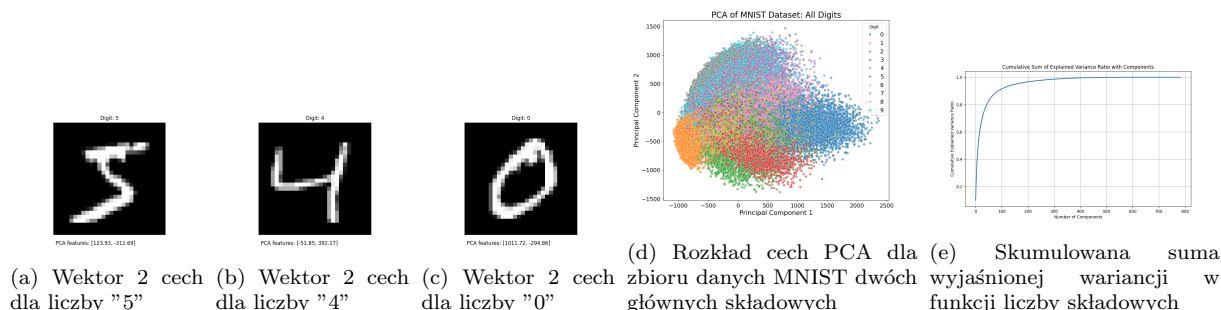


Obliczenia Inteligentne	Projekt 2 — Zadanie 1	
Grupa: Grupa 1	Dzień i czas: Czwartek, 10:00	Rok akademicki: 2023/24
Imię i nazwisko: JAKUB PAWLAK	Imię i nazwisko: MAGDALENA PAKUŁA	

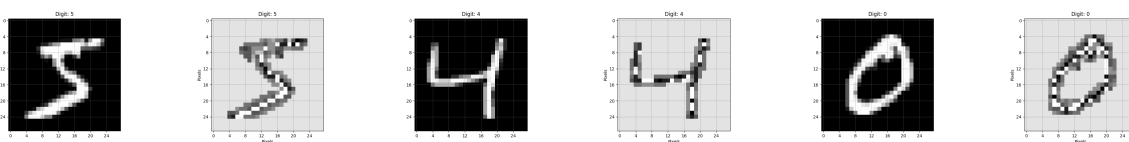
1 Opis ekstrakcji cech — Osoba 1

Principal Component Analysis (PCA) - Analiza głównych składowych technika redukcji wymiarowości, która pomaga w przekształcaniu danych wielowymiarowych w postać niskowymiarową, przy jednoczesnym zachowaniu jak największej oryginalnej zmienności. Metoda PCA standaryzuje dane, oblicza macierz kowariancji rejestrując relacje między cechami, a następnie dekomponuje ją na wartości własne i wektory własne. Wektory własne reprezentują kierunki największej wariancji, a wartości własne jej wielkość. Po posortowaniu wartości własnych malejąco, wybierane są największe z nich i odpowiadające im wektory własne do utworzenia nowej przestrzeni cech. Dane są rzutowane na te przestrzenie, co skutkuje zbiorem danych o zredukowanej liczbie wymiarów, zachowującym najważniejsze informacje. W kontekście danych MNIST, PCA może być stosowana do redukcji liczby wymiarów obrazów cyfr, co ułatwia ich analizę i wizualizację, a także przyspiesza trening modeli uczenia maszynowego poprzez zmniejszenie ilości przetwarzanych danych przy jednoczesnym zachowaniu istotnych informacji.

