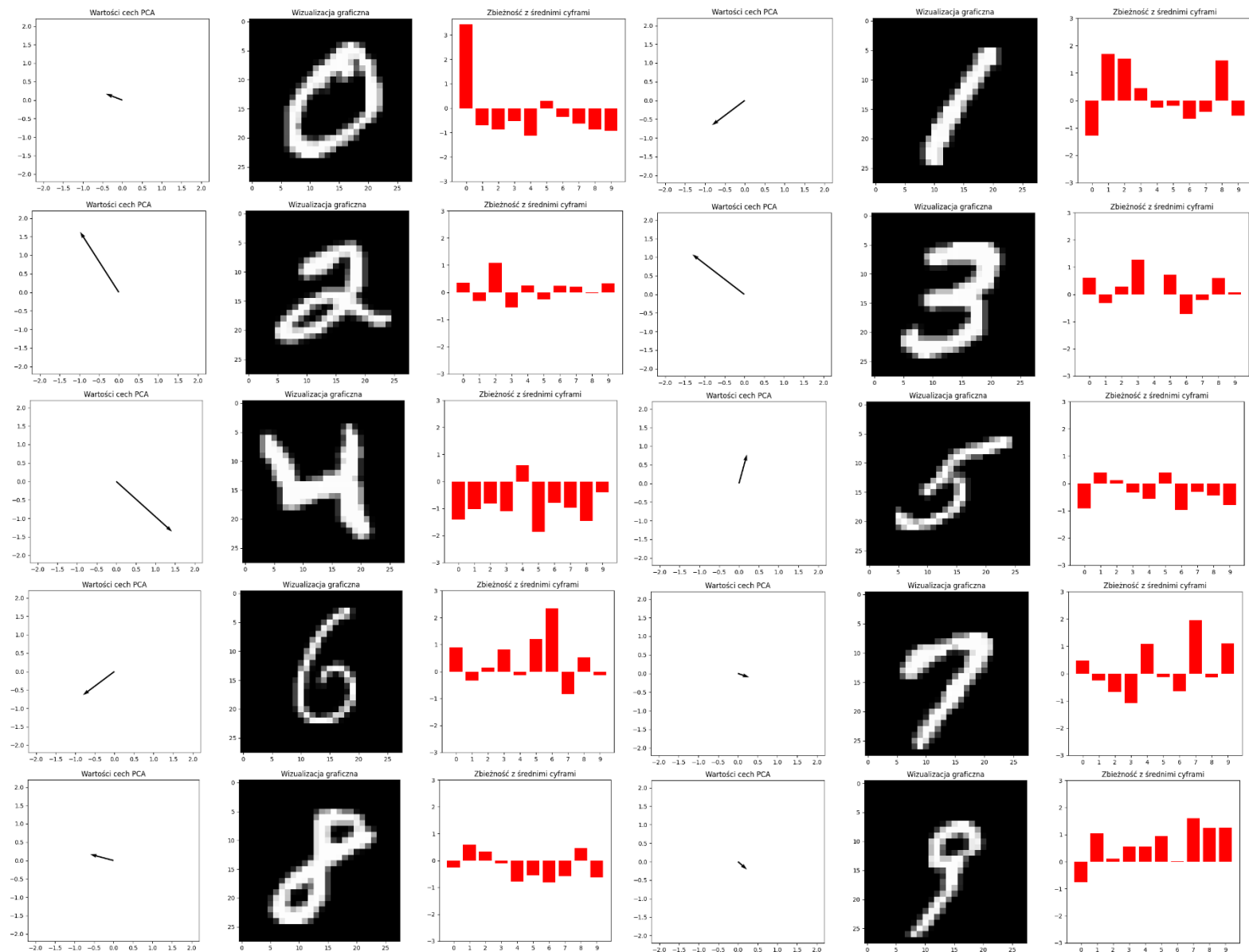


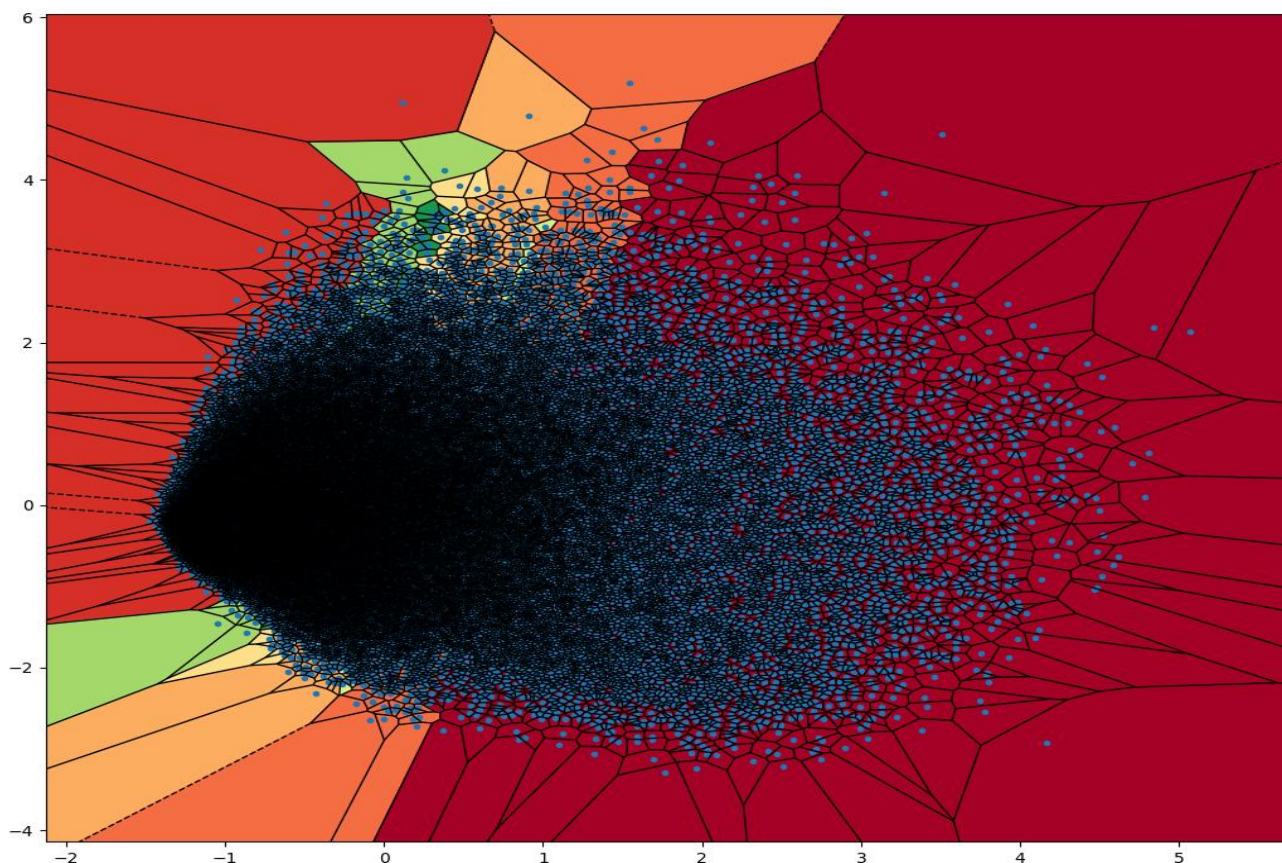
1.1 Opis dwóch sposobów ekstrakcji: algorytm PCA i konwolucja z maskami średnich cyfr



Pierwszy algorytm ekstrakcji, PCA (Principal Component Analysis) z biblioteki scikit learn przypisuje uprzednio znormalizowanemu zbiorowi danych $A_{n,d}$ zredukowany zbiór danych $B_{n,r}$, gdzie n – ilość punktów danych, d – ilość oryginalnych cech, r – żądana ilość zredukowanych cech poprzez scentralizowanie wszystkich cech względem siebie a następnie odnalezienie dekompozycji macierzy $A_{n,d} = USV^*$, gdzie U i V są ortonormalne a S diagonalna nieujemna i zdefiniowanie $B_{n,r} = U\hat{S}V^*$, gdzie \hat{S} jest równe S wszędzie, poza $(d-r)$ najmniejszymi wartościami macierzy S , którym przypisane jest 0. Algorytm ten realizuje zadanie minimalizacji normy Frobeniusa różnicy macierzy $A_{n,d}$ i $B_{n,r}$, przez co wynikowa macierz może być uważana jako aproksymacja macierzy oryginalnej. Dane na koniec są normalizowane dla każdej cechy. Metoda ta została wybrana ze względu na zdolności aproksymacyjne i dowolność wymiaru zredukowanej macierzy dzięki czemu może posłużyć do wizualizacji danych.

Drugi algorytm ekstrakcji wpieryw generuje 10 mask na podstawie zbioru treningowego, poprzez sumowanie obrazów o zgodnej etykiecie dla całego zbioru, a następnie znormalizowanie ich przez podzielenie każdego elementu przez sumę wszystkich elementów maski do której jest przypisany. Następnie każdy punkt danych zostaje przypisany do wektora 10 cech, będących konwolucją kolejnych mask z punktem podzielonym przez sumę elementów punktu. Dane na koniec są normalizowane dla każdej cechy. Metoda ta została wybrana poprzez heurystyczne podejście do problemu klasyfikacji z myślą że cyfra powinna być najbardziej podobna do tych samych cyfr.

