

Recenzja

rozprawy doktorskiej mgr inż. Kamila Bolka pt.: „Wyznaczanie dynamicznej mapy 2D na podstawie wielomodalnych danych rejestrowanych przez urządzenia mobilne”

Promotor rozprawy: dr hab. inż. Adam Świtoński prof. Pol. Śl.

1. Aktualność i znaczenie tematyki rozprawy

Rozprawa dotyczy problemu generacji dwuwymiarowych map terenu. Podjęcie takiej tematyki jest niezwykle trudnym zadaniem wobec bogatej oferty map klasycznych, a przede wszystkim łatwej dostępności takich narzędzi jak Mapy Google, Google Earth, czy Street View. Z drugiej strony połączenie technologii GPS z mapami cyfrowymi pozwala na natychmiastowe umiejscowienie na mapie osoby szukającej informacji i udostępnienie jej szerokiej wiedzy o najbliższym otoczeniu, przede wszystkim zaś GPS i mapa cyfrowa tworzą podstawę powszechnie używanych systemów nawigacji.

Nowym elementem w tym obszarze, którego wprowadzenia podjął się autor rozprawy jest uwzględnienie zmian w otoczeniu prezentowanym na mapie. Współczesne mapy są wprawdzie stale aktualizowane, ale autor proponuje rozwiązania pozwalające na odnotowanie w systemie mapy cyfrowej zdarzeń, których szybkość zmian odpowiada np. dynamice ruchu ulicznego. Uzupełnianie mapy o takie informacje pozwoli jej nadać cechy temporalne, które można wykorzystać do analiz zmian stanu otoczenia prezentowanego na mapie w określonych momentach czasu, w tym np. obecność rowerów lub zatłoczenie ulic w określonych godzinach dnia. Tak stworzona funkcja wizualizacji na mapie zmian, jakie zachodziły na wybranym obszarze, może być np. wykorzystana w planowaniu przestrzennym danego rejonu miasta.

Przedstawione rozwiązanie, uwzględniające dane temporalne w procesie tworzenia i wykorzystania mapy, jest projektem unikalnym, dla którego brak odniesienia w literaturze.

W założeniach projektu przyjęto, że głównym źródłem danych dla projektowanego systemu będą obrazy z kamer urządzeń mobilnych (telefonów komórkowych) razem z danymi z urządzeń GPS i czujników położenia IMU tych telefonów. Realizacja postawionego celu wymagała opracowania algorytmów umożliwiających m.in. detekcję obiektów na przesyłanych obrazach a następnie lokalizację tych obiektów w terenie, integrację informacji o wykrytych obiektach pochodzących z wielu źródeł, a także opracowanie struktury bazy danych przechowującej mapy umożliwiające zapis historii wykrywanych zmian oraz wizualizację danych przestrzennych dla potrzeb analiz temporalnych.

W wykonanym systemie autor zastosował m.in. sieci neuronowe przetwarzające obrazy obiektów na urządzeniach mobilnych, a do przechowywania mapy i analizy gromadzonych danych przestrzennie-czasowych wykorzystał grafową bazę danych.

Podjęcie przez autora opirionwanej rozprawy takiej tematyki badań oceniam jako zadanie ambitne, a przy tym trudne. Tematyka ta wpisuje się w obszar badań dotyczących nowych technologii i zastosowań map cyfrowych. Uważam, że rozprawa dotyczy ważnej i aktualnej problematyki badawczej.

2. Zakres pracy

Praca składa się z pięciu rozdziałów i bibliografii.

W rozdziale pierwszym autor przedstawił motywacje i cel pracy, a następnie przyjęte tezy. Podstawowa teza ma charakter ogólny: „*Możliwe jest stworzenie mapy 2D na podstawie wielomodalnych danych przestrzennych i czasowych, przesyłanych z wielu źródeł, która dostarczy informacji o położeniu obiektów statycznych i dynamicznych (tj. zmieniających swoje położenie) znajdujących się na wybranym obszarze.*”, dlatego autor dodał cztery tezy pomocnicze, uściślające zakres tematyczny podejmowanych badań. W pierwszej z tych tez autor zapowiada możliwość lokalizacji wykrywanych na obrazach obiektów z dokładnością do 10 m, w drugiej przewiduje poprawę dokładności lokalizacji obiektów przez integrację danych z wielu pomiarów, w trzeciej zapowiada uzyskanie, przy zastosowaniu głębokich sieci neuronowych działających na urządzeniach mobilnych, prędkości detekcji obiektów na obrazach nie mniejszej niż 3 klatki na sekundę, w czwartej natomiast przewiduje przyspieszenie wyszukiwania informacji o mapach poprzez przeniesienie stosownej bazy danych z relacyjnego do grafowego systemu zarządzania bazą danych.

