

**A G H**

**AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE**  
**Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej**

**Praca dyplomowa**

*Multimodalne wykrywanie emocji  
A multimodal emotion detection*

Autor: *Zuzanna Maria Zielińska*  
Kierunek studiów: *Automatyka i Robotyka*  
Opiekun pracy: *dr inż. Jaromir Przybyło*

*Kraków, 2024*



# Spis treści

1. Wstęp .....	5
2. Emocje oraz systemy ich rozpoznawania .....	6
2.1. Różne sposoby opisywania emocji.....	6
2.2. Zastosowania i zagrożenia systemu wykrywania emocji.....	7
2.3. Zbieranie informacji o stanie emocjonalnym .....	9
3. Badania literaturowe – klasyfikacja emocji.....	10
3.1. Uczenie maszynowe .....	10
3.1.1. Sieci Rekurencyjne [9] .....	10
3.1.2. Klasyfikator SVM .....	13
3.1.3. Algorytm Kazemiego-Sullivan.....	14
3.2. Algorytm HOG .....	16
4. Opis eksperymentu .....	18
4.1. Plan eksperymentu.....	18
4.2. Dobór bazy bodźców .....	18
4.3. Aplikacja do odtwarzania bodźców oraz nagrywania ścieżki wzrokowej .....	21
4.4. Pomiar pulsu.....	21
4.5. Nagranie mimiki .....	21
4.6. Ankieta .....	21
4.7. Grupa badawcza .....	24
5. Algorytm rozpoznający emocje .....	25
5.1. Środowisko programistyczne .....	25
5.2. Przygotowanie danych do uczenia sieci .....	25
5.2.1. Format surowych danych .....	25
5.2.2. Odrzucenie nieistotnych danych.....	26
5.2.3. Sprawdzenie, czy każda próbka ma wszystkie modalności .....	27
5.2.4. Ekstrakcja punktów charakterystycznych twarzy.....	27
5.2.5. Synchronizacja danych .....	27
5.2.6. Wypełnienie pustych kolumn zerami, zakodowanie słów.....	28
5.2.7. Normalizacja danych .....	28
5.2.8. Zmiana formatu etykiet .....	28
5.2.9. Zmiana sformułowania słownych emocji.....	28
5.3. Prototyp sieci wykrywającej emocje .....	28
5.3.1. Typy sieci, ich architektura oraz strategie uczenia.....	28
5.3.2. Zbalansowanie zbioru sztucznyimi danymi i augmentacja .....	32
6. Wyniki .....	33
6.1. Ankieta .....	33
6.1.1. Wyniki ankiety SAM.....	33
6.1.2. Porównanie wyników ankiety SAM zebranych w badaniu ze zbiorem NAPS	
	36

6.1.3. Analiza słownych emocji odczuwanych przez badanych oraz porównanie ich ze zbiorem NAPS.....	40
6.2. System rozpoznawania emocji.....	44
6.2.1. Modele regresyjne.....	44
6.2.2. Modele klasyfikujące .....	48
6.2.3. Zbiory danych .....	52
6.2.4. Wyniki modeli wytrenowanych na danych z augmentacją.....	56
7.Podsumowanie .....	58
8.Dodatki.....	60
8.1. Wyniki ankiet.....	60
8.2. Nazwy kolumn tabeli wejściowej do sieci .....	69
8.3. Przykładowe wyniki, które dawały sieci neuronowe.....	74
9.Bibliografia .....	77

# **1. Wstęp**

Emocje pełnią ważną rolę w codziennym życiu człowieka oraz w komunikacji z innymi ludźmi. Brak umiejętności identyfikacji emocji przez komputery stanowi pewną barierę w dobrym zrozumieniu ludzkich potrzeb.

Celem pracy jest przeprowadzenie badania zbierającego dane biometryczne uczestników podczas odczuwania emocji oraz analiza wypełnionych przez nich ankiet dotyczących przeżywanych uczuć i na tej podstawie zaprojektowanie i wykonanie prototypu systemu wykrywania emocji. Dodatkowo wykonano analizę danych ankietowych, porównano ją ze zbiorem danych NAPS oraz przetestowano i dokonano ewaluacji prototypu systemu wykrywania emocji.

W rozdziale 2. przedstawione są sposoby opisywania emocji, zastosowania i zagrożenia wynikające z wykorzystywania systemów rozpoznawania emocji oraz możliwe problemy z trafnością danych zebranych podczas eksperimentu.

Rozdział 3. prezentuje działanie sieci rekurencyjnych, algorytmu Kazemiego-Sullivan, służącego do ekstrakcji punktów charakterystycznych twarzy oraz algorytmu HOG i klasyfikatora SVM, służących do detekcji twarzy.

W rozdziale 4. opisano plan eksperimentu, dobór bazy bodźców, przyrządy pomiarowe, program do wypełniania ankiety oraz grupę badawczą.

Rozdział 5. przedstawia środowisko programistyczne, format i sposób obróbki danych oraz ich augmentację. Zawiera także architekturę sieci neuronowej do wykrywania emocji oraz sposób jej uczenia.

Rozdział 6. zawiera analizę wyników ankiet, porównanie ich z wynikami z bazy bodźców, oraz ewaluację wyników sieci.

W rozdziale 7. podsumowano wyniki badań oraz wysunięto propozycje rozwoju pracy.

W rozdziale 8. znalazły się pełne, surowe wyniki ankiet (przed ujednoliceniem stylu zdań, poprawą ortografii itp.) oraz tabela z nazwami kolumn danych wejściowych do sieci neuronowej.

Rozdział 9. zawiera bibliografię.

## 2. Emocje oraz systemy ich rozpoznawania

### 2.1. Różne sposoby opisywania emocji

Ludzie od wieków byli zainteresowani pochodzeniem i funkcją emocji. Jeszcze w XIX wieku Charles Darwin twierdził, że emocje mają charakter ewolucyjny i są uniwersalne dla wszystkich ludzi. Współcześnie tę teorię rozwijał Paul Ekman, który badał, jak ludzie wyrażają emocje w różnych kulturach. Wykazał on, że reakcje autonomicznego układu nerwowego u ludzi z różnych kultur mają związek z odczuwanymi emocjami i są uniwersalne [1].

Emocje są złożonym zespołem zmian cielesnych i psychicznych obejmującym pobudzenie fizjologiczne, uczucia, procesy poznawcze, widoczne sposoby ekspresji oraz specyficzne reakcje behawioralne. Pojawiają się one jako reakcja na sytuację postrzeganą jako ważną dla jednostki [2].

Jest wiele różnych sposobów podziału emocji. Do najbardziej znanych należą ankieta SAM oraz teoria podstawowych emocji.

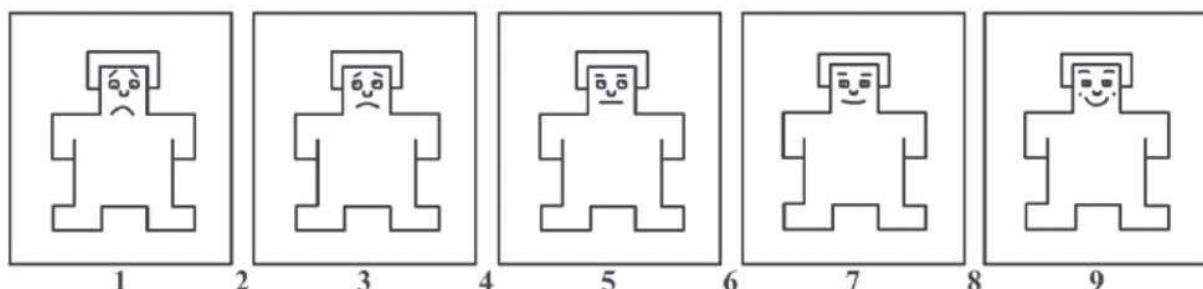
#### 2.1.1. Ankieta SAM [3]

Ankieta SAM (ang. Self-assesment manikin) opiera się na założeniu, że emocje można opisać za pomocą trzech dziewięciostopniowych skali:

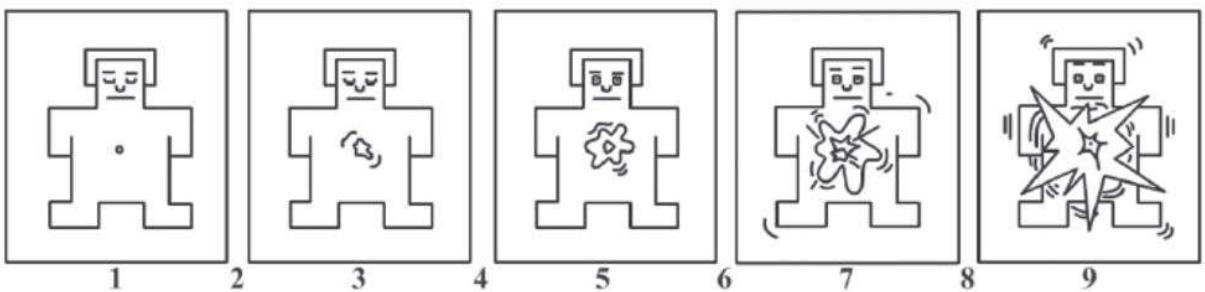
- skali wartościowości – czy badany odczuwa negatywne czy pozytywne uczucia,
- skali pobudzenia – czy badany jest spokojny czy pobudzony,
- skali dominacji – czy badany w pełni kontroluje daną emocję, czy jest przez nią kontrolowany.

Do każdej skali dołączone jest 5 obrazków. Zadaniem badanego jest wybranie tego obrazka, który najbardziej oddaje jego stan emocjonalny lub zaznaczenie wartości między obrazkami (Rys. 2.1, Rys. 2.2, Rys. 2.3).

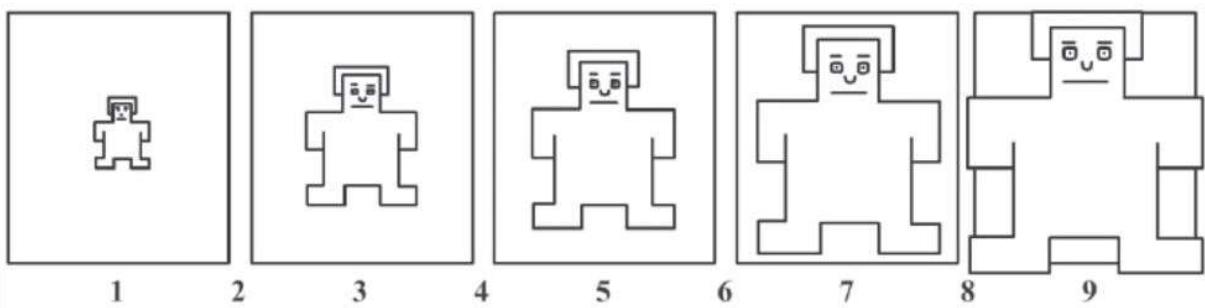
Zaletą ankiety SAM jest to, że można jej używać, nawet gdy badany ma trudności z werbalizacją swoich uczuć lub w ogóle nie zna języka używanego w badaniu.



Rys. 2.1. Skala wartościowości. Negatywne emocje przedstawią cyfra 1, zaś pozytywne – 9 [4].



Rys. 2.2. Skala pobudzenia. Spokój przedstawia cyfra 1, zaś 9 to całkowite pobudzenie [4].



Rys. 2.3. Skala dominacji. Badanego przytłoczonego emocją przedstawia cyfra 1, zaś badanego w pełni kontrolującego swoje emocje – cyfra 9 [4].

### 2.1.2. Teoria podstawowych emocji

Teoria podstawowych emocji odrzuca idee, że odczucia można opisać za pomocą skali. Zakłada ona, że istnieją grupy emocji, nazywanych emocjami podstawowymi, które fundamentalnie różnią się od siebie. Wewnątrz grupy wszystkie są połączone wspólnymi cechami np. ekspresją. Nie istnieje emocja, która nie należy do żadnej z rodzin podstawowych emocji [5]. Do podstawowych emocji należą szczęście, zdziwienie, złość, odraza, strach, smutek.

## 2.2. Zastosowania i zagrożenia systemu wykrywania emocji

System rozpoznawania emocji oznacza system służący do identyfikacji emocji u ludzi na podstawie ich danych biometrycznych [6]. Do tego rodzaju danych należą między innymi: mimika, ton głosu, puls, ścieżka wzrokowa, wielkość żrenicy i impedancja skóry.

Takie systemy mają bardzo wiele zastosowań. Wśród nich można wymienić [7]:

- świadczenie spersonalizowanych usług:
  - analizowanie emocji w celu wyświetlania spersonalizowanych komunikatów w inteligentnych środowiskach,
  - zapewnianie spersonalizowanych rekomendacji, np. dotyczących wyboru muzyki lub materiałów kulturowych,
  - analizowanie wyrazu twarzy w celu przewidywania indywidualnych reakcji na filmy;
- analiza zachowań klientów i reklama:

- analizowanie emocji klientów podczas zakupów, badając zależność między nimi a towarami lub ich rozmieszczeniem w sklepie,
  - reklamy na stacjach kolejowych wykorzystujące system rozpoznawania i śledzenia twarzy do celów marketingowych;
- opieka zdrowotna:
  - wykrywanie autyzmu lub chorób neurodegeneracyjnych,
  - przewidywanie zaburzeń psychotycznych lub depresji w celu identyfikacji użytkowników potrzebujących wsparcia,
  - zapobieganie samobójstwom,
  - wykrywanie depresji u osób starszych,
  - obserwowanie stanu pacjentów podczas leczenia;
- zatrudnienie:
  - wspomaganie podejmowania decyzji przez rekruterów,
  - identyfikowanie niezainteresowanych kandydatów podczas rozmowy kwalifikacyjnej,
  - monitorowanie nastrojów i uwagi pracowników;
- edukacja
  - monitorowanie uwagi studentów,
  - wykrywanie reakcji emocjonalnych użytkowników na program edukacyjny i dostosowywanie ścieżki nauki,
  - projektowanie systemu nauczania opartego na emocjach,
  - wykrywanie zaangażowania w nauczanie online;
- bezpieczeństwo publiczne:
  - wykrywacze kłamstw i inteligentna kontrola graniczna,
  - monitorowanie przestrzeni publicznych w celu identyfikacji emocji wywołujących potencjalne zagrożenie terrorystyczne,
  - analizowanie nagrań z miejsc zbrodni w celu wskazania potencjalnych motywów przestępstwa;
- wykrywanie przestępstw:
  - wykrywanie i zmniejszanie liczby fałszywych roszczeń ubezpieczeniowych,
  - wdrażanie strategii zapobiegania oszustwom,
  - wyłapywanie złodziei sklepowych;
- inne:
  - wykrywanie zmęczenia u kierowców,
  - wykrywanie postaw politycznych.

Istnieje jednak szereg zagrożeń związanych z wykorzystywaniem systemów rozpoznawania emocji. Parlament Europejski w ramach rozporządzenia o sztucznej inteligencji (ang. AI Act) bezwzględnie zakazał stosowania takich systemów w środowisku pracy oraz edukacji. W pozostałych przypadkach są one określane mianem systemów wysokiego ryzyka i podlegają specjalnym restrykcjom przez zagrożenia, jakie mogą wywoływać [6]. Systemy te mogą nie uwzględniać różnic kulturowych, mieć zbyt niską dokładność oraz nie być w stanie wyjaśnić

przyczyn wykrytych emocji. Istnieje również ryzyko, że wynik algorytmu zostanie potraktowany jako niepodważalny fakt, co może prowadzić do błędnych wniosków, zwłaszcza w przypadku osób z zaburzeniami zdrowotnymi, takimi jak paraliż twarzy [7].

### **2.3. Zbieranie informacji o stanie emocjonalnym**

Jednym z kluczowych problemów w badaniach nad emocjami jest tzw. reaktywność, czyli zmiana zachowania osoby badanej na skutek świadomości bycia obserwowaną [8]. Może to prowadzić do zakłócenia wyników badania.

Istnieje wiele strategii minimalizowania reaktywności. Zaliczają się do nich potajemna obserwacja, używanie niezauważalnych urządzeń pomiarowych lub ukrycie prawdziwego celu badania [8]. Niestety w tym przypadku nie są one możliwe do zrealizowania. Zamiast tego zostanie użyta triangulacja danych, czyli korzystanie z wielu źródeł informacji.

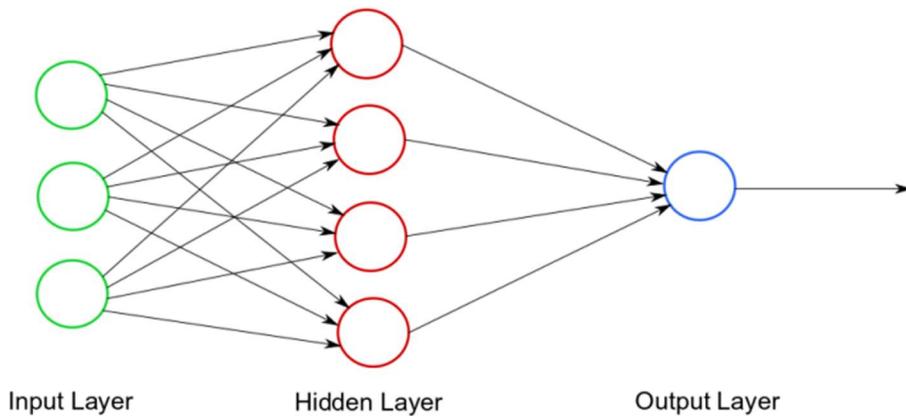
W badaniu zostaną nagrane 3 modalności – ścieżka wzrokowa, mimika i puls. Dodatkowo badani opiszą swoje uczucia w ankiecie.

### 3. Badania literaturowe – klasyfikacja emocji

#### 3.1. Uczenie maszynowe

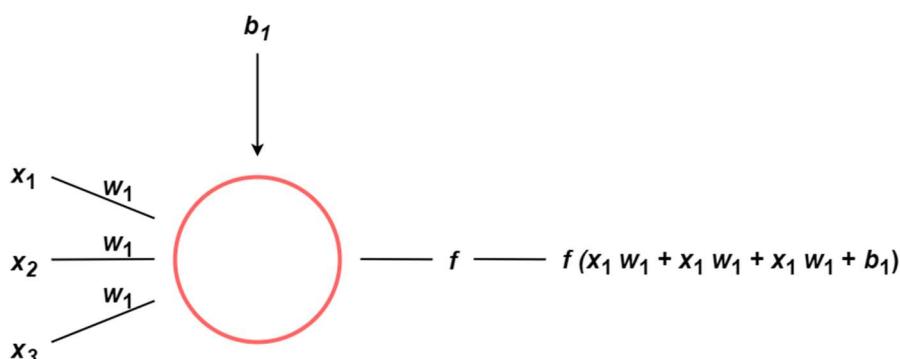
##### 3.1.1. Sieci Rekurencyjne [9]

Sieć neuronowa jest to struktura matematyczna naśladująca działanie ludzkiego mózgu. Składa się ona z elementów nazywanych sztucznymi neuronami, które są grupowane w warstwy. Są trzy rodzaje warstw: wejściowa, ukryta i wyjściowa (Rys. 3.1.).



Rys. 3.1. Schemat sieci neuronowej [10].

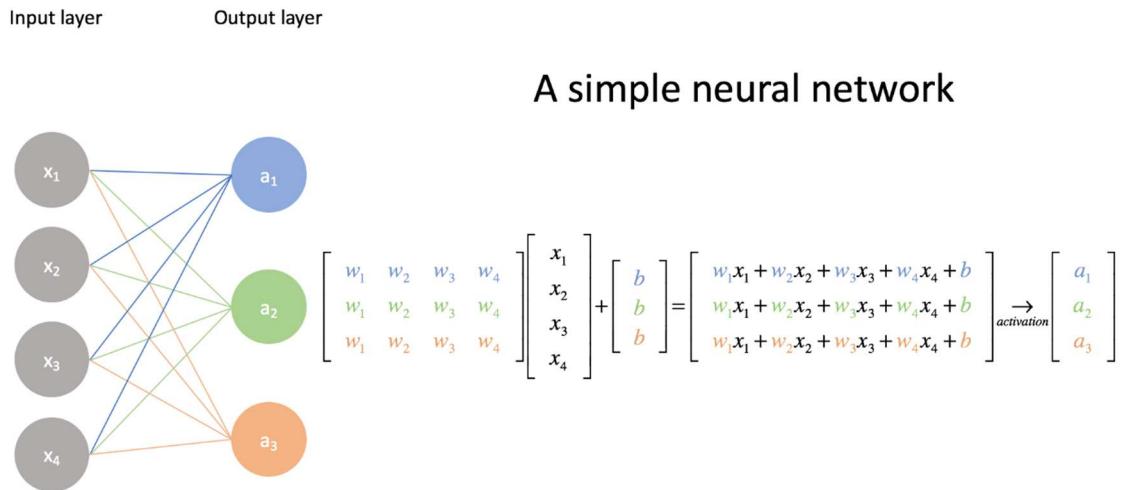
Najprostszą siecią neuronową jest perceptron (Rys. 3.2). Składa się z jednego neuronu. Sumuje on ilorazy wartości wejściowych i ich wag oraz pewną liczbę zwaną obciążeniem (ang. „bias”). Otrzymana suma jest przekazywana do funkcji aktywacji a jej wynik jest wyjściem perceptronu. Do funkcji aktywacji należą m. in. sigmoid, tanh, relu, maxout, elu, softmax. Perceptrony można łączyć w wielowarstwowe sieci (Rys. 3.3.).



Rys. 3.2. Perceptron.

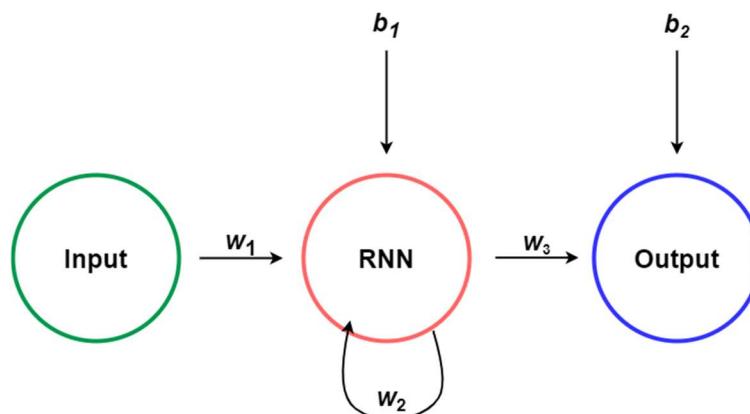
Na początku wartości wag i obciążień są przydzielane losowo. Później są zmieniane przez algorytm wstecznej propagacji błędu (ang. „backpropagation”). Kiedy pewna wartość pojawia się na wejściu sieci neuronowej, wartości wszystkich neuronów są zapisywane i wartość wyjściowa jest porównywana z wartością poprawną. Następnie według pewnej funkcji kosztu, liczony jest błąd warstwy wyjściowej sieci oraz jej gradient według wszystkich

parametrów. Powtarzane jest to dla wszystkich warstw, kończąc na wejściowej. Na koniec wagi są zmieniane zgodnie z wartościami gradientu, minimalizując funkcję kosztu. Ten proces nazywany jest uczeniem się sieci.



Rys. 3.3. Sieć zbudowana z wielu perceptronów [11].

Rekurencyjna sieć neuronowa jest to sieć, której dane wyjściowe jednej warstwy są przekazywane z powrotem na wejście tej samej warstwy razem z nową wartością wejściową. W ten sposób pozwala ona przetwarzanie sekwencji danych (Rys. 3.5.).



Rys. 3.4. Prosta rekurencyjna warstwa sieci neuronowej - zwinięta.