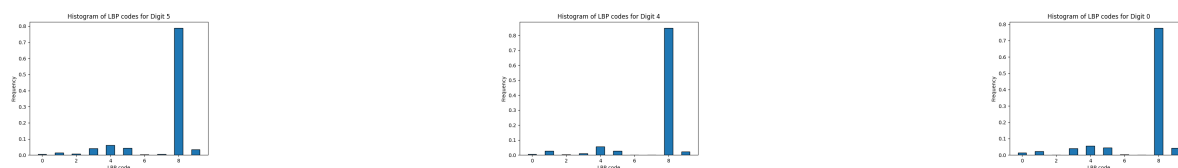


Rysunek 1: Przykłady ekstrakcji cech metoda PCA

Binary Patterns (LBP) - Lokalne wzorce binarne to technika ekstrakcji cech używana głównie w analizie obrazów, która przekształca obraz w zestaw wartości binarnych opisujących teksturę. Proces ten polega na porównywaniu każdego piksela z jego sąsiadami. Piksel centralny jest traktowany jako próg, a każdy sąsiedni piksel jest porównywany do tego progu. Jeśli wartość sąsiada jest większa lub równa wartości centralnego piksela, przypisywana jest wartość 1, w przeciwnym razie 0. Wynikowe wartości binarne są następnie łączone w jednolitą wartość, która reprezentuje wzorec tekstur w danym obszarze obrazu. Metoda LBP jest odporna na zmiany jasności i kontrastu, co czyni ją użyteczną w różnych zastosowaniach, takich jak rozpoznawanie twarzy czy klasyfikacja tekstur. W przypadku danych MNIST, LBP może być używane do ekstrakcji cech, które są następnie wykorzystywane do klasyfikacji obrazów cyfr.



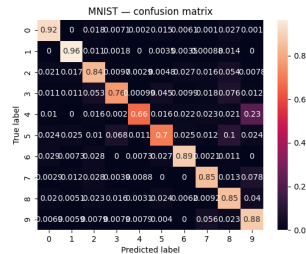
Rysunek 2: Przykłady ekstrakcji cech metoda LBP



Rysunek 3: Histogramy LBP dla cyfry 0

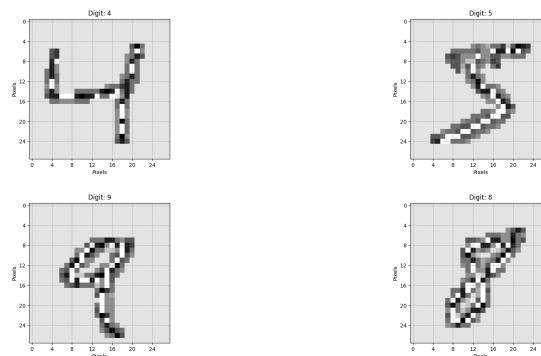
2 Wyniki eksperymentu — Osoba 1

Dla ekstrakcji cech w celu oceny separowalności wytrenowano sieć z małą ilością epok i oceniono macierz pomyłek, która znajduje się na rys ??.



Rysunek 4: Macierz pomyłek na słabo wytrenowanej sieci (ekstakcja LBP)

Zauważyć można, że najgorzej model radzi sobie z rozpoznawaniem cyfr: 3,4,5. Cyfra 4 bardzo często uznawana jest za 9. Często mylone są również 5 i 8 oraz 5 i 3. Wynika to z faktu, że (LBP) opiera się na lokalnych różnicach jasności pikseli w obrazie. W przypadku cyfr o podobnych kształtach, takich jak 4 i 9 lub 5 i 8, te lokalne różnice mogą być niewystarczające do ich jednoznacznego rozróżnienia. Dla przykładu, na rys. ?? przedstawiono różne cyfry, które po ekstrakcji cech wyglądają podobnie. Tak oto przedstawione cyfry 4 oraz 9 mają pionową kreskę pośrodku i podobne zaokrąglenia u góry i u dołu. To podobieństwo może prowadzić do tego, że cechy LBP są dość zbliżone dla tych cyfr, co utrudnia ich rozróżnienie dla klasyfikatora. Obie posiadają w prawej górnej ćwiartce linie ukośną i poziomą. Cyfry 5, jak i 8 mają pojedynczą strukturę zamkniętej petli. Kluczowa różnica polega na orientacji i krzywiźnie linii. Nieco grubsza górna krzywizna w 8 w porównaniu z 5 lub mniej wyraźna dolna krzywizna w 5 może prowadzić do bardzo podobnych wzorców LBP.



(a) Podobne cechy dla różnych cyfr: 4 i 9

(b) Podobne cechy dla różnych cyfr: 5 i 8

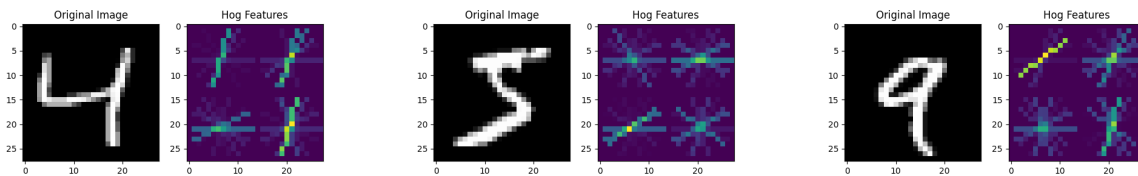
Rysunek 5: Różne cyfry, które posiadają podobne cechy

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.