Oceniając pozytywnie przedstawione tezy wyraziłbym opinię, że bardziej naturalne byłoby sformułowanie tez pomocniczych w nieco ogólniejszej postaci. Podanie w tezach konkretnych wartości (10 m, 3 klatki na sekundę) stwarza bowiem wrażenie, że tezy sformułowano już po przeprowadzeniu badań eksperymentalnych, co odwraca sens idei stawiania tez i wyznaczania celu badań przed ich rozpoczęciem.

Rozdział drugi jest poświęcony omówieniu istniejących rozwiązań. Jak wspomniałem w poprzednim punkcie recenzji, przedstawiany w pracy projekt ma charakter unikalny, dlatego autor nie odwołuje się w tym rozdziale do żadnego systemu o podobnych funkcjach. Natomiast dokonuje przeglądu metod, spośród których musiał dokonać wyboru dla realizacji głównych funkcji zrealizowanego systemu.

Pierwszą grupę diskutowanych metod stanowiły algorytmy badania głębi obrazu, wykorzystane przez autora do oceny odległości między czujnikiem (np. obiektywem kamery), a obiektem na obrazie. W przedstawionym przeglądzie autor wyróżnił podejście oparte na nadzorowanym uczeniu maszynowym, wykorzystujące głębokie sieci neuronowe, oraz podejście oparte na wykorzystaniu samonadzorującego uczenia maszynowego. Metody te określają relacje odległości między obiektami na obrazie. Przedstawioną następnie alternatywą do tych metod jest bezpośredni pomiar odległości z wykorzystaniem lidar lub też szacowanie odległości do obiektów o znanych rozmiarach z wykorzystaniem znajomości parametrów kamery.

Druga grupa rozważanych metod dotyczyła detekcji obiektów na obrazach z kamery. Autor omówił metody wykorzystujące głębokie sieci neuronowe, rolę procesorów graficznych (GPU) w tych metodach oraz technikę kwantyzacji sieci, stosowaną m.in. przy wprowadzaniu sieci głębokich na urządzenia mobilne o ograniczonych zasobach sprzętowych.

Jako rozwinięcie poprzedniego tematu została przedstawiona dyskusja metod detekcji poruszających się obiektów na obrazie z kamery. Ostatnim tematem rozważanym w drugim rozdziale pracy było wykorzystanie języka znaczników KML do wizualizacji danych geograficznych, z uwzględnieniem danych czasowych.

Rozdział trzeci jest głównym rozdziałem pracy i zawiera zarówno opis koncepcji opracowanego systemu wizualizacji mapy na podstawie danych temporalnych

i wielomodalnych, jak i jego realizacji. Zadaniem systemu jest aktualizacja mapy na podstawie nadsyłanych na bieżąco danych zbieranych ze źródeł nazwanych w pracy urządzeniami brzegowymi (telefony komórkowe, mikrokontrolery oraz kamery rozmieszczone na danym obszarze). Mapa jest aktualizowana z zachowaniem starszych danych i czasów ich aktualizacji, co pozwala na wizualizację mapy pod kątem zachodzących w czasie zmian obecności lub położenia obiektów w zadanym obszarze.

System ma budowę wielowarstwową - poszczególne warstwy zapewniają zbieranie danych, ich analizę, integrację i porównanie z danymi obecnymi w bazie danych, aktualizację bazy danych, a wreszcie wyszukiwanie i udostępnianie danych użytkownikom końcowym.

Pierwsza warstwa realizuje pozyskiwanie danych z urządzeń brzegowych. Przy wykorzystaniu sieci neuronowych, instalowanych bezpośrednio na urządzeniach posiadających system operacyjny Android lub iOS, jest realizowana detekcja obiektów w obszarze obrazu z kamery oraz segmentacja tego obrazu w celu identyfikacji nowych klas obiektów.

