

Obliczenia Inteligentne	Projekt 2 — Zadanie 1	
Grupa: Grupa 1	Dzień i czas: Czwartek, 10:00	Rok akademicki: 2023/24
Imię i nazwisko: JAKUB PAWLAK	Imię i nazwisko: MAGDALENA PAKUŁA	

1 Opis ekstrakcji cech — Osoba 1

Analiza głównych składowych (PCA — Principal Component Analysis) to technika redukcji wymiarowości danych powszechnie stosowana do ekstrakcji cech. W kontekście zestawu danych MNIST, PCA może być stosowana do zmniejszenia wymiarowości danych obrazowych, zachowując przy tym większość ich wariancji.

Metoda ekstrakcji cech PCA przekształca dane obrazowe o wysokiej wymiarowości do przestrzeni o niższej wymiarowości, identyfikując główne składowe danych. Te główne składowe to ortogonalne kierunki w przestrzeni cech, które przechwytyują maksymalną wariancję danych.

W naszej implementacji PCA jest stosowana do spłaszczenia każdego obrazu o rozmiarze 28x28 pikseli do wektora o wymiarach 784. Następnie otrzymane wektory są przekształcane do przestrzeni o niższej wymiarowości, zwykle dwóch wymiarów w celu wizualizacji.



Rysunek 1: Przykładowy obraz cyfry “5” z zestawu danych MNIST.

Poniżej znajdują się wygenerowane cechy dla powyższego obrazu “5” za pomocą PCA:

Piksel	PCA cecha 1	PCA cecha 2
1	-248.0374	36.2011
2	-248.0374	36.2011
...
28	-248.0374	36.2011

Tablica 1: Wartości wygenerowanych cech za pomocą PCA dla obrazu “5”.

Binary Patterns (LBP) to deskryptor tekstury używany do ekstrakcji cech w obrazach. W kontekście zestawu danych MNIST, LBP może być stosowany do wydobywania cech tekstury z obrazów.

Metoda ekstrakcji cech LBP działa przez podzielenie obrazu na małe obszary i porównanie każdego piksela z otaczającymi go pikselami. Na podstawie tych porównań dla każdego piksela tworzony jest wzorzec binarny. Poprzez zliczanie wystąpień różnych wzorców binarnych konstruowany jest histogram reprezentujący cechy tekstury obrazu.

Poniżej znajdują się wygenerowane cechy dla powyższego obrazu “5” za pomocą LBP:

LBP cecha	Wartość
1	3.0
2	11.0
3	6.0
4	32.0
5	48.0
6	34.0
7	2.0
8	4.0
9	618.0
10	26.0

Tablica 2: Wartości wygenerowanych cech za pomocą LBP dla obrazu “5”.

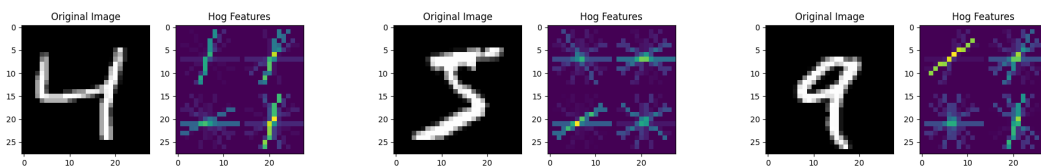
2 Wyniki eksperymentu — Osoba 1

3 Opis ekstrakcji cech — Osoba 2

Histogram of Oriented Gradients (HOG) to metoda ekstrakcji cech używana w przetwarzaniu obrazu. Opiera się ona na zliczaniu gradientów zorientowanych w tym samym kierunku, w określonych fragmentach obrazu. Deskryptor HOG opisuje kształt obiektu na obrazie, więc bardzo dobrze nadaje się do zadania rozpoznawania cyfr, ponieważ następuje ono właśnie na podstawie kształtu. W celu poprawy jakości, wartości gradientów są dodatkowo normalizowane w większych grupach.