1.2 Wyniki eksperymentu pierwszego: algorytm PCA i konwolucja z maskami średnich cyfr



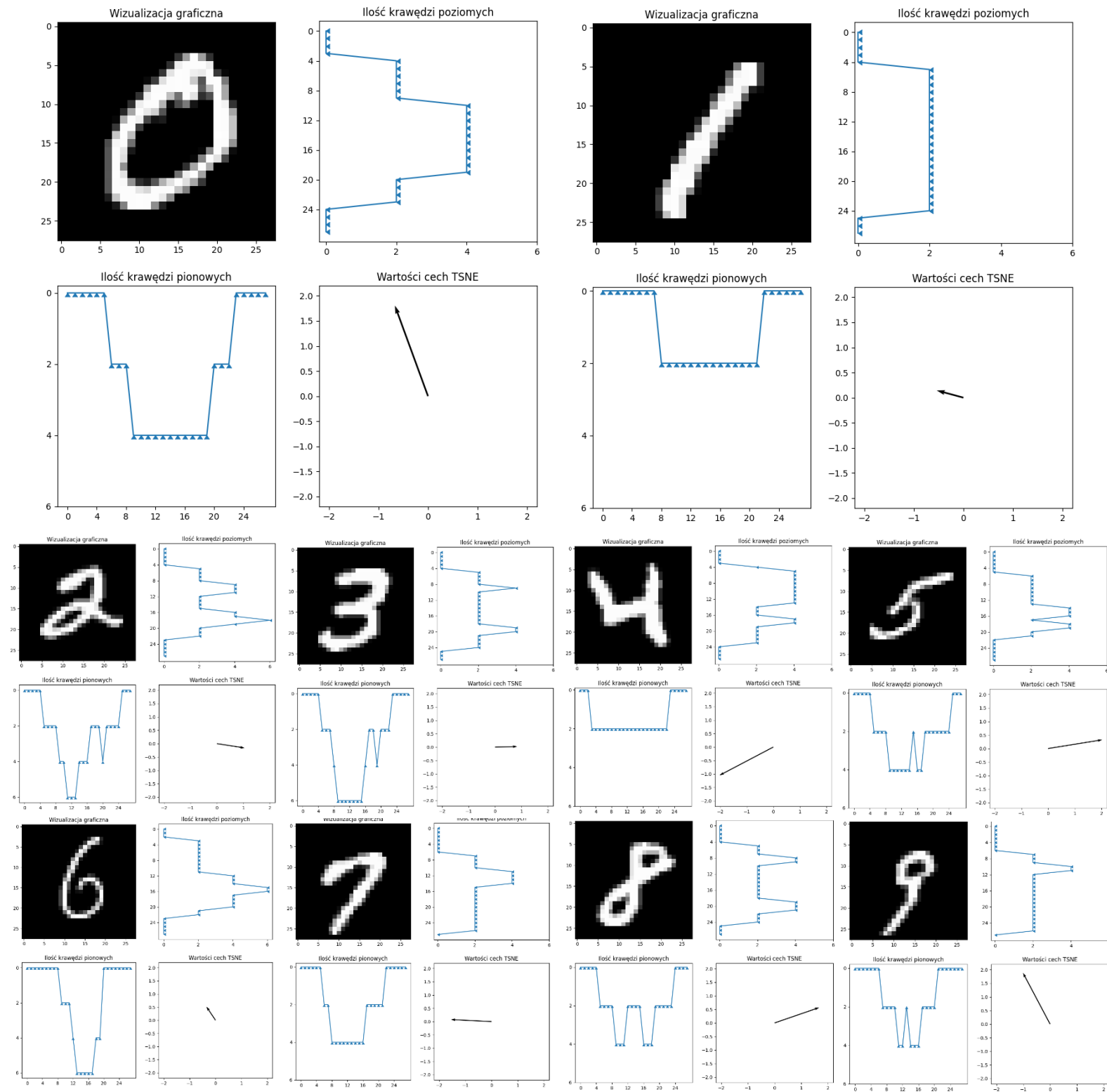
Na diagramie Voronoi'a przestrzeni ekstrahowanych cech metodą PCA i prawdziwych etykiet obserwujemy, że punkty skupione są w centrum o znacznej gęstości, która wraz z odległością spada z prędkością zależącą od kierunku. Po znaczącym stopniu "przemieszania" danych i ogólnym ich rozkładzie w przestrzeni wnioskujemy bardzo niską separowalność danych, co kwantyfikuje niska miara Silhouette Score równa -0.07.

Dla ekstrakcji poprzez zbieżność z średnimi cyframi dla zbioru treningowego wartość miary Silhouette (z wykorzystaniem oryginalnych etykiet) wyniosła 0.11. Sugeruje ona, na podstawie prawdziwych etykiet dla poszczególnych wektorów cech, że są one średnio separowalne – jednakże dobrze wyuczony i przygotowany model będzie w stanie sobie z takim problemem poradzić.

W przypadku metody PCA najczęściej mogą być mylone pary (7,8), (7,9), (2,3) - patrząc na kierunek zobrazowanych wektorów.

W przypadku poprzez zbieżność z średnimi cyframi najczęściej mogą być mylone pary (3,8), (7,9) - patrząc na zbieżności zobrazowane na wykresach słupkowych.

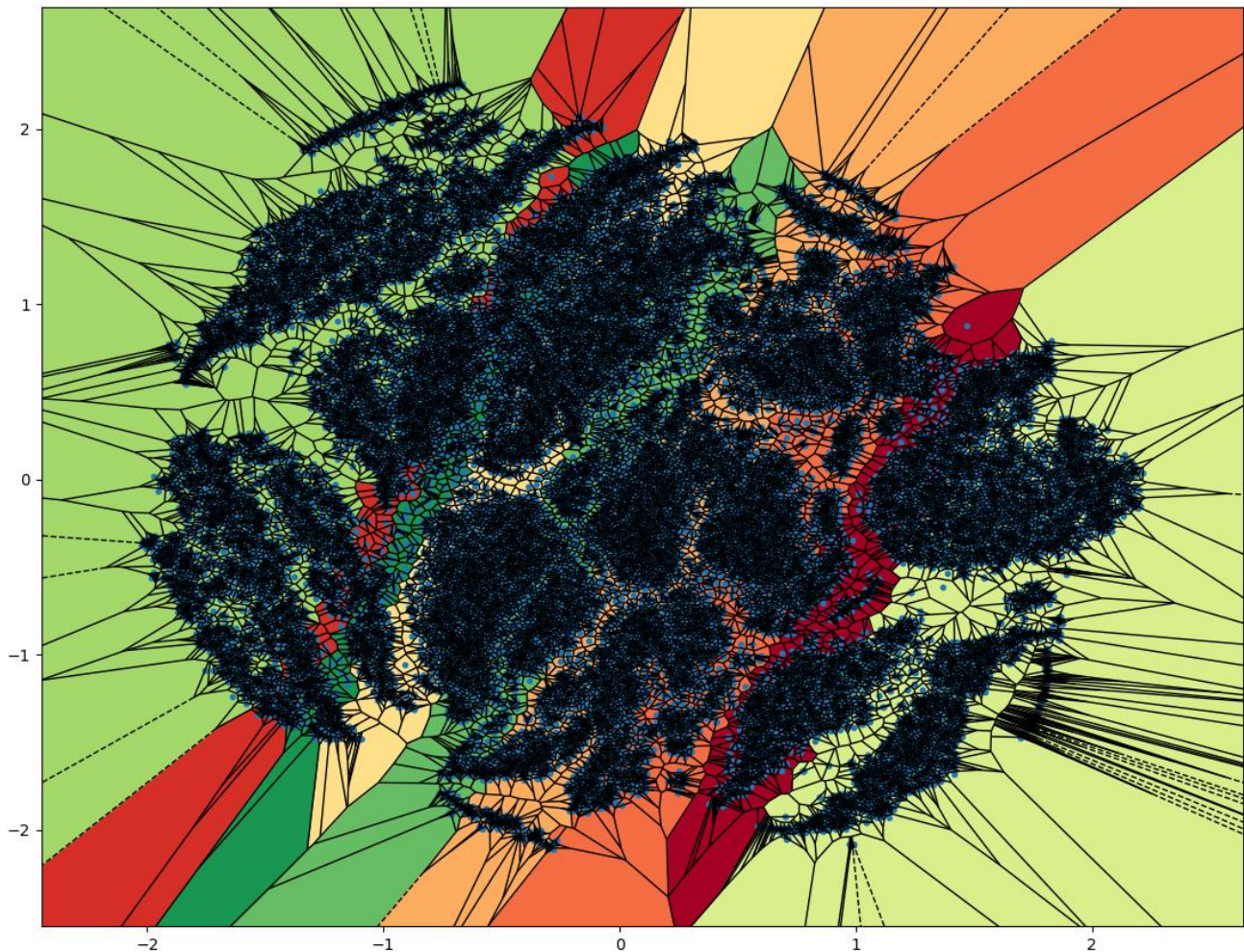
1.3 Opis dwóch sposobów ekstrakcji: wykrywanie krawędzi oraz algorytm TSNE



Pierwszą metodą ekstrakcji jest po uprzedniej normalizacji danych zastosowanie algorytmu t-SNE z biblioteki scikit learn który metodami probabilistycznymi generuje rozkłady w przestrzeni o zredukowanych wymiarach i wybiera ten, który maksymalizuje ilość punktów które zachowają swoje otoczenie z niezredukowanego wymiaru, tj jak najlepiej odzwierciedli odległości pomiędzy punktami. Metoda ta została wybrana ze względu na możliwość reprezentacji względnego grupowania danych.

Drugi algorytm ekstrakcji dla każdego punktu danych przeprowadza progowanie, a następnie zlicza zmiany koloru wzdłuż wszystkich odcinków szerokości i wysokości. Otrzymane 2x28 cechy następnie są normalizowane dla wszystkich danych. Metoda ta została wybrana ponieważ odzwierciedla istotne informacje geometryczne jednocześnie będące relatywnie łatwymi do odnalezienia.

1.4 Wyniki eksperymentu pierwszego: wykrywanie krawędzi oraz algorytm TSNE



W odróżnieniu od diagramu przy metodzie PCA, jesteśmy w stanie wizualnie ocenić kształty potencjalnych klastrów co wraz z rozmieszczeniem etykiet świadczy o obiecującej separowalności biorąc pod uwagę znaczący stopień utraty informacji związany z redukcją 784 wymiarów do 2. Odzwierciedla to wartość Silhouette Score równa 0.23 która nie jest wysoka, jednak znacząco wyższa od 0 reprezentującego bardzo niską separowalność.

Dla sposobu poprzez wykrywanie krawędzi dla zbioru treningowego wartość miary Silhouette (z wykorzystaniem oryginalnych etykiet) wyniosła 0.081. Wartość bliska zero sugeruje możliwość nakładania się na siebie klastrów. Separowalność dla tej metody oceniona została zatem na niską.

W przypadku metody t-SNE najczęściej mogą być mylone pary (3,5), (2,3) oraz (1,7) - patrząc na kierunek zobrazowanych wektorów.

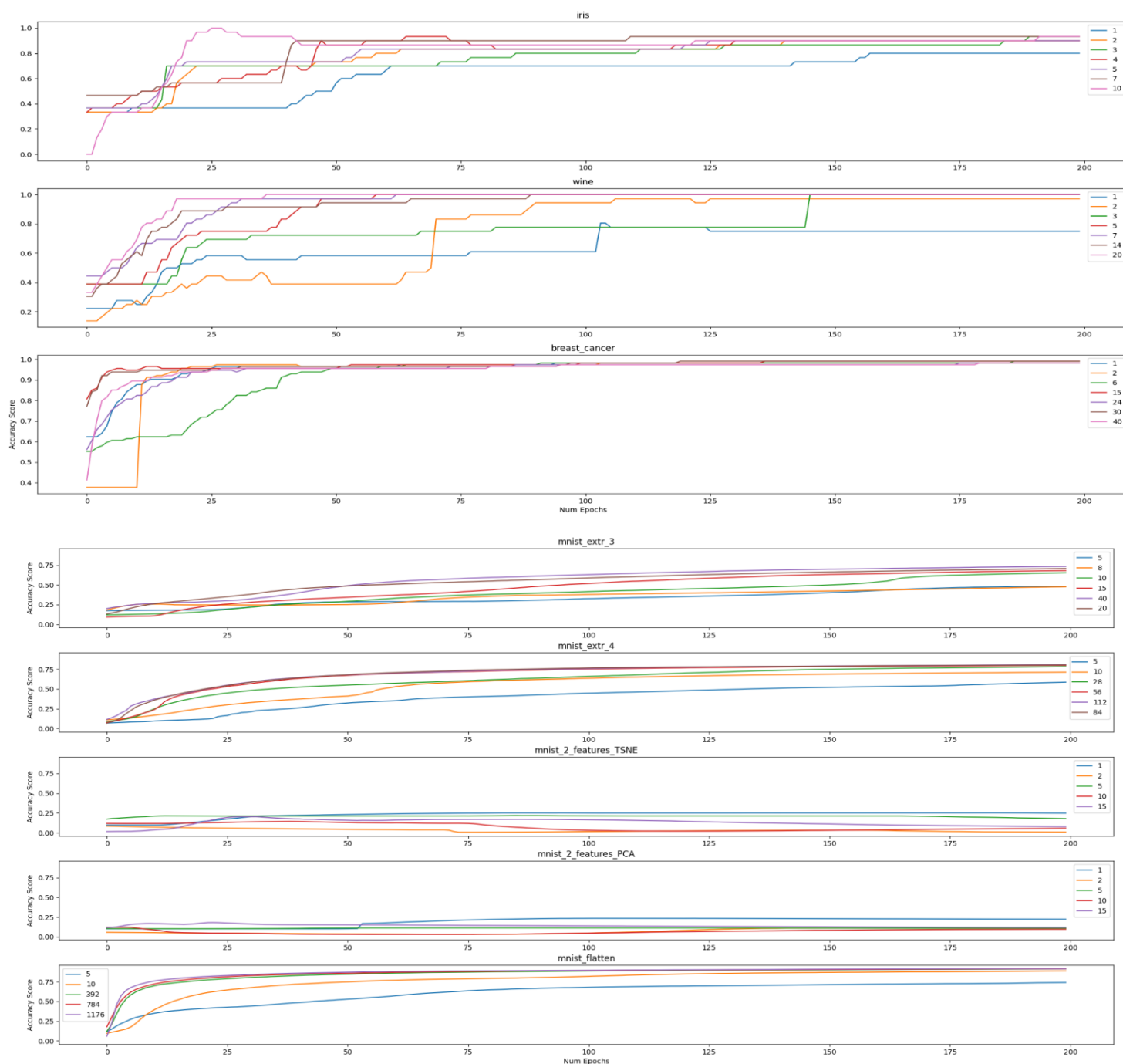
W przypadku metody ekstrakcji poprzez wykrywanie krawędzi najczęściej mogą być mylone pary (2,5), (7,9), (2,3) - patrząc na ilości wykrytych krawędzi w danym miejscu na obrazie.