Głównym zadaniem opracowanej dla tej warstwy aplikacji jest dokładna lokalizacja widocznego w kamerze obiektu. Zastosowano w tym celu algorytm badania głębi obrazu. Oprócz informacji z obrazu do wyznaczenia położenia wykrytych obiektów wykorzystuje się parametry wewnętrzne kamery, dane z czujnika IMU telefonu komórkowego oraz lokalizację i margines błędu urządzenia GPS. Autor szczegółowo przeanalizował dokładność wyznaczania położenia przez systemy GPS i wzbogacił te analizy wykonanymi w terenie eksperymentami pomiarowymi.

Drugim szczegółowo analizowanym parametrem aplikacji pozyskiwania danych była prędkość przetwarzania pojedynczej klatki dla obrazów obiektów ruchomych. Autor przeanalizował szereg implementacji sieci neuronowych na smartfonach, biorąc pod uwagę dostępność na nich procesorów graficznych (GPU). Dla pięciu wybranych sieci przeprowadzono eksperymenty z pomiarem czasu przetwarzania jednej klatki obrazu. Analiza statystyczna wyników pomiarów potwierdziła jedną z tez pracy, dotyczącą wystarczającej prędkości przetwarzania obrazów na smartfonach przy wykorzystaniu sieci neuronowych (przy czym w opisie procesu weryfikacji postawionej hipotezy popełniono błąd formalny, opisany w punkcie 4 tej recenzji) . Oddzielne badania zostały przeprowadzone dla kamer instalowanych na skrzyżowaniach dróg lub w miejscach użyteczności publicznej. Parametry takich kamer ograniczały możliwość zainstalowania na nich poprzednio wykorzystywanych sieci neuronowych.

Następna warstwa opracowanego systemu została zrealizowana na serwerze, stanowiącym centralną część systemu. Zadania tej warstwy to: odbiór danych nadsyłanych z wielu urządzeń brzegowych, porównanie danych z przechowywanymi w bazie danych, zapis nowych danych do bazy danych. Na podstawie przeprowadzonych testów czasu odbioru komunikatów autor dobrał serwerowe systemy obsługujące komunikację z urządzeniami brzegowymi. Minimalizacja czasu zapisu napływających danych do bazy danych, a następnie odczytu tych danych z bazy były również podstawą wyboru systemu zarządzania bazą danych przechowującego mapy oraz dane zbierane z urządzeń brzegowych. O ile skupienie uwagi na bazach NoSQL, w tym na bazach grafowych uważam za wybór interesujący i obiecujący, o tyle przedstawiona argumentacja odrzucenia baz relacyjnych wydaje mi się zbyt powierzchowna, do czego wrócę w uwagach szczegółowych. Po wstępnej selekcji systemów NoSQL z uwagi na takie cechy jak kontrola dostępu do zasobów, wsparcie przez wiele języków programowania, darmowe wykorzystanie z systemu zarządzania bazą danych, autor wybrał dwie bazy grafowe i po przeprowadzeniu eksperymentalnych pomiarów czasów dostępu do tych baz jego wybór padł na system OrientDB.

Jak wspomniano już wyżej, główne zadania opracowanego systemu to: lokalizacja obiektów widocznych w kamerach urządzeń brzegowych, rejestrowanie zmian na mapie

przechowywanej w bazie danych, a następnie wizualizacja mapy dla wybranego regionu w wybranym czasie. Opracowane algorytmy lokalizacji obiektów obejmują dwa etapy: lokalizację po stronie urządzenia brzegowego i lokalizację po stronie serwera.

Dla pierwszego etapu, po detekcji obiektu na urządzeniu brzegowym istotne jest rozstrzygnięcie, czy dany obiekt stanowi element, który już wcześniej był wykryty przez urządzenie brzegowe lub też czy jest to nowy obiekt, który wcześniej nie był rozpoznawany przez sieć neuronową osadzoną na tym urządzeniu. Dla zrealizowania tego zadania przeprowadzono wybór algorytmu śledzenia obiektów na podstawie eksperymentalnej oceny szybkości przetwarzania klatek obrazu. Opracowano też zestaw reguł określających czy dany obiekt jest tym, który znaleziono już wcześniej. Są to oddzielne reguły dla obiektów nieruchomych oraz ruchomych.

