

**Tytuł, stopień, imię i nazwisko**  
prof. dr hab. inż. Bogdan Kwolek

**data** 28.12.2022 r.

**RECENZJA ROZPRAWY DOKTORSKIEJ**  
przygotowana dla Rady dyscypliny informatyki  
Polsko-Japońskiej Akademii Technik Komputerowych

**Tytuł rozprawy: System autonomicznego lądowania bezzałogowego statku powietrznego na podstawie danych wizyjnych**

**Autor rozprawy: mgr inż. Marcin Paszkuta**

**Promotor: prof. dr hab. inż. Konrad Wojciechowski**

1. Przedstawiona mi do recenzji rozprawa doktorska dotyczy automatycznych systemów wizyjnych dla bezzałogowych statków powietrznych. Prace badawcze obejmowały projektowanie systemów wizyjnych wspierających dolot oraz lądowanie bezzałogowego statku powietrznego w sytuacjach awaryjnych. Zasadniczą część prac badawczo-eksperymentalnych dotyczyła konstrukcji i budowy systemów wizyjnych wspomagających lot i awaryjne lądowanie bezzałogowego statku powietrznego w nieznanym terenie, w tym także rozwiązań na potrzeby modelowania lotów i symulacji scen realistycznych w oparciu o wirtualne sensory pokładowe. Zbudowano prototypowy system wizyjny, który zintegrowano z lekkim bezzałogowym statkiem powietrznym typu quadrotor. Oryginalnym wkładem Doktoranta są metody akwizycji danych oraz uczenia sieci neuronowych do segmentacji semantycznej obrazów oraz estymacji map rozbieżności. Autorski system wizyjny zbudowany w oparciu o wspomniane sieci neuronowe umożliwia wyznaczenie bezpiecznego obszaru lądowania oraz detekcję przeszkód w obszarze lądowania. Ewaluacja skuteczności proponowanych rozwiązań na danych rzeczywistych wymagała zaprojektowania systemu wizyjnego, implementacji środowiska symulującego loty w wirtualnej przestrzeni w oparciu o model fizyczny statku oraz sensory do akwizycji niezbędnych danych, opracowania metody generowania stereoskopowych obrazów oraz danych na potrzeby rekonstrukcji głębi w oparciu o estymaty rozbieżności, opracowania efektywnych metod etykietowania danych wielomodalnych na potrzeby uczenia głębokich sieci neuronowych, integracji prototypowego systemu wizyjnego z bezzałogowym statkiem powietrznym typu quadrotor. Autor zaproponował oryginalne rozwiązania oraz zbudował kompletny system wizyjny, który został przebadany na rzeczywistych danych.

Teza rozprawy wyraża się w przekonaniu iż analiza obrazów z użyciem technik głębokiego uczenia, przeprowadzona bezpośrednio na module obliczeniowym zintegrowanym z lekkim bezzałogowym statkiem powietrznym przy zastosowaniu odpowiednich metod uczenia modeli sieci neuronowych realizujących zadanie: i) segmentacji semantycznej, ii) estymacji mapy dysparycji na obrazach stereo umożliwia wyznaczenie bezpiecznego obszaru lądowania oraz detekcję występujących na nim przeszkód. Prawdziwość tezy postawionej w rozprawie została wykazana empirycznie. Tematyka podjęta w rozprawie jest istotna i aktualna, zarówno w aspekcie poznawczym, jak i praktycznym. Od strony poznawczej praca wzbogaca wiedzę w dziedzinie projektowania i implementacji modułów systemów wizyjnych dla bezzałogowych

statków powietrznych, które w razie konieczności umożliwiłyby dołot i wybranie bezpiecznego miejsca do lądowania. Warto wspomnieć, że obecnie dostępne modele komercyjne dronów nie oferują funkcjonalności związanych z percepcją i analizą potencjalnego miejsca awaryjnego lądowania. Na uwagę zasługuje podjęcie interesującego tematu, wyraźnie inspirowanego potrzebami praktycznymi. Opracowane rozwiązania oraz uzyskane w pracy wyniki badań i sprecyzowane na tej podstawie rekomendacje składają się na klarowne metodologie projektowania i wdrażania systemów wizyjnych zdolnych do działania w sytuacjach awaryjnych.