1.5 Opis sposobu wyboru optymalnego modelu

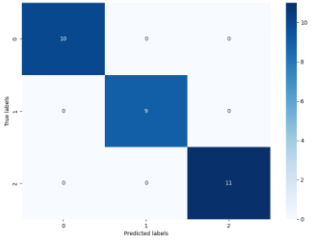
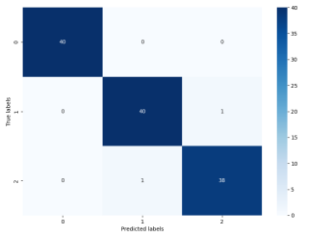
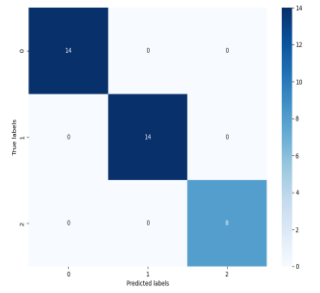
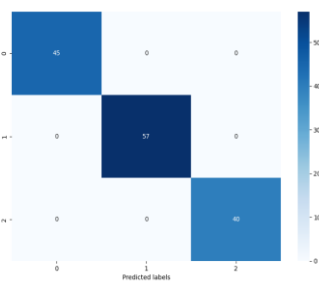
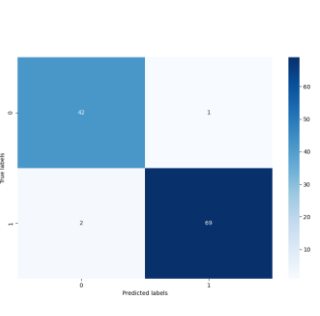
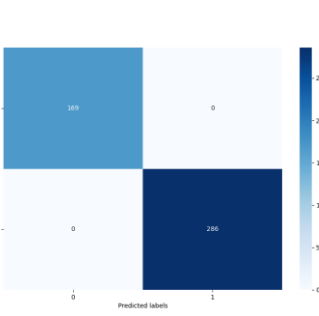
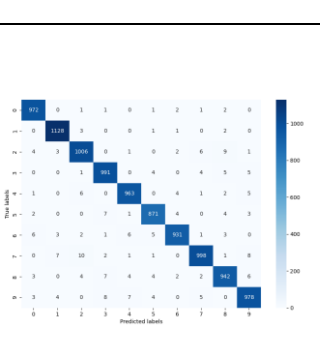
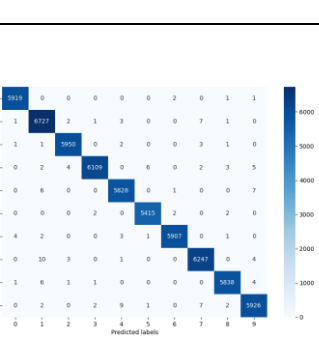
Model został zaprojektowany jako sieć neuronowa z dwoma warstwami liniowymi rozdzielonymi funkcją aktywacji ReLU. Dla warstw liniowych ilość neuronów ukrytych wyznaczono na podstawie trudności zadania klasyfikacji – wzięto pod uwagę liczbę klas do uzyskania oraz liczbę cech dla poszczególnego obiektu.

Do wyboru najlepszego modelu wykorzystano wartość *accuracy* wyliczaną na zbiorze treningowym dla różnych ilości neuronów ukrytych.

Finalnie liczba neuronów ukrytych zbiorów Iris, Wine oraz Breast Cancer została wyznaczona na wartość połowy ilości cech, jednakże z minimalną wartością nie mniejszą niż ilość klas możliwych do uzyskania. W przypadku metod PCA oraz t-SNE zastosowano kryterium maksymalizacji wartości *accuracy*, które wykazało, że najlepszą ilością neuronów jest '1'. Dla ekstrakcji cech ze zbioru MNIST polegającej na spłaszczeniu obrazu przyjęto ilość neuronów ukrytych wynoszącą rozmiar wektora (784). W przypadku ekstrakcji cech poprzez zastosowanie masek przyjęto ilość neuronów ukrytych '40' – taki model posiadał najlepszą wartość dokładności. W przypadku ekstrakcji cech poprzez sprawdzanie krawędzi przyjęto ilość neuronów wynoszącą $1.5 \times \text{ilość cech}$ - sprawdzono również ilość neuronów równą dwukrotności ilości cech, jednakże przyrost dokładności kosztem mocy obliczeniowej nie był opłacalny.



1.6 Opis wyników klasyfikacji dla trzech zbiorów danych oraz zbioru MNIST z pierwszym sposobem ekstrakcji

	TESTOWY		TRENINGOWY		Architektura		
Parametr	Acc	Confusion Matrix	Acc	Confusion Matrix	Ilość neuronów	Ilość cech	Ilość klas
IRIS	1,00		0,98		3	4	3
WINE	1,00		1,00		7	13	3
BREAST CANCER	0,97		1,00		15	30	2
MNIST (ekstrakcja poprzez spłaszczenie obrazu)	0,98		0,99		784	784	10

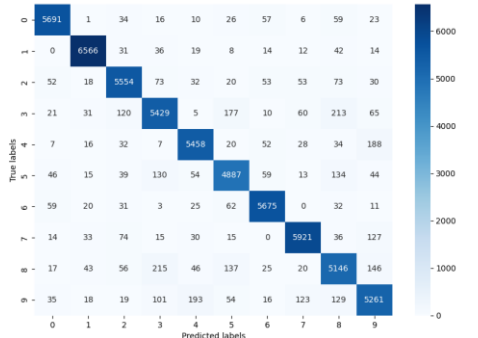
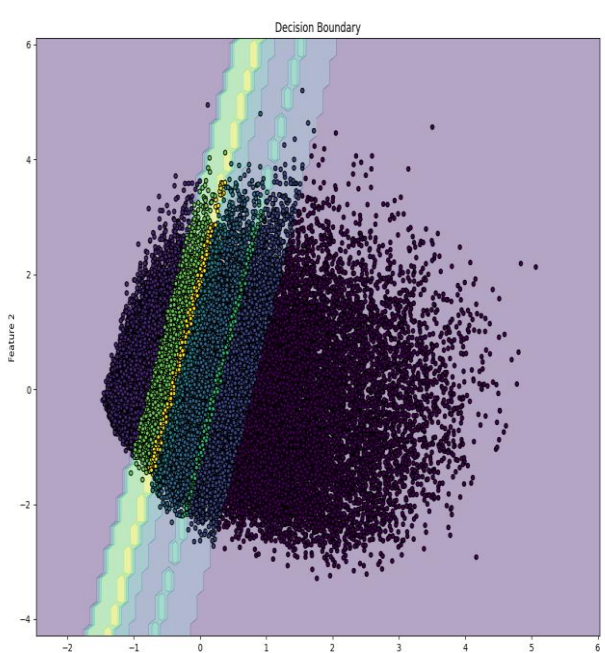
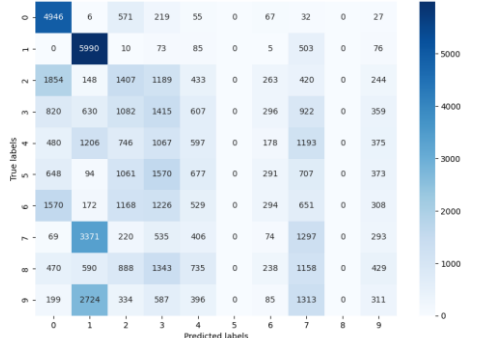
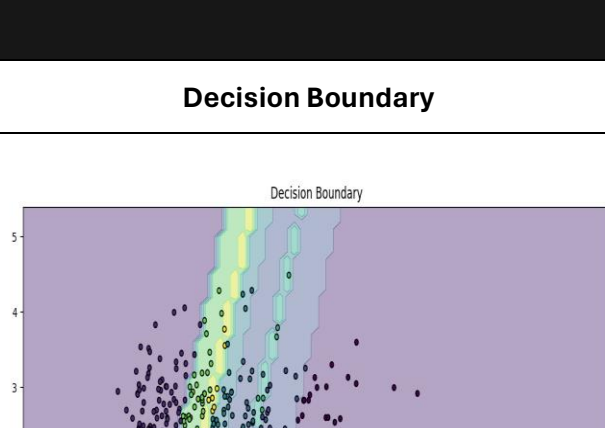
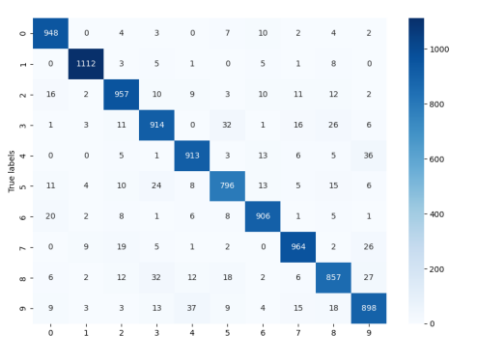
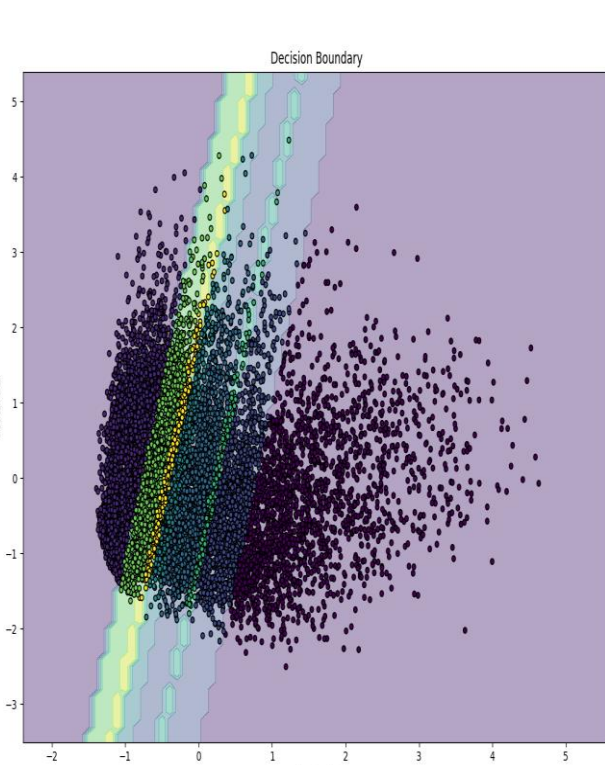
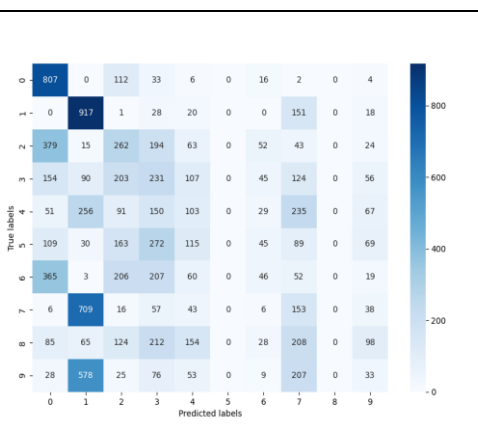
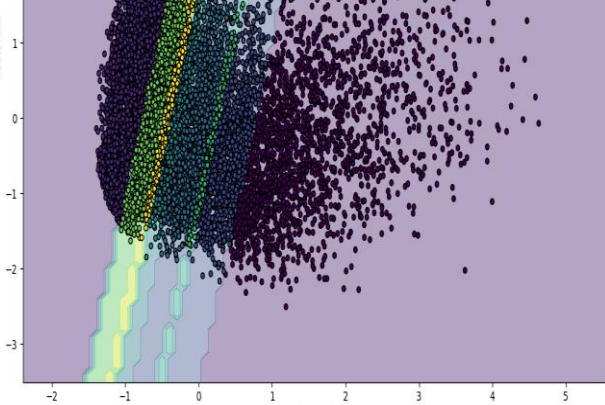
Dla zbioru Iris, posiadającego 3 cechy oraz 3 rodzaje możliwych klasyfikacji (3 różne rodzaje irysów w zbiorze) wybrano ilość neuronów wynoszącą 3.