Lokalizacja danych po stronie serwera skupia się na synchronizacji danych pozyskanych równolegle od wielu użytkowników dotyczących określonego obszaru z danymi przechowywanymi w bazie danych. Opracowano algorytmy synchronizacji danych dla obiektów nieruchomych oraz ruchomych, umożliwiające m.in. odnotowywanie w bazie danych czasów pojawiania się lub znikania na danym obszarze wykrytych obiektów, a także uściślanie położenia tych obiektów.

Ostatnia część trzeciego rozdziału pracy jest poświęcona optymalizacji procesu wyszukiwania danych pochodzących z map systemu Open Street Map, wykorzystywanych w ocenianej rozprawie. Mapy te w oryginale są przechowywane w relacyjnej bazie danych PostgreSQL, wymagającej dość kosztownej konfiguracji sprzętowej serwera bazy danych. Dla obniżenia wymagań sprzętowych, czyli kosztów całego systemu, a zarazem przyspieszenia dostępu do danych, autor zastąpił tabelaryczną strukturę bazy map w relacyjnym systemie PostgreSQL strukturą grafu w bazie danych OrientDB, reprezentującej grafowe bazy danych typu NoSQL. Opracowano nową strukturę bazy danych zawierającą, obok poprzednio istniejących klas opisujących elementy mapy, nowo utworzoną klasę o nazwie Area, obejmującą obszary grupujące w bazie elementy mapy znajdujące się w danej lokalizacji. Przyspiesza to znalezienie elementów mieszczących się na wyświetlanym obszarze. Dalsze przyspieszenie wyświetlania map uzyskano dzięki użyciu biblioteki obsługującej skompresowane pliki binarne bez potrzeby ich konwersji. Dodatkowo wykorzystano w bazie danych OrientDB mechanizm tworzenia klastrów grupujących w sobie węzły mapy znajdujące się w danym obszarze. Pozwala to ograniczyć zakres wyszukiwania elementów mieszczących się na wyświetlanym obszarze, zazwyczaj zawierającym się wewnątrz jakiegoś klastra. Ponadto w procesie klasteryzacji grupy węzłów znajdujących się blisko siebie zastępowano jednym nowym węzłem, co pozwoliło na wyraźne zmniejszenie objętości bazy danych. Nowa baza danych zawiera też dane temporalne, określające momenty czasu, w których wykryto zmiany obiektów na mapie.

W czwartym rozdziale pracy przedstawiono przeprowadzone eksperymenty weryfikujące prawidłowość i szybkość działania wykonanego systemu.

Poprawność detekcji obiektów sprawdzono w eksperymencie rozpoznawania tablic rejestracyjnych. Dokładność algorytm estymacji odległości oceniono poprzez porównanie wyników realizacji algorytmu z pomiarami referencyjnymi przeprowadzonymi w terenie.

Wykonany prototyp aplikacji mobilnej dla urządzeń z systemem Android lub iOS został sprawdzony w zadaniu określenia lokalizacji obiektów. Eksperymenty przeprowadzono w środowisku testowym oraz w warunkach przewidywanego użytkownika systemu, w pracy zaprezentowano przykład skanowania hydrantu na ulicy. Działanie algorytmu zostało ocenione jako poprawne, przy czym interpretacja przedstawionych wyników jest dosyć uboga (np. na rys. 4.5 i 4.6).

Druga grupa eksperymentów miała na celu sprawdzenie poprawności i dokładności algorytmu integrującego pomiary położenia obiektów z wielu źródeł. Poprawność algorytmu

sprawdzono na danych symulowanych, natomiast dokładność na danych pochodzących z ponad tysiąca pomiarów odległości dla 10 klas obiektów (przy czym w opisie procesu weryfikacji postawionej hipotezy popełniono błąd formalny, opisany w punkcie 4 tej recenzji). Przedstawiono też analizę wyników potwierdzających poprawność rejestrowania danych dotyczących obiektów ruchomych. Eksperymenty tej grupy potwierdziły też tezę, że zwiększanie liczby pomiarów i integracja ich wyników poprawia dokładność lokalizacji obiektów.