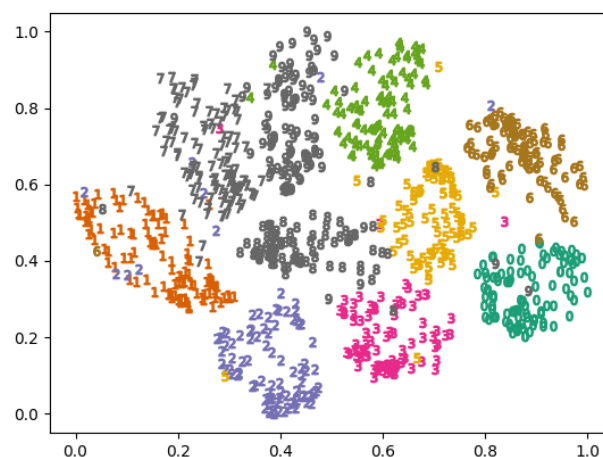
W przypadku zbioru MNIST użyto ekstrakcji HOG rozpatrującej 9 kierunków zorientowania gradientu (w celu wykrycia linii zarówno ortogonalnych jak i ukośnych), a cały obraz został podzielony na 4 komórki o rozmiarze 14×14 px. Histogramy gradientów znormalizowano w odniesieniu do całości obrazu.

W ten sposób obraz 28×28 zredukowano do 36-elementowego wektora.



Rysunek 2: Przykłady ekstrakcji cech metodą HOG

t-Distributed Stochastic Neighbor Embedding (t-SNE) jest stochastyczną metodą redukcji wymiarów. Próbuje ona rozłożyć punkty w przestrzeni docelowej zachowując lokalnych sąsiadów z przestrzeni źródłowej. Metoda ta bada odległości między punktami w przestrzeni źródłowej i przypisuje im rozkład prawdopodobieństwa w oparciu o rozkład standardowy. Następnie wybiera (losowo lub przez PCA) rozkład punktów w przestrzeni docelowej i analizuje ich odległości, przypisując im prawdopodobieństwa oparte o rozkład Cauchy'ego. Następnie za pomocą metody minimalizacji gradientu stara się zminimalizować różnicę pomiędzy rozkładami w przestrzeni źródłowej i docelowej.



Rysunek 3: Wizualizacja docelowej przestrzeni t-SNE

4 Wyniki eksperymentu — Osoba 2

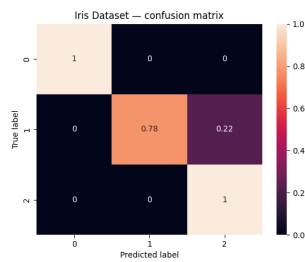
5 Wybór optymalnego modelu

Podczas eksperymentów przeprowadzonych w ramach pierwszego projektu dało się zauważyć, że w miarę zwiększania liczby neuronów w warstwie ukrytej, skuteczność ewentualnie ulegała wypłaszczeniu. W takim przypadku, zwiększanie liczby neuronów skutkowałoby jedynie utrudnieniem obliczeń, bez pozytywnego efektu na skuteczności modelu. Optymalnym modelem jest zatem taki, który maksymalizuje dokładność przy jednoczesnym minimalizowaniu ilości neuronów.

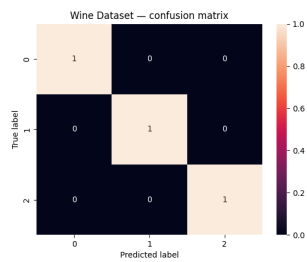
Przeprowadzone eksperymenty nie dostarczyły natomiast jednoznacznego sposobu, aby a priori określić najlepszą liczbę neuronów w warstwie ukrytej. Wybór optymalnego modelu został dokonany poprzez coraz dokładniejsze przeszukiwanie przestrzeni liczb naturalnych, podobne w zasadzie działania do algorytmu *binary search*. Różne wartości liczby neuronów były poddawane przyśpieszonemu procesowi uczenia (na mniejszym zbiorze danych oraz mniejszą ilością epok), którego to wyniki pozwalały na oszacowanie, które wartości są najbardziej obiecujące. Następnie wybierane zostały kolejne ilości neuronów do sprawdzenia — tym razem z sąsiedztwa najlepiej spisujących się w poprzedniej iteracji. Proces powtarzano, aż do znalezienia lokalnego maksimum skuteczności.

Metoda ta pozwoliła uzyskać satysfakcjonujące wyniki, jednak wskazać należy, że jesteśmy świadomi iż nie daje ona gwarancji, że wybrany model jest globalnie optymalny.