Dla zbioru Wine, posiadającego 13 cech oraz 3 rodzaje możliwych klasyfikacji (3 rodzaje w zbiorze Wine) wybrano ilość neuronów wynoszącą 7.

Dla zbioru Breast Cancer, posiadającego 30 cech oraz 2 rodzaje możliwych klasyfikacji (2 rodzaje w zbiorze Breast Cancer) wybrano ilość neuronów wynoszącą 15.

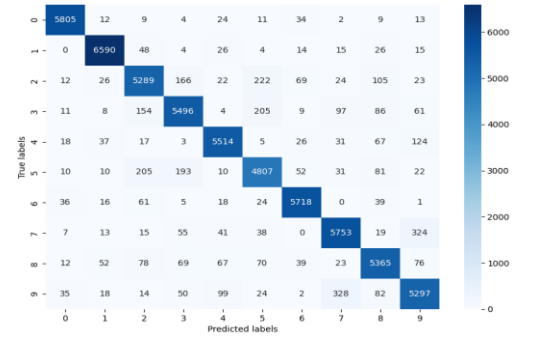
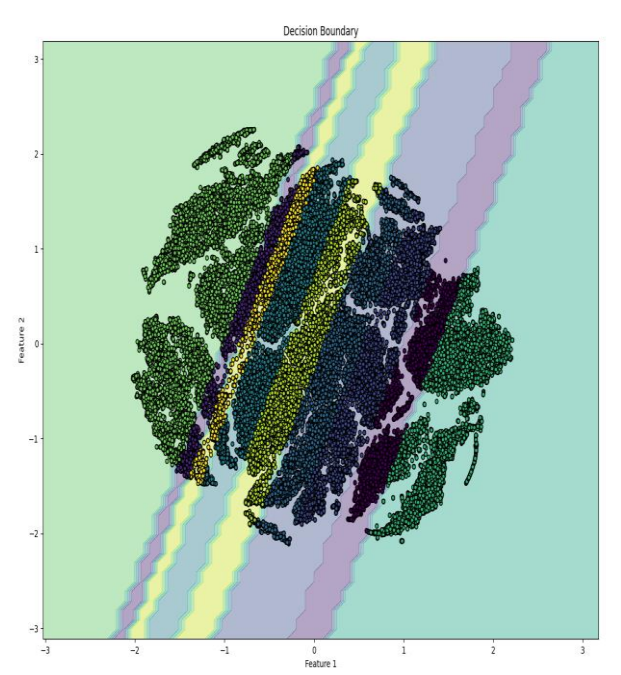
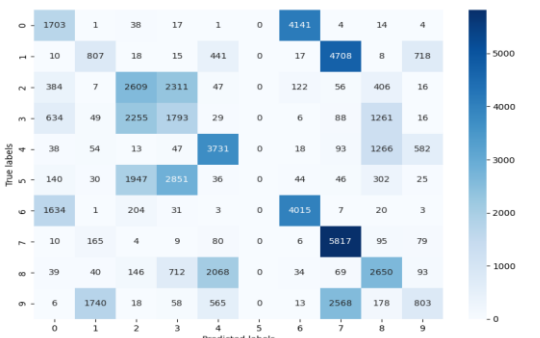
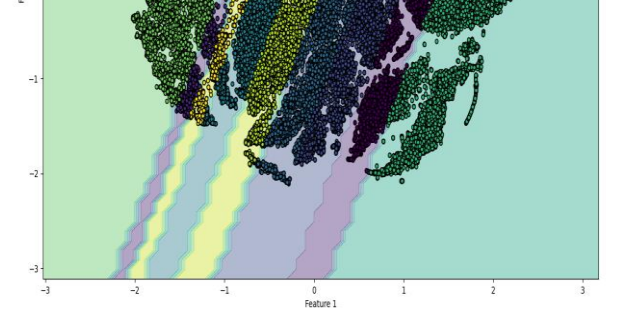
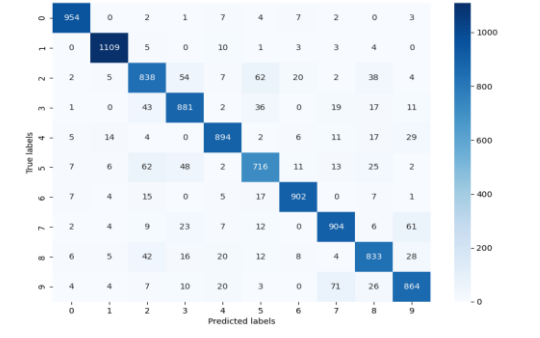
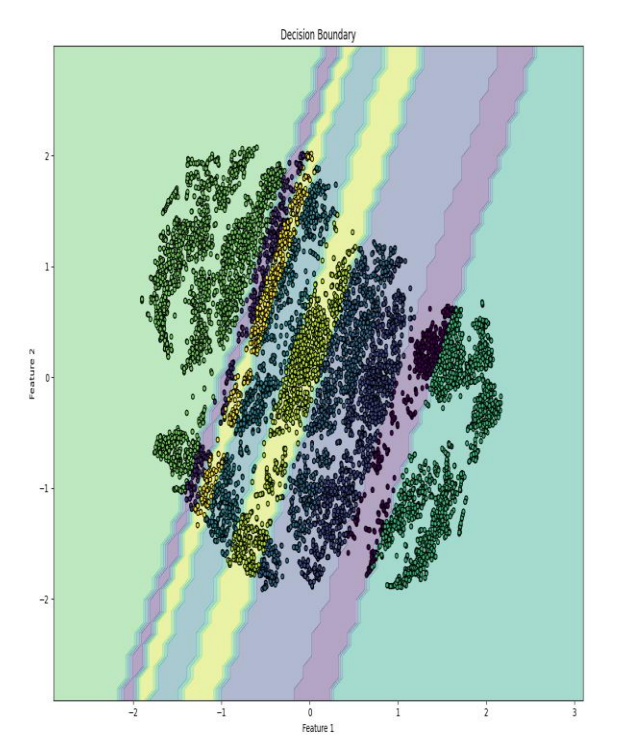
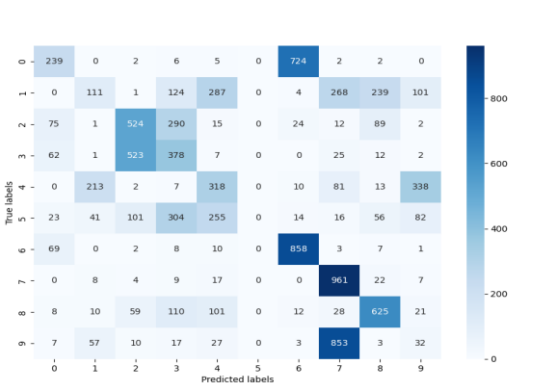
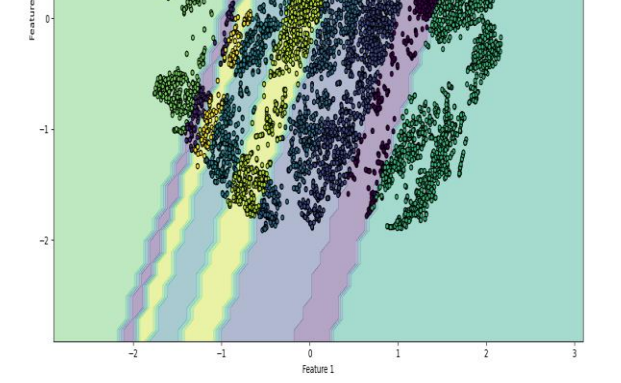
Dla zbioru Mnist, spłaszczonego za pomocą funkcji *flatten()* otrzymano 784 cechy, przy 10 możliwych klasyfikacjach (10 liczb 0-9) wybrano ilość neuronów wynoszącą 784.

1.7 Opis wyników klasyfikacji 'mnist_extr_3' oraz metody PCA

TRAIN	Accuracy Score	Confusion Matrix	Decision Boundary
konwolucja z maskami średnich cyfr	0,93		
PCA	0,27		
TEST	Accuracy Score	Confusion Matrix	Decision Boundary
konwolucja z maskami średnich cyfr	0,93		
PCA	0,26		

Bardzo niska separowalność ekstrakcji PCA sugerowała wyniki klasyfikacji bliskie szansie losowej, jednak wynikowa miara Accuaracy Score okazała się znacząco wyższa, o ponad 15 punktów procentowych. Wynikowe Accuracy Score dla ekstrakcji za pomocą mask okazało się bardzo dobre, jedynie kilka punktów procentowych za surowymi danymi, choć separowalność nie była wysoka. Wyniki te sugerują że nawet prosty model liniowy potrafi radzić sobie z bardzo nachodzącymi na siebie danymi.