2. Rozprawa napisana jest w języku polskim, składa się z sześciu rozdziałów i liczy łącznie 159 stron. Bibliografia liczy 70 pozycji. Streszczenia pracy przedstawiono zarówno w języku polskim, jak i angielskim. W rozdziale pierwszym zawarto ogólne wprowadzenie do tematyki pracy, zaprezentowano znaczenie i umiejscowienie pracy, przedstawiono cel pracy oraz genezę prac, postawiono tezy badawcze, a następnie omówiono układ i treść rozprawy. Rozdział drugi poświęcono omówieniu środowiska symulacyjnego. W pierwszym podrozdziale przybliżono koncepcję cyfrowego bliźniaka. W kolejnym podrozdziale poświęconym prezentacji modelu matematycznego quadrotora szczegółowo omówiono siły i momenty działające na obiekt latający, równania Eulera-Lagrange'a, a następnie omówiono równania stanu systemu dynamicznego. W kolejnym podrozdziale omówiono układ sterowania. W podrozdziale zamykającym drugi rozdział zaprezentowano opracowany system symulacyjny, który oparto na koncepcji cyfrowego bliźniaka. W rozdziale trzecim omówiono opracowany i przebadany eksperymentalnie podsystem detekcji miejsc bezpiecznych do lądowania, który oparto na sieci U-Net operującej na obrazach RGB. W czterech podrozdziałach składających się na rozdział czwarty szczegółowo omówiono podsystem unikania kolizji, który oparto na sieci HITNet operującej na obrazach z kamer w układzie stereoskopowym. Po wprowadzeniu w problematykę prezentowaną w rozdziale szczegółowo omówiono kalibrację kamer w układzie stereoskopowym, geometrię epipolarną, mapy rozbieżności oraz algorytm dopasowania bloków BM, a następnie estymację map głębi, algorytm dopasowania SGM oraz sieć HITNet. Po omówieniu metryk, w ostatnim podrozdziale zaprezentowano algorytm wykrywania wolnej przestrzeni. W piątym rozdziale omówiono prototyp zaprojektowanego i zbudowanego systemu wizyjnego. W kolejnych podrozdziałach omówiono kolejno najważniejsze moduły i układy quadrotora, system wizyjny oparty o moduł Nvidia Jetson oraz protokoły komunikacyjne. W ostatnim podrozdziale omówiono wyniki badań eksperymentalnych. W ostatnim rozdziale Doktorant sformułował wnioski końcowe, wskazał na potencjał opracowanych i zweryfikowanych praktycznie rozwiązań, w tym także na oryginalne aspekty rozprawy. W końcowej części pracy prezentowany jest dorobek naukowy Doktoranta na który składają się cztery publikacje naukowe w czasopismach wymienionych na liście czasopism punktowanych, dziewięć artykułów w monografiach naukowych oraz jeden patent.

Od strony formalnej rozprawa zredagowana jest poprawnie, z właściwą systematyką rozwiązywanych zagadnień. Nastęstwo rozdziałów jest logiczne i przedmiotowo uzasadnione. Tytuły rozdziałów i podrozdziałów dają syntetyczny pogląd na zawarte w nich treści. Wnioski podsumowujące rozprawę mają oparcie w opracowanych rozwiązaniach oraz uzyskanych wynikach empirycznych. Autorskie rozwiązania mają duże znaczenie nie tylko w aspekcie naukowym, ale również w aspekcie praktycznej realizacji bezzałogowych statków powietrznych zdolnych do działania w sytuacjach awaryjnych w oparciu o rozpoznania systemu wizyjnego.