Rozważono prostą rekurencyjną warstwę sieci składającą się z jednego neuronu. Wartości wejściowe są mnożone przez wagę  $w_1$ . Pętla sprzężenia zwrotnego neuronu rekurencyjnego ma wagę  $w_2$  oraz obciążenie  $b_1$ . Wartość wyjściowa jest mnożona przez wagę  $w_3$  oraz dodawane jest obciążenie  $b_2$  (Rys. 3.4). Przyjęto też, że każda wartość przekazywana wstępco do wejścia neuronu nazywana jest stanem ukrytym  $h$ . Przy rozwinięciu sieci oraz po analizie obliczeń dla kilku wartości wejściowych otrzymujemy:

$$h_0 = 0 \quad (3.1)$$

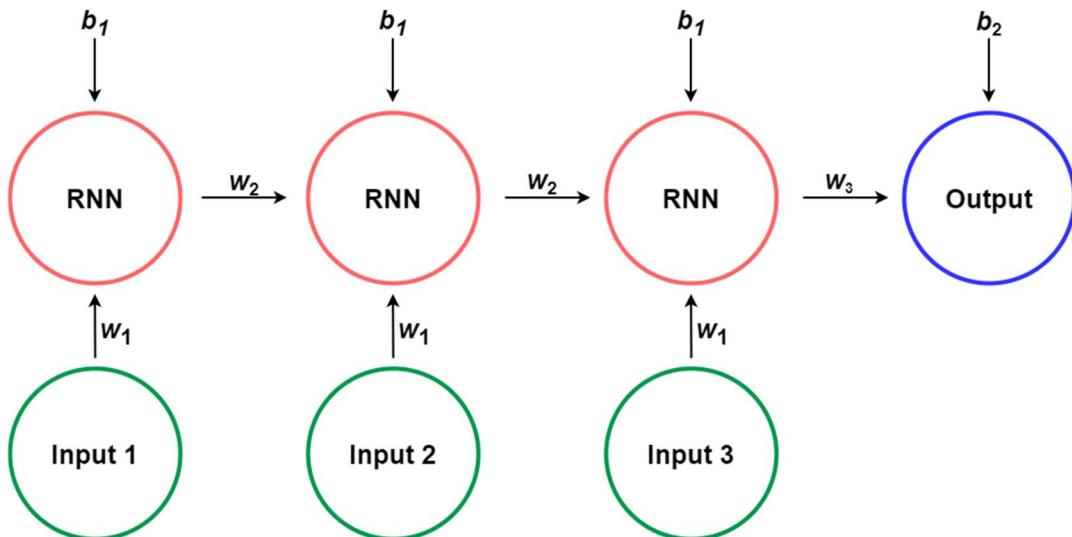
$$h_1 = x_1 w_1 + b_1 \quad (3.2)$$

$$h_2 = x_2 w_1 + b_1 + h_1 w_2 = (x_1 w_1 + b_1) w_2 + x_2 w_1 + b_1 \quad (3.3)$$

$$h_3 = x_3 w_1 + b_1 + h_2 w_2 = (x_1 w_1 + b_1) w_2^2 + (x_2 w_1 + b_1) w_2 + x_3 w_1 + b_1 \quad (3.4)$$

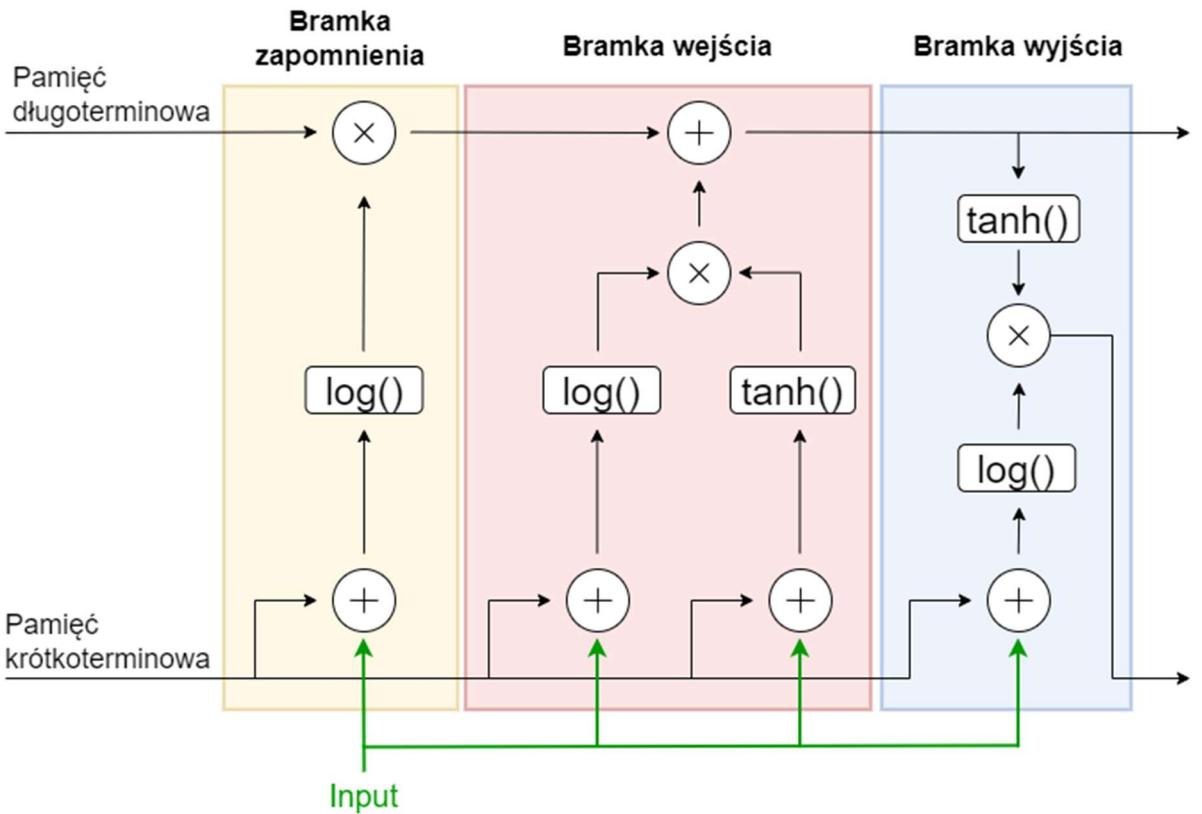
$$\begin{aligned} h_n &= x_n w_1 + b_1 + h_{n-1} w_2 \\ &= (x_1 w_1 + b_1) w_2^{n-1} + (x_2 w_1 + b_1) w_2^{n-2} + \dots + x_n w_1 + b_1 \end{aligned} \quad (3.5)$$

Powyższe obliczenia pokazują (3.1 – 3.5), że waga  $w_2$  ma ogromne znaczenie dla wyników sieci, ponieważ jest podnoszona do potęgi zależnej od długości serii danych. Jeśli waga jest mała i sekwencja wejściowa dłuża, to wartość sieci będzie bardzo mała. Podczas obliczania gradientu błędu, algorytm będzie wykonywał bardzo małe kroki i może nie zbliżyć się wystarczająco do minimum funkcji kosztu. Jest to problem znikającego gradientu. Kiedy waga jest duża i sekwencja wejściowa dłuża, to wartość sieci będzie bardzo duża. Licząc gradient błędu, algorytm będzie wykonywał bardzo duże kroki i może cały czas omijać minimum funkcji. Jest to problem eksplodującego gradientu.



Rys. 3.5. Prosta rekurencyjna warstwa sieci neuronowej - rozwinięta.

Warstwa LSTM (ang. „Long Short-Term Memory”) jest warstwą rekurencyjną sieci neuronowej, która unika problemu eksplodującego i znikającego gradientu. Posiada pamięć krótkoterminową oraz długoterminową, która jest modyfikowana przez poprzednią. LSTM składa się z bramki zapomnienia, bramki wejścia i bramki wyjścia (Rys. 3.6.). Bramka zapomnienia ustala, jaka część pamięci długoterminowej powinna zostać przekazana dalej. Bramka wejścia dodaje nową wartość wejściową do pamięci krótkoterminowej, modyfikuje ją i dodaje do długoterminowej. Bramka wyjścia modyfikuje pamięć krótkoterminową zgodnie z wartością pamięci długoterminowej.



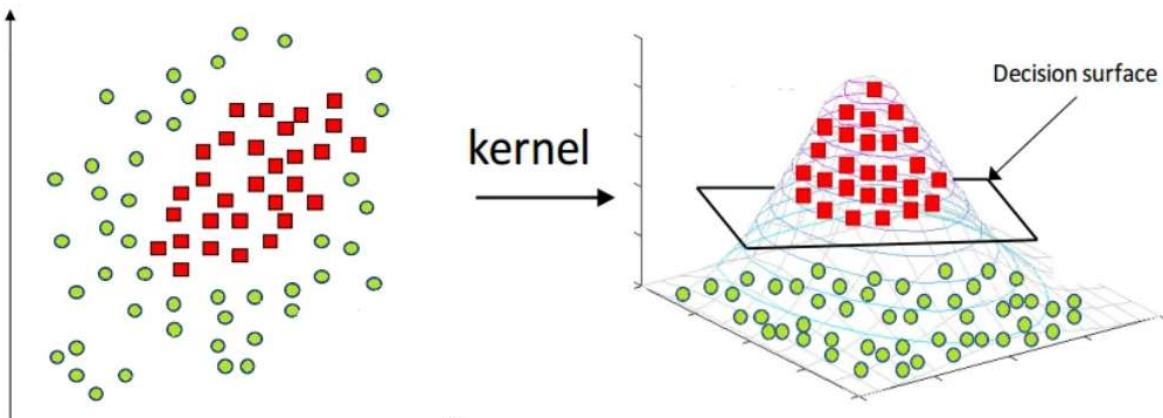
Rys. 3.6. Warstwa LSTM sieci neuronowej.

### 3.1.2. Klasyfikator SVM

Klasyfikator SVM (ang. „Support Vector Machine”) jest popularnym algorytmem uczenia maszynowego wykorzystywanym do klasyfikacji i regresji. Jego zadaniem jest podzielenie zbioru danych na części przez pewną hiperpłaszczyznę. Stawiana jest ona tak, aby odległość od niej do najbliższego punktu z każdej klasy była największa a próbka była po odpowiedającej stronie hiperpłaszczyzny. Odległość od hiperpłaszczyzny do najbliższego punktu dla każdej klasy nazywa się marginesem (ang. „margin”) a przedstawione podejście nazywane jest używaniem klasyfikatora maksymalizacji marginesu (ang. „maximal margin classifier”) [12].

Główną wadą metody maksymalizacji marginesu jest to, że jest ona bardzo wrażliwa na anomalie. Jeśli jest w stanie idealnie dopasować się do danych treningowych i margines jest mały, to dla danych testowych dopasowanie może być znacznie gorsze. Rozwiązaniem tego problemu jest użycie miękkiego marginesu (ang. „soft margin”). Wprowadza on pewien błąd klasyfikacji, który daje możliwość niepoprawnego sklasyfikowania niektórych punktów podczas uczenia klasyfikatora SVM, za to lepiej dopasowuje się do ogólnego trendu danych i daje lepsze rezultaty dla zbioru testowego. To podejście jest nazywane używaniem klasyfikatora marginesu miękkiego (ang. „soft margin classifier”) lub klasyfikatora wektorów wektorów (ang. „support vector classifier”) [13].

Kiedy dane są trudno separowane można je przekształcić do wyższego wymiaru, dzięki czemu staje się możliwe wykonanie klasyfikacji SVM (Rys. 3.7). Żeby zredukować złożoność obliczeniową tej operacji stosuje się przekształcenie do problemu dualnego. Umożliwia to także zastosowanie jąder przekształceń: liniowe, wielomianowe, gaussowskie (ang. Radial Basis Function, RBF) lub sigmoidalne [13].



Rys. 3.7. Przekształcanie danych do wyższego wymiaru [14].

### 3.1.3. Algorytm Kazemiego-Sullivan

Algorytm Kazemiego -Sullivan służy do wykrywania punktów charakterystycznych twarzy. Jest to implementacja kaskady drzew regresyjnych.

Drzewo decyzyjne (ang. „decision tree”) jest to technika uczenia maszynowego używana do klasyfikacji lub regresji (Rys. 3.8.). Tworzy model na podstawie szeregu binarnych decyzji. Pierwsza z nich nazywana jest korzeniem. Kolejne decyzje, nazywane wierzchołkami, połączone są gałęziami. Jeśli wierzchołek jest ostatnim elementem drzewa, czyli nie prowadzi do kolejnych decyzji, jest nazywany liściem [15].

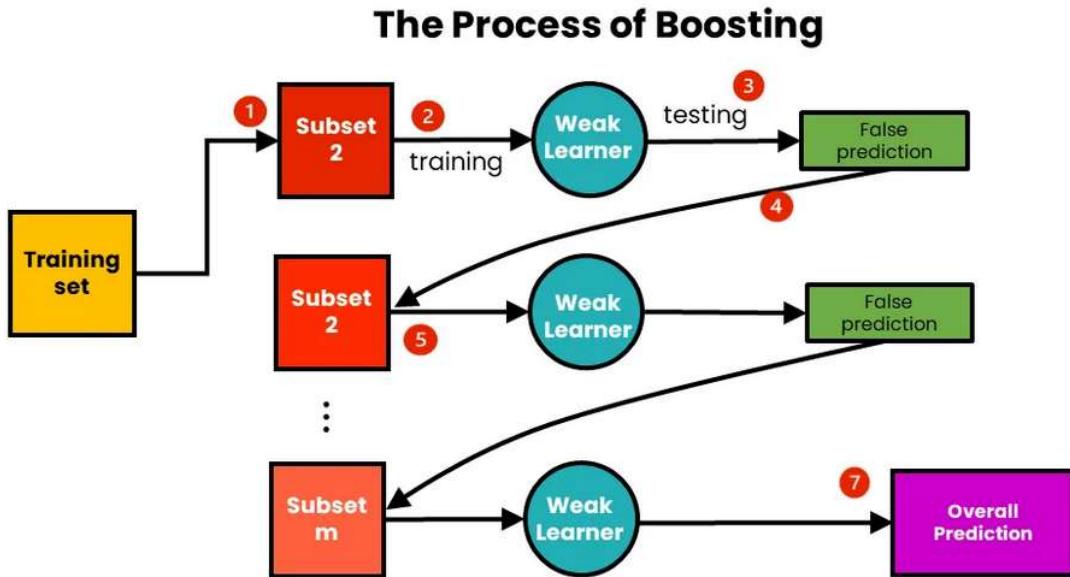
W drzewie regresyjnym liście przyjmują wartości ciągłe, zamiast etykiety klasy [16]. Algorytm dzieli zbiór danych na dwie części. W każdej dokonuje predykcji, że dane można przybliżyć za pomocą średniej wszystkich wartości w grupie. Łączny błąd predykcji jest funkcją kosztu, której minimum determinuje podział, który staje się korzeniem. Analogicznie są wyznaczane każde kolejne węzły. Drzewo jest skończone, jeśli w każdym liściu będzie zbyt mało elementów, żeby dokonać podziału [17].



Rys. 3.8. Drzewo decyzyjne (wykonano na podstawie [18]).

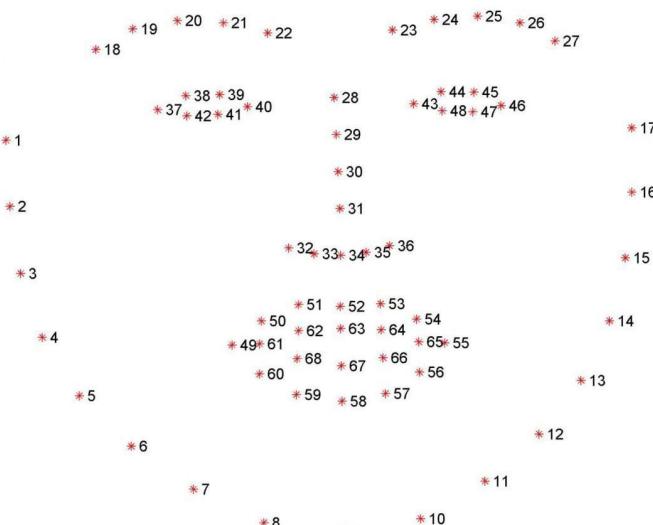
Boosting jest to technika łączenia sekwencyjnego szeregu modeli, nazywanych słabymi klasyfikatorami (ang. „weak learners”), które razem tworzą lepsze predykcje (Rys. 3.9.) [17].

Należy do rodziny metod uczenia zespołowego (ang. „ensemble methods”). W gradient boostingu celem każdego kolejnego modelu jest przewidywanie błędów poprzedniego, minimalizując funkcję kosztu całego zespołu modeli [19].



Rys. 3.9. Przedstawienie boostingu [20].

Algorytm Kazemiego-Sullivan opiera się na kaskadzie drzew regresyjnych trenowanych za pomocą gradient boostingu. Punkty charakterystyczne twarzy (Rys. 3.10.) są zapisywane, jako współrzędne w wektorze nazywanym kształtem. Każdy model jest uczony na danych składających się z obrazu, aktualnego kształtu oraz docelowego kroku aktualizacji kształtu – różnicy między aktualnym kształtem a docelowym. Pierwszy kształt twarzy jest to średnia kształtów z danych treningowych, przeskalowana do ramki otaczającej twarz. Aktualny kształt oraz docelowy krok aktualizacji kształtu jest zmieniany z każdym kolejnym drzewem regresyjnym aż do otrzymania zadowalających rezultatów [21].



Rys. 3.10. Punkty charakterystyczne twarzy [21].

## 3.2. Algorytm HOG

Algorytm HOG (ang. „Histogram of Oriented Gradients”) jest używany do ekstrakcji cech z obrazu, czyli przekształcenia obrazu na wektor cech, który zawiera tylko najistotniejsze informacje o obrazie np. krawędzie.

Gradient funkcji wielowymiarowej jest to wektor pochodnych każdego argumentu. Pochodną funkcji ciągłej  $f$  w punkcie  $x_0$  definiuje się jako [22]:

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}, \quad (3.6)$$

gdzie  $h$ , to pewna mała odległość od kolejnego argumentu funkcji. Pochodną funkcji dyskretniej  $g$  w punkcie  $s_n$  można zapisać jako [23]:

$$g'(s_n) = \frac{g(s_n) - g(s_0)}{n}, \quad (3.7)$$

gdzie  $n$  to odległość między punktami.

Analogicznie gradient obrazu jest to różnica wartości między dwoma sąsiednimi pikselami. Można przekształcić wzór (3.7) do postaci:

$$2g'(s_n) = g(s_{n+1}) - g(s_{n-1}). \quad (3.8)$$

Licząc gradient dla każdego piksela otrzymamy zbiór podwojonych gradientów. Ponieważ celem algorytmu jest porównywanie gradientów między sobą, więc możemy uprościć obliczenia do postaci:

$$g'(s_n) = g(s_{n+1}) - g(s_{n-1}). \quad (3.9)$$

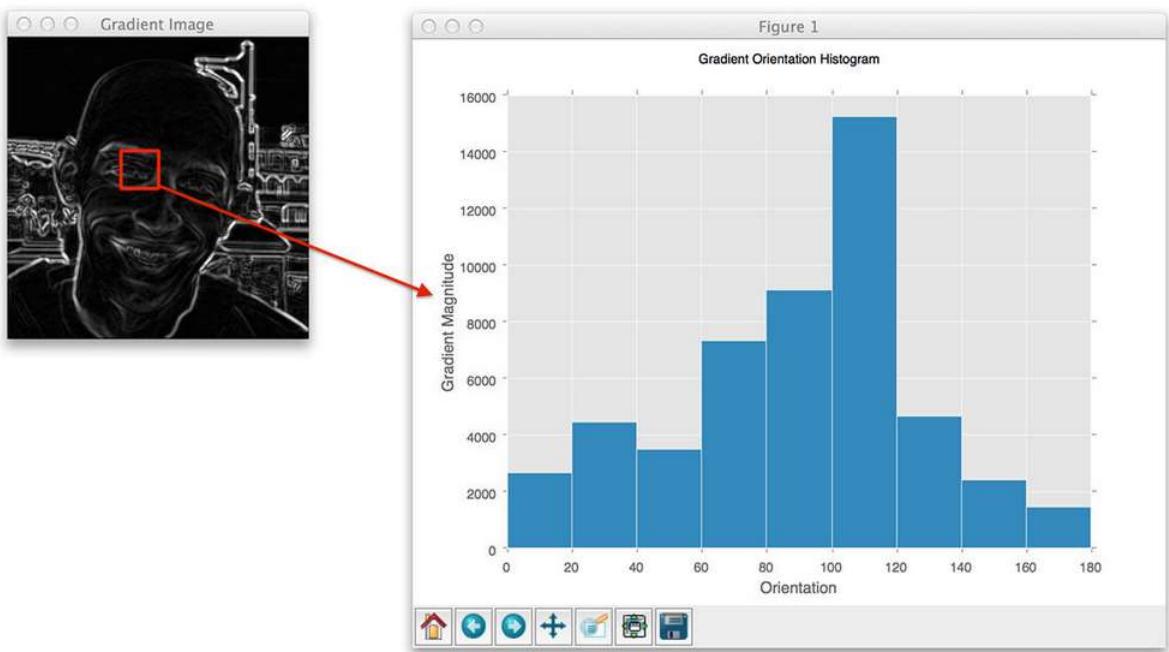
Wzór (3.9) można zaimplementować wykonując splot obrazu z maską  $[-1, 0, 1]$ . Wynikiem będzie gradient poziomy  $g_x$ . Do otrzymania gradientu pionowego  $g_y$  używa się maski  $[-1, 0, 1]^T$ . Następnie można wyliczyć wartość ogólną  $g$  i kierunek gradientu  $\theta$  za pomocą wzoru [24]:

$$g = \sqrt{g_x^2 + g_y^2} \quad (3.10)$$

$$\theta = \arctan \frac{g_y}{g_x} \quad (3.11)$$

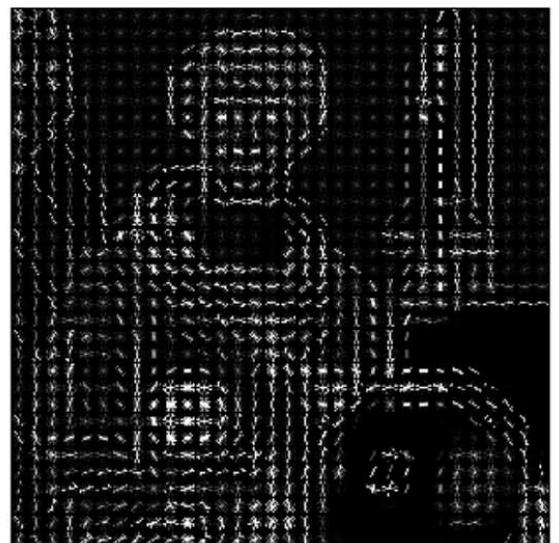
Dla obrazka kolorowego wylicza się oddzielnie gradienty odcieni R, G, B i do dalszych obliczeń przekazywana jest największa wartość gradientu  $g$  i odpowiadający niemu kąt  $\theta$  [24].

Kolejnym krokiem algorytmu jest podzielenie obrazu na kwadraty  $8 \times 8$ . Dla każdego kwadratu tworzony jest histogram dzielący wartości gradientu na 9 kolumn, w zależności od kąta (Rys. 3.11.). Otrzymane histogramy są normalizowane dla bloków  $16 \times 16$ . Na koniec wszystkie histogramy łączone są w jeden wektor cech obrazu [24].



Rys. 3.11. Histogram dla bloku 8x8 [25].

Otrzymany wektor cech (Rys. 3.12.) w połączeniu z wyuczonym klasyfikatorem SVM umożliwia wykrywanie twarzy na obrazie.



Rys. 3.12. Wyniki algorytmu HOG – wizualne przedstawienie wektora cech obrazu [26] .

## 4. Opis eksperymentu

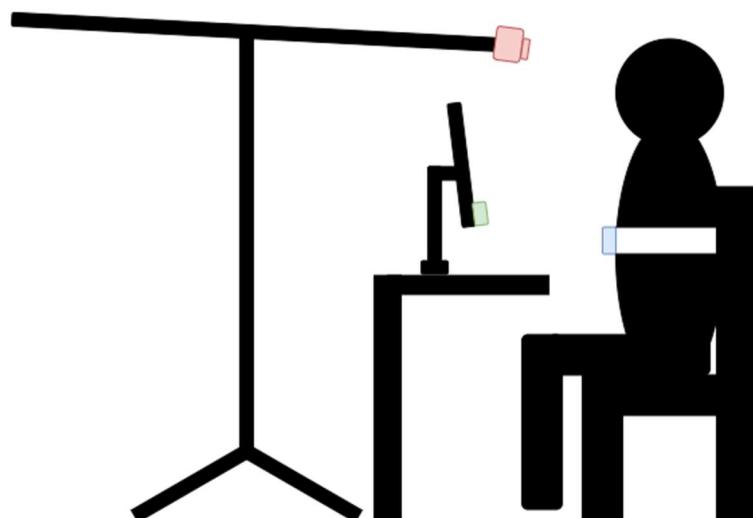
### 4.1. Plan eksperymentu

Na początku eksperymentu następuowało wyjaśnienie osobom badanym celu i przebiegu badania oraz wypełnienie pozwolenia na zbieranie danych osobowych. Następnie wykonywano czynności techniczne: badany zakładał pulsomierz, dopasowywano pozycję kamery oraz kalibrowano eye-tracker. Schemat stanowiska przedstawia Rys. 4.1.