1.8 Opis wyników klasyfikacji ekstrakcji 'mnist Extr_4' oraz metody TSNE

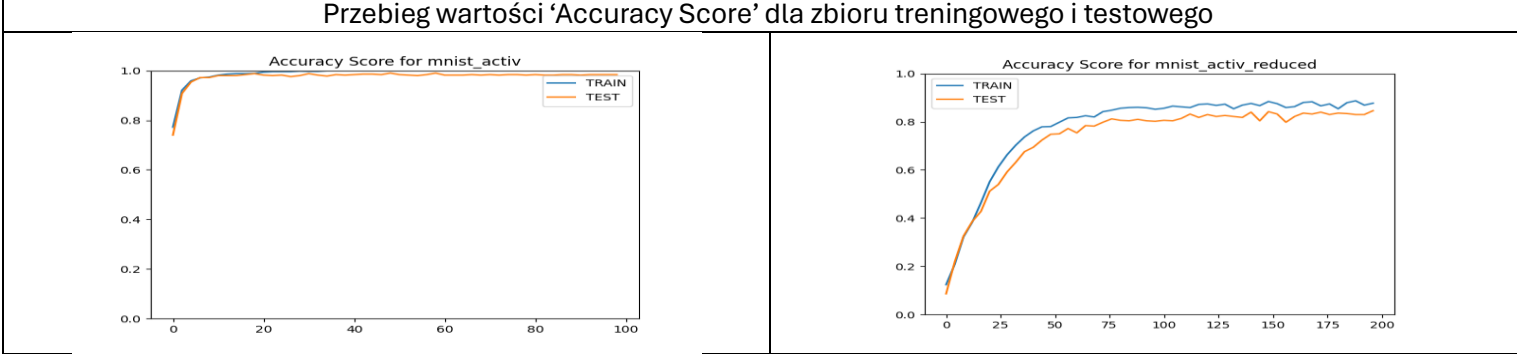
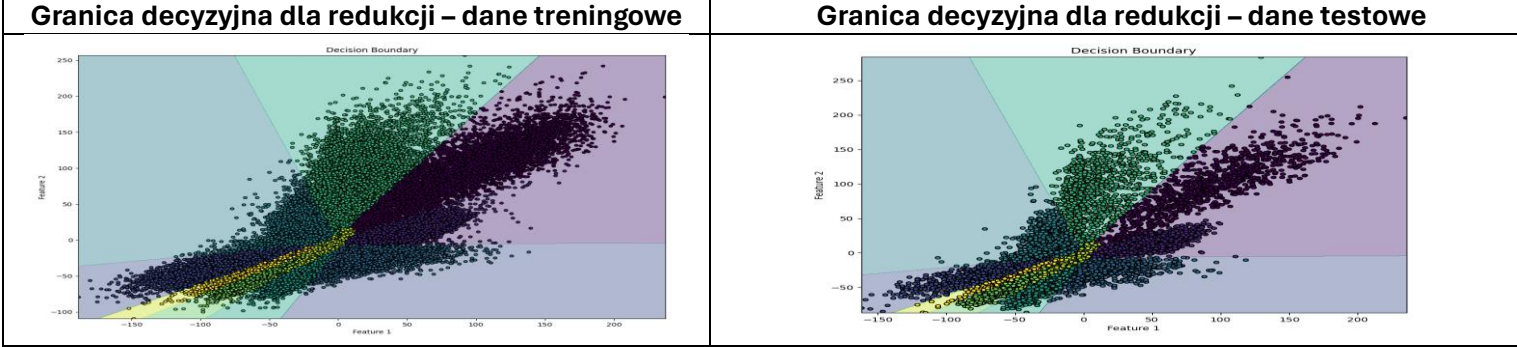
TRAIN	Accuracy Score	Confusion Matrix	Decision Boundary
wykrywanie krawędzi	0,9272		
TSNE	0,40		
TEST	Accuracy Score	Confusion Matrix	Decision Boundary
wykrywanie krawędzi	0,89		
TSNE	0,40		

Zgodnie z oczekiwaniami odnośnie separowalności klas dla metody uwzględniającej algorytm t-SNE, dokładność przypisanych etykiet jest bardzo dobra, biorąc pod uwagę stopień redukcji cech (z 784 na 2), bo wynosząca 0.4. W przypadku metody ekstrakcji cech poprzez wykrywanie krawędzi, dokładność przypisania etykiet jest bardzo wysoka, mimo wcześniej nisko ocenionej separowalności klas, poprzez miarę Silhouette sugerującą możliwość nakładania się na siebie klastrów.

2.1 Opis dwóch architektur sieci splotowej dla zbioru MNIST – badanie wpływu funkcji aktywacji

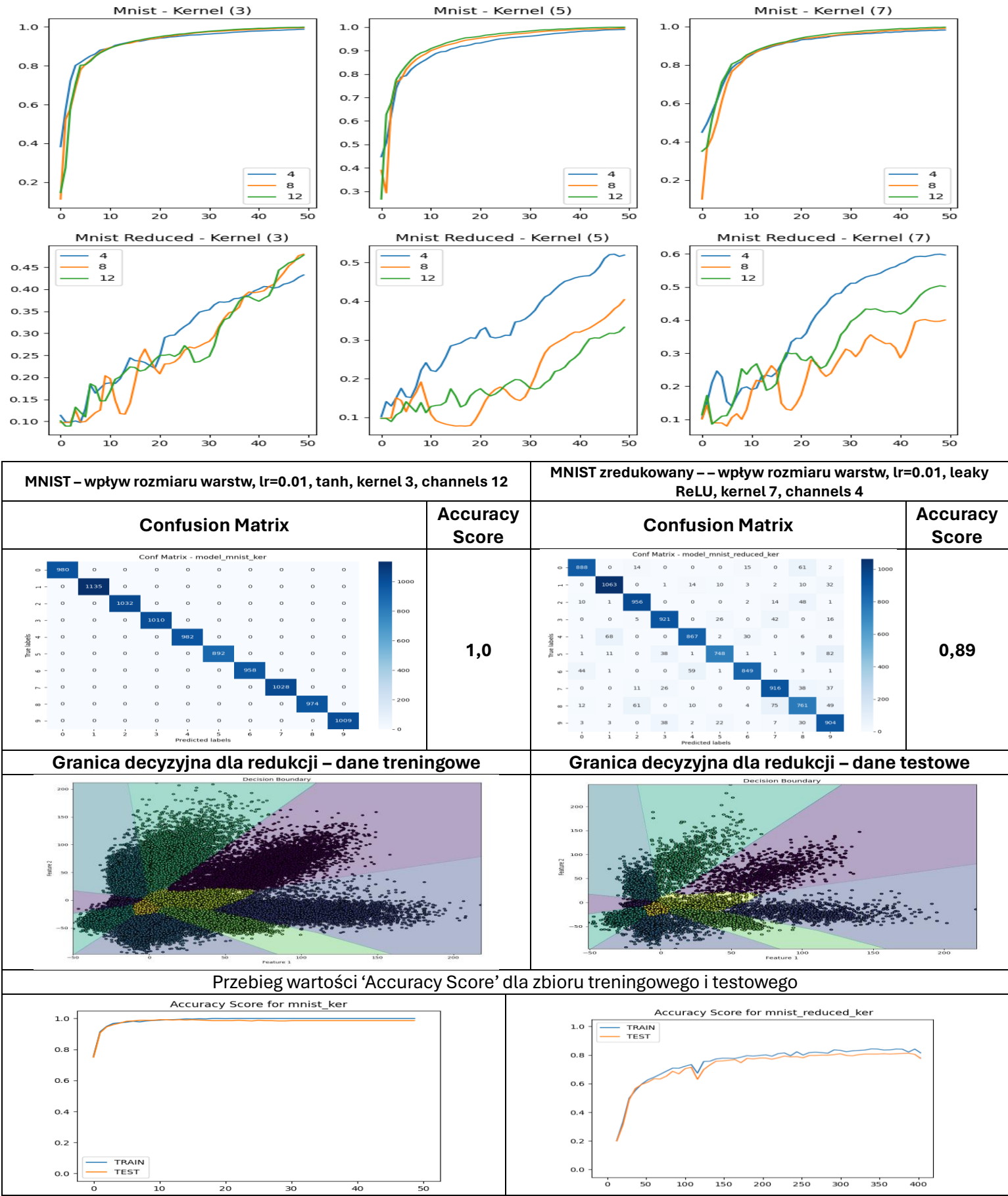


MNIST – funkcja aktywacji tanh, lr=0.01		MNIST zredukowany – funkcja aktywacji leaky ReLU, lr=0.01	
Confusion Matrix	Accuracy Score	Confusion Matrix	Accuracy Score
	1,0		0,80



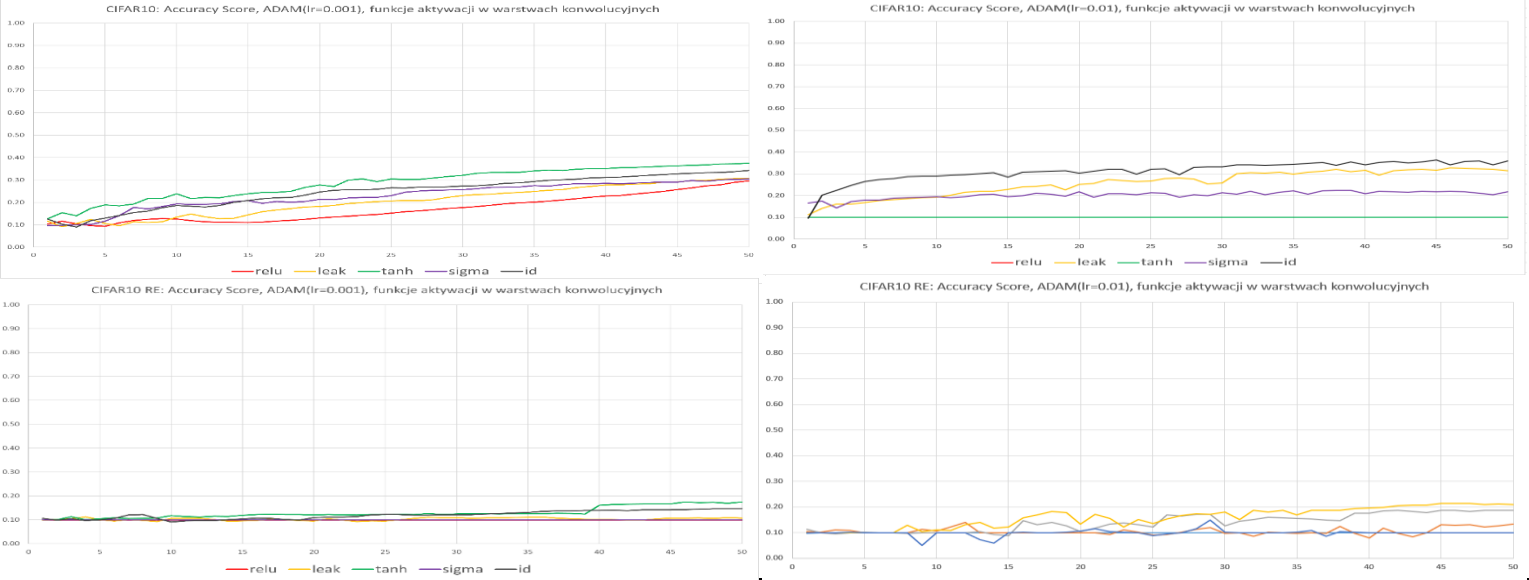
Model bazowy składał się z części ekstrakcyjnej (warstwy konwolucyjnej o jądrze 5 rzutującej na 8 kanałów, wybranej funkcji aktywacji, pooling'u metodą maksymalnej wartości macierzą 2x2, kolejnej warstwy konwolucyjnej rzutującej na 16 kanałów i ponownie funkcji aktywacji i pooling'u) oraz części liniowej (warstwa liniowa z funkcją aktywacji RELU oraz kolejna warstwa liniowa prowadząca do wektora etykiet). W przypadku ekstrakcji do dwóch cech druga warstwa konwolucyjna miała jądro o wymiarach otrzymanych obrazów i rzutowała na dwa kanały, by otrzymać w ten sposób dwie cechy na które wpływa cały obraz. Dla różnych wartości Learning Rate optymalizatora ADAM przetestowano 5 różnych funkcji aktywacji dla warstw konwolucyjnych i wybrano model z najszybszym przyrostem wskaźnika Accuracy Score, odpowiednio TANH i Leaky RELU z nachyleniem wstecznym 0.33 dla zbioru zredukowanego, przy learning rate 0.01. Rozmiary warstw liniowych dobrano na podstawie wyników w części pierwszej projektu.