3. Tematyka badań koncentruje się na zagadnieniach, które były przedmiotem prac badawczo-rozwojowych w grantie NCBIR realizowanym we współpracy z PJATK i dotyczyły autonomicznego lądowania samolotu bezzałogowego w nieznanym terenie na podstawie danych wizyjnych. W kontekście tak sformułowanej problematyki badań istotne znaczenie ma to, że liczba pokrewnych prac nie jest znacząca. Warto podkreślić jest także to, że mimo rozbudowanych systemów detekcji przeszkód w dostępnych obecnie komercyjnych modelach dronów, ciągle niedostępne są funkcjonalności, które w sytuacjach awaryjnych umożliwiałyby rozpoznanie potencjalnie bezpiecznego miejsca do lądowania. Wynika to po części ze złożoności omawianego zagadnienia. Aby zatem badania w tym kierunku były w ogóle możliwe niezbędne jest zgromadzenie odpowiednio poetykietowanych danych do wytrenowania odpowiednich modeli w skali wymaganej dla tej klasy problemów, dysponowanie statkami powietrznymi z modułami o otwartej architekturze i z sensorami odpowiednimi dla tego typu zadań, a także systemami pomiarowymi na potrzeby weryfikacji i oceny rozwiązań. Co więcej, z uwagi na konieczność operowania na danych 3D/mapach głębi niezbędne są odpowiednie moce obliczeniowe oraz synchronizacje danych. Mając na względzie powyższe, w niniejszej rozprawie zbudowano prototypowy system wizyjny, który zintegrowano z lekkim bezzałogowym statkiem powietrznym typu quadrotor, opracowano i zaimplementowano środowisko symulacyjne, które odzwierciedla fizykę lotu bezzałogowego statku powietrznego typu quadrotor z sensorami do akwizycji danych treningowych, a także umożliwia przebadanie systemu wizyjnego w wirtualnej rzeczywistości. Takie całościowe podejście do zagadnienia znacząco podnosi wagę uzyskanych wyników. Niewątpliwym atutem pracy są także zaprezentowane w pracy modele matematyczne na potrzeby modelowania dynamiki lotu i sterowania. Na oryginalny dorobek Autora składa się (i) opracowane i zaimplementowane środowisko symulacyjne w oparciu o silnik gier Unreal Engine dla bezzałogowego statku powietrznego typu quadrotor wraz z symulowanymi sensorami z fizycznego prototypu, (ii) metody generowania wielomodalnych poetykietowanych/referencyjnych danych do trenowania głębokich sieci neuronowych na potrzeby segmentacji semantycznej oraz detekcji potencjalnych obszarów do lądowania, (iii) metody generowania stereowizyjnych par obrazów wraz z danymi referencyjnymi na potrzeby trenowania głębokich sieci neuronowych do rekonstrukcji głębi sceny, (iv) prototypowy system wizyjny wspomagający lądowanie w nieznanym terenie, który zintegrowano z lekkim bezzałogowym statkiem powietrznym.

Model quadrotora opracowany na potrzeby niniejszej pracy wraz z układem sterowania wzorowanym na kontrolerze lotu PX4 o otwartym oprogramowaniu, który odzwierciedla układ sterowania rzeczywistego obiektu, umożliwił przebadanie algorytmów sterowania w środowisku symulacyjnym dla reprezentatywnych przypadków użycia z uwzględnieniem ograniczeń fizycznego quadrotora. Dzięki umiejętnemu wykorzystaniu możliwości oferowanych przez silniki gier nie tylko wiernie odtworzono zachowanie w środowisku wirtualnym fizycznego prototypu statku powietrznego, stworzono możliwość analizy dynamiki ruchu, czy też możliwość użycia тренаżera, ale także stworzono możliwość generowania tanim kosztem obrazów z wirtualnych kamer i danych z wirtualnych sensorów inercyjnych. Dedykowany symulator jest bardzo wartościowym narzędziem dla projektowania i budowy zaawansowanych systemów wizyjnych dla bezpilotowych statków powietrznych.