Alorytm najpierw dzieli obraz na komórki o określonym rozmiarze. W przypadku zbioru MNIST użyto komórek 14×14 , uzyskując podział całego obrazu na 4 komórki. W każdej komórce, oblicza się dla każdego piksela lokalny gradient. Następnie, wewnątrz każdej komórki zlicza się gradienty w określonych kierunkach i tworzy się z nich histogram. Aby umożliwić wykrycie linii zarówno ortogonalnych jak i ukośnych, liczba kierunków została ustawiona na 9. W celu poprawy jakości, wartości gradientów są dodatkowo normalizowane w większych grupach. W tym przypadku, użyto grup o rozmiarze 2×2 komórek, co odpowiada całemu obrazowi.

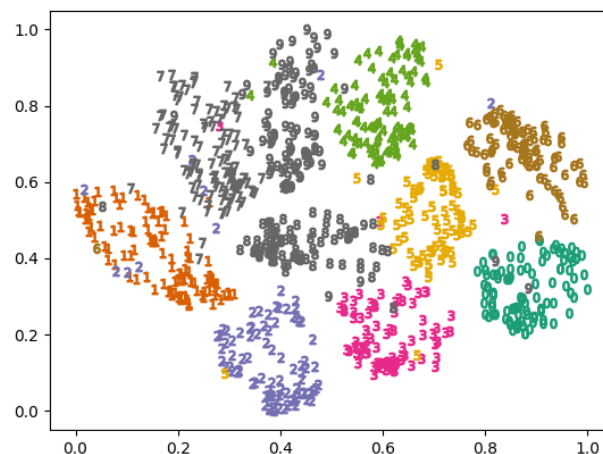
Cały obraz zostaje zatem opisany za pomocą 4 komórek, każda zawierająca wartości dla 9 kierunków. W ten sposób obraz 28×28 zredukowano do 36-elementowego wektora.



Rysunek 6: Przykłady ekstrakcji cech metoda HOG

t-Distributed Stochastic Neighbor Embedding (t-SNE) jest stochastyczna metoda redukcji wymiarów, często wykorzystywana przy tworzeniu wizualizacji.

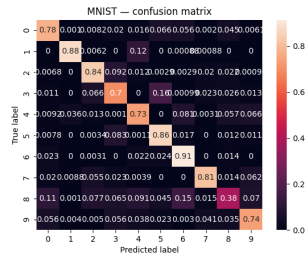
Próbuje ona rozłożyć punkty w przestrzeni docelowej zachowując lokalnych sąsiadów z przestrzeni źródłowej. Metoda ta bada odległości między punktami w przestrzeni źródłowej i przypisuje im rozkład prawdopodobieństwa w oparciu o rozkład standardowy. Następnie wybiera (losowo lub przez PCA) rozkład punktów w przestrzeni docelowej i analizuje ich odległości, przypisując im prawdopodobieństwa oparte o rozkład Cauchy'ego. Następnie za pomocą metody minimalizacji gradientu stara się zminimalizować różnice pomiędzy rozkładami w przestrzeni źródłowej i docelowej.



Rysunek 7: Wizualizacja docelowej przestrzeni t-SNE

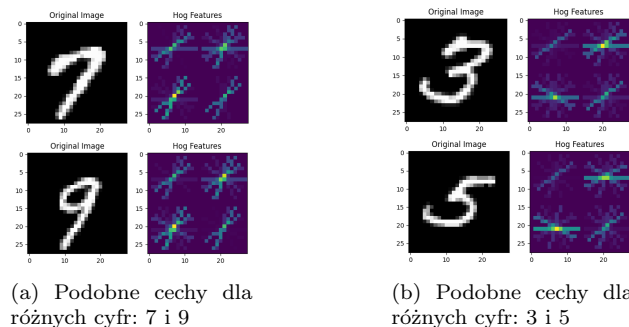
4 Wyniki eksperymentu — Osoba 2

Dla ekstrakcji cech w celu oceny separowalności wytrenowano sieć z małą ilością epok i oceniono macierz pomyłek, która znajduje się na rys ??.