Po wstępnych przygotowaniach na monitorze przed badanym wyświetlała się seria obrazów w formacie:

- krzyż skupiający wzrok na środku ekranu – 5 sekund,
- obraz wywołujący emocje – aż do kliknięcia przez badanego,
- plansza z napisem „ankieta”, podczas której badany opisywał swoje emocje – aż do kliknięcia przez badanego,
- pusta plansza pozwalająca badanemu na wyciszenie się przed kolejnym obrazkiem – 10 sekund.

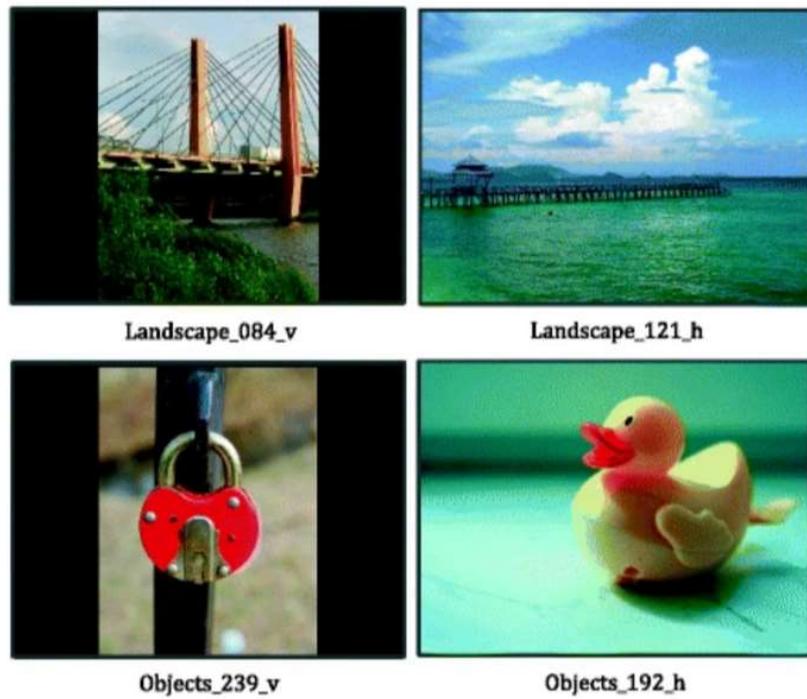
Cały eksperyment dla jednego badanego trwał od 30 minut do godziny.



Rys. 4.1. Schemat stanowiska pomiarowego. Na czerwono zaznaczona jest kamera. Na niebiesko zaznaczony jest pulsomierz. Na zielono zaznaczony jest eye-tracker.

### 4.2. Dobór bazy bodźców

W badaniu została użyta baza NAPS (ang. Nencki Affective Picture System) (Rys. 4.2.) [27]. Jest to baza danych składająca się z 1356 obrazów wywołujących emocje oraz ich średnie wyniki wartościowości, pobudzenia i dominacji. Każdy obraz jest podzielony na 5 kategorii: ludzie, twarze, krajobrazy, zwierzęta, przedmioty. Dodatkowo spośród wszystkich zdjęć wybrano 510 i zbadano jakie, z sześciu podstawowych, wywołują emocje [28].



Rys. 4.2. Przykładowe obrazy z bazy NAPS [29].

Spośród ww. 510 zdjęć zostały wybrane dwa zestawy obrazów o najbardziej zróżnicowanych wynikach według ankiety SAM. Pierwszy zawiera 20 obrazów (Tabela 4.1.), drugi 30 (Tabela 4.2.). W mniejszym ułożono zdjęcia pod względem wartościowości w kolejności rosnącej i wybrały 5 obrazów równooddalonych od siebie. Powtórzono czynność dla pobudzenia i dominacji. Pozostałe obrazy dobrano tak, aby w zestawie znajdowały się obrazy wywołujące każdą z sześciu podstawowych emocji. Większy zbiór został wybrany analogicznie, tylko zamiast dla 5, to dla 7 równooddalonych od siebie obrazów.

Tabela 4.1. Zestaw 20 obrazów. Kolumna „W” przedstawia średnie wyniki wartościowości zbadane w [27] dla obu płci, „D” dla dominacji i „P” dla pobudzenia.

Nr	Nazwa	Opis	Emocja	W	D	P
1	Animals 006 v	wąż	strach	5,13	4,83	6,23
2	Animals 179 h	psy	szczęście	7,51	7,34	5,38
3	Faces 235 v	uśmiechnięta kobieta	szczęście	7,40	6,81	4,81
4	Objects 328 h	świnia	mieszane	3,78	3,69	6,27
5	Faces 135 h	uśmiechnięty chłopiec	szczęście	7,35	6,71	4,84
6	Landscapes 102 h	skały	szczęście	7,18	7,44	3,90
7	Landscapes 075 h	szyba	szczęście	6,15	6,00	4,17
8	People 025 v	rzeźnia	odraza	3,43	3,27	6,31
9	Landscapes 120 v	łąka	szczęście	8,31	8,35	2,04
10	Landscapes 024 v	blok mieszkalny	mieszane	4,80	4,71	4,36
11	Objects 201 v	ubrania	szczęście	6,00	5,88	3,91
12	People 198 h	okaleczona noga	odraza	1,58	1,43	7,56
13	Objects 311 h	autobus	mieszane	5,04	4,88	4,98
14	People 127 h	napaść	złość	2,15	2,65	7,62
15	Objects 093 h	grzyby	zdziwienie	5,78	5,61	4,72

16	People_075_v	swastyka	złość	2,75	3,32	6,63
17	People_217_h	operacja	mieszane	3,14	2,82	6,31
18	Objects_115_h	węgiel	zdziwienie	4,74	4,91	5,04
19	People_074_v	żołnierz	strach	3,88	4,59	6,42
20	Animals_101_v	kot	szczęście	7,42	6,94	4,39

Tabela 4.2. Zestaw 30 obrazów. Kolumna „W” przedstawia średnie wyniki wartościowości zbadane w [27] dla obu płci, „D” dla dominacji i „P” dla pobudzenia.

Nr	Nazwa	Opis	Emocja	W	D	P
1	Animals_221_h	dżdżownice	odraza	3,63	3,27	6,48
2	Animals_006_v	waż	strach	5,13	4,83	6,23
3	Objects_093_h	grzyby	zdziwienie	5,78	5,61	4,72
4	Animals_195_h	kot	szczęście	5,83	5,78	4,88
5	Animals_060_h	chora pantera śnieżna	smutek	3,02	3,29	6,59
6	Animals_208_h	mały kotek	szczęście	7,41	7,39	4,29
7	Faces_227_v	uśmiechnięta kobieta	szczęście	7,25	6,75	4,21
8	Faces_356_h	uśmiechnięta para	szczęście	7,65	7,09	5,34
9	Faces_318_v	para	szczęście	6,06	5,63	4,49
10	People_127_h	napaść	złość	2,15	2,65	7,62
11	Faces_149_v	okaleczona twarz	smutek	2,30	2,19	7,61
12	Faces_182_h	starszy mężczyzna	szczęście	5,83	5,51	4,78
13	Faces_115_h	uśmiechnięty chłopiec	szczęście	7,06	6,51	3,89
14	Faces_218_h	kobieta	mieszane	4,96	4,96	4,85
15	Faces_037_h	płaczący chłopiec	mieszane	4,14	5,18	5,82
16	Faces_283_h	płacząca starsza kobieta	smutek	2,48	4,27	7,26
17	Landscapes_165_h	góry	szczęście	8,22	8,15	2,57
18	Landscapes_120_v	iąka	szczęście	8,31	8,35	2,04
19	Landscapes_091_h	okno	mieszane	5,16	5,54	4,57
20	Objects_239_v	zamek	mieszane	5,00	5,15	4,74
21	People_198_h	okaleczona nogą	odraza	1,58	1,43	7,56
22	Objects_128_h	płytki	mieszane	4,04	3,96	5,20
23	Objects_136_h	przedmioty	mieszane	3,96	4,49	5,39
24	Objects_115_h	węgiel	zdziwienie	4,74	4,91	5,04
25	People_075_v	swastyka	złość	2,75	3,32	6,63
26	People_074_v	żołnierz	strach	3,88	4,59	6,42
27	Faces_037_h	płaczący chłopiec	mieszane	4,14	5,18	5,82
28	People_190_h	kierowca	szczęście	8,22	8,24	4,68
29	People_044_h	dziecko	szczęście	6,85	6,56	3,77
30	Animals_136_h	kot	szczęście	6,81	6,67	3,33

### **4.3. Aplikacja do odtwarzania bodźców oraz nagrywania ścieżki wzrokowej**

Eye-tracker *Tobii Pro Fusion* [30] to urządzenie służące do zbierania danych o ruchu gałek ocznych, wielkości żrenicy oraz stopnia otwarcia oka badanego. Składa się z dwóch kamer rejestrujących dane z maksymalną częstotliwością 250 próbek/s. Jest umieszczone na dolnej krawędzi monitora (Rys. 4.1.) i połączone z komputerem kablem USB. Podczas pomiaru wysyła dane do programu *Tobii Pro Lab*.

*Tobii Pro Lab* [31] jest to oprogramowanie służące do projektowania i przeprowadzania eksperymentów, które polegają na wyświetlaniu obrazów lub filmów oraz rejestrowaniu reakcji na nie przez eye-tracker.

Oprócz dodawania obrazów lub filmów, użytkownik może ustalić ich kolejność oraz sposób przechodzenia między nimi (np. czasowo lub po kliknięciu). W czasie trwania eksperymentu informacje o ruchu gałek ocznych są mapowane na konkretne fragmenty wyświetlanego obrazu tworząc ścieżkę wzrokową.

Ponieważ każdy badany ma inny wzrost oraz rozstaw oczu, dlatego przed badaniem należało indywidualnie skalibrować eye-tracker.

### **4.4. Pomiar pulsu**

Pulsomierz *Polar H10* [32] jest to urządzenie, które zbiera dane o aktualnym tętnie. Zakłada się je na klatkę piersiową badanego (Rys. 4.1.). Wykorzystuje ono optyczny pomiar tętna i przesyła zebrane dane do komputera. Prędkość próbkowania to 1 próbka/s.

### **4.5. Nagranie mimiki**

Na początku do nagrywania mimiki użyta została kamera *Intel Depth Camera D415* [33], która nagrywa z częstotliwością 60 klatek/s. Z powodu nieoczekiwanych problemów technicznych nie była ona możliwa do późniejszego użycia. Zamiast niej korzystano z kamery *Logitech QuickCam Pro 9000* [34], której częstotliwość nagrywania to 15 klatek/s. Informacje o tym, która kamera została użyta, zostały zawarte w Tabela 4.4.

Pozycje obu kamer były takie same - na statywie nad monitorem (Rys. 4.1.). Wysokość kamery była ustawiana na nowo dla każdego badanego. Filmy były rejestrowane z rozdzielczością 640 x 480 pikseli i zapisywane do pliku w formacie AVI (ang. „Audio Video Interleave”).

### **4.6. Ankieta**

Do przeprowadzenia ankiety użyto usługi *Formularze Google* [35]. Ankieta była wypełniana po objerzeniu każdego obrazu. Badany wpisywał słownie emocję, jaką odczuwał (np. strach, nostalgia, podziw, irytacja), następnie wypełniał kwestionariusz SAM w skali od 1 do 9. Wszystkie wyniki ankiet są w Tabela 8.1.

Pytania dla jednego obrazka składały się z 15 elementów. Przedstawia je Tabela 4.3. oraz Rys. 4.3.-Rys. 4.5.

Tabela 4.3. Elementy wyświetlające się na raz na jednej stronie ankiety.

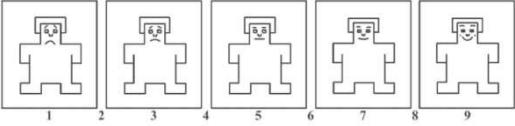
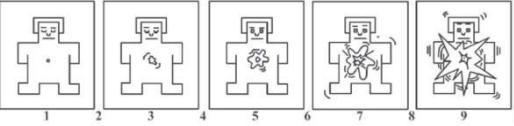
Nr	Wyświetlane pole
1	Numer obrazka
2	Pytanie: „jaką emocję teraz odczuwasz?”
3	Pole do wpisania tekstu
4	Napis: „wartościowość”
5	Piktogram wartościowości
6	Pytanie: „czujesz negatywną emocję (1) czy pozytywną (9)?”
7	Skala od 1 do 9, co jeden
8	Napis: „pobudzenie”
9	Piktogram pobudzenia
10	Pytanie: „czujesz się spokojny (1) czy pobudzony (9)?”
11	Skala od 1 do 9, co jeden
12	Napis: „dominacja”
13	Piktogram dominacji
14	Pytanie: „czujesz się przytłoczony emocją (1) czy w pełni ją kontrolujesz (9)?”
15	Skala od 1 do 9, co jeden

Obraz 1

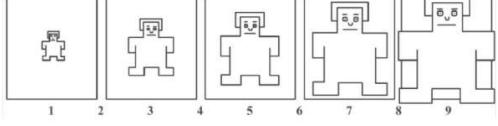
Jaką emocję teraz odczuwasz? \*

Twoja odpowiedź \_\_\_\_\_

Rys. 4.3. Pola 1 - 3 ankiety. Dotyczą numeru obrazka oraz wpisywania słownej emocji.

<p><b>Wartościowość</b></p>  <p>1      2      3      4      5      6      7      8      9</p>	<p><b>Pobudzenie</b></p>  <p>1      2      3      4      5      6      7      8      9</p>
<p>Czujesz negatywną emocję (1) czy pozytywną (9)? *</p> <p><input type="radio"/> 1  <input type="radio"/> 2  <input type="radio"/> 3  <input type="radio"/> 4  <input type="radio"/> 5  <input type="radio"/> 6  <input type="radio"/> 7  <input type="radio"/> 8  <input type="radio"/> 9</p> <p>Czujesz się spokojny (1) czy pobudzony (9)? *</p> <p><input type="radio"/> 1  <input type="radio"/> 2  <input type="radio"/> 3  <input type="radio"/> 4  <input type="radio"/> 5  <input type="radio"/> 6  <input type="radio"/> 7  <input type="radio"/> 8  <input type="radio"/> 9</p>	

Rys. 4.4. Pola 4 – 11 ankiety. Dotyczą podawania wartościowości i pobudzenia.

<p><b>Dominacja</b></p>  <p>1      2      3      4      5      6      7      8      9</p>	<p>Czujesz się przytłoczony emocją (1) czy w pełni ją kontrolujesz (9)? *</p> <p><input type="radio"/> 1  <input type="radio"/> 2  <input type="radio"/> 3  <input type="radio"/> 4  <input type="radio"/> 5  <input type="radio"/> 6  <input type="radio"/> 7  <input type="radio"/> 8  <input type="radio"/> 9</p>
---	--

Rys. 4.5. Pola 12 – 15 ankiety. Dotyczą podawania dominacji.

## 4.7. Grupa badawcza

W badaniu wzięło udział 17 osób. Przy badaniu trzech z nich wystąpiły błędy, których rezultatem były niekompletne dane. Nie zostały one użyte do stworzenia prototypu systemu rozpoznawania emocji, ale ich odpowiedzi są zawarte w analizie wyników ankiet w rozdziale 6.1.

Tabela 4.4. Dane uczestników badania oraz eksperymentu. Obrazy z każdego zestawu obrazów zostały opisane w Tabela 4.1. i Tabela 4.2. Kamera *Logitech QuickCam Pro 9000* nagrywała z częstotliwością 60 klatek/s a kamera *Intel Depth Camera D415* z częstotliwością 15 klastek/s.

Nr	Wiek	Płeć	Komplet danych	Zestaw obrazów	Kamera
1	23	M	Tak	30 obrazów	Logitech QuickCam Pro 9000
2	23	M	Tak	30 obrazów	Intel Depth Camera D415
3	19	K	Nie	30 obrazów	Intel Depth Camera D416
4	20	M	Tak	20 obrazów	Intel Depth Camera D417
5	24	M	Tak	30 obrazów	Intel Depth Camera D418
6	23	M	Tak	20 obrazów	Intel Depth Camera D419
7	21	M	Nie	30 obrazów	Intel Depth Camera D420
8	24	M	Tak	20 obrazów	Intel Depth Camera D421
9	24	M	Tak	20 obrazów	Intel Depth Camera D422
10	24	M	Nie	20 obrazów	Intel Depth Camera D423
11	23	M	Tak	30 obrazów	Intel Depth Camera D424
12	23	M	Tak	20 obrazów	Intel Depth Camera D425
13	23	M	Tak	30 obrazów	Logitech QuickCam Pro 9000
14	24	M	Tak	30 obrazów	Logitech QuickCam Pro 9001
15	27	K	Tak	30 obrazów	Logitech QuickCam Pro 9002
16	27	K	Tak	30 obrazów	Logitech QuickCam Pro 9003
17	24	M	Tak	30 obrazów	Logitech QuickCam Pro 9004

## 5. Algorytm rozpoznający emocje

### 5.1. Środowisko programistyczne

Do implementacji systemu rozpoznawania emocji oraz obróbki danych został użyty język *Python 3.11* oraz następujące jego biblioteki:

- *Numpy 1.26.1* [36] – rozszerza możliwe do użycia funkcje matematyczne,
- *Matplotlib 3.8.0* [37] – umożliwia tworzenie wykresów,
- *Pandas 2.1.2* [38] – pozwala na zaawansowane operacje na słownikach,
- *Tensorflow 2.14.0* [39] – ułatwia implementację metod uczenia maszynowego,
- *Keras 2.10* [40] – umożliwia tworzenie architektur sieci neuronowych,
- *Skmultilearn 0.2.0* [41] – zawiera funkcję dokonującą podziału danych na zbiór treningowy i testowy w sposób umożliwiający równomierny podział klas, dla wielowymiarowych etykiet.

Do trenowania wielu różnych architektur sieci neuronowych został użyty *Google colab* [42]. Korzystano na nim przede wszystkim z wersji:

- *Tensorflow 2.18* [39],
- *Keras 3.6* [40].

Dodatkowo do detekcji twarzy i ekstrakcji punktów charakterystycznych wykorzystano język *Python 3.9* oraz biblioteki:

- *Open CV 4.9.0.80* [43] – zawiera algorytmy do wczytywania i przetwarzania obrazów,
- *Dlib 19.22.99* [44] – zawiera wiele algorytmów i już wytrenowanych modeli, w szczególności algorytm HOG i klasyfikator SVM, pozwalający na detekcję twarzy oraz algorytm ekstrakcji punktów charakterystycznych,
- *Imutils 0.5.4* [45] – dodaje kolejne algorytmy przetwarzania obrazów.

### 5.2. Przygotowanie danych do uczenia sieci

Surowe dane o tętnie, ścieżce wzrokowej i mimice zostały poddane procesowi przekształcenia na zsynchronizowaną tabelę skupiającą wszystkie te informacje ( Rys. 5.1.).

#### 5.2.1. Format surowych danych

##### Mimika

Każdemu filmowi z nagraniem mimiki była przyporządkowana tabela w pliku CSV, która w każdym wierszu zapisywała numer klatki i odpowiadający jej znacznik czasowy.

##### Puls

W każdym wierszu była zarejestrowana przez pulsomierz wartość tętna oraz odpowiadający jej znacznik czasowy.

## Ścieżka wzrokowa

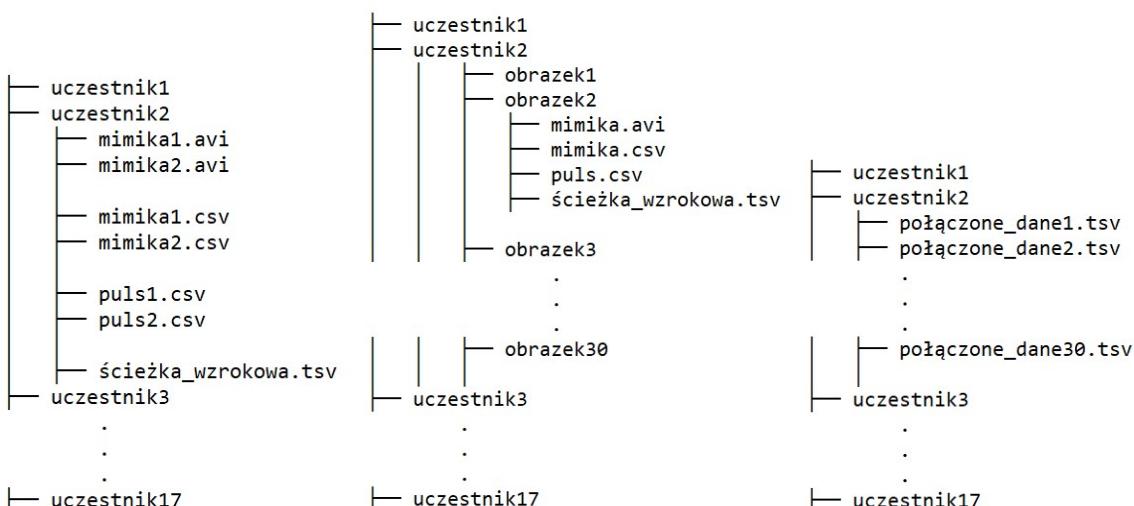
Nagrania ścieżki wzrokowej były importowane z programu *Tobii Pro Lab* w pliku TSV. Tabela miała 102 kolumny. Zawierały one kompleksowe informacje dotyczące eksperymentu (np. znaczniki czasowe, nazwa sensora, czas rozpoczęcia eksperymentu), uczestnika (np. nazwa, płeć, wiek), wyświetlonego obrazu (np. nazwa, rozdzielcość), wyników kalibracji, ścieżki wzrokowej, wielkości źrenicy oraz stopnia otwarcia oka badanego.

### 5.2.2. Odrzucenie nieistotnych danych

Dane surowe to jednostajne nagrania danych biometrycznych zawierające np. ustawianie kamery, odpowiadanie na ankietę czy skupianie wzroku na środku ekranu. Do dalszego procesu przekształcania danych pozostawiono tylko te fragmenty, gdy badany reagował na obraz wywołujący emocje.

Program *Tobii Pro Lab* zapisywał znaczniki czasowe przechodzenia między poszczególnymi obrazkami. Na ich podstawie pozostawiono tylko te fragmenty tabel pulsu, mimiki oraz ścieżki wzrokowej zawierające wyłącznie reakcje na odpowiednie obrazy. Z nagrania mimiki wyekstraktowano klatki w formacie PNG, odpowiadające pozostawionym wierszom tabeli ze znacznikami czasowymi, wielkości 640 x 480 pikseli. Dzięki temu utworzono strukturę plików, gdzie reakcje jednego badanego na jeden obraz dla wszystkich modalności znajdowały się w tym samym folderze.

Dodatkowo zmodyfikowano tabele z zapisem ścieżki wzrokowej. Wcześniej zawierały one dużo informacji istotnych do analizy badania, ale nieważnych dla systemu rozpoznawania emocji np. nazwa badanego czy data rozpoczęcia eksperymentu. W nowych tabelach pozostawiono między innymi znaczniki czasowe, ścieżkę wzrokową, wielkość źrenicy i otwartość oka. Wszystkie nazwy kolumn końcowych danych znajdują się w Tabela 8.2.



Rys. 5.1. Po lewej - poglądowa struktura surowych danych. Po środku - poglądowa struktura po odrzuceniu nieistotnych danych. Po prawej – dane po synchronizacji.

### 5.2.3. Sprawdzenie, czy każda próbka ma wszystkie modalności

Ponieważ nagrania dla wszystkich modalności zostały podzielone, pozwoliło to na sprawdzenie kompletności danych. Jeśli dana reakcja nie zawierała pulsu, ścieżki wzrokowej lub mimiki, odrzucano cały komplet nagrań.