W proponowanym podejściu do rozpoznawania miejsca lądowania w pierwszej kolejności realizowana jest segmentacja semantyczna obrazu w oparciu o sieć U-Net. Po wygenerowaniu masek reprezentujących obszary potencjalnie nienadające się do lądowania, w pozostały obszar wpasowywany jest prostokąt o wymiarach nie mniejszych od minimalnych rozmiarów lądowiska.

Zdefiniowano szereg klas obszarów z obszarami nienadającymi się do lądowania, obszarami nadającymi się do lądowania, obszary typu drogi i parkingi, które to mogą być wybrane przy braku innych możliwości, a także klasy reprezentujące ludzi i samochody. Z uwagi na różnorodność obszarów zurbanizowanych, do wytrenowania sieci U-Net generującej maski o odpowiedniej reprezentatywności i szczegółowości wymagane było zgromadzenie reprezentatywnych danych. Co więcej, oprócz wygenerowania i zarejestrowania repozytorium o odpowiedniej reprezentatywności i wielkości niezbędne było zapewnienie odpowiedniego zbalansowania danych trenujących, a w szczególności poetykietowania obszarów obrazów. Z Głównego Urzędu Geodezji i Kartografii pozyskano zbiór ortofotomap z odpowiednio zróżnicowanymi obszarami. Celem uzyskania widoków odpowiadających rejestracji obrazów z odpowiedniej wysokości z wysokorozdzielczych obrazów wycinano podobrazy, które następnie skalowano. W finalnym zbiorze treningowym, który zbudowano w oparciu o wybrane dane ze zbiorów UAVid, Aerospace, TU-GRAZ landing, Cityscape oraz własnej bazy obrazów syntetycznych wyodrębniono siedem klas. Skonfigurowano sieci U-Net dla obrazów o rozmiarach 320x240 oraz 512x256, które następnie wytrenowano dla funkcji straty opartej o IoU. Z uwagi na zniekształcenia obrazu wnoszone przez układ optyczny oraz wpływ perspektywy, przejście ze współrzędnych lokalnych obrazu do współrzędnych globalnych sceny realizowano z uwzględnieniem parametrów modelu kamery wyznaczonych na etapie kalibracji sensora oraz pozycji statku powietrznego w chwili akwizycji obrazu. Działanie podsystemu, którego zasadniczym celem jest wybór miejsca do lądowania bezzałogowego statku powietrznego zilustrowano na przykładowych obrazach zarejestrowanych przez prototypowy quadrotor. Prezentowane wyniki prac eksperymentalnych pokazują użyteczność podsystemu do rozpoznawania miejsca lądowania. Jednocześnie wyniki prac pokazały, że podjęty problem nie jest trywialny. Ich użyteczność wynika z poprawnych założeń i decyzji projektowych.