2.2 Opis dwóch architektur sieci splotowej dla zbioru MNIST – badanie wpływu rozmiaru warstw

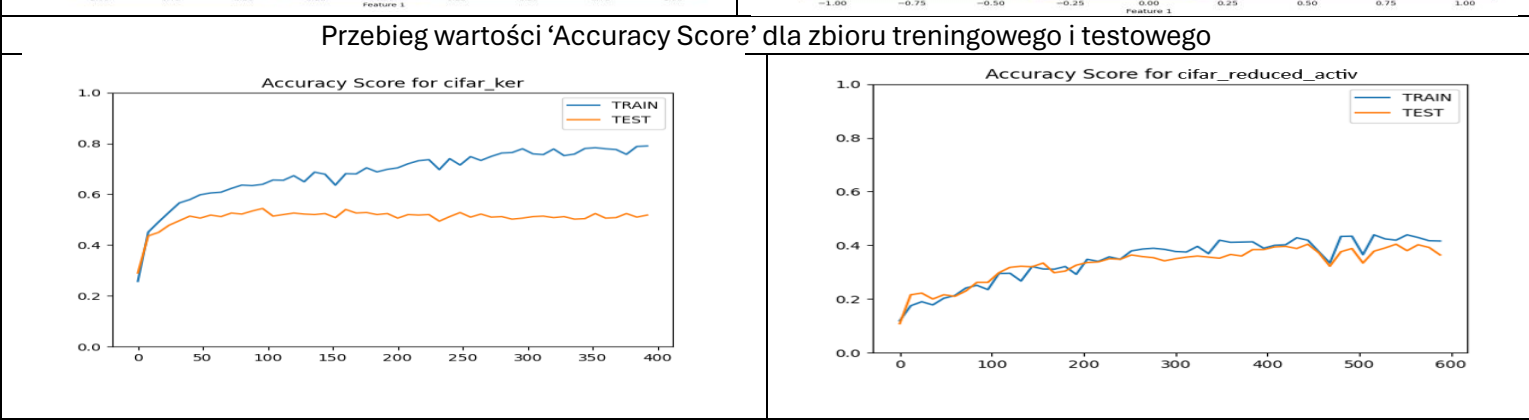
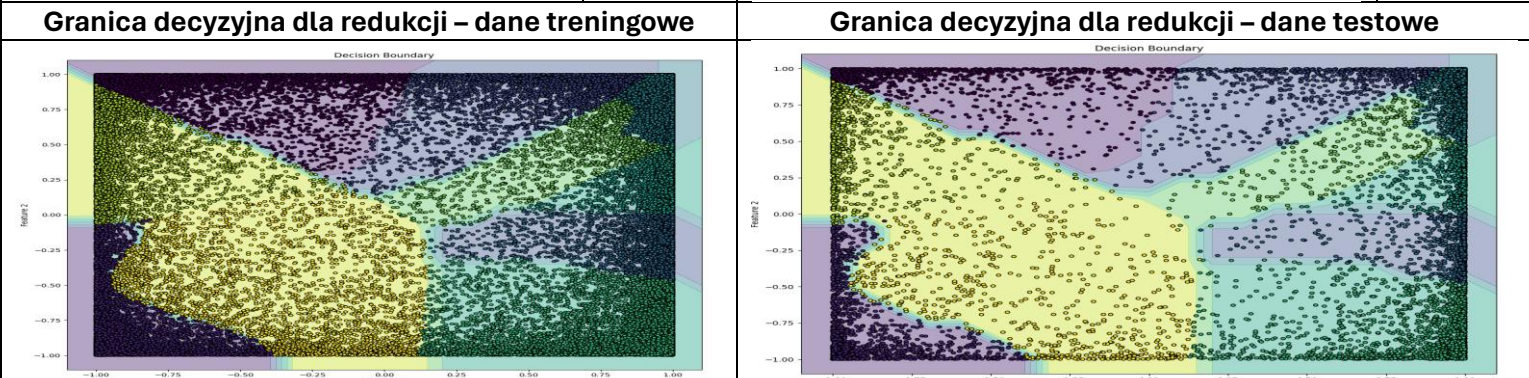


Korzystając z wyników otrzymanych w poprzednim eksperymencie, ponownie przetestowano modele zmieniając rozmiary jądra pierwszej warstwy konwolucyjnej oraz ilość jej wyjściowych kanałów i zbadano ich wpływ na jakość klasyfikacji. Jako najlepszy model wybrano ten o najszybszym przyroście wskaźnika Accuracy Score. Wybrano odpowiednio rozmiar jądra równy 3 i 12 kanałów oraz rozmiar jądra 7 i 4 kanały dla zbioru zredukowanego.