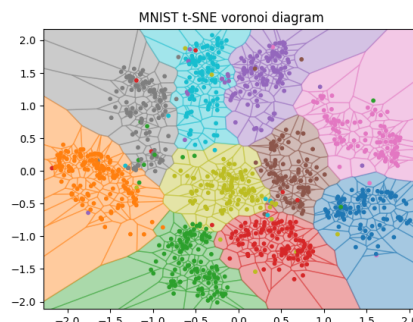
Rysunek 8: Macierz pomyłek na słabo wytrenowanej sieci (ekstrakcja HOG)

Zauważyć można, że najgorzej model radzi sobie z rozpoznawaniem cyfr 8,3,4, 9 i 0. Cyfra 8 bardzo często uznawana jest za 0,2,3,4 lub 6. Często mylone są również 3 i 5. Wynika to z faktu, że algorytm HOG wewnątrz komórek uwzględnia tylko ilość gradientów, a nie ich położenie względem siebie. Dla przykładu, na rys. ?? przedstawiono różne cyfry, które po ekstrakcji cech wyglądają podobnie. Tak oto przedstawione cyfry 3 oraz 5 obie posiadają w prawej górnej ćwiartce linię ukośną i pozioma. W cyfrze 3 linia pozioma idzie w prawo, a w cyfrze 5 — w lewo, natomiast dla algorytmu HOG nie ma to znaczenia, ponieważ obie linie znajdują się w tej samej komórce. Podobnie z cyframi 7 i 9, różniacymi się jedynie poziomym domknięciem, co jednak nie znajduje odzwierciedlenia w deskrypcji HOG, ponieważ pozioma linia już występuje gdzieś indziej w tej samej ćwiartce.



Rysunek 9: Różne cyfry, które posiadają podobne cechy

W metodzie t-SNE można ocenić separowalność wizualnie, na podstawie rys. ??, jak również diagramu Woronoja na rys. ?. Od razu widać, że poszczególne klasy są separowalne, choć zdarzają się pola do pomyłek, zwłaszcza przy cyfrach 3,5 i 8, oraz 7,9 i 4.



Rysunek 10: Diagram Woronoja dla zbioru po przekształceniu t-SNE

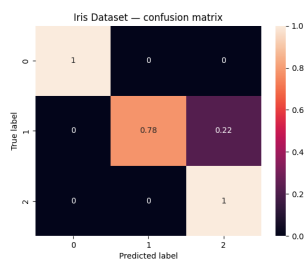
5 Wybór optymalnego modelu

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.

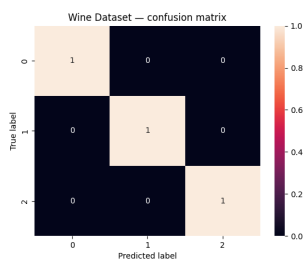
Przeprowadzone eksperymenty nie dostarczyły natomiast jednoznacznego sposobu, aby a priori określić najlepszą liczbę neuronów w warstwie ukrytej. Wybór optymalnego modelu został dokonany poprzez coraz dokładniejsze przeszukiwanie przestrzeni liczb naturalnych, podobne w zasadzie działania do algorytmu *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.

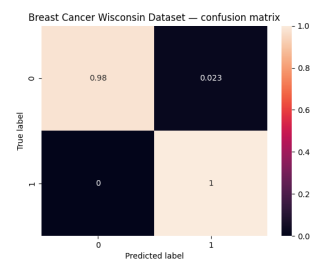
6 Wyniki klasyfikacji dla pierwszego sposobu ekstrakcji cech



