

## Streszczenie

Technika lotniczego skanowania laserowego pozwala na bardzo szybkie pozyskanie szczegółowych informacji dotyczących ukształtowania terenu oraz obiektów znajdujących się na jego powierzchni. Pozyskiwane zbiory danych posiadają ogromny wolumen, zarówno ze względu na duży obszar opracowania, jak i na wciąż rosnącą gęstość pomiaru. Dane pozyskane za pomocą techniki lotniczego skanowania laserowego znajdują zastosowanie w wielu dziedzinach takich jak modelowanie powodziowe czy planowanie urbanistyczne. W zależności od pożądanego zastosowania, dane lotniczego skaningu mogą zostać opracowane w różny sposób, np. na etapie filtracji zidentyfikowane mogą zostać punkty reprezentujące powierzchnię terenu, a na etapie klasyfikacji wszystkie punkty pochodzące z pomiaru mogą zostać przypisane do klas, reprezentujących zarówno obiekty przedstawiające powierzchnię jak i elementy pokrycia terenu, takie jak drzewa i budynki. Manualne wykonanie filtracji czy klasyfikacji jest zadaniem niezwykle pracochłonnym, długotrwałym i kosztownym, w związku z czym podejmowane są próby usprawnienia procesu opracowania danych lotniczego skaningu laserowego, poprzez opracowanie automatycznych metod filtracji i klasyfikacji danych.

W ramach niniejszej pracy opracowano metodę klasyfikacji danych lotniczego skaningu laserowego z wykorzystaniem algorytmu Random Forests. Opracowana metoda bazuje na podejściu klasyfikacji nadzorowanej, a do opracowania skutecznego klasyfikatora konieczne jest dostarczenie prawidłowo sklasyfikowanej próbki uczącej. W ramach opisanych w niniejszej pracy badań opracowano szereg cech, obliczanych na podstawie właściwości punktów lotniczego skaningu laserowego, jak i cech bazujących na wzajemnych relacjach między punktami. Cechy te obliczane są dla każdego z punktów zbioru danych. Wartości poszczególnych cech dla punktów pola treningowego oraz informacja o przynależności punktu pola treningowego do klasy wykorzystywana jest do budowy klasyfikatora Random Forests. Wytrenowany klasyfikator może zostać użyty do klasyfikacji pozostałych, niesklasyfikowanych danych, dla których również obliczane są wartości poszczególnych cech.

W praktyce, podczas klasyfikacji danych niezwykle istotną informacją jest wysokość nad powierzchnią terenu, obliczana na podstawie numerycznego modelu terenu. Z uwagi na szczególną wagę tej cechy, w ramach pracy opracowana została metoda filtracji, oparta o metodę rosnących regionów.

W ramach badań przeprowadzono szereg eksperymentów numerycznych, wykorzystujących zbiory danych lotniczego skaningu laserowego o różnej gęstości oraz opisanych różnymi zestawami klas. Zbadana została również skuteczność algorytmu w zakresie klasyfikacji zbiorów danych, pozyskanych za pomocą nowoczesnych skanerów wielospektralnych.

Otrzymane rezultaty eksperymentów numerycznych potwierdzają wysoką skuteczność algorytmu w zakresie klasyfikacji większości przeanalizowanych zbiorów danych testowych. Jednocześnie należy podkreślić, iż opracowana metoda pozwala na wykonanie klasyfikacji na zróżnicowane zestawy klas, co czyni ją bardzo elastyczną. Należy również zaznaczyć, iż algorytm Random Forests jest bardzo popularnym algorytmem klasyfikacji danych, wykorzystywanym w wielu dziedzinach. Fakt ten sprawia, iż jest on zaimplementowany w wielu bibliotekach obliczeniowych, wliczając w to narzędzia takie jak Apache Spark, pozwalające na wykonywanie obliczeń w sposób równoległy i rozproszony. Fakt ten sprawia, iż możliwa jest implementacja opracowanego podejścia w taki sposób, aby produkcyjne środowisko obliczeniowe skalowalne było horyzontalnie, a automatyczne procesy opracowania wielkich zbiorów danych wykonywane były w bardzo krótkim czasie.

Słowa kluczowe: Lotniczy skaningu laserowy, klasyfikacja, filtracja, Random Forests, ocena dokładności