Ostatni z przeprowadzonych eksperymentów dotyczył porównania czasów wyświetlania losowo wybieranych obszarów z map przechowywanych w bazie danych PostgreSQL, bazie OrientDB oraz bazie OrientDB po wcześniej wykonanej klasteryzacji. Wyniki podzielone zostały na trzy grupy: najlepszy, średni i najgorszy czas odpowiedzi, przy czym nie podano dokładnych kryteriów takiego podziału. W dwóch pierwszych grupach najkrótsze czasy uzyskano dla opracowanej przez autora struktury bazy wykorzystującej model grafowy (system OrientDB) z zastosowaniem klasteryzacji.

Ostatni rozdział rozprawy zawiera podsumowanie uzyskanych wyników i prezentację wniosków z realizacji pracy.

Autor wskazał w podsumowaniu szereg cech i zastosowań opracowanego systemu, klasyfikując go jako produkt komercyjny. Dokumentacja rozprawy nie daje jednak, moim zdaniem, podstaw do oceny komercyjnych cech opracowanej aplikacji, gdyż nie zawiera ona (i słusznie) takich elementów jak np. specyfikacja zewnętrzna (instrukcja użytkownika), wymaganych w prezentacji produktów komercyjnych.

3. Ocena pracy

Recenzowana rozprawa doktorska wnosi poważny wkład w rozwój narzędzi i metod tworzenia map 2D przy wykorzystaniu danych zbieranych przy pomocy urządzeń mobilnych.

Głównym osiągnięciem autora jest, w mojej ocenie, opracowanie samej koncepcji budowy mapy z uwzględnieniem temporalnych zmian w prezentowanym na mapie otoczeniu, a następnie zaprojektowanie i zbudowanie systemu zbierania i przetwarzania danych przestrzennych oraz obsługi takiej mapy.

Szczegółowe osiągnięcia autora, związane z realizacją poszczególnych etapów tego systemu, obejmują:

- 1) Opracowanie algorytmu lokalizacji obiektu wykrytego przez sieć neuronową w obrazie uzyskanym w kamerze urządzeń mobilnych – lokalizacja jest wyznaczana na podstawie informacji o wykrytych na obrazie obiektach, parametrach kamery, danych z czujnika położenia urządzenia mobilnego oraz lokalizacji urządzenia GPS i jego dokładności.
- 2) Opracowanie algorytmu integracji danych o lokalizacji obiektów, uzyskanych z wielu źródeł – algorytm zapewnia poprawienie dokładności lokalizacji, umożliwia też wykrywanie zmian w obecności i lokalizacji obiektów na mapie oraz powiązanie tych zmian z danymi temporalnymi zapisywanymi do bazy danych.
- 3) Opracowanie struktur danych umożliwiających przechowywanie map według wymagań grafowego modelu bazy danych, z uwzględnieniem danych temporalnych, a następnie przeniesienie do takiej bazy mapy z relacyjnej bazy danych.
- 4) Opracowanie koncepcji przyspieszenia wyszukiwania informacji w bazie danych map, poprzez utworzenie klastrów grupujących węzły znajdujące się w wyróżnionych obszarach map.

Autor zaimplementował wszystkie algorytmy i przeprowadził ich eksperymentalną ocenę, która potwierdziła spełnienie tez postanowionych w pierwszym rozdziale pracy. W pracy nie przedstawiono przy tym technicznego opisu zbudowanego systemu.

Realizacja przedstawionych algorytmów i rozwiązań programowych świadczy o głębokiej znajomości zagadnień przetwarzania obrazów i danych przestrzennych oraz organizacji map cyfrowych. Specyfiką realizowanego projektu była konieczność dopasowania opracowanych lub też wybieranych algorytmów i metod do możliwości sprzętowych urządzeń mobilnych, służących nie tylko jako podstawowe źródło danych przestrzennych, ale również miejsce pierwszego, a zarazem bardzo złożonego etapu przetwarzania tych danych.

Moje uwagi krytyczne odnoszą się głównie do strony redakcyjnej pracy. Pierwsza uwaga dotyczy zawartości trzeciego rozdziału pracy. Autor umieścił w nim opis zarówno koncepcji jak i realizacji całego projektu. Interesujące byłoby podjęcie próby rozdzielania tych etapów, a w szczególności oddzielenia opisu algorytmów i metod od opisu badań oceniających te metody, a także opisu programowej implementacji wybranych funkcji bądź formatów danych.