### 5.2.4. Ekstrakcja punktów charakterystycznych twarzy

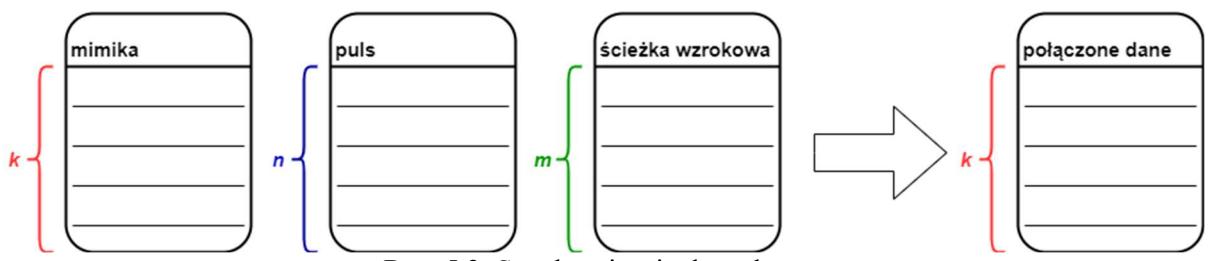
W momencie kiedy wszystkie dane były podzielone i kompletne, można było przeprowadzić detekcję twarzy za pomocą algorytmu HOG z liniowym klasyfikatorem SVM. Następnie przeprowadzono ekstrakcję punktów charakterystycznych twarzy za pomocą algorytmu Kazemego-Sullivan (Rys. 5.2.). Użyto do tego celu model z biblioteki *dlib* o nazwie *shape\_predictor\_68\_face\_landmarks*. Zmieniono w ten sposób każdą klatkę filmu w wektor punktów charakterystycznych oraz zapisano je w tabeli razem ze znacznikami czasowymi.



Rys. 5.2. Przykładowe wyniki ekstrakcji punktów charakterystycznych twarzy [21].

### 5.2.5. Synchronizacja danych

Wszystkie modalności były zapisywane z różną częstotliwością. Nagrania mimiki pozostawiono bez zmian i do ich znaczników czasowych przymierzały pozostałe modalności. Nagrania ścieżki wzrokowej miały większą częstotliwość, więc wytypowano te wiersze, które miały najmniejszą różnicę czasu i zapisano je do nowego pliku. Zapisy pulsu miały mniejszą częstotliwość, więc powielono wiersze. Na koniec połączono punkty charakterystyczne twarzy, puls i ścieżkę wzrokową w jedną tabelę (Tabela 8.2.) (Rys. 5.3).



#### 5.2.6. Wypełnienie pustych kolumn zerami, zakodowanie słów

Pliki importowane z programu *Tobii Pro Lab* zawierały wiele pustych kolumn. Zastąpiono je zerami. Niektóre kolumny zawierały ograniczoną liczbę wartości słownych. Każde słowo zakodowano pewną cyfrą i zapisano w tabeli końcowej.

#### 5.2.7. Normalizacja danych

Na koniec dane zostały znormalizowane. Połączono wszystkie próbki w jedną tablicę o 208 kolumnach (Tabela 8.2.). Dla każdej kolumny danych obliczono średnią  $\bar{x}$  i odchylenie standardowe  $\sigma$  oraz wykonano normalizację danych według wzoru:

$$y = \frac{x - \bar{x}}{\sigma}, \quad (5.1)$$

gdzie  $x$  oznacza dane wejściowe a  $y$  dane wyjściowe.

#### 5.2.8. Zmiana formatu etykiet

Format danych ankietowych z importowanych formularzy google był taki, że wszystkie odpowiedzi uczestnika znajdowały się w jednym wierszu. Zmieniono go w taki sposób, że w jednym wierszu są wyniki ankiet jednego uczestnika dla jednego obrazka.

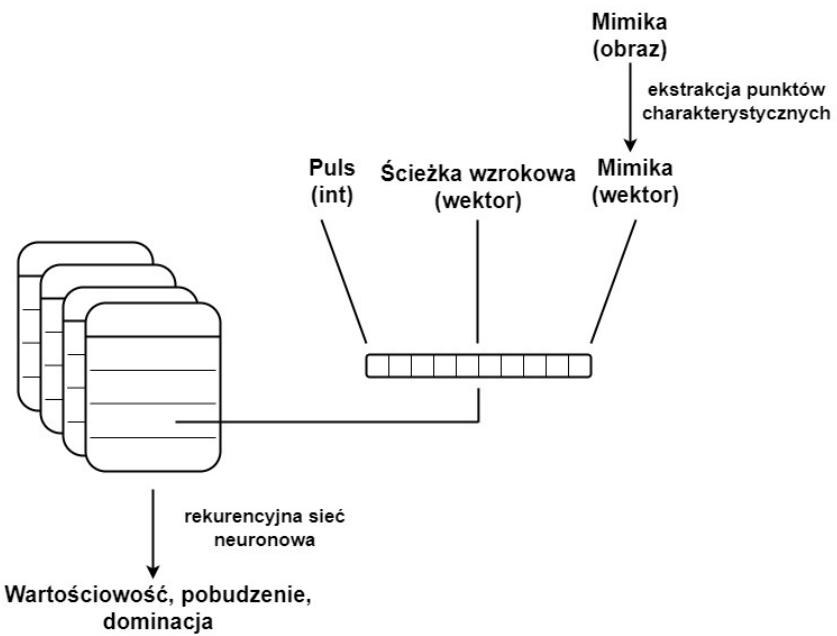
#### 5.2.9. Zmiana sformułowania słownych emocji

Słowne emocje nie były używane w uczeniu sieci neuronowej. Mimo to, dla łatwiejszej ich analizy, zmieniono niektóre sformułowania (np. neutralnie na neutralną emocję), ujednoliciono na te same części mowy (np. spokojny na spokój), sprawiono, że wszystkie wyrazy są napisane małymi literami i poprawiono błędy ortograficzne. W ten sposób otrzymano ujednoliczoną tablicę danych bez straty ich znaczenia.

### 5.3. Prototyp sieci wykrywającej emocje

#### 5.3.1. Typy sieci, ich architektura oraz strategie uczenia

Celem sieci jest podanie wartości ankiety SAM na podstawie tabeli z nagraniami wyrażającymi emocje (Rys. 5.4.). Wartościami wejściowymi są tabele etykiet oraz tabele danych o 208 kolumnach (Tabela 8.2.) i różnej liczbie wierszy. Wynika to z tego, że badany sam decydował, kiedy przejść z obrazka wywołującego emocje do odpowiadania na ankietę. Dane zostały podzielone na zbiór treningowy (258 próbek), testowy (54 próbki) oraz walidacyjny (45 próbek).

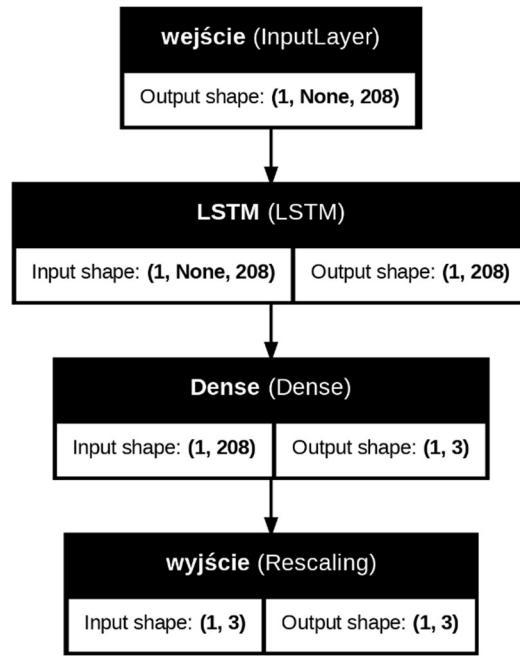


Rys. 5.4. Schemat informacji zawartych w danych wejściowych do sieci oraz danych wyjściowych.

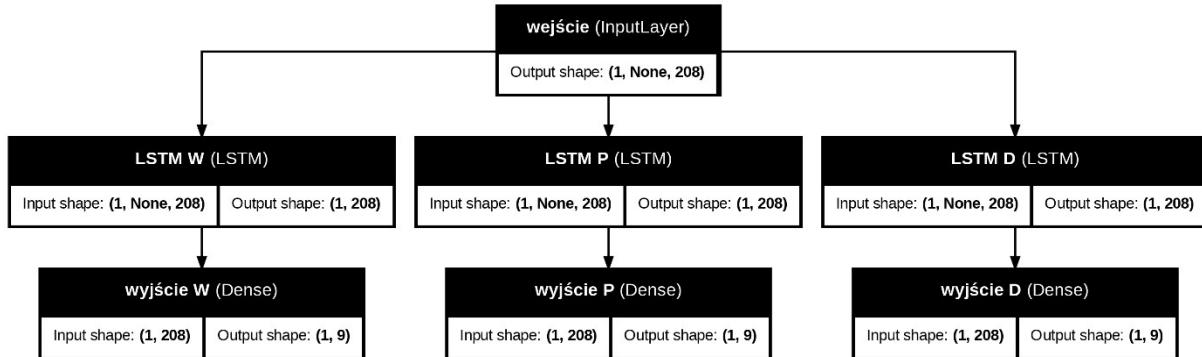
Ustawiono rozmiar paczki danych (ang. batch size) na 1 przez różną długość tablic wejściowych oraz sposób napisania biblioteki *tensorflow*.

Każda tablica miała różną liczbę wierszy. Żeby to skompensować, użyto ważonych metryk podczas treningu. Waga była odwrotnie proporcjonalna do długości tablicy.

Wartościami wyjściowymi sieci są 3 parametry od 1 do 9. Ten problem można potraktować na dwa sposoby. Wartościowość, pobudzenie i dominacja przyjmują dyskretnie wartości, dlatego można stworzyć sieć do klasyfikacji. Z drugiej strony każda z tych wartości reprezentuje punkt na skali, do przewidywania czego odpowiednia jest regresja. Dodatkowo jedna sieć może przewidawać wszystkie 3 wartości lub można stworzyć 3 sieci przewidujące tylko jeden parametr. Biorąc to pod uwagę wytrenowano 20 architektur sieci regresyjnej, które przewidywały trzy wartości (Rys. 5.5.) oraz 20 architektur trzech sieci klasyfikujących które przewidywały jedną wartość. Dla uproszczenia uczenia i testowania sieci, te ostatnie były połączone jednym wejściem, które rozgałęziało się w trzy sekwencje takich samych warstw, które nie były później ze sobą połączone (Rys. 5.6.).



Rys. 5.5. Struktura najprostszej sieci regresyjnej.



Rys. 5.6. Struktura najprostszej sieci klasyfikacyjnej.

Modele regresyjne były skompilowane z parametrami:

- optymalizator – *Adam*,
- współczynnik uczenia –  $8 \cdot 10^{-7}$ ,
- funkcja kosztu – błąd średnio-kwadratowy,
- ważona metryka – błąd średnio-kwadratowy.

Modele klasyfikujące były skompilowane z parametrami:

- optymalizator – *Adam*,
- współczynnik uczenia –  $8 \cdot 10^{-7}$ ,
- funkcja kosztu – rzadka kategoryczna entropia krzyżowa,
- ważona metryka – dokładność.

Oba typy modeli miały na początku warstwę wejściową. Modele regresyjne były zakończone warstwą w pełni połączoną, zawierającą trzy neurony i sigmoidalną funkcję aktywacji oraz warstwę skalującą, która przekształca dane z przedziału  $<0, 1>$  do  $<1, 9>$  (Rys. 5.5.) (Tabela 8.3.). Wynikiem tych sieci były bezpośrednio wyniki ankiety SAM. Ostatnią warstwą modeli klasyfikujących były trzy warstwy w pełni połączone zawierające dziewięć

neuronów i sigmoidalną funkcję aktywacji (Rys. 5.6.). Warstwa ta dawała prawdopodobieństwo, że dane należą do jednej z klas (Tabela 8.4.). Jako odpowiedź sieci była traktowana klasa z największym prawdopodobieństwem wystąpienia.

Środkowe warstwy różniły się między modelami (Tabela 5.1.). Składały się z przynajmniej jednej warstwy LSTM oraz potencjalnie z różnych kombinacji kolejnych warstw LSTM, warstwy dropout, która porzuca 0,2 połączeń między jej wejściem a wyjściem, warstwy konwolucyjnej z jądrem o wielkości 5 oraz warstwy dense – kolejnej warstwy w pełni połączonej, której funkcją aktywacji jest tangens hiperboliczny. Parametry wszystkich warstw przedstawia Tabela 5.2.

Tabela 5.1. Lista badanych modeli wraz z ich architekturą. Wykorzystano warstwy LSTM, w pełni połączone, konwolucyjne oraz dropout, który jest aplikowany na warstwę go poprzedzającą.

Nazwa	Architektura
model1	lstm
model2	lstm, dense, dense
model3	dense, dense, lstm
model4	dense, dense, lstm, dense, dense
model5	dense, dense, lstm, dense, dropout, dense
model6	conv, lstm
model7	conv, conv, lstm
model8	conv, dense, conv, lstm
model9	conv, dense, conv, dropout, lstm
model10	conv, lstm, dense, dense
model11	conv, conv, lstm, dense, dense
model12	lstm, lstm
model13	lstm, lstm, dense, dense
model14	dense, dense, lstm, lstm
model15	dense, dense, lstm, lstm, dense, dense
model16	dense, dense, lstm, lstm, dense, dropout, dense
model17	dense, dense, lstm, dense, lstm, dense, dense
model18	dense, dense, lstm, dense, lstm, dense, dropout, dense
model19	conv, lstm, lstm
model20	lstm, lstm, lstm

Celem warstwy LSTM jest skumulowanie predykcji dla całej tabeli danych w jeden wektor cech. Jeśli sieć zawierała więcej niż jedną warstwę LSTM, to zamiast tego poprzednie warstwy przekazywały dalej swoje predykcje dla każdego wiersza tabeli.

Parametr wielkości w bibliotece *tensorflow* dla warstwy wejściowej i środkowych przed i włącznie z ostatnią warstwą LSTM wynosił (1, None, 208). Po niej rozmiar warstw środkowych wynosił (1, 208). Wyjściowe warstwy dla regresji miały wielkość (1, 3) a dla klasyfikacji (1, 9). Wartość *None* oznacza różną dla każdej próbki liczbę wierszy tabeli. Dodatkowo długość tabeli etykiet musi odpowiadać długości tabeli wejściowej, dlatego przy wczytywaniu danych powielano jej wiersze.

Każdy model trenowano dla 50 epok. Po każdej z nich były wywoływanie dwie funkcje. Jeśli wyniki sieci dla zbioru walidacyjnego były lepsze od poprzednich według odpowiadającej metryki, to wagi sieci były zapisywane. Metryką dla modeli regresyjnych był błąd średnio-kwadratowy a dla klasyfikacji wartość funkcji kosztu. Jeśli wyniki dla zbioru walidacyjnego nie poprawiły się od 3 iteracji, uczenie sieci było przedwcześnie przerywane. Uczenie każdego modelu było powtarzane 5 razy i zapisywany był ten, z najlepszymi wynikami.

Dodatkowo te same architektury wytrenowano dla 50 epok, ale bez funkcji wcześniejszego przerywania uczenia. Uczenie jednego modelu wykonywano tylko raz.

Tabela 5.2. Liczba parametrów różnych warstw używanych w sieciach.

<b>Warstwa</b>	<b>Liczba parametrów sieci</b>
Input	0
LSTM	346,944
Dense (1, None, 208)	43,472
Dense (1, 208)	43,472
Dense (1, 9)	1,881
Dense (1, 3)	627
Dropout	0
Conv	216,528
Rescale	0

### 5.3.2. Zbalansowanie zbioru sztucznymi danymi i augmentacja

Mała ilość danych stanowi problem w poprawnym uczeniu sieci. Potencjalnym jego rozwiązaniem jest augmentacja danych, czyli powielanie danych i późniejsze ich zniekształcenie.

Po podziale danych na zbiór treningowy, walidacyjny i testowy, dokonano augmentacji danych na zbiorze treningowym. Tworzoneo trzy kopie danych i dodawano do nich szum. Szum miał postać losowych wartości z rozkładu gaussowskiego o odchyleniu standardowym 0.07 i średniej w punkcie 0. Składał się z wektora długości 208, który był dodawany do każdego wiersza tabeli. Wektor dla tej samej próbki nie ulegał zmianie. Wynika to z tego, że to właśnie różnice w poszczególnych wierszach mogą wskazywać na odczuwane emocje.

W ten sposób powiększono zbiór testowy z 258 próbek do 1032. Dokonano ponownego uczenia na nowym zbiorze danych dla najlepszego modelu regresyjnego oraz najlepszego modelu klasyfikującego.

W zbiorze testowym niektóre klasy były znacznie bardziej liczne od innych, przez co sieć najbardziej się na nich skupiała podczas treningu. Podjęto próbę zaradzenia temu problemowi, poprzez dokonanie augmentacji zbioru treningowego, ale tylko dla wartości mniej spotykanych. Jeśli etykieta wartościowościami była różna od 5 i 6 lub etykieta pobudzenia była większa od 4 lub etykieta dominacji była mniejsza od 7, to tworzoneo dwie kopie danych i dodawano do nich szum. Otrzymano zbiór danych treningowych z 702 próbками. Nauczono na nim najlepsze sieci.

# 6. Wyniki

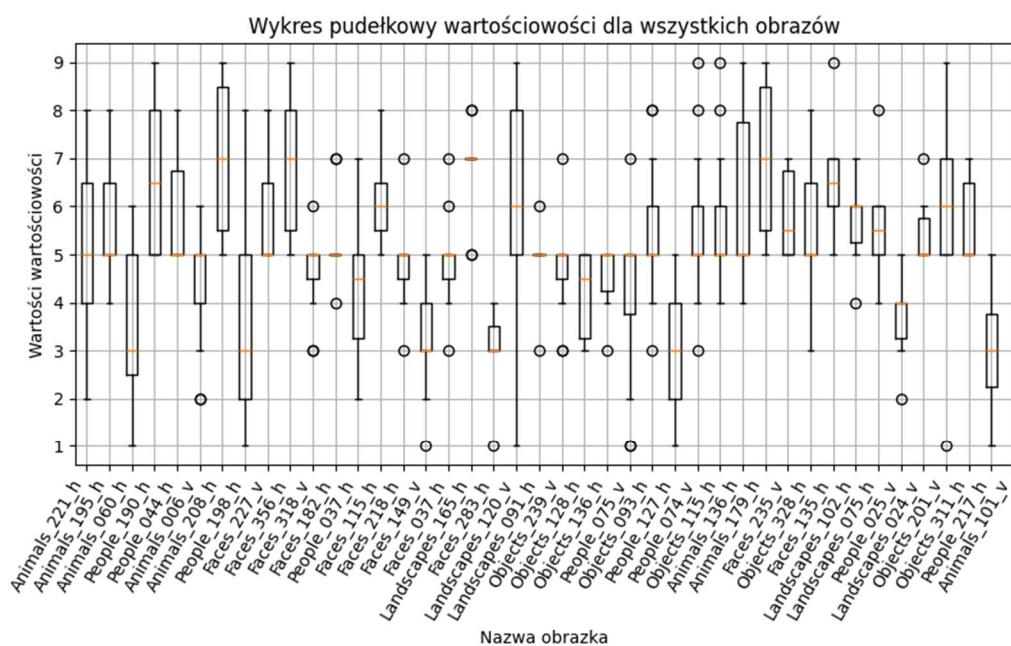
## 6.1. Ankieta

### 6.1.1. Wyniki ankiety SAM

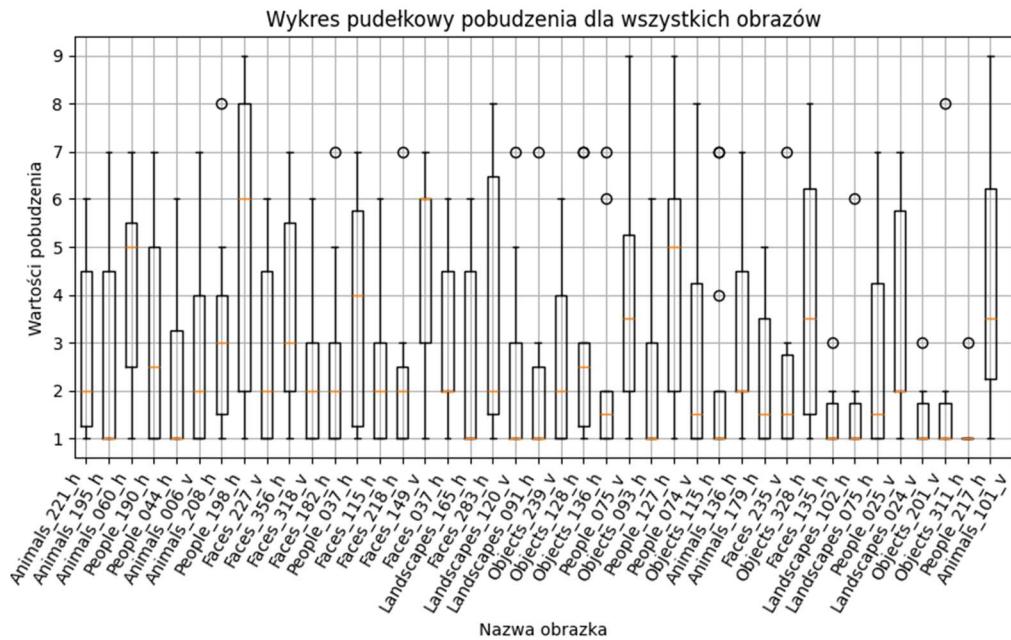
W założeniach eksperymentu było uzyskanie zbioru obrazów, które wywołują najbardziej zróżnicowane emocje. Ten cel został tylko po części osiągnięty. Rys. 6.1. przedstawia wyniki wartościowości dla wszystkich obrazów. Widać na nim, że zgodnie z oczekiwaniemi obrazy wywoływały emocje pozytywne i negatywne, pokrywając całą skalę od 1 do 9.

Nieco gorzej prezentują się wyniki pobudzenia (Rys. 6.2.). Dane są wyraźnie przesunięte w kierunku mniejszych wartości. Znalazły się osoby zaznaczające najwyższe pobudzenie, ale mediana dla większości obrazków nie przekroczyła 3.

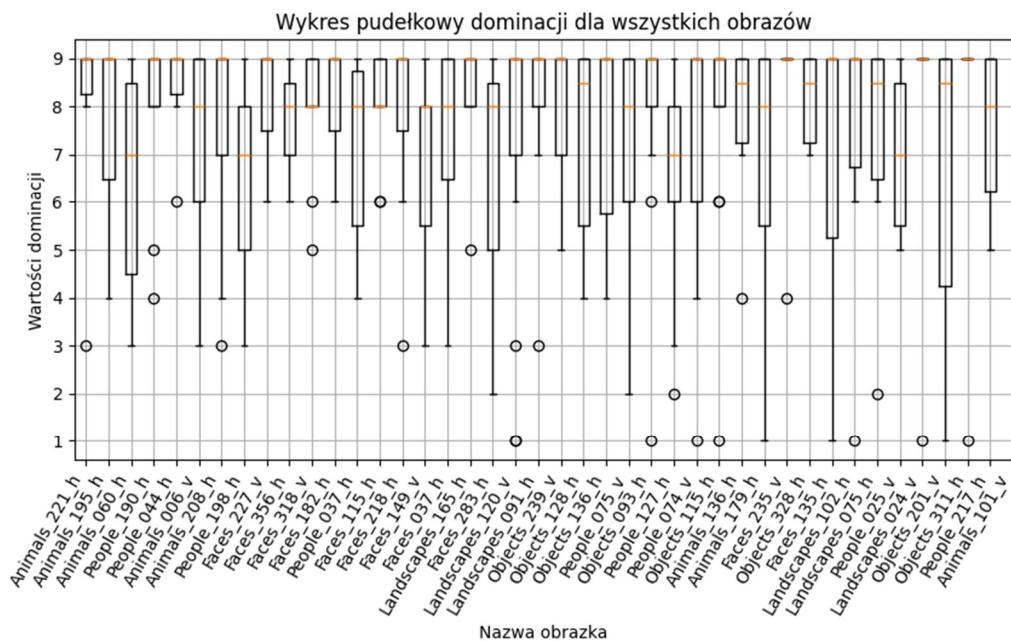
Najgorzej ma się sytuacja dla dominacji (Rys. 6.3.). Trend danych jest bardzo przesunięty w kierunku większych wartości. Mediana ponad połowy obrazów wynosi 9, co oznacza pełną kontrolę nad emocjami. Dla czterech obrazów wynosi 7 i jest to najmniejsza jej wartość. Znowu znalazły się osoby zaznaczające również niższe wartości, ale należą one do mniejszości.



Rys. 6.1. Wykres pudełkowy wartościowości dla wszystkich obrazów. Mediana jest zaznaczona kolorem pomarańczowym. Kwartył dolny i górny ustalają granice kwadratu. Pionowe linie prowadzą do minimalnej i maksymalnej wartości zbioru. Wartości odstające oznaczone są za pomocą kół.



Rys. 6.2. Wykres pudełkowy pobudzenia dla wszystkich obrazów. Mediana jest zaznaczona kolorem pomarańczowym. Kwartyl dolny i górny ustalają granice kwadratu. Pionowe linie prowadzą do minimalnej i maksymalnej wartości zbioru. Wartości odstające oznaczone są za pomocą kół.

