

ZAAWANSOWANE ALGORYTMY KLASYFIKACJI DANYCH RASTROWYCH

Zaawansowane algorytmy klasyfikacji danych rastrowych

- 30 godzin
- Oddzielne oceny z wykładu i ćwiczeń
- Ocena z przedmiotu to 55% oceny z wykładu i 45% oceny z ćwiczeń
- Wykład – zaliczany na podstawie egzaminu
- Ćwiczenia – sprawdziany + projekty + punkty za aktywność
- Możliwe dwie nieobecności

Zaawansowane algorytmy klasyfikacji danych rastrowych

1. Klasyfikacja – wprowadzenie
2. Ocena dokładności
3. Pola treningowe i weryfikacyjne
4. Klasyfikacja/detekcja obiektów
5. Korekcje danych wysokorozdzielczych
6. Wstępne przetwarzanie danych
7. Analiza poklasyfikacyjna
8. Wykorzystanie różnych rodzajów danych do klasyfikacji
9. Łączenie danych

Metody klasyfikacji

- Support Vector Machine
- Random Forest
- Sztuczne sieci neuronowe
- Spectral Angle Mapper
- Spectral Unmixing
- Drzewo decyzyjne
- Matched Filtering
- Mixed Tuned Matched Filtering

Etapy klasyfikacji cyfrowej

1. Wstępna analiza i wybór metod
2. Wybór danych wejściowych
3. Wzmocnienie obrazu
4. Klasyfikacja
5. Procesy poklasyfikacyjne
6. Ocena dokładności

Etapy klasyfikacji cyfrowej

1. **Wstępna analiza i wybór metod**
2. Wybór danych wejściowych
3. Wzmocnienie obrazu
4. Klasyfikacja
5. Procesy poklasyfikacyjne
6. Ocena dokładności

Wstępna analiza i wybór metod

- Cel klasyfikacji
- Ograniczenia posiadanych danych (rozdzielczości)
- Ograniczenia i wymagania poszczególnych metod klasyfikacji
- Wybór metody

Cel klasyfikacji

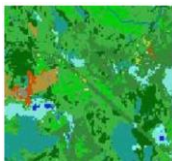
- Po co wykonywana jest klasyfikacja
- Jaki obszar jest analizowany
- Ile i jakiego rodzaju klasy?
- Jaki rodzaj danych mamy?
- Oczekiwania odbiorcy
- Koszty
- Czas

Ograniczenia danych

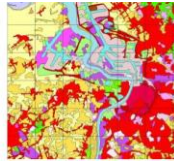
1. Charakterystyka używanych danych
 - rozdzielczości
2. Specyfikacja finalnej mapy
 - Format mapy wyjściowej (raster/wektor)
 - Najmniejsza jednostka
3. Specyfika analizowanego obszaru
 - Strefa klimatyczno-roślinna,
 - rodzaj pokrycia terenu
4. Dostępność dodatkowych danych
 - Mapy
 - Dane do wprowadzenia poligonów testowych i referencyjnych
 - Dane terenowe

Ograniczenia posiadanych danych

(MMU) Minimalna jednorodna powierzchnia na materiałach źródłowych i finalnej mapie



Raster = piksel



Wektor = najmniejszy obiekt (szerokość, powierzchnia)

Np. Corine Land Cover
skala robocza 1:100 000

Najmniejsza powierzchnia 25 ha, minimalna szerokość 100 m
– wynik klasyfikacji wizualnej danych rastrowych Tematic Mapper z satelity Landsat o rozdzielczości przestrzennej 30 m.

Specyfika analizowanego obszaru

- Zróżnicowanie geograficzne
 - Urbanizacja
 - Strefa klimatyczna
 - Szerokość geograficzna
- Podział na podregiony?
- Oplacalność analizy
 - Szybkość klasyfikacji (przetwarzania i przygotowania)
- Pokrycie terenu – wstępne rozpoznanie
 - Mapy/obrazy/zdjęcia
- Miksele

Etapy klasyfikacji cyfrowej

1. Wstępna analiza i wybór metod
2. **Wybór danych wejściowych**
3. Wzmocnienie obrazu
4. Klasyfikacja
5. Procesy poklasyfikacyjne
6. Ocena dokładności

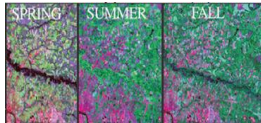
Wybór danych wejściowych

- Dopasowanie danych wejściowych do celu klasyfikacji
- Określenie rozdzielczości, jakie będą najlepsze
- Wybór między klasyfikacją nadzorowaną a nienadzorowaną (określenie potencjalnych miejsc pól treningowych)
- Do ustalenia:
 - Czy dane pozwolą na zidentyfikowanie założonych obiektów (czy są kanały, gdzie są różnice w odbiciu między obiektami)
 - Korekcja obrazu
 - Dodatkowa informacja do określania pól treningowych
 - Pomiary terenie:
 - Najlepiej wykonywane w tym samym czasie, co obraz
 - Kosztowne
- Czasem najlepszym rozwiązaniem jest redukcja liczby klas.