2.3 Opis dwóch architektur sieci splotowej dla zbioru CIFAR10 – badanie wpływu funkcji aktywacji

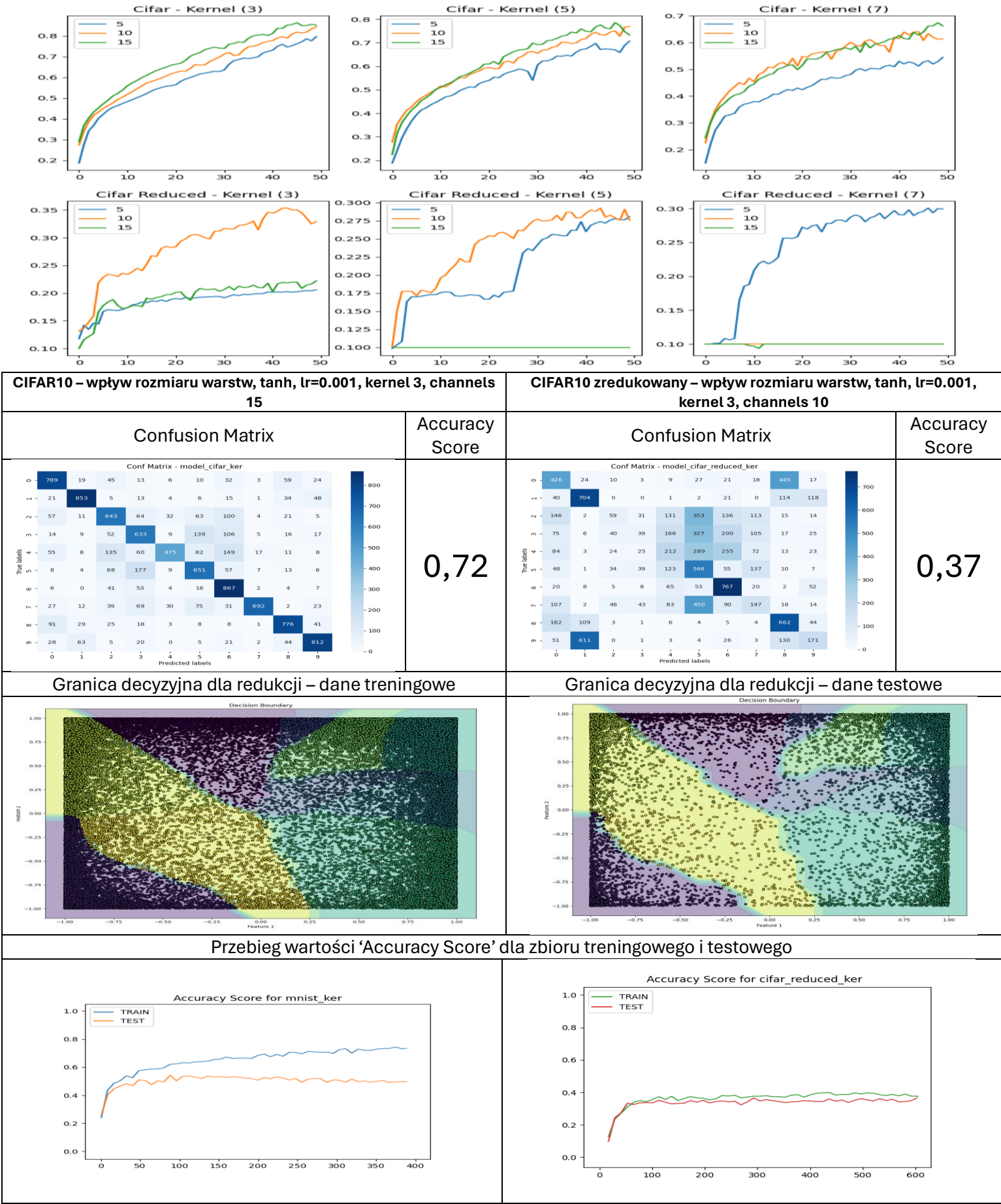


CIFAR10 – funkcja aktywacji tanh, lr=0.001										CIFAR10 zredukowany – funkcja aktywacji tanh, lr=0.001																																																																																																																																																																																																																																																													
Confusion Matrix										Accuracy Score	Confusion Matrix										Accuracy Score																																																																																																																																																																																																																																																		
<div>Conf Matrix - model_cifar_activ</div> <table><tr><td>0</td><td>818</td><td>29</td><td>38</td><td>4</td><td>14</td><td>1</td><td>7</td><td>8</td><td>68</td><td>13</td></tr><tr><td>1</td><td>17</td><td>841</td><td>6</td><td>7</td><td>4</td><td>3</td><td>8</td><td>3</td><td>34</td><td>77</td></tr><tr><td>2</td><td>60</td><td>15</td><td>693</td><td>52</td><td>58</td><td>55</td><td>20</td><td>21</td><td>18</td><td>8</td></tr><tr><td>3</td><td>19</td><td>16</td><td>52</td><td>648</td><td>32</td><td>131</td><td>45</td><td>18</td><td>15</td><td>24</td></tr><tr><td>4</td><td>34</td><td>11</td><td>65</td><td>39</td><td>706</td><td>34</td><td>37</td><td>45</td><td>17</td><td>12</td></tr><tr><td>5</td><td>7</td><td>9</td><td>40</td><td>115</td><td>24</td><td>731</td><td>31</td><td>26</td><td>10</td><td>7</td></tr><tr><td>6</td><td>9</td><td>8</td><td>33</td><td>63</td><td>28</td><td>42</td><td>791</td><td>10</td><td>5</td><td>11</td></tr><tr><td>7</td><td>8</td><td>7</td><td>14</td><td>25</td><td>37</td><td>15</td><td>8</td><td>867</td><td>8</td><td>11</td></tr><tr><td>8</td><td>65</td><td>32</td><td>14</td><td>9</td><td>9</td><td>4</td><td>4</td><td>5</td><td>837</td><td>21</td></tr><tr><td>9</td><td>15</td><td>83</td><td>4</td><td>18</td><td>6</td><td>8</td><td>12</td><td>10</td><td>34</td><td>810</td></tr><tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr></table>										0	818	29	38	4	14	1	7	8	68	13	1	17	841	6	7	4	3	8	3	34	77	2	60	15	693	52	58	55	20	21	18	8	3	19	16	52	648	32	131	45	18	15	24	4	34	11	65	39	706	34	37	45	17	12	5	7	9	40	115	24	731	31	26	10	7	6	9	8	33	63	28	42	791	10	5	11	7	8	7	14	25	37	15	8	867	8	11	8	65	32	14	9	9	4	4	5	837	21	9	15	83	4	18	6	8	12	10	34	810		0	1	2	3	4	5	6	7	8	9	0,77	<div>Conf Matrix - model_cifar_reduced_activ</div> <table><tr><td>0</td><td>577</td><td>18</td><td>71</td><td>9</td><td>44</td><td>10</td><td>21</td><td>13</td><td>225</td><td>12</td></tr><tr><td>1</td><td>20</td><td>615</td><td>3</td><td>7</td><td>2</td><td>1</td><td>44</td><td>1</td><td>50</td><td>257</td></tr><tr><td>2</td><td>65</td><td>4</td><td>188</td><td>41</td><td>454</td><td>85</td><td>74</td><td>59</td><td>25</td><td>5</td></tr><tr><td>3</td><td>10</td><td>6</td><td>72</td><td>195</td><td>172</td><td>130</td><td>361</td><td>37</td><td>7</td><td>10</td></tr><tr><td>4</td><td>38</td><td>4</td><td>123</td><td>60</td><td>525</td><td>109</td><td>79</td><td>47</td><td>12</td><td>3</td></tr><tr><td>5</td><td>9</td><td>1</td><td>88</td><td>152</td><td>328</td><td>164</td><td>213</td><td>37</td><td>5</td><td>3</td></tr><tr><td>6</td><td>3</td><td>7</td><td>19</td><td>113</td><td>74</td><td>36</td><td>719</td><td>11</td><td>3</td><td>15</td></tr><tr><td>7</td><td>14</td><td>2</td><td>175</td><td>66</td><td>462</td><td>101</td><td>81</td><td>82</td><td>9</td><td>8</td></tr><tr><td>8</td><td>411</td><td>64</td><td>34</td><td>7</td><td>11</td><td>5</td><td>17</td><td>3</td><td>406</td><td>42</td></tr><tr><td>9</td><td>33</td><td>537</td><td>11</td><td>8</td><td>3</td><td>1</td><td>66</td><td>5</td><td>31</td><td>305</td></tr><tr><td></td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr></table>										0	577	18	71	9	44	10	21	13	225	12	1	20	615	3	7	2	1	44	1	50	257	2	65	4	188	41	454	85	74	59	25	5	3	10	6	72	195	172	130	361	37	7	10	4	38	4	123	60	525	109	79	47	12	3	5	9	1	88	152	328	164	213	37	5	3	6	3	7	19	113	74	36	719	11	3	15	7	14	2	175	66	462	101	81	82	9	8	8	411	64	34	7	11	5	17	3	406	42	9	33	537	11	8	3	1	66	5	31	305		0	1	2	3	4	5	6	7	8	9	0,38
0	818	29	38	4	14	1	7	8	68	13																																																																																																																																																																																																																																																													
1	17	841	6	7	4	3	8	3	34	77																																																																																																																																																																																																																																																													
2	60	15	693	52	58	55	20	21	18	8																																																																																																																																																																																																																																																													
3	19	16	52	648	32	131	45	18	15	24																																																																																																																																																																																																																																																													
4	34	11	65	39	706	34	37	45	17	12																																																																																																																																																																																																																																																													
5	7	9	40	115	24	731	31	26	10	7																																																																																																																																																																																																																																																													
6	9	8	33	63	28	42	791	10	5	11																																																																																																																																																																																																																																																													
7	8	7	14	25	37	15	8	867	8	11																																																																																																																																																																																																																																																													
8	65	32	14	9	9	4	4	5	837	21																																																																																																																																																																																																																																																													
9	15	83	4	18	6	8	12	10	34	810																																																																																																																																																																																																																																																													
	0	1	2	3	4	5	6	7	8	9																																																																																																																																																																																																																																																													
0	577	18	71	9	44	10	21	13	225	12																																																																																																																																																																																																																																																													
1	20	615	3	7	2	1	44	1	50	257																																																																																																																																																																																																																																																													
2	65	4	188	41	454	85	74	59	25	5																																																																																																																																																																																																																																																													
3	10	6	72	195	172	130	361	37	7	10																																																																																																																																																																																																																																																													
4	38	4	123	60	525	109	79	47	12	3																																																																																																																																																																																																																																																													
5	9	1	88	152	328	164	213	37	5	3																																																																																																																																																																																																																																																													
6	3	7	19	113	74	36	719	11	3	15																																																																																																																																																																																																																																																													
7	14	2	175	66	462	101	81	82	9	8																																																																																																																																																																																																																																																													
8	411	64	34	7	11	5	17	3	406	42																																																																																																																																																																																																																																																													
9	33	537	11	8	3	1	66	5	31	305																																																																																																																																																																																																																																																													
	0	1	2	3	4	5	6	7	8	9																																																																																																																																																																																																																																																													



Model bazowy składał się z części ekstrakcyjnej (warstwy konwolucyjnej o jądrze 7 rzutującej na 10 kanałów, wybranej funkcji aktywacji, poolingu metodą maksymalnej wartości macierzą 2x2, kolejnej warstwy konwolucyjnej rzutującej na 20 kanałów i ponownie funkcji aktywacji i poolingu) oraz części liniowej (warstwa liniowa z funkcją aktywacji RELU oraz kolejna warstwa liniowa prowadząca do wektora etykiet). W przypadku ekstrakcji do dwóch cech druga warstwa konwolucyjna miała jądro o wymiarach otrzymanych obrazów i rzutowała na dwa kanały, by otrzymać w ten sposób dwie cechy na które wpływa cały obraz. Dla różnych wartości Learning Rate optymalizatora ADAM przetestowano 5 różnych funkcji aktywacji dla warstw konwolucyjnych i wybrano model z najszybszym przyrostem wskaźnika Accuracy Score, w obu przypadkach zastosowano aktywację TANH oraz learning rate 0.001 ze względu na lepsze wyniki przy dłuższym uczeniu. Rozmiary warstw liniowych dobrano na podstawie wyników w części pierwszej projektu.

2.4 Opis dwóch architektur sieci splotowej dla zbioru CIFAR – badanie wpływu rozmiaru warstw



2.5 Wyniki eksperymentu drugiego dla najlepszych architektur dla zbioru MNIST

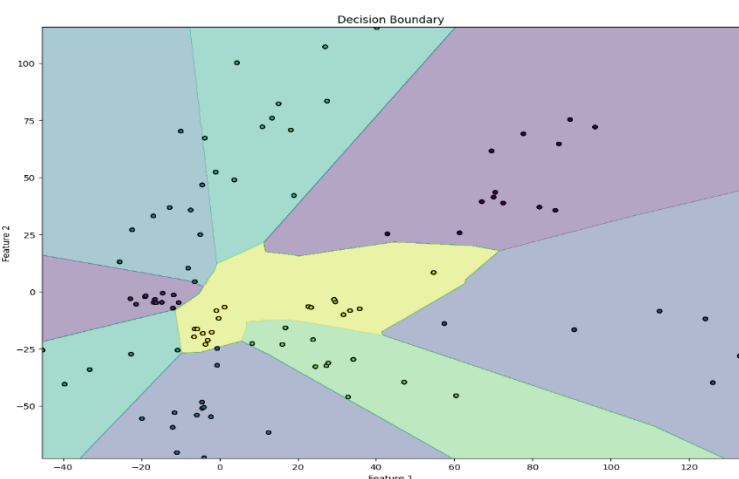
Średnia wartość 'Accuracy Score' z 10 treningów oraz 'Odchylenie standardowe' dla 10 treningów na zbiorze MNIST.

augmentacja	Średnia - MNIST				Odchylenie standardowe - MNIST			
	all	100	200	1000	all	100	200	1000
bez	0.99	0.76	0.86	0.95	0.00072	0.0084	0.0155	0.0024
rotacja	0.99	0.80	0.85	0.95	0.00079	0.0088	0.0220	0.0017
kontrast itd.	0.99	0.78	0.86	0.95	0.00070	0.0323	0.0191	0.0025

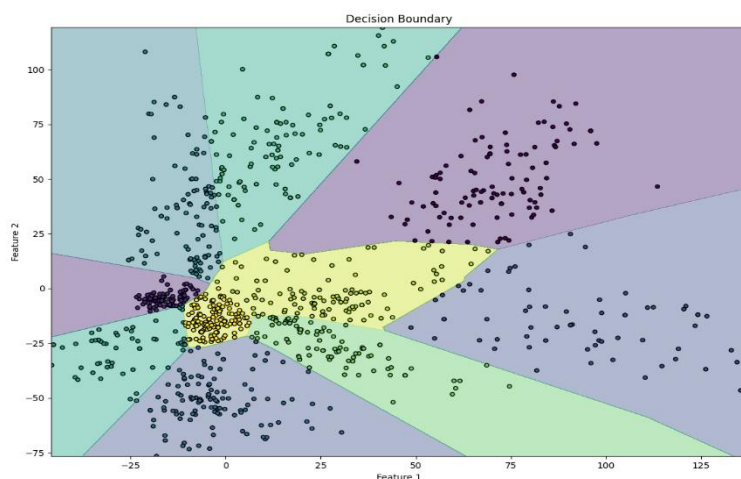
augmentacja	Średnia - MNIST reduced				Odchylenie standardowe - MNIST reduced			
	all	100	200	1000	all	100	200	1000
bez	0.82	0.40	0.39	0.64	0.0063	0.066	0.13	0.021
rotacja	0.82	0.33	0.32	0.66	0.0058	0.113	0.18	0.027
kontrast itd.	0.82	0.35	0.36	0.66	0.0077	0.085	0.12	0.022

Rozkład danych treningowych dla 100 przykładów oraz przypadek, gdy dla każdego ze 100 przykładów zastosowana została augmentacja – zmiana wartości saturacji, kontrastu itd.

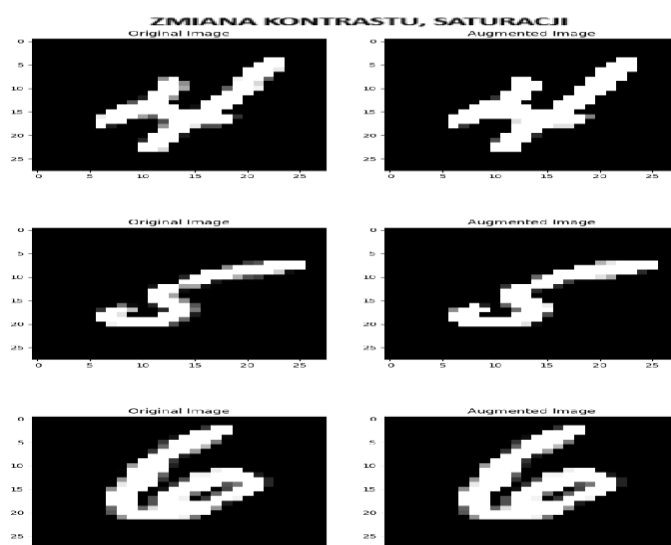
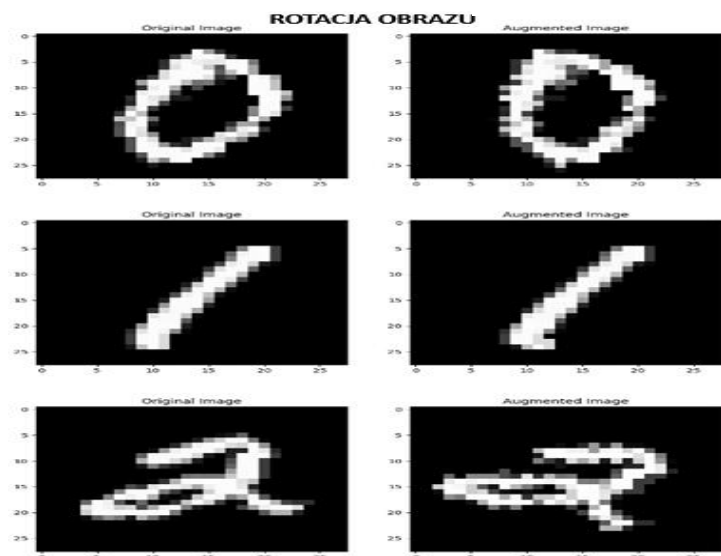
100 przykładów



100 przykładów z zastosowaną 10 razy augmentacją rotacji o kąt w przedziale <0,20>



Przykładowe obrazy przed i po augmentacji



Augmentacje dla zbioru MNIST to: rotacja o kąt w przedziale <0,15> stopni oraz zmiana saturacji, kontrastu i jasności zapisanej liczby – wybrane zostały za najbardziej sensowne, ponieważ liczba może zostać zapisana pod pewnym kątem, nie zawsze idealnie prosto oraz mogą również zostać zapisane grubszym pociągnięciem, co zmiana jasności piksela pozwoli nam wychwycić.