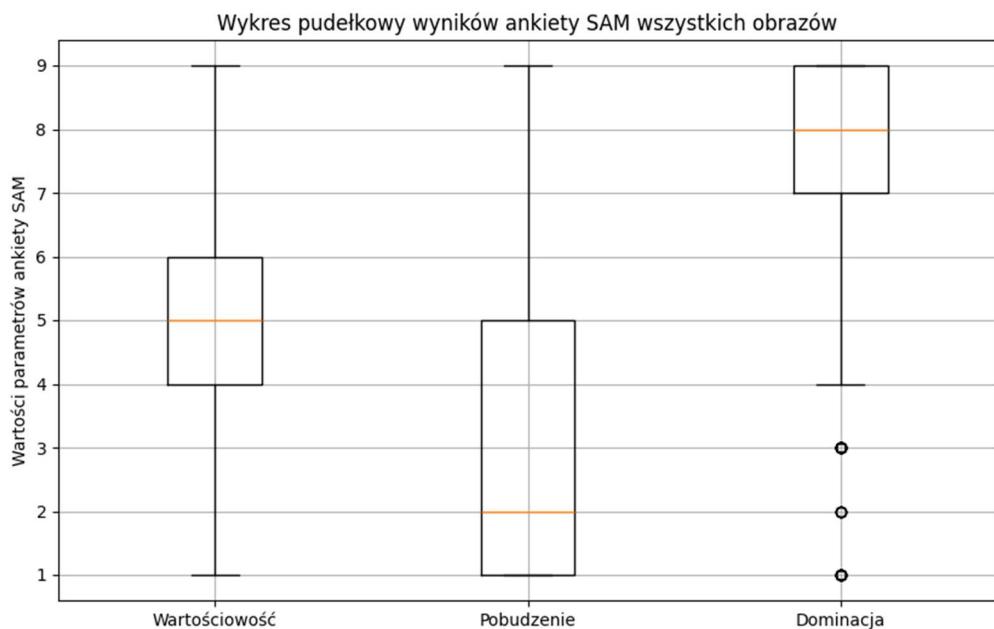


Rys. 6.3. Wykres pudełkowy dominacji dla wszystkich obrazów. Mediana jest zaznaczona kolorem pomarańczowym. Kwartyl dolny i górny ustalają granice kwadratu. Pionowe linie prowadzą do minimalnej i maksymalnej wartości zbioru. Wartości odstające oznaczone są za pomocą kół.

Do podobnych wniosków można dojść patrząc na Rys. 6.4, gdzie wyniki dla wszystkich obrazków są skumulowane w jednym wykresie. Wyniki wartościowe dla połowy próbki są między wartościami 4 i 6. Oznacza to, że większość próbek przedstawia neutralne emocje.

Połowa próbek przedstawia wartości 1 i 2 dla pobudzenia a  $\frac{3}{4}$  równe lub niższe niż 5. Najbardziej przesunięte wyniki są dla dominacji, gdzie  $\frac{3}{4}$  próbek jest dla wartości większej lub równej 7.

Zbadano liczbę występowania każdej kombinacji parametrów wartościowości, pobudzenia i dominacji, nazywanych dalej wyjątkowymi trójkami. Wykryto, że jest ich 172. Kombinacja {5, 1, 9} pojawiła się aż 119 razy. {5, 2, 9} i {5, 2, 8} pojawiły się 17 razy. Spośród kolejnych wyników w tabeli, tylko czterech nie można określić jako neutralne emocje ({7, 4, 9}, {9, 1, 9}, {7, 1, 9}, {7, 3, 9}).



Rys. 6.4. Wykres pudełkowy, gdzie są skumulowane odpowiedzi dla wszystkich obrazów. Mediana jest zaznaczona kolorem pomarańczowym. Kwartył dolny i górny ustalają granice kwadratu. Pionowe linie prowadzą do minimalnej i maksymalnej wartości zbioru. Wartości odstające oznaczone są za pomocą kół.

Tabela 6.1. Dziesięć najczęściej występujących wyników ankiety SAM. Kolumna „W, P, D” oznacza kombinację wartościowości, pobudzenia i dominacji.

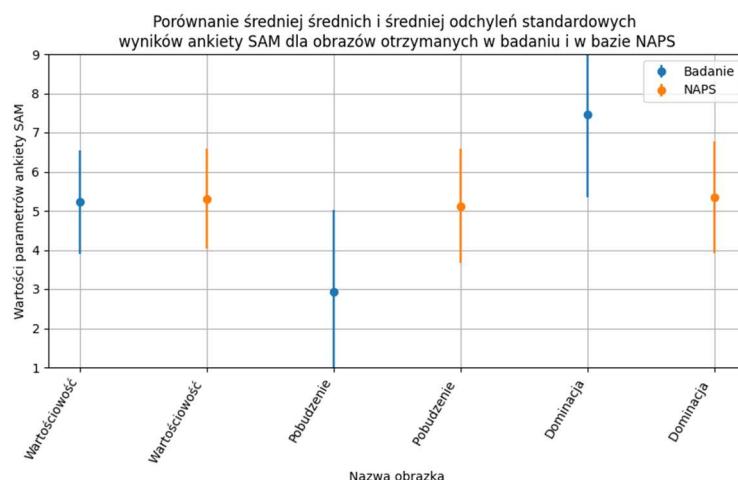
Ranking popularności	W, P, D	Liczba wystąpień
1	5, 1, 9	119
2	5, 2, 9	17
3	5, 2, 8	17
4	6, 1, 9	11
5	4, 2, 9	8
6	7, 4, 9	7
7	9, 1, 9	6
8	7, 1, 9	5
9	7, 3, 9	5
10	5, 3, 8	4

### 6.1.2. Porównanie wyników ankiety SAM zebranych w badaniu ze zbiorem NAPS

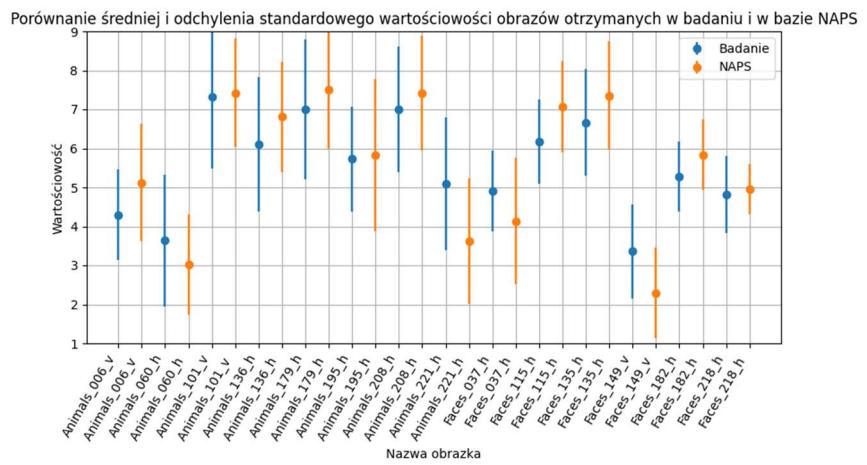
Porównano średnie wyniki wartościowości, pobudzenia i dominacji oraz ich odchylenia standardowe dla wszystkich obrazków z danymi pochodzącyymi ze zbioru NAPS (Rys. 6.5. - Rys. 6.14.). Średnia wartościowość obu zbiorów była podobna. Dla większości obrazków różnica nie przekraczała 1. Odchylenie standardowe dla obu zbiorów było relatywnie małe, jego średnia dla danych w badaniu wynosiła 1,32 a dla zbioru NAPS 1,27.

Wyniki pobudzenia różniły się drastycznie. Tylko w ośmiu przypadkach średnia była podobna, w pozostałych punkty miały znaczącą odległość lub wręcz były po dwóch przeciwnych stronach wykresu. Większość średnich miała mniejsze wartości w badaniu, niż w bazie NAPS. Odchylenia standardowe również były duże. Ich średnia wyniosła 2,11 w porównaniu do 1,46 dla zbioru NAPS.

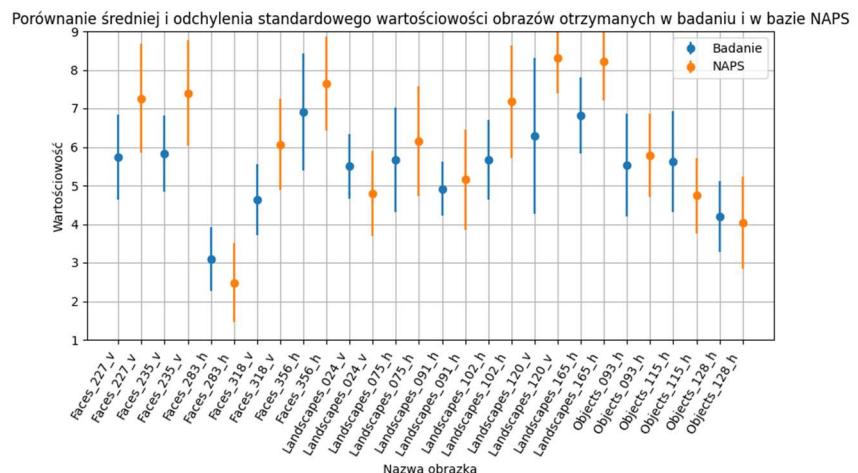
Analogicznie, jak w przypadku pobudzenia, prezentują się wyniki dominacji. Różnice między średnimi są duże, chociaż zdają się przypadki, gdzie osiągają podobne wartości. Średnie zwykle są większe, tak samo jak odchylenia standardowe, których średnia to 2,12 w badaniu i 1,42 w bazie NAPS.



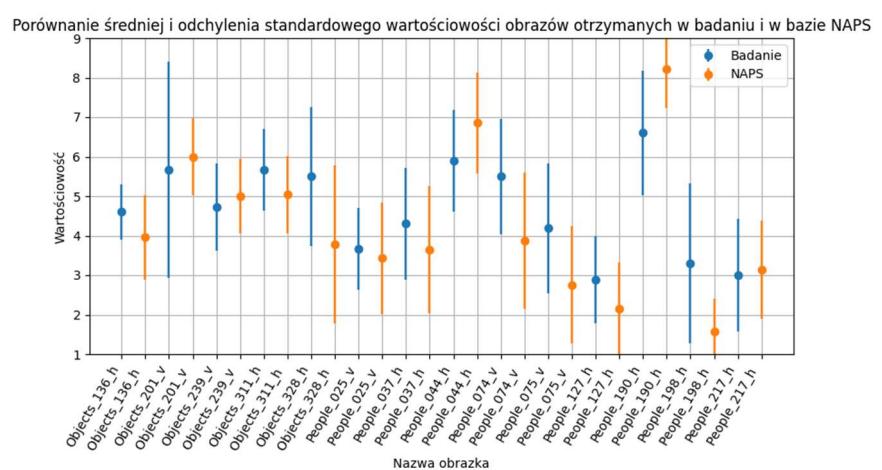
Rys. 6.5. Porównanie średniej średnich i średniej odchyleń standardowych wyników ankiety SAM zebranych w badaniu i w bazie NAPS dla wszystkich obrazków.



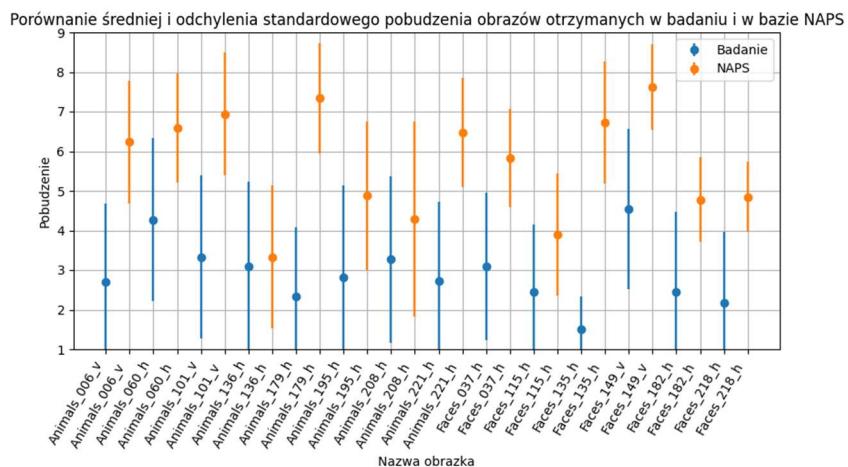
Rys. 6.6. Porównanie średniej i odchylenia standardowego wartościowej zebranych w badaniu i w bazie NAPS dla 14 obrazów.



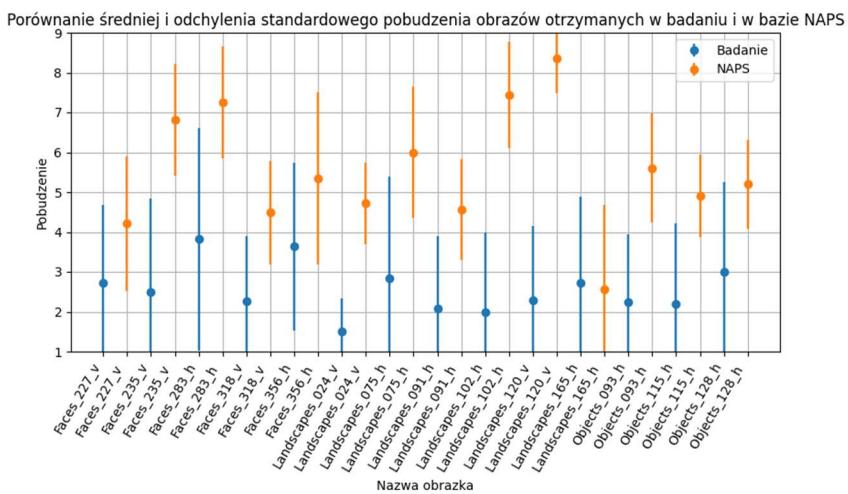
Rys. 6.7. Porównanie średniej i odchylenia standardowego wartościowej zebranych w badaniu i w bazie NAPS dla 14 obrazów.



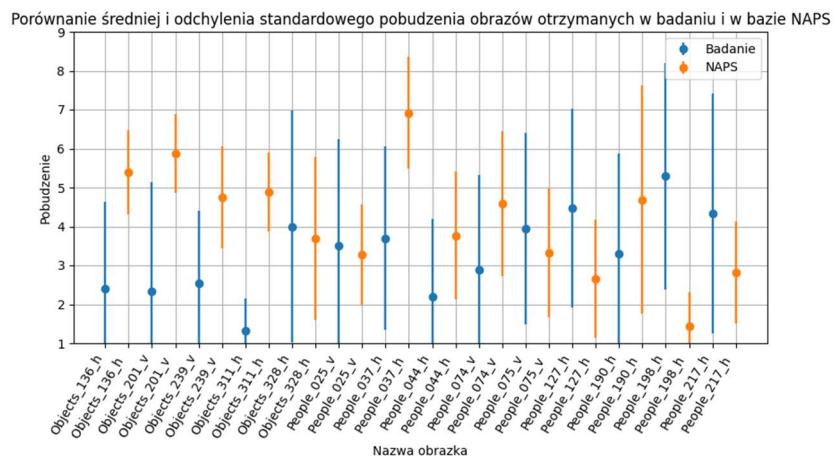
Rys. 6.8. Porównanie średniej i odchylenia standardowego wartościowej zebranych w badaniu i w bazie NAPS dla 14 obrazów.



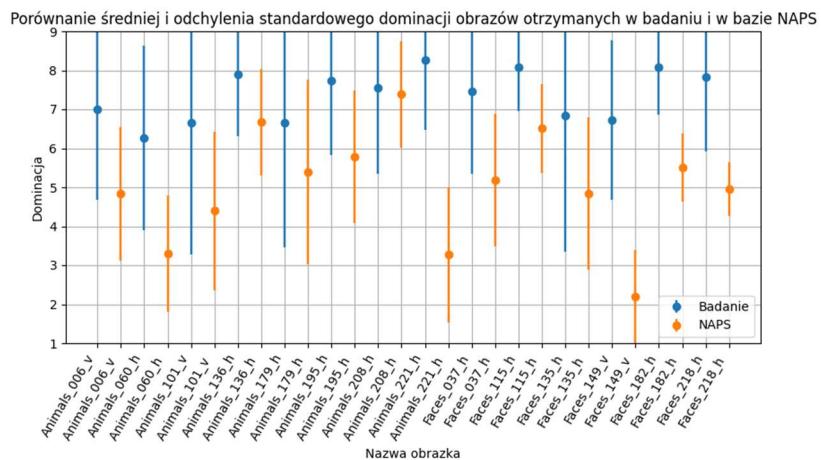
Rys. 6.9. Porównanie średniej i odchylenia standardowego pobudzenia zebranych w badaniu i w bazie NAPS dla 14 obrazów.



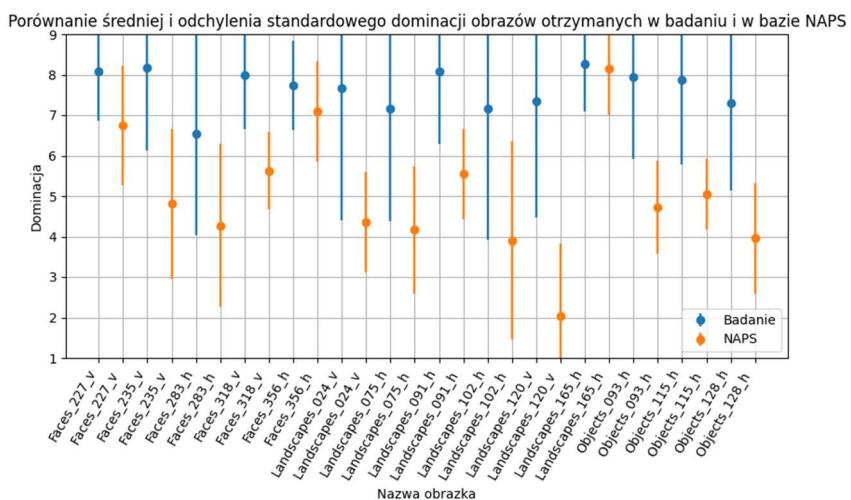
Rys. 6.10. Porównanie średniej i odchylenia standardowego pobudzenia zebranych w badaniu i w bazie NAPS dla 14 obrazów.



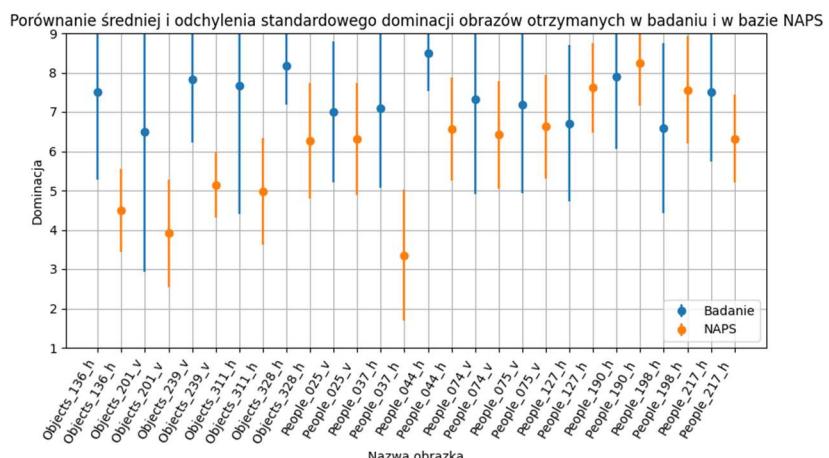
Rys. 6.11. Porównanie średniej i odchylenia standardowego pobudzenia zebranych w badaniu i w bazie NAPS dla 14 obrazów.



Rys. 6.12. Porównanie średniej i odchylenia standardowego dominacji zebranych w badaniu i w bazie NAPS dla 14 obrazów.



Rys. 6.13. Porównanie średniej i odchylenia standardowego dominacji zebranych w badaniu i w bazie NAPS dla 14 obrazów.



Rys. 6.14. Porównanie średniej i odchylenia standardowego dominacji zebranych w badaniu i w bazie NAPS dla 14 obrazów.

### 6.1.3. Analiza słownych emocji odczuwanych przez badanych oraz porównanie ich ze zbiorzem NAPS

Badani w większości wpisywali proste nazwy emocji np. radość, nostalgia, zniesmaczenie. Czasami jednak nie wystarczało to do opisu tego, co czuli i sięgali po sprecyzowanie przyczyny występowania emocji. Zdanie „nienawidzę suszonych grzybów” wprowadza inne wyobrażenie emocji w głowie czytelnika, niż słowo „nienawiść”. Innym przykładem jest sformułowanie „uroczość”. Nie jest to nazwa emocji, jednak daje jasny przekaz uczuć, które wzburza patrzenie się na coś uroczego.

Nie można porównać bezpośrednio emocji wpisywanych przez badanych z grupami emocji w zbiorze NAPS. Można jednak sprawdzić czy wszystkie słowne emocje należą do grupy emocji (Tabela 6.2.).

Dla obrazka *Landscapes\_120\_v* jeden z badanych opisał swoje emocje słowami „nostalgia, szczęście, tęsknota”, kolejny „tęsknota, spokój”. Przedstawiają one spójny obraz stanu emocjonalnego badanych. Nie można go powiązać z jedną grupą emocji, ponieważ jest on między szczęściem i smutkiem. Można to też zauważyć dla innych obrazków, np. do obrazka *People\_198\_h* jeden badany napisał „wstręt, ciekawość, rozbawienie”. Sugeruje to, że badany ma czarny humor. Jego emocje można by było przypisać do odrazy, szczęścia i zdziwienia.

Większość obrazków wywoływała różne emocje u różnych badanych. Na przykład dla obrazka *Animals\_060\_h* jeden badany napisał „spokój” a drugi „trwoga, stres, zaniepokojenie”, mimo że w zbiorze NAPS nazwą grupy emocji był smutek.

To jakie emocje wywołuje obrazek, zależy od indywidualnych skojarzeń i doświadczeń. Obrazek *People\_074\_v* przedstawiający żołnierza u jednego badanego wzburzał „spokój, bezpieczeństwo”, u drugiego „zaciekawienie”, u kolejnego „zdenerwowanie przez ryzyko poboru”.

Większość obrazków w przynajmniej jednym badanym wywoływała neutralne emocje. Tylko dwa obrazki nie zawierały odpowiedzi ze zbioru {spokój, nic, brak, neutralna emocja}. Łącznie te emocje pojawiły się w odpowiedziach 212 razy. Oprócz nich najbardziej licznymi odpowiedziami były radość, obrzydzenie, nostalgia, smutek, ciekawość, niepokój (Tabela 6.3.). Łącznie w badaniu pojawiło się 99 różnych nazw emocji.

Tabela 6.2. Porównanie emocji słownych zebranych w badaniu z typem grupy emocji w bazie NAPS. Uczestnik miał możliwość wypisania wielu emocji naraz. Nawiasami kwadratowymi są zaznaczone odpowiedzi różnych badanych a przecinkami różne emocje jednego badanego. Dla jednego obrazka zostały usunięte powtórki tych samych emocji.