Zaawansowane algorytmy klasyfikacji danych rastrowych

Termin danych

- Termin jednego obrazu powinien być wybrany w zależności od klas, jakie mają być rozpoznane (np. w miesiące)
- Różne dane powinny pochodzić z tego samego okresu (np. obraz i mapa, na podstawie której wyznaczane są pola treningowe)
- Dane wieloczasowe – ułatwiają eliminację problemu zmian fenologicznych
- Np.:
 - Obszary nawadniane i nienawadniane
 - Lasy liściaste i wiecznie zielone
 - Rozpoznawanie upraw
 - Utrudniona identyfikacja roślinności na początku sezonu wegetacyjnego



The production of the US National Land Cover Database (NLCD) – 2001 is based on a multi-temporal approach

Zaawansowane algorytmy klasyfikacji danych rastrowych

Dodatkowe dane wejściowe

Ciągłe:

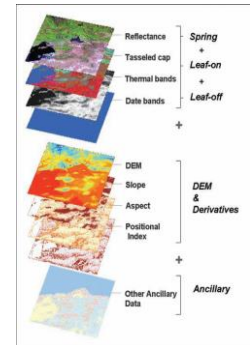
- wysokość,
- nachylenie,
- wskaźniki

Klasy:

- gleba,
- Użytkowanie terenu

Rozdzielczość przestrzenna

US National Land Cover Database 2001



Zaawansowane algorytmy klasyfikacji danych rastrowych

Etapy klasyfikacji cyfrowej

1. Wstępna analiza i wybór metod
2. Wybór danych wejściowych
3. **Wzmocnienie obrazu**
4. Klasyfikacja
5. Procesy poklasyfikacyjne
6. Ocena dokładności

Zaawansowane algorytmy klasyfikacji danych rastrowych

Wzmocnienie klasyfikowanego obrazu

- Zwiększenie rozdzielczości przestrzennej
- Filtrowanie danych
- Resampling
- Zmiana rozdzielczości spektralnej
 - Selekcja kanałów
 - PCA, MNF
- Likwidacja szumów = filtracja (dane radarowe)

Zaawansowane algorytmy klasyfikacji danych rastrowych

Etapy klasyfikacji cyfrowej

1. Wstępna analiza i wybór metod
2. Wybór danych wejściowych
3. Wzmocnienie obrazu
4. **Klasyfikacja**
5. Procesy poklasyfikacyjne
6. Ocena dokładności

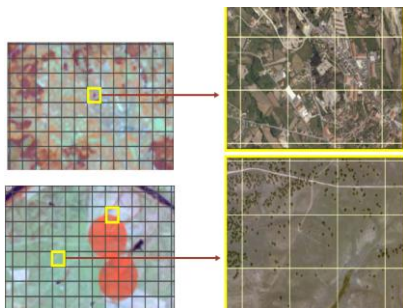
Zaawansowane algorytmy klasyfikacji danych rastrowych

Klasyfikacja

1. **Wybór metody klasyfikacji**
2. Opracowanie i wybór maski
3. Faza treningowa – selekcja danych wzorcowych
4. **Klasyfikacja obrazu przy zastosowaniu różnych parametrów (najczęściej wielokrotna)**

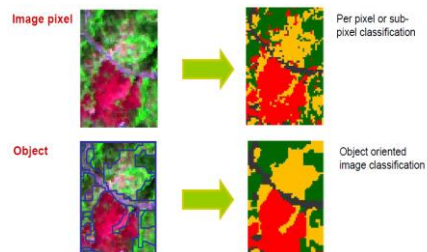
Zaawansowane algorytmy klasyfikacji danych rastrowych

Miksele



Zaawansowane algorytmy klasyfikacji danych rastrowych

Klasyfikacja pikselowa a obiektowa



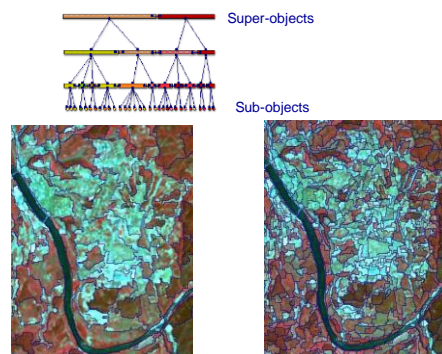
Zaawansowane algorytmy klasyfikacji danych rastrowych