Druga ogólna uwaga krytyczna dotyczy języka i stylu pisania tekstu. W pracy rażą wyrażenia żargonowe, nadużywanie terminów w języku angielskim, bardzo niedbała interpunkcja (brak przecinków, zwłaszcza w zadaniach złożonych podrzędnie), niedbały styl. Wydaje się, że tekst pracy był pisany pośpiesznie, bez należytej samokontroli.

Trzecia uwaga jest pytaniem o publikowanie wyników pracy. W bibliografii nie znalazłem żadnej publikacji z nazwiskiem autora, nie zauważyłem też informacji na ten temat w tekście pracy. Czy autor podjął próby opublikowania wyników swej rozprawy?

Dalsze szczegółowe uwagi merytoryczne, a potem redakcyjne, zamieściłem w kolejnym punkcie recenzji.

4. Uwagi do pracy

Lektura tekstu pracy nasunęła mi kilka kolejnych uwag:

- 1) Str. 3 i dalsze: Nie znalazłem w pracy objaśnienia użytego w tytule i wielokrotnie w tekście rozprawy określenia „wielomodalne dane”. Ponieważ nie jest to termin powszechnie używany, proponuję skomentować jego znaczenie.
- 2) Str. 44: Odrzucenie z rozważań relacyjnych baz danych zostało uzasadnione dość powierzchownie następującym zdaniem: „Z uwagi na specyfikę projektu zapytania kierowane do bazy będą dotyczyły obszarów dzięki czemu wydajniejszą metodą jest sięganie do konkretnych wierzchołków grafu niż przeszukiwanie i łączenie tabel.” Ta sugestia jest słuszna, jeśli opis wyszukiwanych obszarów wymaga użycia kilku tabel w relacyjnej bazie danych. Pełniejsze uzasadnienie mogłoby więc zawierać opis struktur danych tworzących mapę w relacyjnym modelu danych.
- 3) Str. 35: W opisach procesów weryfikacji hipotez w testach przedstawionych na stronach 35 oraz 82 popełniono błąd formalny, gdyż obliczone wartości statystyk testowych U mieszczą się w wyznaczonych zbiorach krytycznych $(-1, -u_{(1-\alpha)}]$ tj. $(-1, -1,6449]$.
- 4) Str. 47 i dalsze: Celowe byłoby objaśnienie, co obejmuje klasteryzacja baz danych OrientDB oraz Neo4j, gdyż pojęcie klastra i klasteryzacji ma w informatyce kilka nieco różniących się znaczeń.

- 5) Już na pierwszych stronach pracy pojawia się określenie „serwer” bez bliższego objaśnienia tego terminu. Ponieważ ten termin może się odnosić zarówno do samego sprzętu (komputer o określonych parametrach i funkcjach), jak i oprogramowania (oprogramowanie o specjalnych zadaniach, np. serwer bazy danych), dlatego byłoby celowe objaśnienie zakresu znaczenia tego terminu przy pierwszym jego użyciu w pracy.
- 6) Na stronach 28, 48, 78 pojawiają się określenia „długość i szerokość geograficzna”, natomiast na wielu innych stronach pracy (np. strony 49, 56, 58, 66, ...) używane są określenia „wysokość i szerokość geograficzna” – skąd te różnice?
- 7) Str. 34-35: Użyte sformułowanie „wartość oczekiwana czasu przetwarzania prędkości głębokich sieci neuronowych” nie jest jednoznaczne.
- 8) Str. 57: W zapisie reguły 3 pojawia się zdanie: „Jeśli na podstawie tej reguły określono, że obiekt jest tym samym co zapisany w bazie to następuje doprecyzowanie lokalizacji obiektu w bazie danych.” – skąd pewność, że nowa lokalizacja jest bardziej dokładna?
- 9) Str. 3: Przedstawiając ogólnie cel pracy autor zapowiedział m.in. takie jego cechy: „Docelowe rozwiązanie będzie oferowało API do łatwej integracji z zewnętrznymi systemami oraz graficzny interfejs użytkownika (GUI) pozwalające na wizualizacje przetwarzanych przez system danych.” Trudno potwierdzić zrealizowanie tych zapowiedzi, gdyż te elementy opracowanego systemu nie zostały przedstawione.
- 10) W pracy nie zamieszczono spisu używanych skrótów wraz z objaśnieniami.