Mimo znaczącej uwagi poświęconej detekcji przeszkód w oparciu kamery w układzie stereo-pary, wymagane jest znaczące udoskonalenie technik, a w szczególności wyjście poza scenariusze ze statycznymi scenami jak w zbiorze Middlebury, czy wręcz nawet nowszym KITTI stereo 2015 z referencyjnymi danymi głębi. Do unikania kolizji Autor zaproponował rozwiązanie oparte na stereoskopowym postrzeganiu głębi. W proponowanym podejściu wykorzystywane są kamery w układzie stereoskopowym jako źródło informacji o przeszkodach, które mogą stwarzać zagrożenia w trakcie lotu lub w podczas nienadzorowanego awaryjnego lądowania. W opracowanym podejściu do rekonstrukcji głębi, mapy rozbieżności wyznaczane są w oparciu o konwolucyjną sieć neuronową. Z uwagi na konieczność zapewnienia odpowiedniej reprezentacji referencyjnej rozbieżności w zbiorze treningowym, a także rzeczywistej reprezentacji obserwowanych scen, sieć wytrenowano wstępnie w oparciu o obrazy pozyskane z otwartych repozytoriów oraz własne repozytorium z obrazami wygenerowanymi przy użyciu autorskiego symulatora, a następnie dotrenowano w oparciu o dane pobrane ze zbudowanego systemu stereowizyjnego. Po skalibrowaniu kamer w układzie kanonicznym przedmiotem prac eksperymentalnych były algorytmy dopasowania bloków BM, metoda półglobalnego dopasowania SGM oraz sieci neuronowe PSMNet. Z uwagi na uwarunkowania obliczeniowe, mapy rozbieżności wyznaczano w oparciu o sieć HITNet. W pracach badawczo-eksperymentalnych analizowano wyniki dla wybranych miar podobieństwa map rozbieżności, wartości metryk do oceny błędów korespondencji, metryki do oceny błędów rekonstrukcji głębi oraz jakość wyznaczonych map rozbieżności przez wspomniane algorytmy, zaś dla wybranych algorytmów analizowano użyteczność danych rzeczywistych w zbiorach trenujących. Warto przy tym wspomnieć, że wyznaczenie referencyjnych map rozbieżności lub map głębi dla



dynamicznych rzeczywistych obrazów jest stosunkowo złożonym zadaniem. W ramach niniejszej pracy podjęto próbę przygotowania obrazów referencyjnych z naniesionymi mapami głębi, które wyznaczano przy pomocy 64 wiązkowego skanera LIDAR Ouster OS1. Wykrywanie wolnej przestrzeni odbywa się w oparciu o algorytm programowania dynamicznego, który operuje na wyznaczonych z góry mapach głębi. Wynik działania algorytmu detekcji wolnej przestrzeni zaprezentowano na obrazach ze zbioru KITTI. Do unikania przypadkowych zderzeń zaimplementowano prosty algorytm, który w oparciu o wyznaczone wolne obszary oraz dostępne mechanizmy w PX4 umożliwia skierowanie drona w zadany punkt przestrzeni.

System wizyjny przebadano na zaprojektowanej i wykonanej platformie do badań eksperymentalnych. Quadrotor wyposażono w niezbędne układy napędowe oraz kontroler lotu, sensor IMU, sensor wysokości, odbiornik GNSS, moduły do komunikacji i sterowania, system wizyjny zaimplementowany na module Nvidia Jetson oraz dwa zestawy kamer stereoskopowych. W końcowej części pracy na kilku rzeczywistych obrazach zaprezentowano wyniki jakościowe ilustrujące wyznaczone mapy dysparycji, mapy głębi odniesione do referencyjnych map głębi, a także wyniki detekcji wolnej przestrzeni odniesione do ręcznie wyznaczonych masek wolnej przestrzeni.

Podjęta w pracy problematyka jest ważką i istotną, nie tylko w aspekcie naukowym, ale przede wszystkim w aspekcie praktycznej realizacji systemów dla autonomicznego lądowania bezzałogowych statków powietrznych. Rozprawa wzbogaca naszą wiedzę i stanowi bardzo dobry materiał do dalszych prac. Autor wykazał się dobrą znajomością zagadnień związanych z przedmiotem rozprawy. Wyniki prac potwierdziły prawdziwość stawianych w pracy hipotez.