Podejście obiektowe

- Pierwszym etapem klasyfikacji jest segmentacja.
- Segmentacja jest wieloetapowa.
- Zasady:
 - Przy segmentacji brane jest pod uwagę otoczenie piksela.
 - Możliwe jest łączenie segmentów.
 - Segmentacja bierze pod uwagę: kolor, spójność, gładkość obiektu i wiele innych.
- Na podstawie segmentów generowane są klasy.
- Przy klasyfikacji obiektowej analizowane jest otoczenie segmentów przez przydzieleniem do klasy.
- Znaczną część procedur klasyfikacji można przeprowadzić na segmentach.
- Przy prawidłowo wykonanej klasyfikacji wzrasta dokładność i rozpoznanie klas.
- Długotrwała procedura

Zaawansowane algorytmy klasyfikacji danych rastrowych

Podejście obiektowe



Zaawansowane algorytmy klasyfikacji danych rastrowych

Różnice między klasyfikacjami

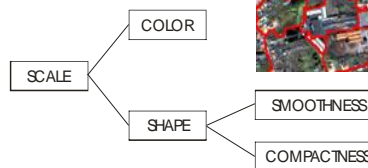
Pikselowa	Obiektowa
bazuje na wartościach pikseli	oprócz wartości pikseli użytkownik może uwzględnić np. sąsiedztwo klas, kształt segmentów
Jeden etap	Wieloetapowa
klasyfikowanie małych grup pikseli (sól i pieprz), często wymaga użycia filtrów	Brak wyizolowanych pikseli

Zaawansowane algorytmy klasyfikacji danych rastrowych

Podejście obiektowe

Podobne do interpretacji wizualnej

- Rozmiar i kształt
- Jednorodność obiektów
- Relacje z otoczeniem

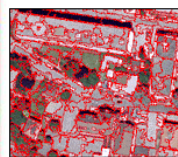


Zaawansowane algorytmy klasyfikacji danych rastrowych

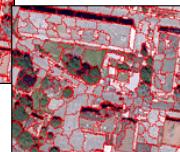
Cechy klasyfikacyjne

- średnia wartość pikseli
- odchylenie standardowe
- kształt i rozmiar
- sąsiedztwo
- związek między pod-obiektem i super-obiektem

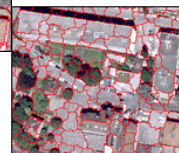
Zaawansowane algorytmy klasyfikacji danych rastrowych



color 0,9
shape 0,1



color 0,5
shape 0,5



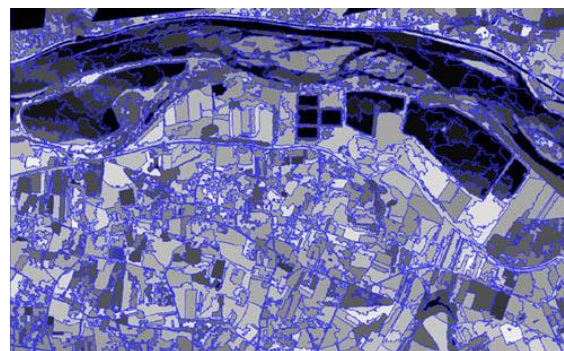
color 0,1
shape 0,9

Harayama A., Jaquet J.M., 2004, *Multi-source object-oriented classification of landcover using very high resolution imagery and digital elevation model*, Materiały XVIII Konferencji Informatics for Environmental Protection – EnviroInfo, Geneva, Szwecja.



obraz Loary, satelita Terra

<http://www.terraimaging.nl/index.php?id=42>



obraz po segmentacji

<http://www.terraimaging.nl/index.php?id=42>



rezultat klasyfikacji (woda, roślinność i budynki)

<http://www.terraimaging.nl/index.php?id=42>

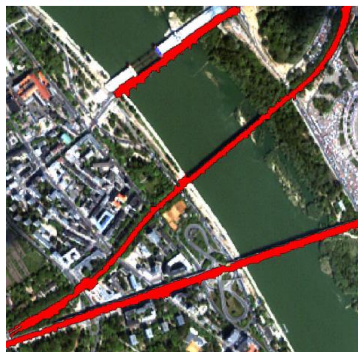
Zaawansowane algorytmy klasyfikacji danych rastrowych

- odbicie spektralne
- sąsiedztwo (cieni)



Lewiński S., 2004, *Examples of object-oriented classification performed in high-resolution satellite images*, Miscellanea Geographica, Vol. II, Warszawa, str. 349-358.