2.6 Wyniki eksperymentu drugiego dla najlepszych architektur dla zbioru CIFAR10

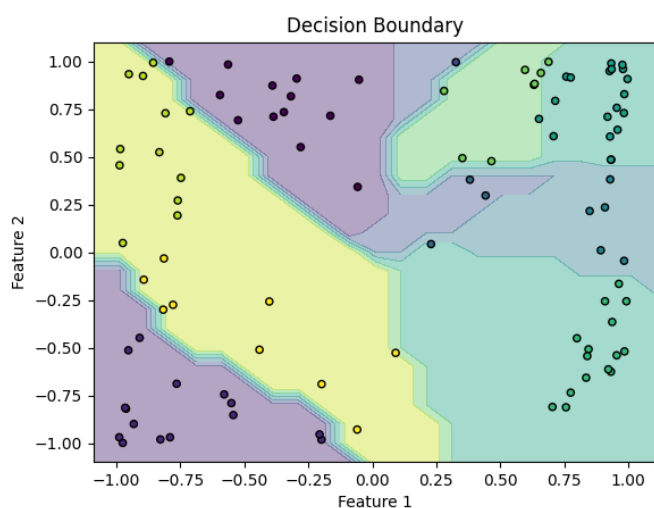
Średnia wartość 'Accuracy Score' z 10 treningów oraz 'Odchylenie standardowe' dla 10 treningów na zbiorze CIFAR10.

augmentacja	Średnia - CIFAR10 reduced				Odchylenie standardowe - CIFAR10 reduced			
	all	100	200	1000	all	100	200	1000
bez	0.31	0.13	0.10	0.25	0.00035	0.030	1.0E-04	0.0026
rotacja	0.31	0.12	0.10	0.24	0.00505	0.022	0.0012	0.0088
kontrast itd.	0.31	0.15	0.10	0.20	0.00175	0.037	0	0.035

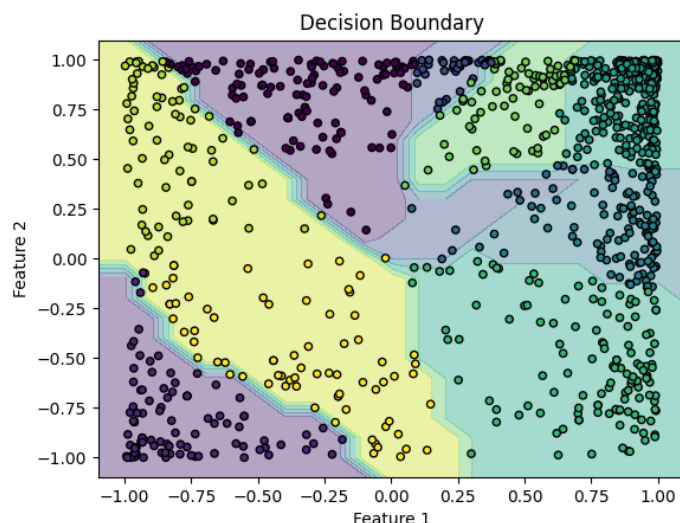
augmentacja	Średnia - CIFAR10				Odchylenie standardowe - CIFAR10			
	all	100	200	1000	all	100	200	1000
bez	0.55	0.21	0.23	0.38	0.008	0.007	0.049	0.0059
rotacja	0.54	0.23	0.21	0.36	0.007	0.008	0.043	0.0086
kontrast itd.	0.55	0.22	0.21	0.37	0.005	0.010	0.058	0.0070

Rozkład danych treningowych dla 100 przykładów oraz przypadek, gdy dla każdego ze 100 przykładów zastosowana została augmentacja – zmiana wartości saturacji, kontrastu itd.

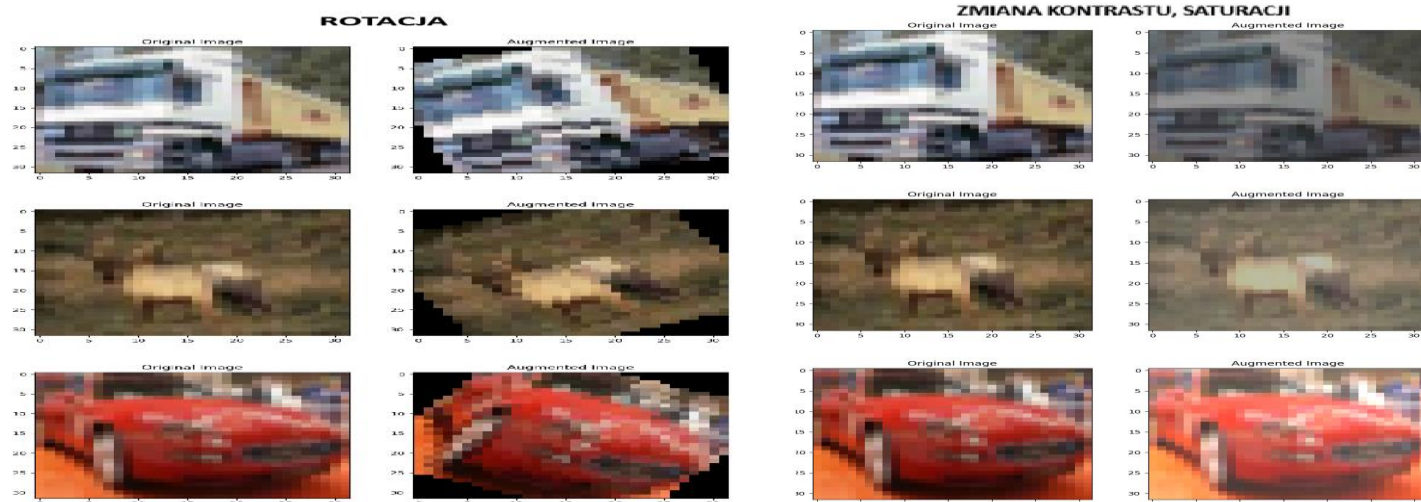
100 przykładów



100 przykładów z zastosowaną 10 razy augmentacją rotacji o kąt w przedziale <0,20>



Przykładowe obrazy przed i po augmentacji



Augmentacje dla zbioru CIFAR10 to również: rotacja o kąt w przedziale <0,10> stopni oraz zmiana saturacji, kontrastu i jasności zapisanego obrazka – wybrane zostały za najbardziej sensowne, ponieważ zrobione zdjęcie obiektu czy zwierzęcia może być czasami niedokładne, może mieć źle odwzorowane kolory czy może być zrobione pod kątem przez niepoprawnie trzymany obiektyw. Dodatkowo, zmiana kolorów zwierząt na losowe wydaje się kontrproduktywna.

2.7 Analiza i wnioski

Wyniki z eksperymentu pierwszego sugerują, że parametr Learning Rate optymalizatora ADAM jest ściśle związany ze skomplikowaniem zbioru. Zbiory MNIST, które są dobrze odseparowane i mają relatywnie niewiele zmiennych pracują lepiej na wyższych wartościach tego parametru, podczas gdy zbiór CIFAR10 miał znaczny problem przy jego wyższych wartościach. Przy zbyt dużym zwiększeniu optymalizator wykazywał niestabilne zachowanie, natomiast przy zbyt małym postęp w uczeniu był niezauważalny. Odpowiedni dobór parametru Learning Rate może znacząco wpłynąć na efektywność naszego modelu.

Funkcje aktywacji warstw konwolucyjnych znacząco wpływają na proces uczenia. Brak funkcji, czy aktywacja tożsamościowa wykazywał znaczną niestabilność, szczególnie przy zbyt dużych wartościach parametru Learning Rate, jednak na początku procesu uczenia wykazywały się szybkim przyrostem dokładności. Podobnie jak aktywacja Leaky RELU, której właściwości matematyczne są bardzo podobne do tożsamościowej, obie szybko osiągały swoje ekstrema z których nie mogły się posunąć do przodu. Funkcje o nieujemnym zbiorze wartości jak RELU czy Sigmoida wykazały się dużo gorszymi osiągnięciami, co może sugerować potrzebę korekt map konwolucyjnych. Aktywacja TANH w większości przypadków okazywała się być najlepszą funkcją aktywacji, oferowała najlepsze wyniki uczenia. Posiada jednocześnie część ujemną, oraz jest ograniczona przez co znacznie trudniej jest jej wpaść w niestabilność w procesie uczenia.

Ekstrakcja do dwóch cech poprzez uczące się warstwy konwolucyjne dla zbioru MNIST pozwoliła na wytrenowanie modeli klasyfikujących z celnością bliską 90%. Jest to wynik bardzo dobry, znacznie przewyższający metody ekstrakcji zastosowane w pierwszej części projektu. Pokazuje to, że nawet czynność ekstrakcji cech może być wykonana przez model który wymaga jedynie parametryzacji i może dorównać i przewyższać metodom projektowanym przez człowieka. Przy zbiorach CIFAR10 zaprojektowanie takiej metody wydaje się bardzo trudnym zadaniem, kiedy model potrafi osiągnąć celność przewyższającą 50% na własnych cechach.

Na zbiorze CIFAR10 klasyfikacja okazała się zadaniem trudnym, nawet model nie forsujący redukcji informacji dla celów wizualizacyjnych klasyfikował w okolicach 90%. Po macierzy pomyłek i odnosząc się do etykiet zaobserwować można że model głównie myli między sobą zwierzęta. Zdjęcia są w bardzo słabej jakości i odróżnienie ich jest zadaniem trudnym dla człowieka, wyniki takie więc wydają się bardzo zadowalające.

Augmentacja danych treningowych pozwala nam na osiągnięcie ich większej ilości, poprzez niewielkie zmiany w obrazach źródłowych. Poprawnie i dobrze dobrana augmentacja pozwala na zwiększenie różnorodności w zbiorze treningowym, co może przyczynić się do poprawy dokładności wytrenowania modelu. Jednakże w przeprowadzonych eksperymentach zaobserwowano, że zaprojektowane augmentacje rotacji oraz zmiany kontrastu czy nasycenia obrazu nie wpływały znacząco na zmianę wartości accuracy. Szczególnie zbiór MNIST ze względu na niskie skomplikowanie osiągał świetne wyniki i augmentacja wydaje się nie mieć możliwości mu pomóc ze względu na osiągnięcie limitu informacji jakie model może interpretować. W przypadku zbioru CIFAR10, gdyby obrazy były w większej rozdzielczości, a nie 32x32 piksele, augmentacja poprzez np. random crop, aby przybliżyć znajdujący się na zdjęciu obiekt, być może pozwoliłaby na poprawę w rozpoznawaniu zwierząt między sobą.