Nazwa obrazka	Emocje NAPS	Emocje
Animals_221_h	odraza	[neutralna emocja]; [neutralna emocja, brak]; [radość, przyjemność]; [obrzydzenie]; [nic]; [lekkie rozbawienie]; [obrzydzenie]; [lekkie zdegustowanie]; [spokój];
People_198_h	odraza	[neutralna emocja]; [zaskoczenie]; [złość]; [odczucie głodu]; [radość, obojętność]; [szczęście]; [uroczość]; [raczej neutralne zaskoczenie];

People_037_h	odraza	[neutralna emocja]; [ciekawość, współczucie]; [trwoga, stres, zaniepokojenie]; [brak]; [lekka przykrość]; [współczucie, niezadowolenie]; [obrzydzenie, smutek]; [smutek]; [odraza]; [spokój];
People_025_v	odraza	[neutralna emocja]; [nic]; [podziw]; [ciekawość, lekkie zaniepokojenie]; [nostalgia, tęsknota]; [radość, ciekawość]; [spokój];
Landscapes_024_v	odraza	[neutralna emocja]; [nic]; [brak]; [radość, sympatia]; [lekkie rozbawienie]; [uroczość]; [neutralna emocja, spokój]; [spokój];
Animals_060_h	smutek	[neutralna emocja]; [ciekawość, obojętność]; [nic]; [zaskoczenie, niepokój]; [strach, niepokój]; [brak]; [niepokój]; [dyskomfort]; [zaniepokojenie]; [spokój]; [strach];
Faces_149_v	smutek	[neutralna emocja]; [skupienie]; [zadowolenie]; [obrzydzenie]; [uroczość]; [radość]; [zaciekawienie, nostalgia]; [brak albo delikatne ciepło na sercu];
Faces_283_h	smutek	[neutralna emocja]; [współczucie bólu]; [wstręt, ciekawość, rozbawienie]; [nic]; [brak, delikatne współczucie]; [obrzydzenie, odraza]; [obrzydzenie]; [niepokój]; [odraza, zniesmaczenie]; [spokój];
Animals_006_v	strach	[neutralna emocja]; [brak]; [zdziwienie]; [spokój]; [szczęście]; [radość]; [entuzjazm]; [obojętność];
People_074_v	strach	[neutralna emocja]; [nic]; [brak, neutralna emocja]; [pozytywna emocja, lekkie rozbawienie]; [ciekawość, zdziwienie]; [radość, sympatia]; [zaciekawienie]; [radość]; [spokój];
Animals_195_h	szczęście	[neutralna emocja]; [podejrliwość, niechęć, ciekawość]; [nic]; [zdziwienie, zaniepokojenie]; [współczucie, smutek]; [brak]; [zazdrość]; [spokój];
People_190_h	szczęście	[neutralna emocja]; [nic]; [brak]; [radość, spokój]; [nostalgia]; [spokój];
People_044_h	szczęście	[neutralna emocja]; [nic]; [brak]; [radość, ciekawość, współczucie]; [smutek]; [obrzydzenie]; [spokój]; [strach, obrzydzenie];
Animals_208_h	szczęście	[neutralna emocja]; [szczęście, zmartwienie]; [nic]; [brak]; [radość, sympatia]; [zadowolenie]; [spokój]; [wesołość];
Faces_227_v	szczęście	[neutralna emocja]; [zaskoczenie]; [nic]; [brak sympathii]; [brak]; [nostalgia, szczęście, spokój]; [obojętność, znudzenie]; [niepokój];
Faces_356_h	szczęście	[neutralna emocja]; [nic]; [współczucie, obrzydzenie, przerażenie]; [spokój]; [obrzydzenie]; [brak, lekki smutek]; [obrzydzenie, zaskoczenie]; [ciekawość, wstręt, współczucie]; [współczucie]; [litość, lekkie przerażenie, zdziwienie];
Faces_318_v	szczęście	[neutralna emocja]; [nostalgia, szczęście]; [brak]; [smutek, współczucie, rozbawienie]; [niepokój]; [irytagia]; [spokój];

Faces 182 h	szczęście	[neutralna emocja]; [nic]; [szczęście, odkrywczość, podniecenie]; [ekscytacja]; [smutek]; [radość, spokój]; [nostalgia]; [spokój]; [wolność, beztroska, spokój];
Faces 115 h	szczęście	[neutralna emocja]; [strach, zaskoczenie]; [lekki niepokój, irytacja]; [lekka przykrość]; [lekki smutek]; [smutek]; [smutek, współczucie]; [współczucie, smutek]; [spokój]; [lekki niepokój];
Landscapes 165 h	szczęście	[neutralna emocja]; [nic]; [spokój, szczęście]; [tęsknota, spokój]; [nostalgia, szczęście, tęsknota]; [radość, spokój]; [radość]; [spokój]; [nostalgia];
Landscapes 120 v	szczęście	[neutralna emocja]; [nic]; [zdziwienie, dezorientacja]; [ciekawość]; [brak]; [niepokój]; [zdziwienie, neutralna emocja]; [spokój];
Animals 136 h	szczęście	[neutralna emocja]; [nic]; [rozbawienie]; [brak]; [tęsknota]; [niepokój]; [spokój];
Animals 179 h	szczęście	[neutralna emocja]; [nic]; [przytłoczenie, poczucie biedy]; [brak]; [obrzydzenie]; [rozczarowanie]; [spokój];
Faces 235 v	szczęście	[neutralna emocja]; [zaskoczenie]; [niepewność]; [neutralna emocja, ignorancja]; [brak]; [lekkie zobojętnienie]; [niepokój]; [irytacja]; [obojętność];
Faces 135 h	szczęście	[neutralna emocja]; [zaskoczenie]; [nic]; [śmiech]; [złość]; [rozbawienie absurdem, lekkie ziryutowanie idiotyzmem]; [spokój]; [obrzydzenie]; [zażenowanie]; [zażenowanie]; [rozbawienie, rozczarowanie];
Landscapes 102 h	szczęście	[neutralna emocja]; [złość, irytacja]; [opanowanie]; [zdezorientowanie, ciekawość]; [brak]; [obrzydzenie]; [smutek]; [zaciekawienie]; [zachwyt]; [irytacja]; [nostalgia]; [obojętność];
Landscapes 075 h	szczęście	[neutralna emocja]; [skupienie]; [ciekawość, złość]; [zażenowanie głupotą, ciekawość]; [złość, smutek]; [strach, zażenowanie]; [irytacja]; [lekki niepokój]; [niepokój]; [obrzydzenie, zaniepokojenie]; [wrogość]; [zażenowanie]; [przerażenie]; [spokój]; [strach];
Objects 201 v	szczęście	[neutralna emocja]; [spokój, zainteresowanie]; [zdenerwowanie przez ryzyko poboru]; [brak]; [dezorientacja, co to jest miecz czy skarpeta?]; [zaciekawienie]; [niepokój]; [dyskomfort]; [neutralna, nic]; [zaniepokojenie, niepokój]; [spokój, bezpieczeństwo]; [spokój];
Animals 101 v	szczęście	[neutralna emocja]; [nic]; [neutralna emocja, smutek, niepewność]; [nostalgia, radość, satysfakcja]; [ciekawość]; [zadowolenie]; [brak]; [spokój]; [niepokój]; [zaciekawienie]; [lekkie zaciekawienie]; [radość]; [obojętność];
Objects 093 h	zdziwienie	[neutralna emocja]; [rozbawienie]; [lekki strach]; [brak]; [uroczość]; [zauroczenie]; [spokój];
Objects 115 h	zdziwienie	[swoboda]; [brak]; [szczęście]; [radość]; [spokój];
People 075 v	złość	[satysfakcja]; [radość]; [spokój]; [neutralna emocja];
People_127 h	złość	[satysfakcja]; [neutralna emocja]; [gód]; [śmiech, zdziwienie]; [obrzydzenie];

People 217 h	odraza, smutek, strach, zdziwienie	[neutralna emocja]; [zadowolenie]; [radość]; [radość, spokój]; [spokój]; [nostalgia];
Faces 218 h	odraza, złość, smutek, strach, szczęście , zdziwienie	[zdziwienie]; [zachwyt]; [spokój]; [brak];
Objects 239 v	odraza, złość, smutek, strach, szczęście , zdziwienie	[neutralna emocja]; [szok]; [brak, lekkie obrzydzenie, nienawidzę suszonych grzybów]; [ciekawość]; [zadowolenie]; [obojętność];
Objects 311 h	odraza, złość, smutek, strach, szczęście , zdziwienie	[strach]; [brak]; [obrzydzenie]; [smutek]; [zaniepokojenie]; [lekki niepokój];
Objects 328 h	odraza, złość, smutek, strach, szczęście, zdziwienie	[neutralna emocja]; [obojętność]; [nostalgia]; [niepokój];
Objects 128 h	odraza, złość, smutek, strach, zdziwienie	[neutralna emocja]; [ból]; [tęsknota]; [nostalgia, radość]; [nostalgia]; [spokój];
Landscapes 091 h	smutek, szczęście, zdziwienie	[neutralna emocja]; [obojętność]; [nostalgia, radość]; [radość]; [spokój];
Objects 136 h	smutek, zdziwienie	[niepokój, lekki stres]; [ból, strach]; [obrzydzenie]; [współczucie, smutek]; [zaniepokojenie]; [wstret];
Faces 037 h	smutek, złość	[szczęście]; [delikatna irytacja]; [radość]; [optimizm];

Tabela 6.3. Emocje, których liczba wystąpień była większa od 4.

Nr	Nazwa emocji	Liczba wystąpień
1	neutralna emocja	108
2	spokój	56
3	radość	31
4	obrzydzenie	28
5	brak	28
6	nic	20
7	nostalgia	16
8	smutek	15
9	ciekawość	15
10	niepokój	14
11	obojętność	11
12	współczucie	11
13	szczęście	10
14	strach	9
15	zaniepokojenie	8
16	zdziwienie	8
17	zaskoczenie	7
18	irytacja	6
19	złość	5
20	tęsknota	5
21	zaciekawienie	5

22	zadowolenie	5
23	rozbawienie	5
24	zażenowanie	5

## 6.2. System rozpoznawania emocji

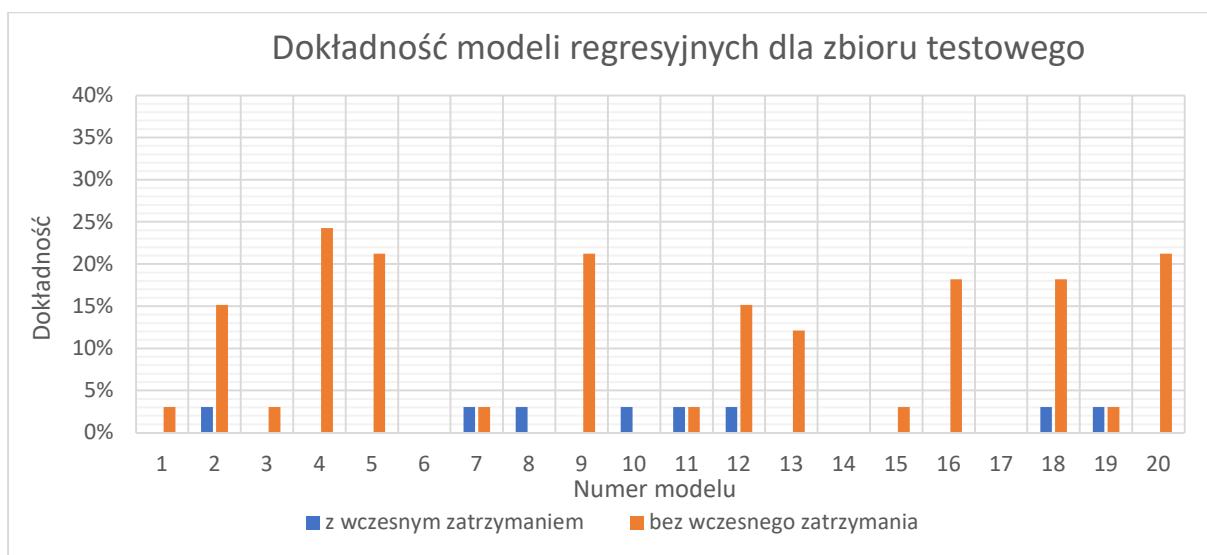
### 6.2.1. Modele regresyjne

W tabeli Tabela 6.4. zostały przedstawione wyniki dla zbioru testowego wyuczonych modeli. Przyjęto pięć kryteriów jakości: dokładność, dokładność przybliżoną, dokładność przewidywania wartościowości, pobudzenia i dominacji oraz liczba wyjątkowych trójk. Przy dokładności modelu i parametrów ankiety SAM próbka była uznawana za poprawnie przewidzianą, jeśli wartość bezwzględna różnicy jej i etykiety była w granicy błędu zaokrąglenia (mniejsza od 0,5). W dokładności przybliżonej uznaje się próbkę za dobrą, jeśli suma wartości bezwzględnych różnic jej i etykiety była mniejsza od 1. Liczba wyjątkowych trójk to jest liczba wyjątkowych kombinacji wartości wartościowości, pobudzenia i dominacji. Spośród wyszkolonych modeli wyróżnia się model 4, który ma największą dokładność, mając jednocześnie 4 unikalne trójki.

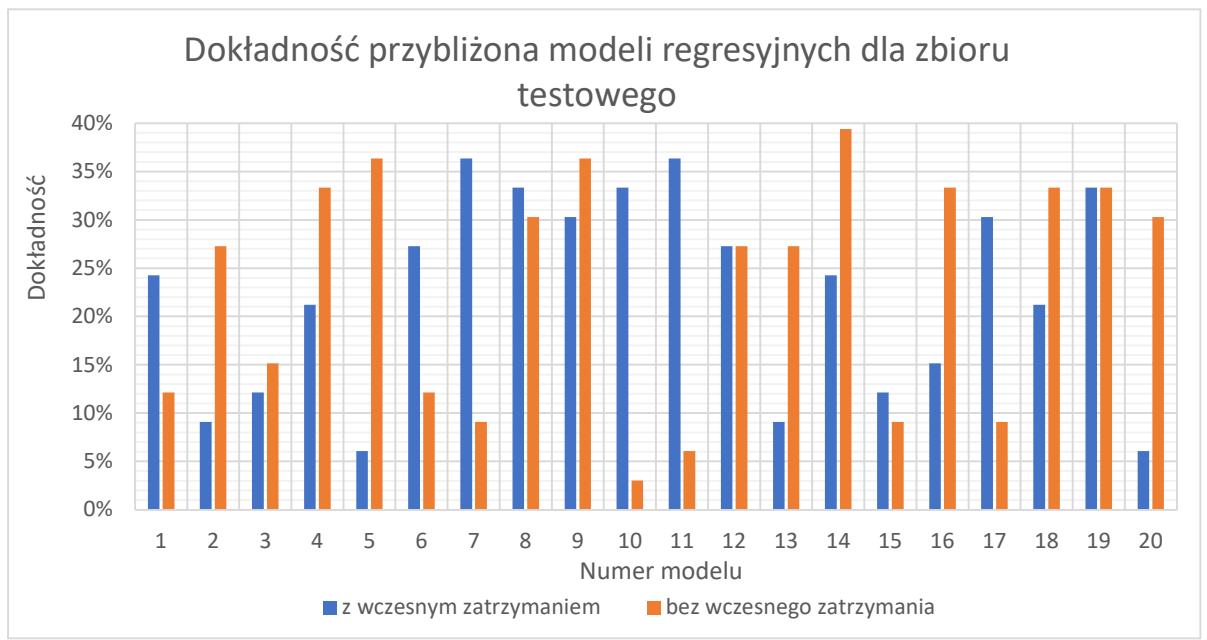
Tabela 6.4. Parametry wyuczonych modeli regresyjnych dla zbioru testowego: dokładności, dokładności przybliżonej, dokładności przewidywania wartościowości, pobudzenia i dominacji oraz liczba wyjątkowych trójk. Kolumny „z” oznaczają trenowanie z wcześnieym zatrzymywaniem. Kolumny „bez” oznaczają modele trenowane bez wcześniego zatrzymywania.

Nr	Dokładność		Otoczenie dokładności		Dokładność W		Dokładność P		Dokładność D		Liczba trójk	
	z	bez	z	bez	z	bez	z	bez	z	bez	z	bez
1	0%	3%	24%	12%	18%	21%	9%	6%	6%	0%	9	8
2	3%	15%	9%	27%	30%	33%	12%	18%	15%	21%	11	7
3	0%	3%	12%	15%	18%	33%	9%	12%	3%	3%	6	7
4	0%	24%	21%	33%	27%	33%	9%	6%	0%	30%	4	6
5	0%	21%	6%	36%	21%	33%	9%	18%	3%	30%	7	4
6	0%	0%	27%	12%	33%	33%	6%	3%	0%	0%	3	5
7	3%	3%	36%	9%	33%	33%	12%	9%	0%	6%	1	2
8	3%	0%	33%	30%	33%	33%	12%	12%	0%	0%	1	1
9	0%	21%	30%	36%	33%	33%	9%	12%	0%	39%	1	2
10	3%	0%	33%	3%	33%	33%	12%	9%	0%	0%	2	3
11	3%	3%	36%	6%	33%	6%	12%	9%	0%	12%	1	1
12	3%	15%	27%	27%	33%	30%	12%	21%	3%	36%	5	8
13	0%	12%	9%	27%	27%	33%	6%	15%	0%	21%	7	10
14	0%	0%	24%	39%	33%	33%	9%	12%	3%	9%	5	3
15	0%	3%	12%	9%	33%	33%	9%	9%	0%	12%	5	2
16	0%	18%	15%	33%	33%	33%	12%	18%	3%	21%	5	4
17	0%	0%	30%	9%	33%	24%	12%	12%	6%	3%	4	4
18	3%	18%	21%	33%	33%	30%	12%	21%	0%	33%	4	8
19	3%	3%	33%	33%	33%	33%	15%	6%	0%	15%	2	3
20	0%	21%	6%	30%	33%	33%	6%	18%	3%	36%	4	9

Jak przedstawiono na rysunku Rys. 6.15. największą dokładność miał model 4 bez wczesnego zatrzymania. Można zauważyć, że wczesne zatrzymanie powodowało znaczący spadek dokładności modelu, która nie przekraczała poziomu 5%. Model 14 w obu przypadkach uzyskał dokładność równą 0%. Kontrastuje to z danymi przedstawionymi na wykresie Rys. 6.16, gdzie przedstawiono dokładność przybliżoną modeli regresyjnych dla tego samego zbioru. Tam model 14 uzyskał najwyższą wartość tego parametru, co wskazuje, że choć dokładnie nie potrafił przewidywać emocji, to najlepiej wykrywał jej otoczenie. Nie widać zależności pomiędzy rozpoznawaniem otoczenia emocji a wcześniejszym zatrzymywaniem uczenia. Modele w miarę dobrze radziły sobie z jego predykcją, uzyskując wyniki ok. 25-40%

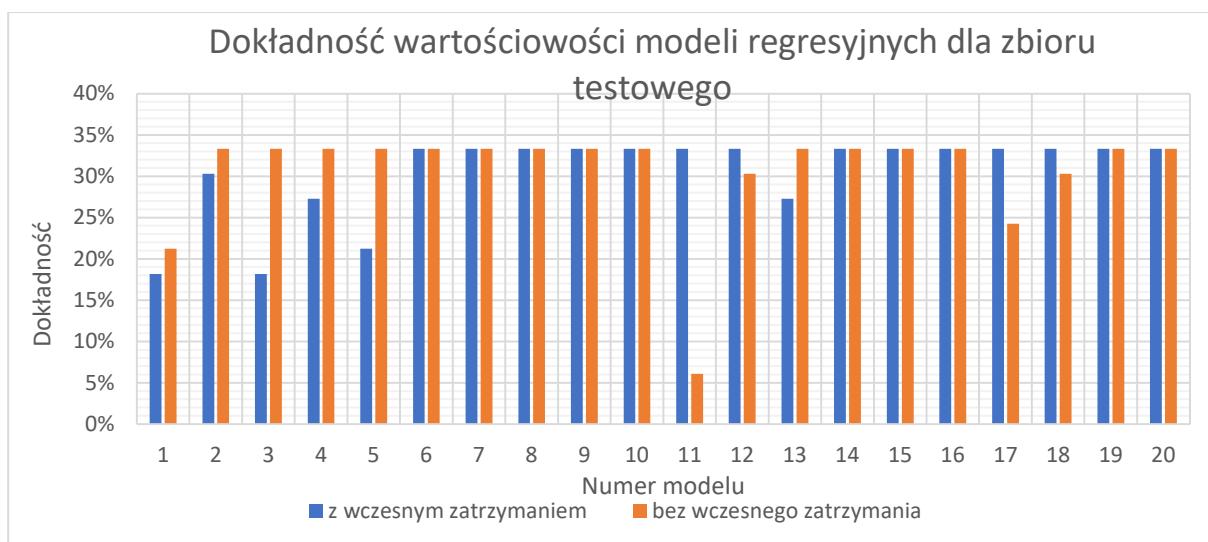


Rys. 6.15. Dokładność dla zbioru testowego badanych modeli regresyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

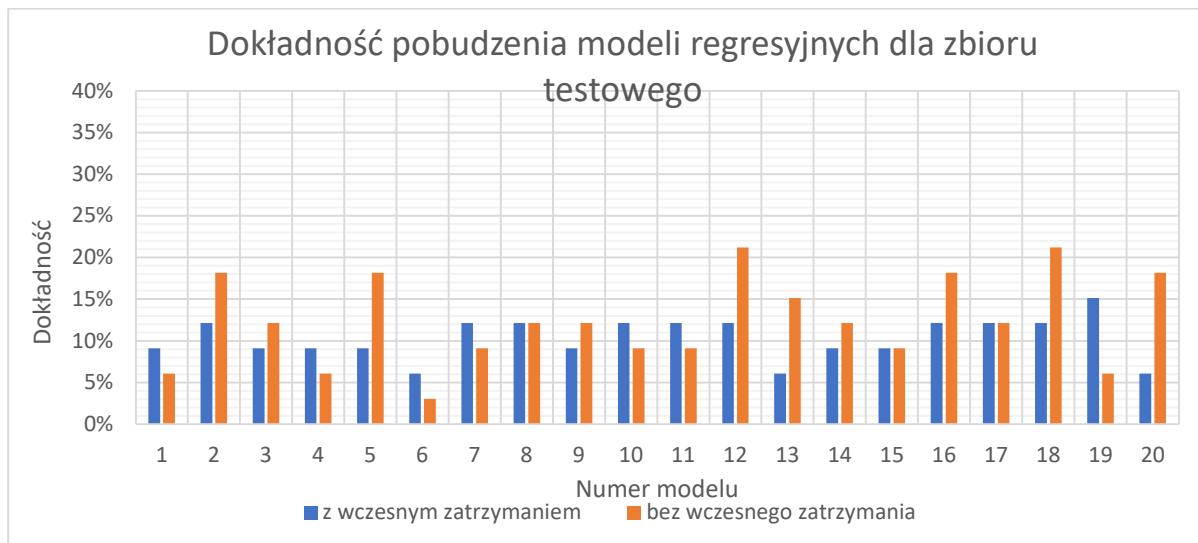


Rys. 6.16. Dokładność przybliżona dla zbioru testowego badanych modeli regresyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

Jak przedstawiono na rysunku Rys. 6.17. dokładność wartościowości modeli regresyjnych była bardzo podobna dla wszystkich modeli i wynosiła ona ok. 33%. Inaczej jednak było w przypadku przewidywania pobudzenia, jak przedstawiono na rysunku Rys. 6.18, gdzie widać było zdecydowanie lepsze wyniki modeli 2, 5, 12, 16, 18 i 20 dla uczenia bez wcześniejszego zatrzymywania. Osiągały one dokładności rzędu 15-22%. Gorzej od nich radziły sobie modele z wcześniejszym zatrzymywaniem, osiągając średnio wyniki na poziomie 10-15%. Najgorzej poradził sobie model 6, w którym w obu przypadkach dokładność była mniejsza niż 7%.