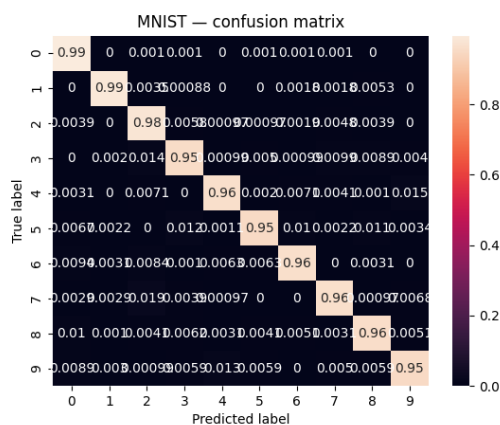
(a) Zbiór Iris



(b) Zbiór Wine



(c) Zbiór Breast Cancer Wisconsin



(d) Zbiór MNIST

Rysunek 11: Wyniki klasyfikacji dla 1 sposobu ekstrakcji cech

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

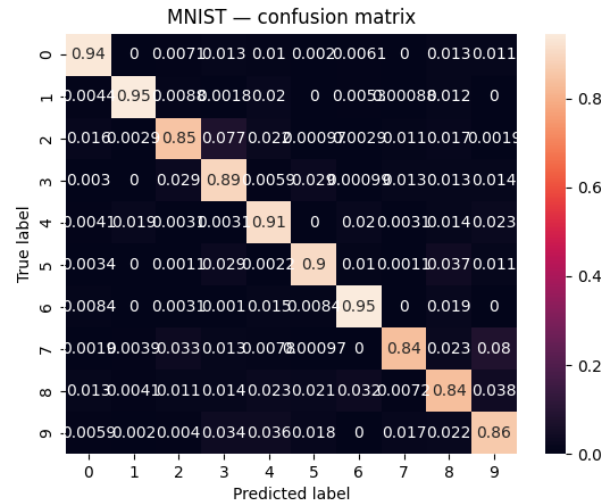
Tabela 1: Wartości accuracy wytrenowanego modelu

7 Wyniki klasyfikacji — Osoba 1

8 Wyniki klasyfikacji — Osoba 2

8.1 HOG

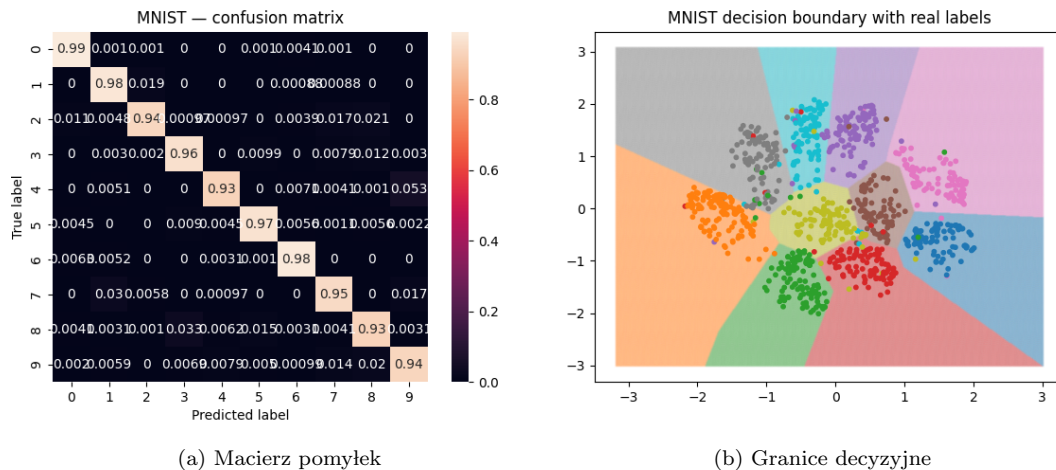
Wyniki klasyfikacji z metoda HOG przedstawiono na rys. ???. Widać znacząco poprawę w stosunku do wstępnych oczekiwań, choć tak jak przewidziano, z cyframi 2,3,7,8,9 model radzi sobie słabiej niż z resztą. Tak jak przewidywano, często mylone są cyfry pary (2,3), (3,5), (7,9), (8,6), (8,9). Ogólna wartość accuracy osiągnęła 89.21%.



Rysunek 12: Wyniki klasyfikacji dla ekstrakcji cech metoda HOG (Accuracy: 89.21%).

8.2 t-SNE

Na rys. ??? przedstawiono macierz pomyłek oraz przebieg granicy decyzyjnej. Model poradził sobie bardzo dobrze, jednak można zauważyć drobne pomyłki. Zgodnie z przewidywaniami, 4 i 7 były często uznawane za 9, jak również mylone były cyfry z grupy 3,5,8. Pomimo mniejszej liczby cech, model osiągnął lepszą niż w poprzedniej metodzie wartość accuracy wynoszącą 95.67%. Wskazać jednak należy istotną wadę tej metody, mianowicie t-SNE operuje na całym zbiorze danych, więc użycie tej metody ekstrakcji cech nie będzie można zastosować w przypadku pojawienia się nowych danych.



(a) Macierz pomyłek

(b) Granice decyzyjne

Rysunek 13: Wyniki klasyfikacji dla ekstrakcji cech metoda t-SNE (Accuracy: 95.67%)