

## Podsumowanie

Zaproponowana metoda klasyfikacji wysokorozdzielczych danych hiperspektralnych – cechujących się ciągłym widmem elektromagnetycznym w całym zakresie spektrum (VIS-SWIR) – zapisanych w dużej liczbie stopni szarości, pozwoliła na klasyfikację roślinności Tatr Wysokich. Istota zaproponowanego algorytmu polegała na:

- 1) eliminacji błędów wynikających z ruchu samolotu podczas rejestracji (parametryczna korekcja geometryczna);
- 2) eliminacji wpływu atmosfery i rzeźby terenu na obrazy (korekcja atmosferyczna);
- 3) eliminacji najmniej informacyjnych kanałów zobrażenia DAIS 7915;
- 4) terenowego przygotowania wzorców do uczenia sieci neuronowych oraz wzorców do weryfikacji danych poklasyfikacyjnych;
- 5) wykonaniu kilku wstępnych klasyfikacji w celu optymalnego doboru parametrów wejściowych symulatora sieci neuronowych;
- 6) wykonaniu właściwych klasyfikacji;
- 7) opracowaniu statystyk poklasyfikacyjnych;
- 8) końcowym przygotowaniu map poklasyfikacyjnych.

Jednym z ważniejszych osiągnięć pracy jest porównanie jakości klasyfikacji skompresowanych (transformacja MNF) oraz oryginalnych danych. Na uwagę zasługuje także analiza spadku błędu klasyfikacji wraz ze wzrostem liczby klasyfikowanych kanałów oraz liczby iteracji (ryc. 26).

Użycie sieci neuronowych, a w szczególności analiza danych rozmytych pozwoliła wykorzystać obok cech parametrycznych (np. jasność, odpowiadającą fototonowi danego piksela) – cechy nieparametryczne obrazu (struktura i tekstura, sąsiedztwo innych obiektów). Dzięki temu układ analityczny symulatora jest zasilany dodatkowymi, ważnymi informacjami – poszczególne

obiekty występujące w środowisku cechują się powtarzalnymi przestrzennie układami, np. sąsiedztwo, współwystępowanie innych obiektów (takich jak cienie, woda wokół zbiorowisk hydrofilnych, czy skały wokół porostów naskalnych). W zdecydowanej większości standardowych klasyfikacji cech nieparametryczne są pomijane, gdyż podczas klasyfikacji wykorzystuje się odległości spektralne od zadanego piksela, pełniące rolę wzorca.

Zaletą symulatorów sztucznych sieci neuronowych jest to, że potrafią wyeliminować ze wzorca danej klasy błędne sygnały (np. pojedyncze drzewa, kamienie). Jest to przydatne na etapie analizy wzorców i prowadzenia pierwszych iteracji klasyfikacji. Na obszarach heterogenicznych, gdzie występuje silne mieszanie sygnałów pochodzących od różnych spektralnie obiektów, możliwe staje się szybkie odseparowanie właściwych sygnałów i wykonanie właściwej klasyfikacji.

Dzięki wykonanej korekcji atmosferycznej, możliwe było porównywanie sygnałów poszczególnych obiektów na różnych obrazach lotniczych z rejestrowanymi na poziomie terenu. Dodatkowa korekcja dwukierunkowości odbicia promieniowania (BRDF) pozwala na porównywanie sygnałów odbitych z różnych ekspozycji i nachylenia terenu. Takie działania należy uznać za istotne w analizie obszarów wysokogórskich. Zastosowanie Numerycznego Modelu Terenu (NMT) w klasyfikacji dostarczyło informacje o położeniu zbiorowisk roślinnych w piętrach wysokościowych, co ułatwiło ich prawidłową identyfikację.

Dane wzorcowe i referencyjne zostały pozyskane w trakcie badań terenowych. Do stworzenia wzorca weryfikującego dokładność klasyfikacji wykorzystano metodę klasyfikacji SAM, która pozwoliła na wybór obszarów najbardziej zbliżonych spektralnie do poszukiwanych wzorców, eliminując jednocześnie obiekty obce

dla danej klasy (np. pojedyncze drzewa, kępy wyraźnie różniące się roślinnością, ścieżki, wody).

Ze względu na dużą liczbę zbiorowisk roślinnych i ich kompleksów przestrzennych oraz chęć wykonania klasyfikacji o wysokiej dokładności, zarówno do uczenia sieci, jak i analizy poklasyfikacyjnej stworzone zostały wzorce składające się z dużej liczby pikseli (starano się, by wszystkie klasy były reprezentowane przez próbki większe niż 400 pikseli).