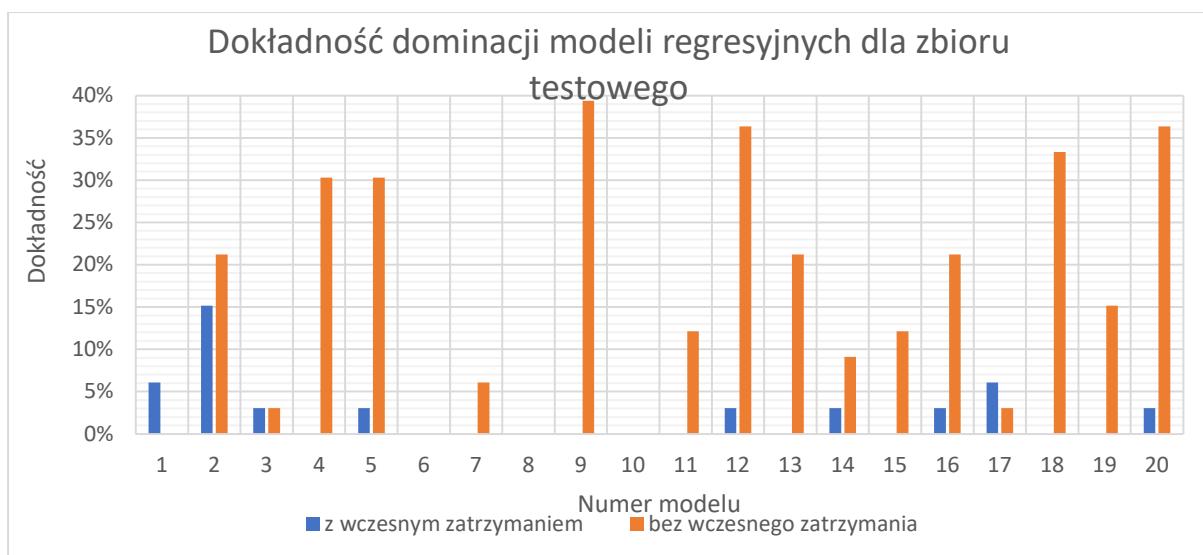


Rys. 6.17. Dokładność wartościowości dla zbioru testowego badanych modeli regresyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

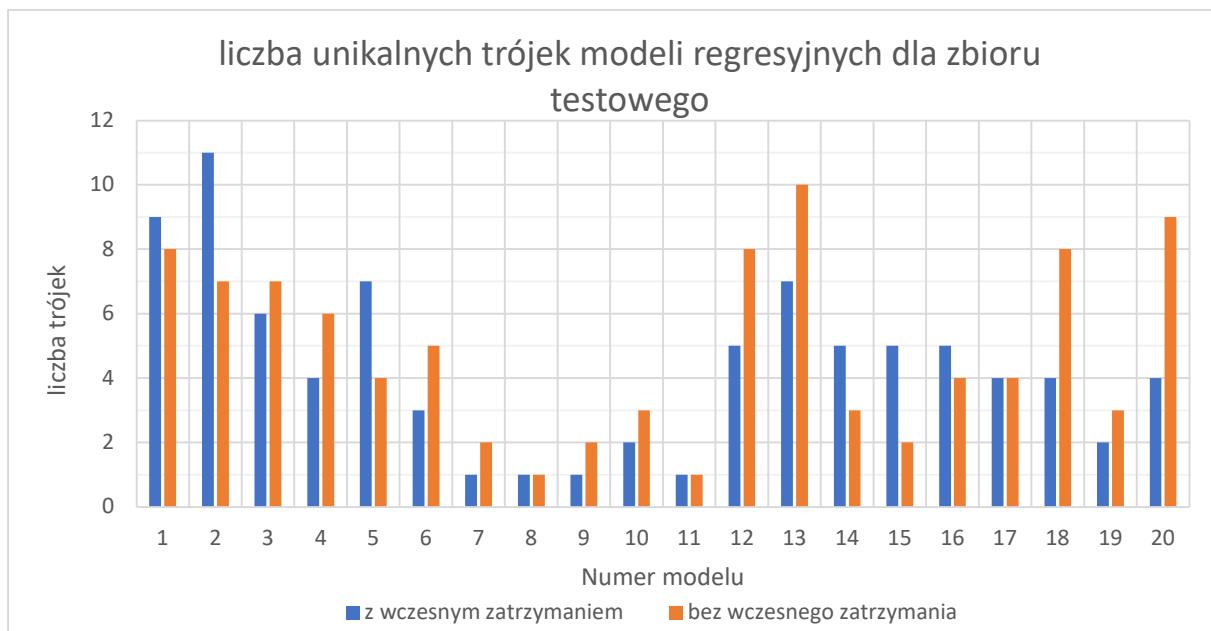


Rys. 6.18. Dokładność pobudzenia dla zbioru testowego badanych modeli regresyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

Z rysunku Rys. 6.19. wynika, że pobudzenie było najtrudniejsze do zaklasyfikowania. Wczesne zatrzymanie powodowało, że modele nie były w stanie go poprawnie zaklasyfikować. Jednak widać bardzo dużą różnorodność dokładności dla modeli bez wczesnego zatrzymania – od 0% do nawet 40% dla modelu 9. Zarówno model 6, 8 jak i 10 w obu przypadkach nie były w stanie w ogóle poprawnie go wykryć. Na rysunku Rys. 6.20. przedstawiono natomiast liczbę wyjątkowych trójk. Można zauważyć, że modele te w obu przypadkach miały porównywalną ich liczbę. Modele 7-11 oraz model 15 i 19 potrafiły poprawnie wykryć ich jedynie od 1 do 3.



Rys. 6.19. Dokładność dominacji dla zbioru testowego badanych modeli regresyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.



Rys. 6.20. Liczba wyjątkowych trójk dla zbioru testowego badanych modeli regresyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

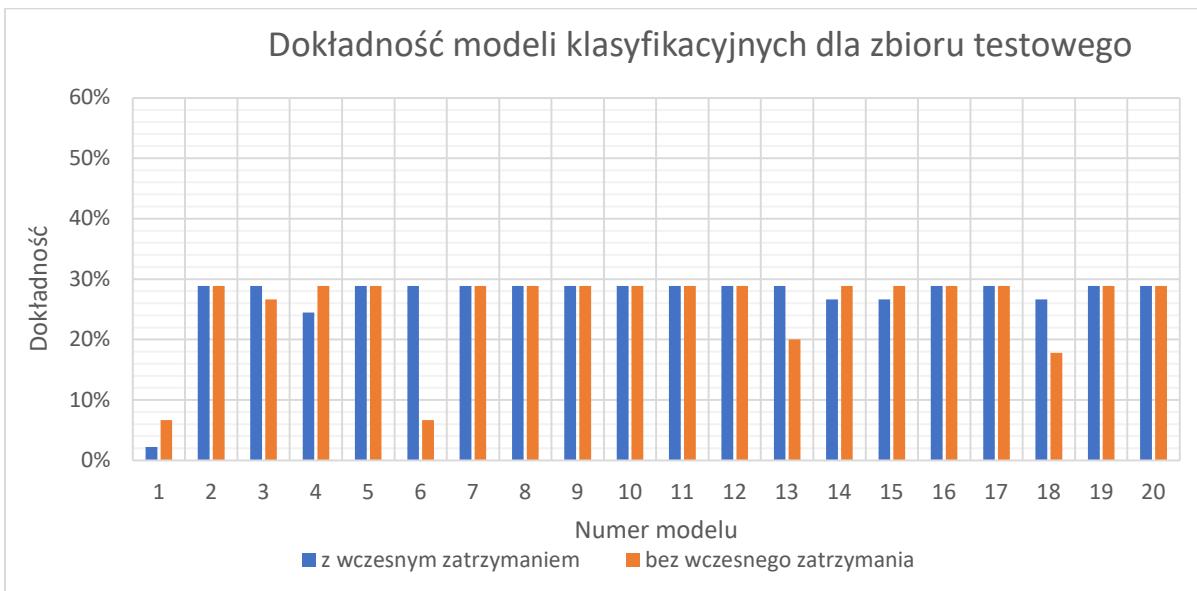
### 6.2.2. Modele klasyfikujące

W Tabeli 6.5. zostały przedstawione wyniki dla zbioru testowego wyuczonych modeli klasyfikujących. Przyjmowano te same kryteria jakości, jak do modeli regresyjnych. Spośród wyszkolonych modeli wyróżnia się model 3, który ma jedną z największych dokładności, mając jednocześnie 6 wyjątkowych trójkę w przypadku trenowania bez wcześniego zatrzymywania.

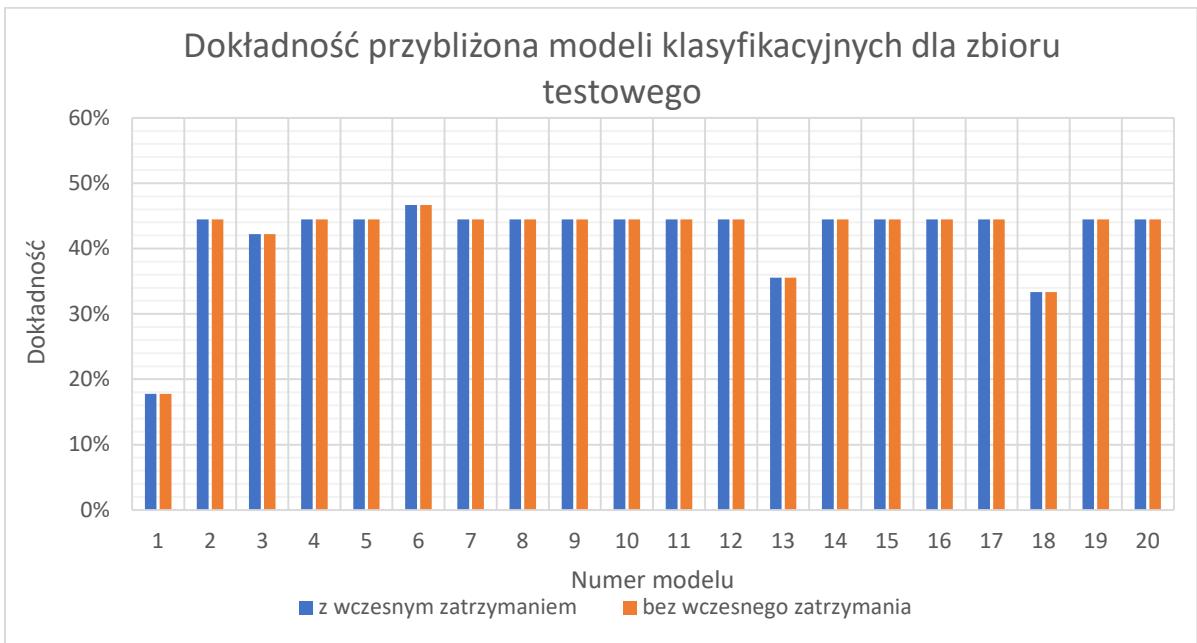
Tabela 6.5. Parametry wyuczonych modeli klasyfikujących dla zbioru testowego: dokładności, dokładności przybliżonej, dokładności przewidywania wartościowości, pobudzenia i dominacji oraz liczba wyjątkowych trójkę. Kolumny „z” oznaczają trenowanie z wcześnieym zatrzymywaniem. Kolumny „bez” oznaczają modele trenowane bez wcześniego zatrzymywania.

Nr	Dokładność		Dokładność przybliżona		Dokładność W		Dokładność P		Dokładność D		Liczba wyjątkowych trójkę	
	bez	z	bez	z	bez	z	bez	z	bez	z	bez	z
1	7%	2%	18%	18%	49%	11%	16%	16%	38%	29%	14	15
2	29%	29%	44%	44%	47%	47%	29%	38%	36%	40%	1	2
3	27%	29%	42%	44%	40%	49%	31%	33%	40%	40%	5	2
4	29%	24%	44%	44%	38%	42%	22%	24%	31%	33%	1	6
5	29%	29%	44%	44%	42%	49%	24%	27%	27%	33%	1	2
6	7%	29%	47%	44%	36%	49%	27%	24%	29%	24%	3	1
7	29%	29%	44%	44%	44%	40%	36%	24%	38%	33%	1	1
8	29%	29%	44%	44%	49%	38%	40%	31%	36%	31%	1	1
9	29%	29%	44%	44%	49%	47%	31%	33%	36%	38%	1	1
10	29%	29%	44%	44%	53%	47%	33%	31%	40%	36%	1	1
11	29%	29%	44%	44%	38%	40%	29%	24%	31%	27%	1	1
12	29%	29%	44%	44%	47%	49%	31%	31%	31%	31%	1	3
13	20%	29%	36%	44%	42%	44%	20%	27%	27%	29%	3	1
14	29%	27%	44%	42%	44%	56%	29%	33%	31%	40%	1	2
15	29%	27%	44%	42%	36%	53%	24%	36%	27%	42%	1	4
16	29%	29%	44%	44%	47%	40%	31%	27%	36%	27%	1	1
17	29%	29%	44%	44%	44%	44%	31%	29%	33%	36%	1	1
18	18%	27%	33%	42%	27%	44%	29%	33%	36%	36%	8	2
19	29%	29%	44%	44%	47%	47%	31%	36%	36%	38%	1	1
20	29%	29%	44%	44%	47%	49%	31%	36%	36%	42%	1	1

Jak przedstawiono na rysunku Rys. 6.21. większość modeli osiągnęła dokładność równą 27-29%. Można zauważyć, że wcześnie zatrzymanie nie powodowało znaczącego spadku dokładności modeli. Model 1 w obu przypadkach uzyskał bardzo niską dokładność mniejszą niż 8%. Podobne wnioski można wysnuć z Rys. 6.22, gdzie przedstawiono dokładność przybliżoną modeli klasyfikujących dla tego samego zbioru. Tam model 1 uzyskał również najniższą wartość tego parametru, co wskazuje na nieadekwatność tego modelu do przedstawionego problemu. Nie widać zależności pomiędzy rozpoznawaniem otoczenia emocji a wcześniejszym zatrzymywaniem uczenia. Modele w miarę dobrze radziły sobie z jego predykcją, uzyskując wyniki ok. 40-45%.

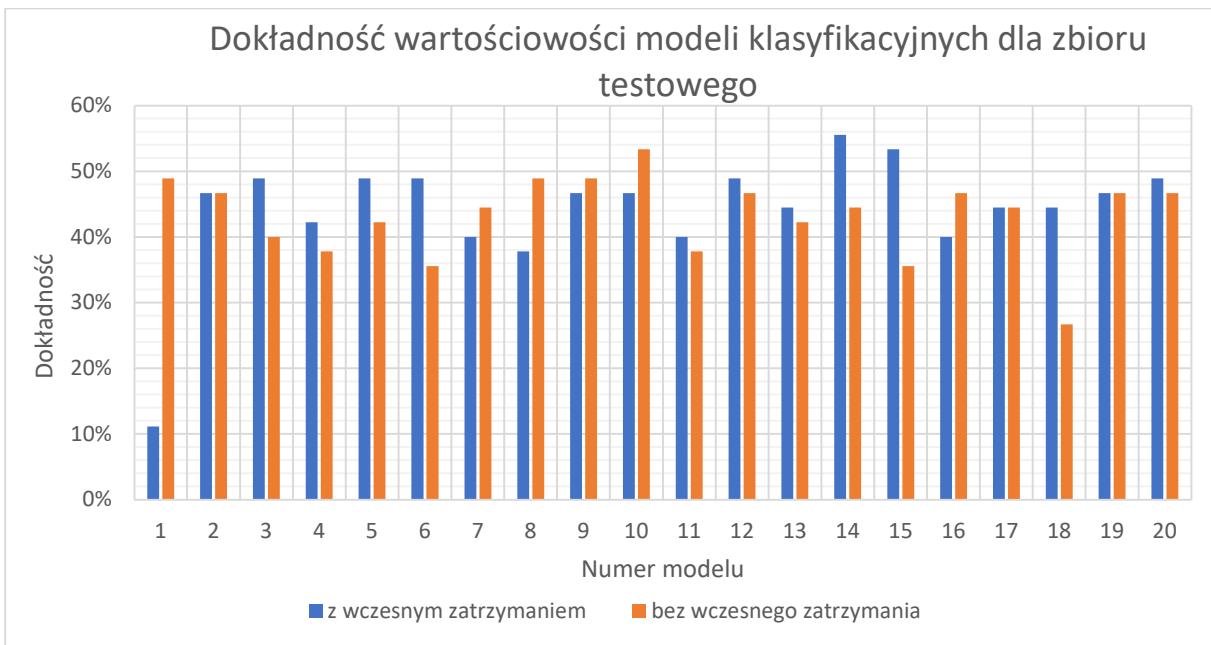


Rys. 6.21. Dokładność dla zbioru testowego badanych modeli klasyfikacyjnych z uwzględnieniem możliwego użycia wcześniego zatrzymania uczenia.

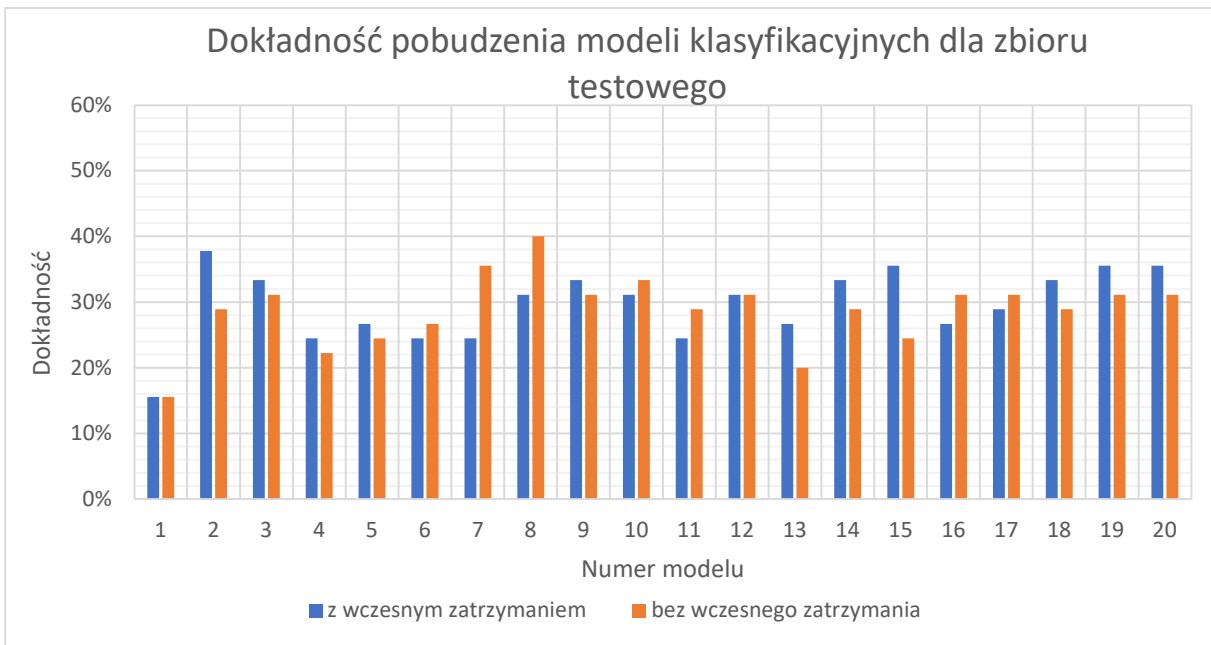


Rys. 6.22. Dokładność przybliżona dla zbioru testowego badanych modeli klasyfikacyjnych z uwzględnieniem możliwego użycia wcześniego zatrzymania uczenia.

Jak przedstawiono na rysunku Rys. 6.23. dokładność parametru W modeli regresyjnych była podobna dla wszystkich modeli i miała ona wartość ok. 40-50%. Ponownie najgorszy okazał się model 1 z wcześnieym zatrzymaniem. Osiągając dokładność ok. 11%. Nie inaczej było w przypadku przewidywania pobudzenia, jak przedstawiono na rysunku Rys. 6.24, gdzie uzyskana dokładność wynosiła ok 20-35%. Wyróżnia się tu jednak model 8, który uzyskał dokładność podczas uczenia bez zatrzymywania na poziomie 40%.

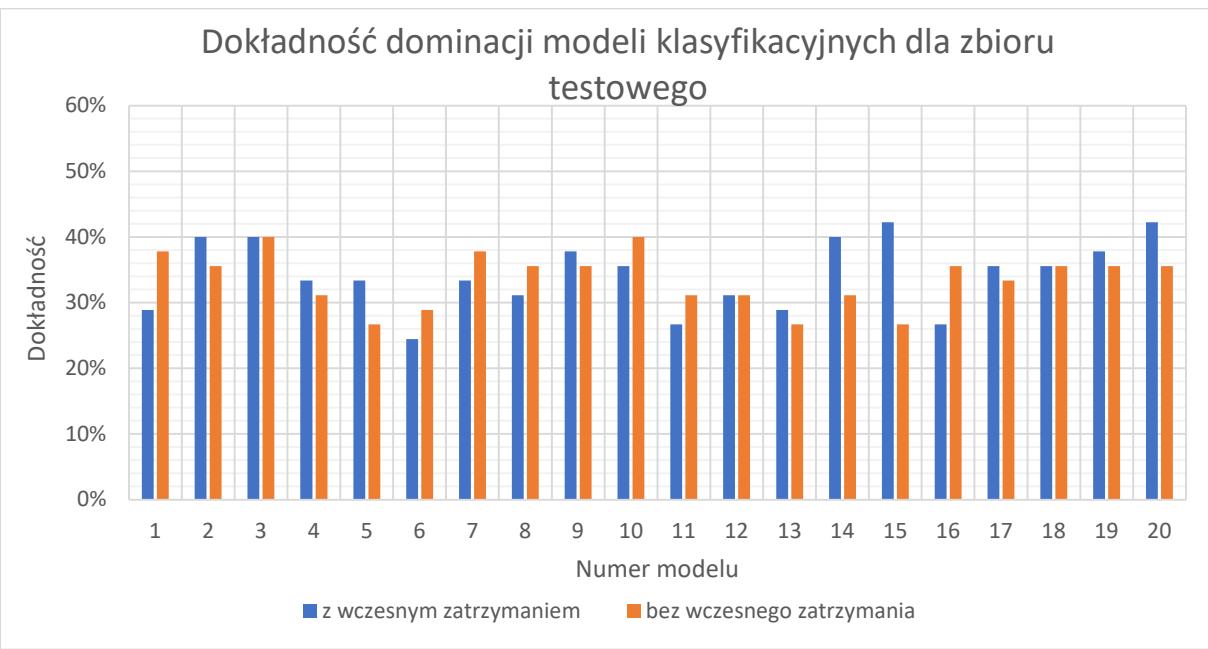


Rys. 6.23. Dokładność wartościowości dla zbioru testowego badanych modeli klasyfikacyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

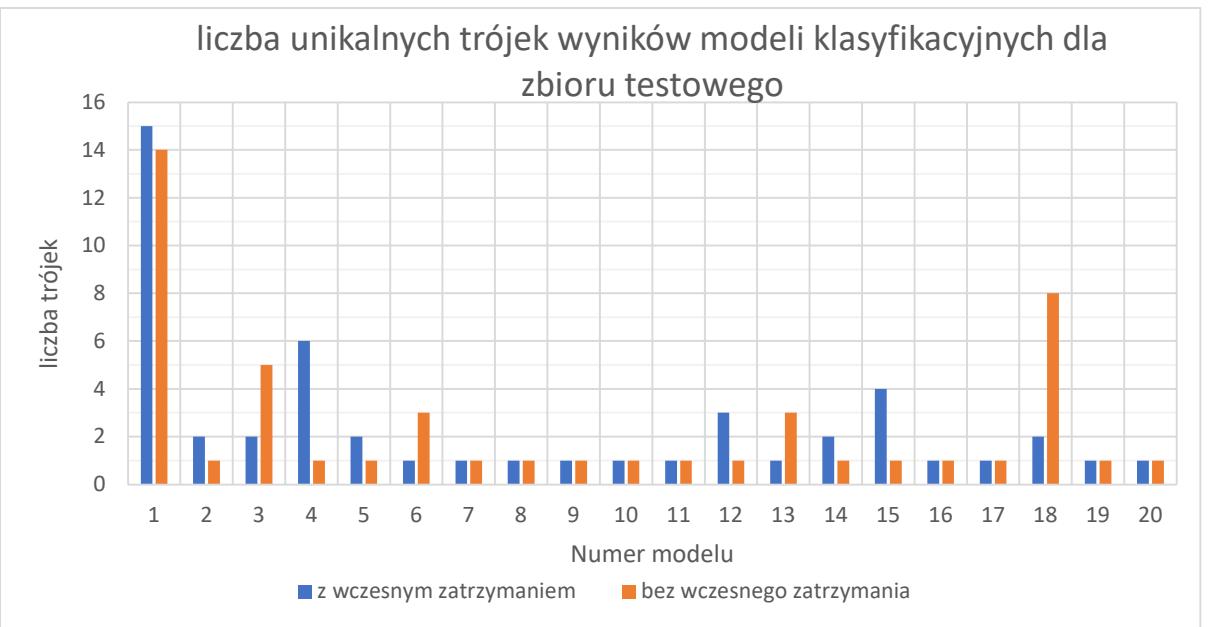


Rys. 6.24. Dokładność pobudzenia dla zbioru testowego badanych modeli klasyfikacyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

Z rysunku Rys. 6.25. wynika, że dominacja, w przeciwieństwie do regresji, nie była najtrudniejsza do zaklasyfikowania. Nie widać wpływu wcześniejszego zatrzymywania na skuteczność klasyfikacji. Najlepszy wynik uzyskały modele 15 i 20, uzyskując 42% poprawnych przewidywań. Na rysunku Rys. 6.26. przedstawiono natomiast liczbę wyjątkowych trójkę. Można zauważać, że modele te w obu przypadkach miały porównywalną ich liczbę. Jednakże bardzo dużo modeli podczas predykcji uzyskiwało od jednego do trzech wyjątkowych wyników. Najwięcej uzyskał model 1 – ok. 14 niepowtarzalnych, lecz skutkowało to obniżeniem dokładności modelu.