4. Poniższe uwagi, które mają polemiczny, a nie krytyczny charakter, odnoszą się do zagadnień, które mogłyby stanowić inspirację do dyskusji naukowej. Od strony poznawczej pracę wzbogaciłoby bardziej szczegółowe analizy i prezentacje zbiorów danych do trenowania finalnych sieci neuronowych, w tym także w kontekście technik balansowania danych oraz augmentacji danych. Interesujące byłoby także bardziej dogłębne przeanalizowanie ograniczeń proponowanych rozwiązań, w tym także scenariuszy lub sytuacji w których uzyskane rozwiązania byłyby niezadowalające, m.in. wskutek oślepienia, pogorszenia się warunków oświetleniowych, w scenariuszach z dynamicznymi zwrotami i rotacjami drona, drganiem kamer w wyniku uszkodzenia wirnika, itp. Mając na względzie to, że podsystem detekcji miejsc bezpiecznych do lądowania wykorzystuje dane reprezentujące aktualną pozycję, ale także orientację oraz pomiary wysokości interesujące byłoby przestudiowanie technik fuzji danych sensorycznych, w tym także wpływu jakości estymat na końcowe decyzje. Interesujące byłoby przestudiowanie możliwości stwarzanych przez techniki lokalizacji w oparciu o sygnały GPS i mapy z reprezentacjami 3D, m.in. w kontekście rozwiązań dla pojazdów autonomicznych. Mając za uwagę to, że w ostatnich latach uzyskano znaczące udoskonalenie sieci do detekcji obiektów, w tym także sieci typu mask R-CNN, interesujące byłoby odniesienie uzyskanych wyników do wyników uzyskiwanych przez wspomniane sieci, w szczególności po dotrenowaniu ich na potrzeby detekcji pożądanego obiektu, w tym także w powiązaniu z mapami wolnych obszarów, czy też czasów działania takich sieci na systemach wbudowanych.

Powyższe uwagi mają przede wszystkim charakter dyskusyjny i nie umniejszają w najmniejszym stopniu wartości pracy, która stanowi oryginalny wkład w dziedzinę informatyki i inteligencji maszynowej.

Pracy nie zaszkodziłaby staranniejsza korekta językowa, zadbanie o odwołania w tekście do wszystkich rysunków. W wielu miejscach rozprawy brakuje odniesień do literatury przedmiotu.

5. Przewieziona mi do recenzji rozprawa doktorska jest wartosciowym studium naukowym, ktore wnosi solidna wiedze oraz know-how w zakresie projektowania i realizacji systemow wizyjnych dla robotow latajacych. Rozprawa potwierdza znaczaca wiedze Doktoranta w obszarze informatyki oraz szeroko rozumianej robotyki, umiejetnosc samodzielnego prowadzenia badan naukowych, a takze duze zdolnosci w zakresie projektowania, realizacji i ewaluacji zlozonych systemow wizyjnych. Tematyka badawcza jest aktualna i ma znaczacy potencjal aplikacyjny. Do rozwiazania postawionych problemow niezbedne bylo uzycie zaawansowanych technik percepcji sceny, wytrenowanie glubokich sieci neuronowych o szybkościach dzialania wymaganych dla statków powietrznych, a takze zaproponowanie technik generowania i pozyskiwania obrazow z danymi odniesienia dla pojedynczych pikseli/obszarow obrazow. W podsumowaniu stwierdzam, ze recenzowana rozprawa mgra inz. Marcina Paszkuta spelnia wszystkie wymagania ustawowe i zwyczajowe stawiane rozprawom doktorskim. Stwierdzam niniejszym, ze przedstawiona mi do recenzji rozprawa doktorska spelnia kryteria określone w art. 13 ust. 1 Ustawy. Wnosze o dopuszczenie Pana mgr. inz. Marcina Paszkuta do kolejnych etapow postepowania o nadanie stopnia doktora nauk inzynieryjno-technicznych w dyscyplinie informatyka. Ponadto, wysoki poziom merytoryczny i potencjal praktyczny zaprojektowanych i zrealizowanych podsystemow wizyjnych, ktore poparto publikacja w czasopiśmie ze znaczaca liczba punkow (140 pkt.) na liście czasopism punktowanych MEiN, sa przeslanką do wnioskowania do Rady Informatyki Polsko-Japonskiej Akademii Technik Komputerowych o wyroznienie rozprawy doktorskiej Pana mgr. inz. Marcina Paszkuta.

Bogdan Kwiatek