Zaawansowane algorytmy klasyfikacji danych rastrowych

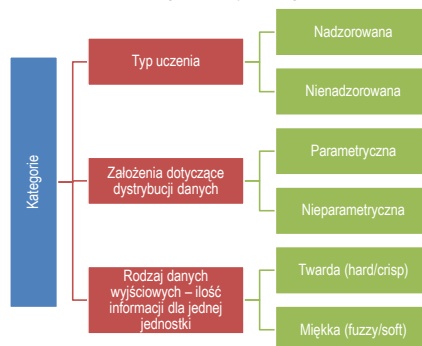


- odbicie spektralne
- kształt
- sąsiedztwo (wody)

Lewiński S., 2004, *Examples of object-oriented classification performed in high-resolution satellite images*, Miscellanea Geographica, Vol. II, Warszawa, str. 349-358.

Zaawansowane algorytmy klasyfikacji danych rastrowych

Rodzaje klasyfikacji



Zaawansowane algorytmy klasyfikacji danych rastrowych

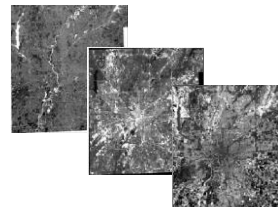
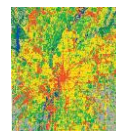
Klasyfikacja parametryczna a nieparametryczna

- Parametryczna wykorzystuje w procesie klasyfikacji informacje o parametrach statystycznych (średnia, miara rozproszenia w klasie).
Przykłady:
 - Odległości Mahalanobisa,
 - Największego prawdopodobieństwa
- Nieparametryczna klasyfikacja nie wykorzystuje parametrów statystycznych, bazuje na innych algorytmach. Przykłady:
 - Metoda drzewa decyzyjnego
 - Support vector machines
 - Najbliższego sąsiada

Zaawansowane algorytmy klasyfikacji danych rastrowych

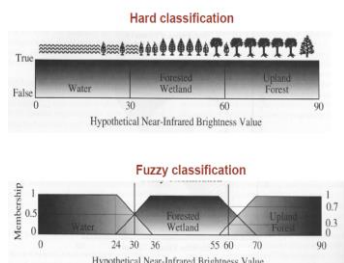
Metody klasyfikacji: twarda i miękka

- Twarda:
 - Obrazem wynikowym zawsze jest jeden obraz z określonymi klasami.
 - Jeden piksel przynależy do jednej klasy
- Mięka/fuzzy:
 - Każdy piksel ma określony procent prawdopodobieństwa przynależności do danej klasy.
 - Każdy piksel może należeć do więcej niż jednej klasy.
 - Aby otrzymać mapę klas należy określić wartości progowe dla poszczególnych klas.



Zaawansowane algorytmy klasyfikacji danych rastrowych

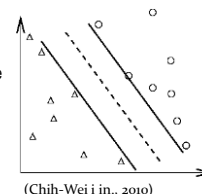
Klasyfikacja miękka i twarda



Zaawansowane algorytmy klasyfikacji danych rastrowych

Support Vector Machines

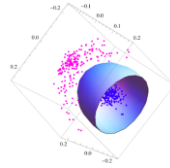
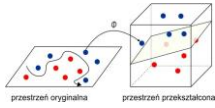
- Maszyny wektorów nośnych/wspierających,
- Pierwszym etapem jest proces uczenia.
- Konstruowana jest optymalna hiperpłaszczyzna w wielowymiarowej przestrzeni cech, która dzieli dane należące do przeciwnych klas z możliwie największym marginesem zaufania.
- Utworzone pomocnicze wektory określają najkrótszą odległość od hiperpłaszczyzny.



(Chih-Wei i in., 2010)

Support Vector Machines

- Algorytm sprawdza się dla małej liczby próbek.
- Duża liczba pikseli poddawanych klasyfikacji na obrazach teledetekcyjnych stanowi utrudnienie dla metody SVM, bo tworzonych jest bardzo wiele pomocniczych wektorów.
- Jeśli dane treningowe nie pozwalają na przeprowadzenie prostej płaszczyzny, możliwe jest utworzenie płaszczyzny nieliniowej.
- Główny element transformacji – wybór funkcji jądra (kernel function) → linear, polynomial, radial basis function, sigmoid)
- Nie jest brana pod uwagę cała krzywa odbicia spektralnego łącznie, co może prowadzić do błędów (błędy w przypadku, gdy ten sam obiekt jest inaczej oświetlony).



Marcinkowska-Ochtyra 2017