Uwagi redakcyjne

Str. vi: w spisie treści jest punkt „Załączniki”, nie znalazłem ich jednak na wskazanej stronie.

Str. 7: jest: „... na skutek ich rozmycia” – powinno być: „... na skutek ich rozmycia”; podobna uwaga do str. 12.

Str. 24: jest: „...obciążenie każdego z nodów...” – żargon.

Str. 45: jest: „... a następnie sekwencyjne wykonywanie pojedynczego zapytania wykonującego operacje zapisu i odczytu elementów wewnątrz.” – niejasne.

Str. 47: jest: „... optymalizuje się query wyszukując wymaganych informacji...” – żargon.

Str. 50 (i dalsze): Autor stosuje skrajnie uproszczoną formę odwołań do rysunków, którą trudno zaakceptować. Sugeruję przyjęcie bardziej typowej formy odwołań.

Str. 54: przyporządkowanie podpisu do rysunku 3.10 nie jest jednoznaczne.

Str. 56: W opisie algorytmu lokalizacji po stronie serwera sformułowanie „Do wskazanej wartości dodaje się wartość 0.8 precyzji sygnału GPS od nadesłanej pozycji ...” proponuję zapisać bardziej formalnie w postaci wyrażenia arytmetycznego.

Str. 60: zapis kroku 3 prezentowanego algorytmu jest niekompletny lub wymaga korekty: „Określ na podstawie zdefiniowanych reguł decyzyjnych (tj. reguł określających jaka odległość pomiędzy obiektem zapisanym w bazie oraz nadesłanym powoduje jego scalenie np. dla roweru jest to odległość 30m w ciągu 10s od chwili zgłoszenia detekcji przez innego użytkownika w skanowanym obszarze).”

Str. 60: używane w pracy określenie „nadpisany” jest żargonem utrudniającym w niektórych przypadkach zrozumienie tekstu; np. „Zakwalifikowanie obiektu jako tożsamy powoduje nadpisanie jego poprzedniej lokalizacji bez względu na dokładność uzyskanej lokalizacji”.

Str. 71: jest: „Algorytm do estymacji odległości ...”, proponuję: „Algorytm estymacji odległości ...”

Str. 72: jest: „...wykorzystując transfer learning ogólnie dostępnego modelu...” – żargon.

Str. 73: w tabeli 4.1 nie zdefiniowano opisu kolumny IoU.

Str. 73: jest: „...centralny punkt bounding boxu był zbliżony do...” – żargon.

Str. 87, rys. 4.13, rys. 4.15: wielkość czcionki praktycznie uniemożliwia odczytanie danych, poza tym treść podpisu pod rysunkiem nie jest jasna.

Przedstawione uwagi są w części dyskusyjne i nie zmieniają mojej pozytywnej oceny rozprawy.

5. Wniosek końcowy

Podsumowując stwierdzam, że w recenzowanej rozprawie został sformułowany, a następnie poprawnie rozwiązany problem naukowy dotyczący organizacji budowy mapy umożliwiającej wizualizację zmian prezentowanego otoczenia. Przedstawiono realizację systemu zbierania i przetwarzania danych będących podstawą budowy takiej mapy, a następnie prezentacji gromadzonych informacji.

Autor wykazał się głęboką wiedzą w zakresie wykorzystania sieci neuronowych do przetwarzania obrazów, przetwarzania wielomodalnych danych w celu skutecznej lokalizacji obiektów przestrzennych, a także organizacji baz danych przechowujących mapy cyfrowe. Równocześnie autor wykazał swe umiejętności i doświadczenie praktyczne poprzez zaprojektowanie i programową implementację systemu tworzącego przedstawioną mapę.

Stwierdzam, że oceniana praca pt.: „Wyznaczanie dynamicznej mapy 2D na podstawie wielomodalnych danych rejestrowanych przez urządzenia mobilne” spełnia wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy. Wnoszę o dopuszczenie Pana mgr inż. Kamila Bolka do dalszych etapów przewodu doktorskiego.