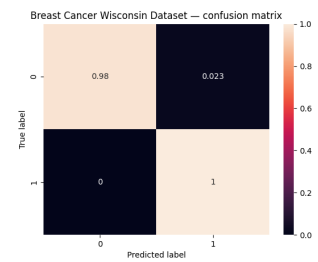
6 Wyniki klasyfikacji dla pierwszego sposobu ekstrakcji cech



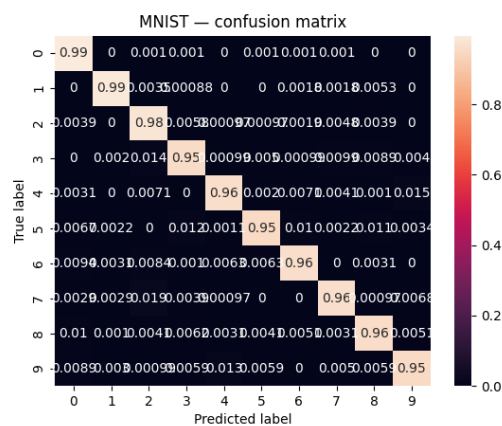
(a) Zbiór Iris



(b) Zbiór Wine



(c) Zbiór Breast Cancer Wisconsin



(d) Zbiór MNIST

Rysunek 4: Wyniki klasyfikacji dla 1 sposobu ekstrakcji cech

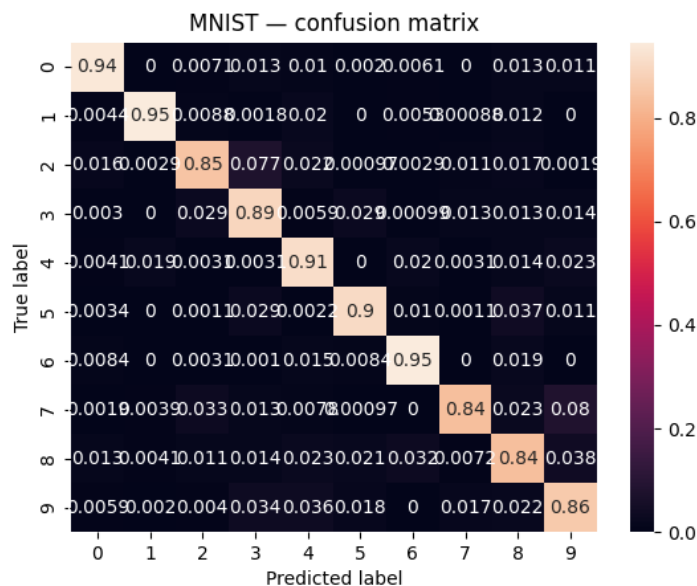
Zbiór danych	Wartość Accuracy
Iris	93.33%
Wine	100%
Breast Cancer Wisconsin	99.12%
MNIST	97.16%

Tablica 3: Wartości accuracy wytrenowanego modelu

7 Wyniki klasyfikacji — Osoba 1

8 Wyniki klasyfikacji — Osoba 2

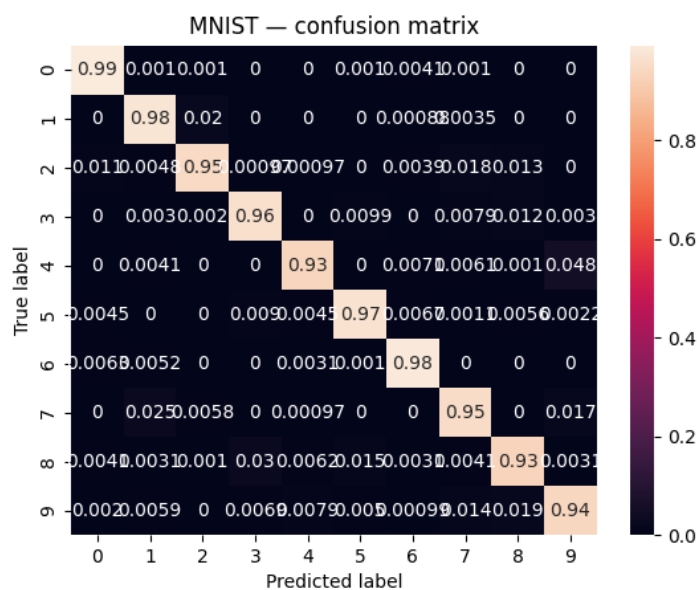
8.1 HOG



Rysunek 5: Wyniki klasyfikacji dla ekstrakcji cech metodą HOG

Accuracy: 89.21%.

8.2 t-SNE



Rysunek 6: Wyniki klasyfikacji dla ekstrakcji cech metodą HOG

Accuracy: todo