, Bedrich Benes, and Sören Pirk. "Learning to reconstruct botanical trees from single images." *ACM Transactions on Graphics (TOG)* 40, no. 6 (2021): 1-15, bez odniesienia się do niej. Natomiast w powyższym artykule nie ma podanego podziału pracy na siedmiu autorów.
Rozdział 3 jest przetłumaczonym artykułem Kaluzny, Jacek, Yannik Schreckenberg, Karol Cyganik, Peter Annighöfer, Sören Pirk, Dominik L. Michels, Mikołaj Cieslak, Farhah Assaad-Gerbert, Bedrich Benes, and Wojciech Palubicki. "LAESI: Leaf Area Estimation with Synthetic Imagery." *CoRR* (2024) bez odniesienia się do niej. W artykule nie ma podanego podziału pracy na wszystkich dziesięciu autorów.
Prosiłbym o podanie podziału pracy w obu przypadkach.
- Rozdział 2 i 3 sprawiają wrażenie tłumaczenia maszynowego z powyższych publikacji (bezosobowo, z tłumaczeniem np. „coarser resolution” jako „zgrubszą rozdzielczość” lub „atrous rates” jako „Wartości stóp atrous”, użyciem kalek z oryginału, np. „intensywne testy”), a rozdział 4 pisanego oryginalnie po polsku (w pierwszej osobie i obfitującego w

wiele błędów literowych i ortograficznych („rozszeżenie”, „wykożystanie”, „osiągnieća”).

- Metody opisane w rozprawie używają bardzo rozbudowanych technik i modeli, jak Stable Diffusion, ControlNet, LoRA czy Depth Anything. Poprzez ograniczenie do bezpośredniego przetłumaczenia wspomnianych wyżej publikacji, brak jest dokładnego omówienia użytych metod i osadzenia ich idei w całości rozwiązania.
- Praca [5] Author, U. (2024). Depth anything: Scaling depth estimation for computer vision. Journal of Artificial Intelligence Research, pages 100–120 chyba nie istnieje.
- Sieć DeepLab-V3 użyta jest do segmentacji semantycznej czy segmentacji instancji (obiektów)?
- Na stronie 55 jest zdanie “Program ten umożliwia proceduralne generowanie drzew w oparciu o L-Systemy, co pozwala na tworzenie złożonych, realistycznych struktur roślinnych, które są następnie używane do generowania syntetycznych obrazów i map głębokosci 2.5.”
- Na stronie 57 przy omawianiu sieci U-Net jest referencja do pracy [18] Eigen, D. and Fergus, R. (2014). Depth map prediction from a single image using a multi-scale deep network. Advances in Neural Information Processing Systems, 27, która chyba nie dotyczy zupełnie tego modelu.
- W rozdziale 4 nie jest jasne umiejscowienie sieci U-Net w testowaniu metody. Dokładnie umiejscowiona jest tylko sieć ControlNet na schemacie na Fig. 4.1.
- Rysunki nazywane są „Fig.”
- Skrót PBR na stronie 12 nie jest wyjaśniony.
- Rysunek 2.9 po przetłumaczeniu z oryginalnej wersji ma poprzesuwane niektóre znaki.

4. Wnioski końcowe recenzji

Podsumowując recenzję stwierdzam, że rozprawa doktorska „Generowanie i walidacja syntetycznych zbiorów danych obrazów do trenowania modeli sztucznej inteligencji w wizji komputerowej” prezentuje oryginalne rezultaty stanowiące rozwiązanie problemu naukowego oraz wkład w rozwój dyscypliny informatyka. Pan Jacek Kałużny wykazał się umiejętnością samodzielnej pracy badawczej, znajomością literatury światowej i wiedzą w zakresie uczenia maszynowego, grafiki komputerowej oraz wizji komputerowej. Rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie informatyka. 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. Wnoszę o jej przyjęcie i dopuszczenie do dalszych etapów postępowania doktorskiego.