

# Laboratorium Rozpoznawania Obrazów – Ćwiczenie #3 & #4

## Rozpoznawanie cyfr pisanych ręcznie

Termin oddawania: **8.04.2019, 11.04.2019, 15.04.2019, 25.04.2019**

W tym ćwiczeniu Państwa zadaniem będzie rozpoznawanie cyfr pisanych ręcznie za pomocą zestawu klasyfikatorów liniowych. W zadaniu wykorzystamy dane z bazy danych MNIST (<http://yann.lecun.com/exdb/mnist>). Zbiory uczący i testowy można pobrać spod podanego wyżej adresu lub z serwera Galera (<http://galeranew.ii.pw.edu.pl/~rkz/rob/mnist.zip>).

Warto wiedzieć, że obrazy cyfr wydzielone z zeskanowanej strony są normalizowane w następujący sposób:

1. Prostokąt zawierający czarno-biały obraz znaku zeskanowanego w rozdzielczości 300 dpi jest skalowany proporcjonalnie do prostokąta o większym z wymiarów równym 20. W trakcie skalowania obraz jest zamieniany na skalę szarości (proporcjonalnie do liczby oryginalnych pikseli pierwszego planu przypadających na jeden piksel obrazu po przeskalowaniu).
2. Jest wyznaczany środek ciężkości przeskalowanego znaku, a znak jest umieszczany w obrazie 28x28 pikseli tak, żeby środek ciężkości znalazł się na środku tego większego obrazu.

Państwa zadaniem, jest przygotowanie klasyfikatora korzystającego z **klasyfikatorów liniowych** rozróżniających konkretne cyfry. Oprócz jakości klasyfikacji na zbiorze testowym należy podać **macierz pomyłek**. Dodatkowe zadanie polega na poprawieniu jakości klasyfikacji w stosunku do klasyfikacji zrealizowanej przez zwykłe głosowanie klasyfikatorów elementarnych. Metodą realizacji tego celu będzie poprawa jakości klasyfikacji **jednego** klasyfikatora elementarnego (można rozszerzyć to rozwiązanie dodając kolejne usprawnione klasyfikatory).

Rozwiązaniem referencyjnym, jest klasyfikator głosujący (wykorzystujący głosy 45 klasyfikatorów liniowych *one vs. one*) wskazujący jako wynik konkretną cyfrę, o ile zostało oddanych na nią 9 głosów. W pozostałych przypadkach jest podejmowana decyzja wymijająca. Wyniki klasyfikacji podsumowuje poniższa tabela (choć ciekawy wgląd w klasyfikację może dać analiza macierzy pomyłek); klasyfikatory działały na 40 składowych głównych.

|                     | Zbiór uczący MNIST |       |            | Zbiór testowy MNIST |       |            |
|---------------------|--------------------|-------|------------|---------------------|-------|------------|
|                     | OK.                | Błąd  | Odrzucenie | OK.                 | Błąd  | Odrzucenie |
| Jakość klasyfikacji | 91.34%             | 5.72% | 2.94%      | 91.55%              | 5.49% | 2.96%      |

Zadanie można podzielić na kilka części:

1. Przygotowanie podstawowego algorytmu wyznaczającego parametry płaszczyzny decyzyjnej dla zadanego zbioru uczącego, zawierającego dwie klasy. W zupełności wystarczy użycie tutaj algorytmu uczenia perceptronu, ale możecie Państwo poeksperymentować i z innymi metodami wyznaczania płaszczyzny decyzyjnej.
2. Testy algorytmu dla wielowymiarowych danych cyfr. Przeprowadzenie tego testu ma na celu sprawdzenie wydajności algorytmu oraz jego zachowania w przestrzeni wielowymiarowej. Dane źródłowe mogą być użyte bezpośrednio (tzn. każdy piksel

oryginalnych obrazów jest traktowany jako cecha), lub po redukcji wymiarowości algorytmem PCA.

3. Rozwiązanie „kanoniczne” polega na przygotowaniu 45 klasyfikatorów – po jednym dla każdej pary cyfr – i przeprowadzeniu głosowania klasyfikatorów na poszczególne cyfry. Dostarczony kod przeprowadza głosowanie jednomyślne (tzn. wszystkie 9 klasyfikatorów „znających” konkretną cyfrę musi być zgodnych co do klasy). Możecie Państwo zmienić tę zasadę, ale w kolejnych eksperymentach trzymajcie się takiej samej metody głosowania.

Na podstawie analizy jakości głosowania klasyfikatorów elementarnych (można także wziąć pod uwagę macierz pomyłek na zbiorze uczącym) typujecie Państwo klasyfikator(y) elementarne do poprawy.

4. Ostatni krok, to usprawnienie wybranego klasyfikatora elementarnego. Jedynym ograniczeniem jest używanie klasyfikatorów liniowych w przestrzeni cech. Nadzieję na poprawę daje użycie do rozróżniania cyfr zespołu klasyfikatorów. Wynik pracy tego zespołu należy traktować dokładnie tak, jak wyniki pozostałych klasyfikatorów z zespołu referencyjnego.

Pierwszym pomysłem (niekoniecznie najlepszym), który przychodzi mi do głowy, jest wykonanie grupowania problemowych cyfr i skorzystanie z gotowej infrastruktury do zrobienia zespołu, który będzie rozróżniać tylko dwie cyfry. Parę kwestii zostaje otwartych: na ile grup dzielić cyfry? jak ustalić, czy grupowanie dało satysfakcjonujący wynik? czy klasyfikatory uczyć na pojedynczych grupach, czy na ich agregacjach? jak w zwięzły i kompletny sposób pokazać właściwości zespołu?

W sprawozdaniu proszę zamieścić:

1. Opis metody klasyfikacji znaków przy użyciu klasyfikatorów liniowych.
2. Opis algorytmu wyznaczania parametrów płaszczyzny decyzyjnej.
3. Dane dotyczące jakości klasyfikacji każdego z wykorzystywanych klasyfikatorów liniowych.
4. Dane dotyczące jakości klasyfikacji cyfr, z wnioskami wynikającymi z analizy **macierzy pomyłek**.
5. Opis niestandardowego podejmowania decyzji klasyfikacyjnej i porównanie jej wyników z rozwiązaniem kanonicznym.

Tu należy podać poprawę jakości zastępowanego klasyfikatora (nie od rzeczy będzie udokumentować jakość klasyfikatorów składowych; jak to dobrze zrobić?) oraz efekt końcowy zastosowania tego klasyfikatora w rozwiązaniu referencyjnym. Warto porównać nie tylko współczynniki klasyfikacji, ale i zmiany w macierzy pomyłek.

Do pakietu sprawozdanie + kod proszę **nie dołączać danych** (koszt dołączenia danych wynosi 2 punkty)!

Osoby zainteresowane mogą sprawdzić, jakie wyniki klasyfikacji osiągną na zbiorze danych przygotowanym ze znaków pisanych lokalnie (tzn. w Warszawie). Zestaw danych w formacie MNIST jest dostępny pod adresem:

<http://galeranew.ii.pw.edu.pl/~rkz/rob/pldigits.zip>. O skutkach użycia zbioru uczącego pochodzącego z innej populacji, niż dane faktycznie klasyfikowane mówiłem, ale warto zobaczyć skutki na własne oczy.