Jednym z kluczowych elementów w klasyfikacjach danych hiperspektralnych jest optymalny dobór kanałów, gdyż każdy dodatkowy kanał rozbudowuje architekturę sieci klasyfikującej, co wydłuża czas samej klasyfikacji. W ramach niniejszego opracowania zastosowano dwie równoległe procedury: (1) dekorrelację szumu i analizę głównych składowych (MNF) i (2) analizę informacyjności poszczególnych kanałów. W drugim przypadku bazowano na algorytmach dostępnych w pakiecie SVM, a analiza informacyjności została wykonana w Interdyscyplinarnym Centrum Modelowania (ICM) UW. Uzyskane wyniki wskazały, że dane po transformacji MNF uzyskiwały dokładność o kilka procent niższą, ale z przyczyn czysto operacyjnych są to dobre wyniki, gdyż czas klasyfikacji jest znacznie krótszy. Wydaje się, że dobrym rozwiązaniem jest przeprowadzenie wstępnej klasyfikacji (w tym dobór parametrów sieci) na zbiorach MNF, a następnie – po ustaleniu optymalnych wskaźników uczenia sieci – wykonanie właściwej klasyfikacji na danych oryginalnych.

Wyniki klasyfikacji poszczególnych linii zobrazowania są w znacznym stopniu porównywalne (tab. 22). Należy jednak pamiętać, że zbiorowiska roślinne w poszczególnych okresach fenologicznych cechują się różną morfologią i parametrami biochemicznymi. Wykorzystane obrazy pozyskano w optimum rozwoju fenologicznego (rejestracja obrazów odbyła się 4.08.2002).

Uzyskane obrazy poklasyfikacyjne, liczba wyodrębnionych klas oraz poklasyfikacyjne dane statystyczne można uznać za satysfakcjonujące. Analiza wyników wskazuje, że spośród 42 analizowanych klas tylko dwie nie zostały rozpoznane na żadnej linii zobrazowania. 5-6 zbiorowisk cechowało się niezbyt dużą dokładnością poklasyfikacyjną. Dotyczy to głównie zbiorowisk powypasowych na Hali Gąsienicowej, które przekształcają się ze względu na zaprzestanie wypasu owiec.

Przegląd literatury potwierdził, że zdecydowana większość prac koncentruje się na klasyfikacji roślinności

obszarów chronionych (najczęściej parków narodowych), co jest przydatne np. w analizie zmian zasięgu gatunków inwazyjnych. Drugim dość popularnym kierunkiem jest analiza roślinności uprawnej. Obiecująco wygląda wykorzystanie danych hiperspektralnych do analizy składu gatunkowego zbiorowisk, jednakże największy potencjał do tego typu badań zawarty jest w spektralnym rozmieszaniu pikseli i analizach subpikselowych, które nie były przedmiotem zainteresowań niniejszej pracy, gdyż wyniki takich analiz przedstawiają miary statystyczne, np. prawdopodobieństwo wystąpienia gatunku na danym obszarze lub też prawdopodobieństwo wystąpienia określonego gatunku w danym zbiorowisku (np. ryc. 43).

Najlepsze średnie wyniki klasyfikacji 40 kanałów wyniosły 92,8% (dokładność producenta) oraz 84,2% (dokładność użytkownika), natomiast najgorsze odpowiednio 84,2% oraz 67,5%. W przypadku klasyfikacji 20 kanałów MNF uzyskano odpowiednio 86,1% i 79,9%, najgorsze zaś 74,7% i 63,7% (tab. 24). Najlepsze wyniki (dokładność producenta i użytkownika >90%) uzyskano w odniesieniu do zbiorowisk naturalnych: murawa alpejska typowa (#8), murawa alpejska torfowcowa (#11), murawa alpejska postać subalpejska (#16), murawa alpejska w kompleksie z wyleżyskami (#17), murawa alpejska w kompleksie z traworoślami (#18) oraz zarośla kosodrzewiny na podłożu bezwęglanowym (#38) (tab. 24, ryc. 36).

Uzyskane wyniki (ryciny 28-35 oraz tabele 5-24) wyraźnie potwierdzają postawioną na wstępie hipotezę, że wykorzystanie wysokorozdzielczych spektralnie i radiometrycznie danych hiperspektralnych oraz zastosowanie symulatorów sztucznych sieci neuronowych pozwala na identyfikację i kartowanie zbiorowisk roślinnych obszarów wysokogórskich, które są jednym z najtrudniejszych obiektów badawczych, nie tylko z teledetekcyjnego punktu widzenia.

Zaproponowana metoda kartowania roślinności powinna być rozwijana na innych zestawach danych hiperspektralnych, pozyskanych w różnych okresach sezonu wegetacyjnego, co umożliwi dokładniejszy monitoring środowiska obszarów chronionych i rolniczych. Przytoczone zagraniczne badania potwierdzają także przydatność danych hiperspektralnych i sieci neuronowych do powszechnego, aplikacyjnego stosowania.