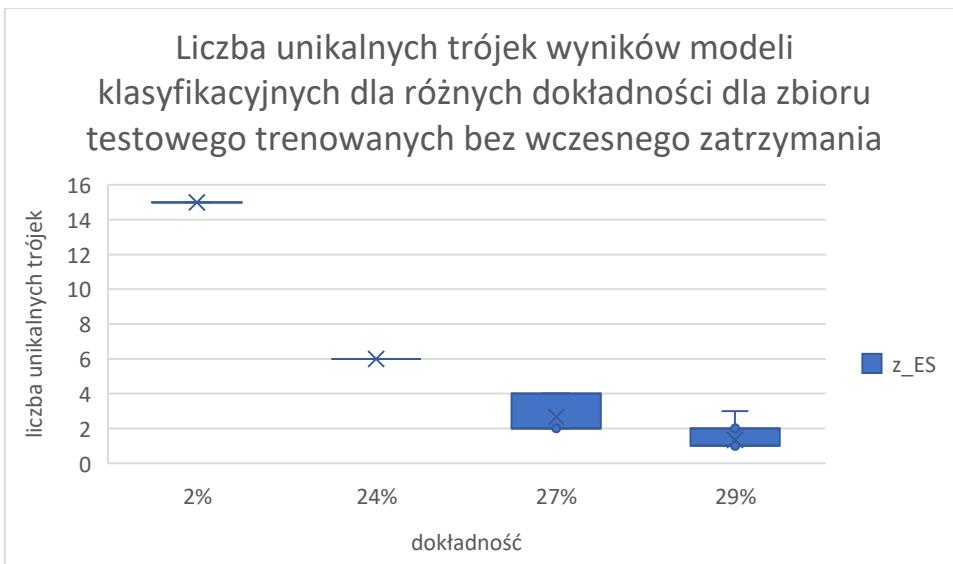


Rys. 6.25. Dokładność dominacji dla zbioru testowego badanych modeli klasyfikacyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.



Rys. 6.26. Liczba wyjątkowych trójk dla zbioru testowego badanych modeli klasyfikacyjnych z uwzględnieniem możliwego użycia wczesnego zatrzymania uczenia.

Jak wynika z rysunku Rys. 6.27. wraz ze wzrostem dokładności modeli klasyfikujących zaobserwowano spadek liczby wyjątkowych trójk. Jest to problem, gdyż dla najwyższej dokładności – 29%, nie uzyskano więcej, niż 3 unikalne trójki.



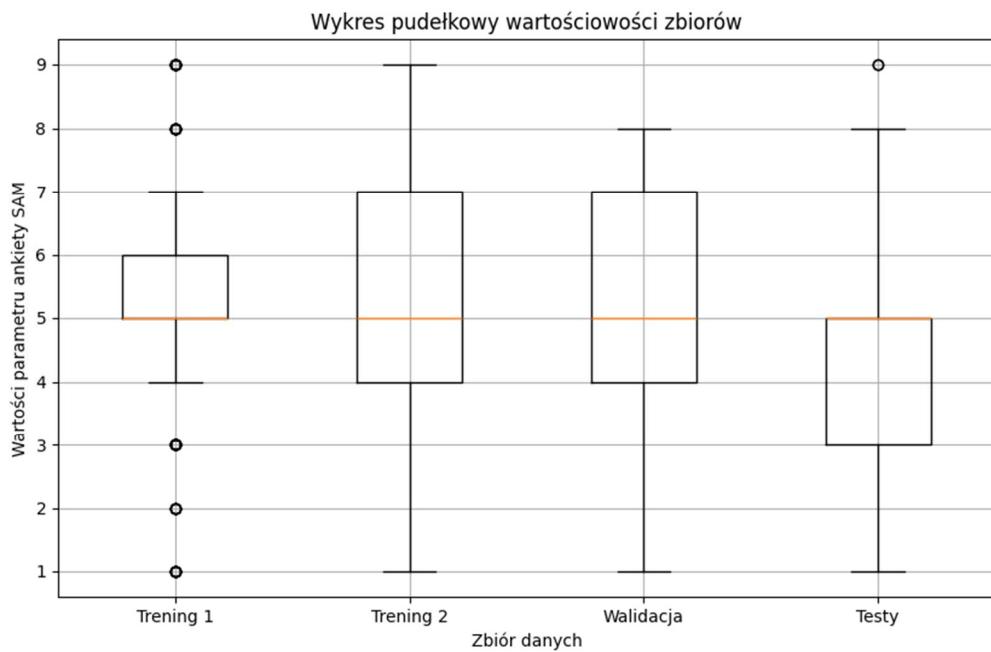
Rys. 6.27. Liczba wyjątkowych trójk wyników modeli klasyfikacyjnych dla różnych dokładności dla zbioru testowego trenowanych bez wczesnego zatrzymania.

### 6.2.3. Zbiory danych

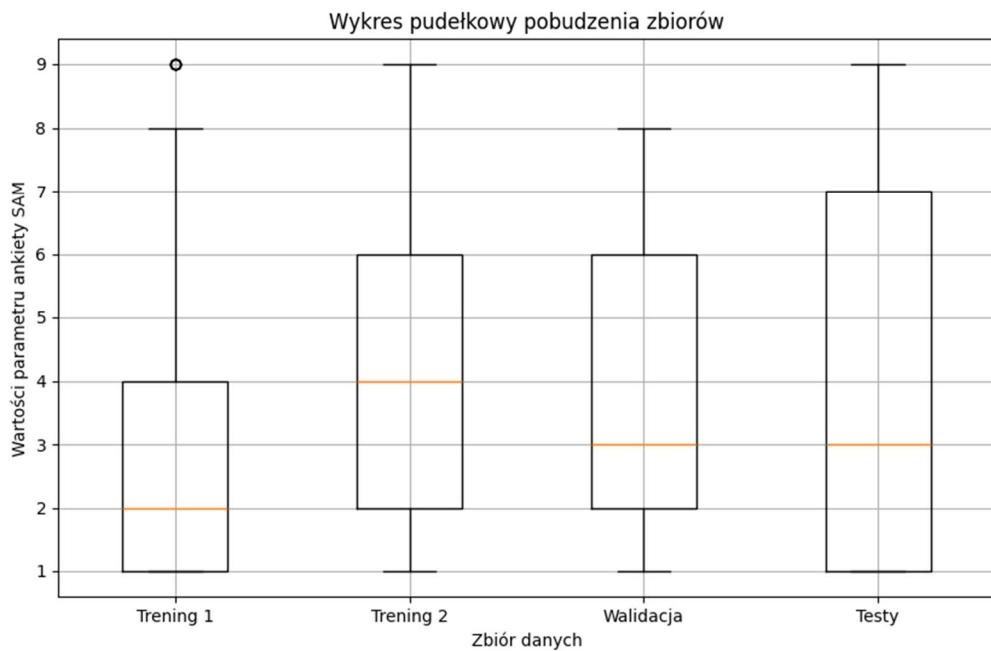
Zbiór treningowy przed powieleniem danych z rzadziej spotykanymi etykietami miał większość próbek rozmiieszczoną na bardzo małym zakresie etykiet (Rys. 6.28.). Najbardziej skumulowany był rozkład wartościowości, w którym 3 kwartyle miały wartość 5 lub 6. Wartość 4 jest oznaczona jako minimalna a 7, jako maksymalna. Pozostałych jest tak niewiele, że są oznaczone jako anomalie. Po powieleniu mediana wciąż znajduje się w punkcie 5, lecz pozostałe wartości pokrywają całą skalę i rozkład jest podobny do rozkładu zbioru walidacyjnego. W zbiorze testowym mediana i trzeci kwartyl mają wartość 5. Pozostałe liczby jednak rozkładają się po reszcie zakresu.

Dla pobudzenia powielenie rzadkich danych znowu poszerzyło rozmiesczenie próbek na wykresie (Rys. 6.29.). Nie ma jednak większego rozkładu, niż rozkład danych testowych. Rozkład zbioru treningowego po przekształceniu i walidacyjnego wyglądają podobnie.

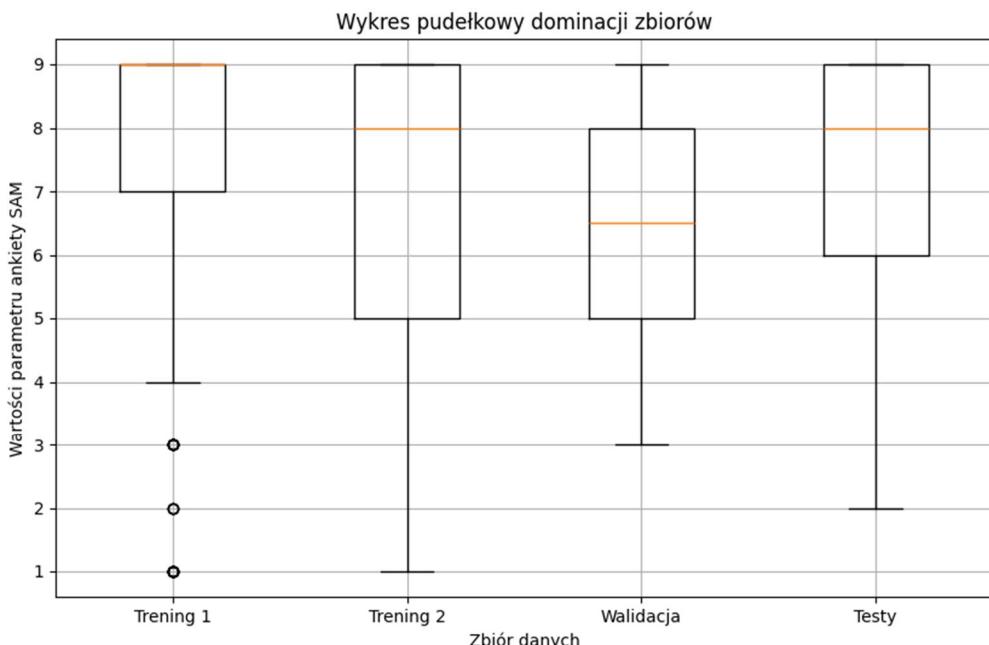
Dla dominacji wartości zbioru testowego są przesunięte w kierunku większych wartości (Rys. 6.30.). Maksymalna wartość, gorny kwartyl i mediana mają wartość 9. Wartość minimalna to 4 a niższe wartości zostały oznaczone jako anomalie. Powielenie danych rzadkich poprawiła ten stan rzeczy i sprawiła, że ten zbiór ma najszerszy rozkład wartości. Rozkład zbioru walidacyjnego ma najmniejszą wartość w punkcie 3 a dla zbioru testowego w punkcie 2.



Rys. 6.28. Wykres pudełkowy wartościowej dla różnych zbiorów danych. „Trening 1” oznacza dane przed powieleniem próbek rzadkich a „Trening 2” po nim.



Rys. 6.29. Wykres pudełkowy pobudzenia dla różnych zbiorów danych. „Trening 1” oznacza dane przed powieleniem próbek rzadkich a „Trening 2” po nim.

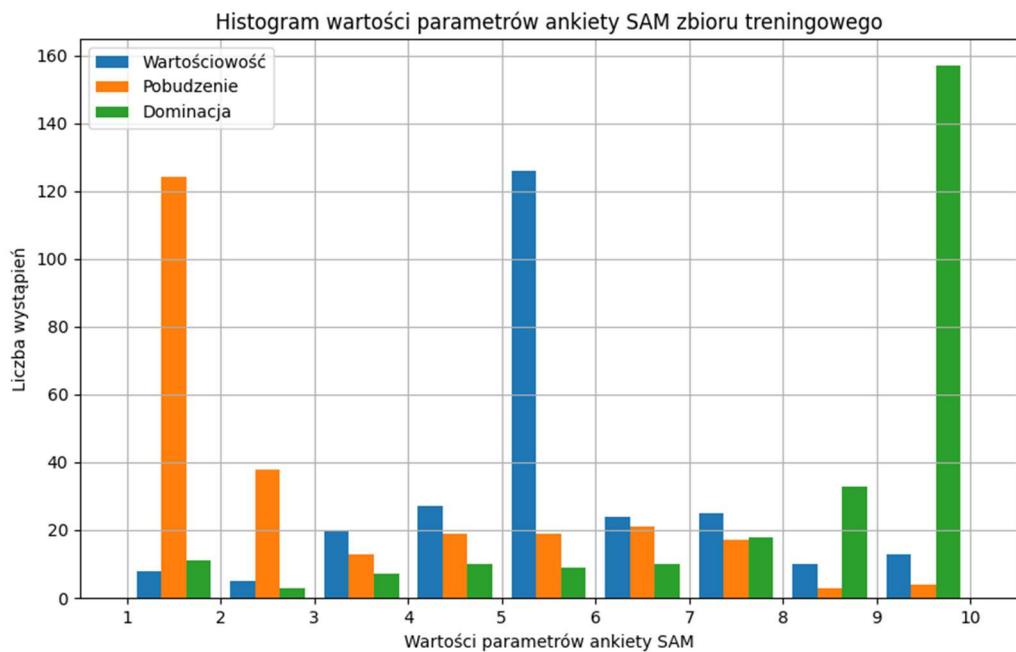


Rys. 6.30. Wykres pudełkowy dominacji dla różnych zbiorów danych. „Trening 1” oznacza dane przed powięleniem próbek rzadkich a „Trening 2” po nim.

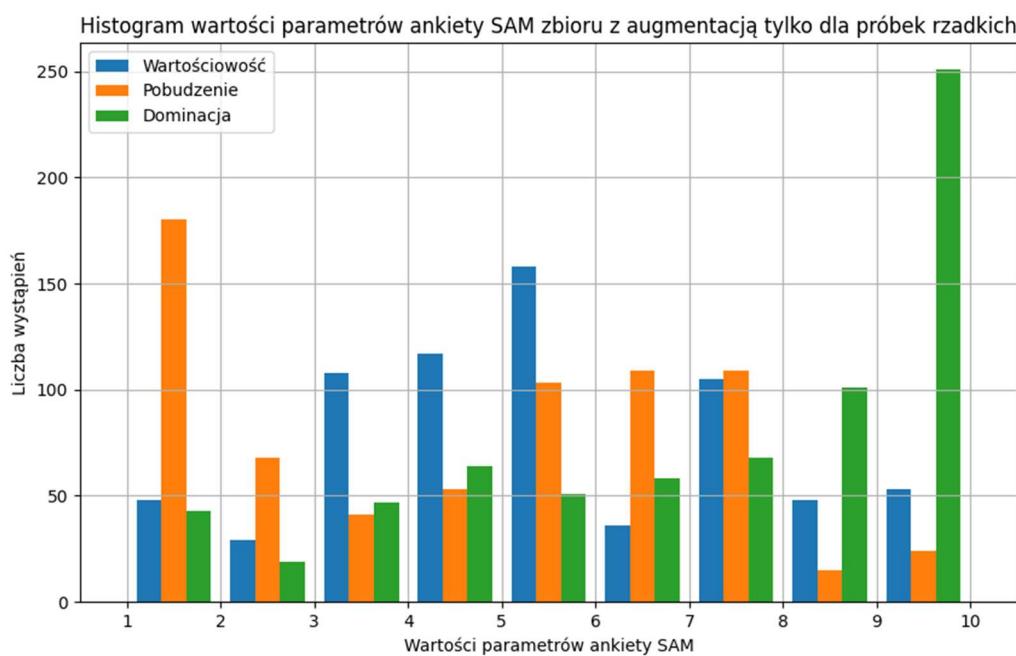
Przyglądając się histogramowi zbioru testowego można zauważyć, że większość próbek jest skumulowana w jednej wartości dla każdego parametru ankiety SAM. Najbardziej ten efekt jest widoczny dla dominacji, gdzie większość klas nie przekracza 10 próbek (Rys. 6.31.). Sztuczne dodanie próbek rzadkich poprawiło występowanie innych klas, ale wciąż się zdarza niska liczba trzech klas (Rys. 6.32.).

Histogram zbioru walidacyjnego nie przypomina histogramu zbioru treningowego. O ile duża liczba wartościowości 5, pobudzenia 1 i dominacji 9 się powtarza, to różnica jest znacznie mniejsza. Powodem tego może być to, że zbiór ten jest 5 razy mniejszy od treningowego. Funkcja rozdzielająca dane skupiła się na tym, żeby pojawiły się w tym zbiorze dane wielu klas. Mimo to brakuje w nim wartości 1 i 2 dominacji oraz wartości 9 wartościowości i pobudzenia (Rys. 6.33.).

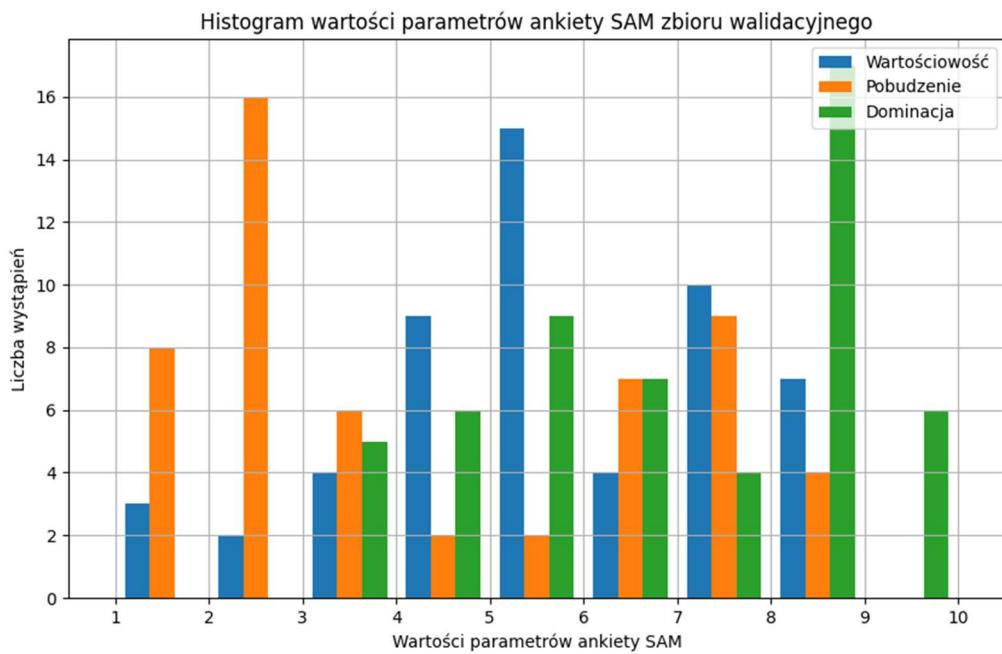
Histogram zbioru testowego bardzo przypomina histogram zbioru treningowego (Rys. 6.34.). Powodem dobrego dopasowania zbioru testowego i złego zbioru walidacyjnego jest to, że podział na zbiór treningowy i testowy odbył się pierwszy. Później zbiór walidacyjny był wydzielony z treningowego, więc miał mniejszą liczbę próbek do wyboru. Podziału dokonano za pomocą funkcji *iterative\_train\_test\_split* z biblioteki *Skmultilearn*.



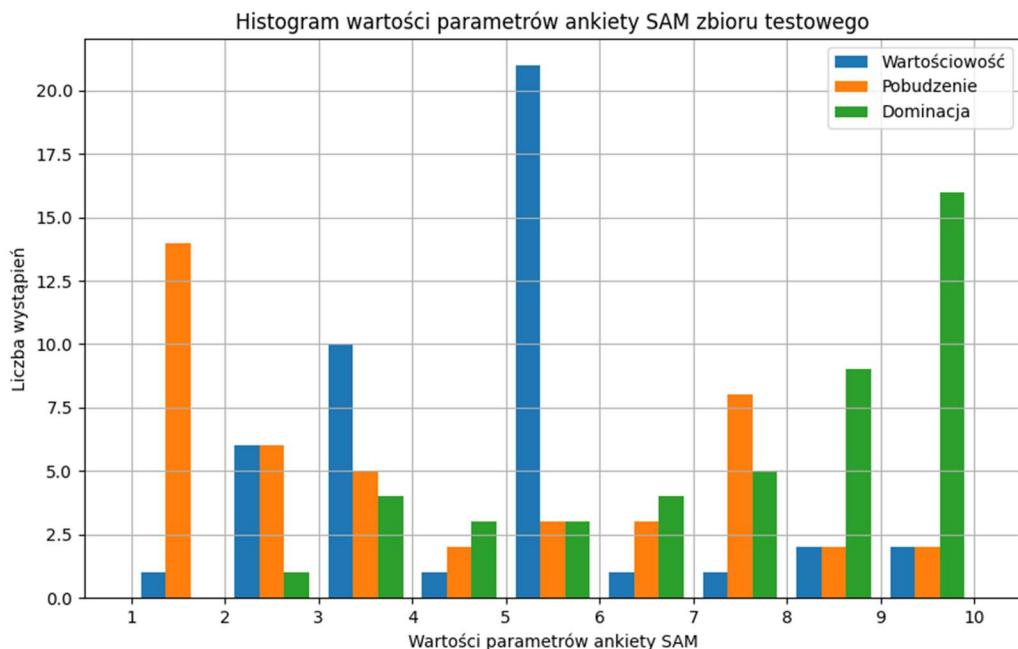
Rys. 6.31. Histogram zbioru treningowego.



Rys. 6.32. Histogram zbioru treningowego po sztucznym zwiększeniu danych rzadkich.



Rys. 6.33. Histogram zbioru walidacyjnego.



Rys. 6.34. Histogram zbioru testowego.

#### 6.2.4. Wyniki modeli wytrenowanych na danych z augmentacją

Wybrano najlepsze modele i wytrenowano je na zbiorze treningowym z augmentacją danych. Wytypowanymi modelami były uczone raz, bez wcześniego zatrzymywania 4. model regresyjny, którego warstwy ukryte to: dense, dense, lstm, dense, dense oraz 3. model klasyfikujący, którego warstwy ukryte to: dense, dense, lstm.

Uczenie modelu regresyjnego na powiększonym zbiorze danych nie przyniosło lepszych rezultatów. Wynik dokładności to 0%, dokładności przybliżonej to 6%, dokładności

wartościowości, pobudzenia i dominacji to odpowiednio 26%, 6% i 6%. Są one dalekie od tych osiąganych na mniejszym zbiorze danych. Liczba wyjątkowych trójk wynosi 10.

Podobnie wygląda sytuacja dla modelu klasyfikującego. Wynik dokładności to 15%, dokładności przybliżonej to 28%, dokładności wartościowości, pobudzenia i dominacji to odpowiednio 44%, 42% i 42%. Liczba wyjątkowych trójk wynosi 11. Liczba wyjątkowych trójk jest duża, ale pozostałe kryteria są porównywalne lub gorsze od tych osiąganych na mniejszym zbiorze.

Uczenie modelu regresyjnego na danych, gdzie została przeprowadzona augmentacja tylko dla danych rzadkich, przyniosły lepsze rezultaty, niż dla augmentacji na całym zbiorze. Mimo to daleko im do wartości otrzymanych dla danych przed obróbką. Wynik dokładności to 2%, dokładności przybliżonej to 8%, dokładności wartościowości, pobudzenia i dominacji to odpowiednio 46%, 8% i 13%. Jedynie wartośćowość przykuwa uwagę, ponieważ jest najlepsza dla wszystkich modeli regresyjnych. Liczba wyjątkowych trójk wynosi 8.

Wyniki modelu klasyfikującego dla zbioru ze zwiększoną liczbą klas rzadkich przyniosły porównywalne rezultaty, jak dane z pełną augmentacją oraz dane przed obróbką. Wynik dokładności to 13%, dokładności przybliżonej to 26%, dokładności wartościowości, pobudzenia i dominacji to odpowiednio 35%, 44% i 40%. Liczba wyjątkowych trójk wynosi 12.

## **7. Podsumowanie**

Zostały osiągnięte podstawowe cele pracy dyplomowej. Przygotowano i przeprowadzono eksperyment zbierający pomiary emocji za pomocą pulsomierza, kamery i eye-trackera oraz dane ankietowe. Dokonano obróbki i analizy zebranych danych. Zaprojektowano i skonstruowano prototyp systemu rozpoznającego emocje, poddano go testom i dokonano ewaluacji jego wyników.

Ankieta składała się trzech dziewięciostopniowych skali: wartościowości, pobudzenia i dominacji. Dodatkowo badani słownie wpisywali emocje, jakie czuli.

Wyniki ankiet były nieoczekiwane. Dobrano bazę bodźców zgodnie z dołączonymi do niej danymi w taki sposób, żeby wywoływała różnorodne emocje. Wyniki wartościowości takie właśnie były. Wykresy rozkładu danych dla wszystkich obrazków łącznie pokrywały całą skalę danych. Średnia i odchylenie standardowe były porównywane do tego w zbiorze NAPS. Niestety dane zebrane podczas eksperymentu dla pobudzenia i dominacji odbiegały znacząco od tych opisanych w bazie bodźców. Różnica średnich dla tego samego obrazka dla większości bodźców była większa od 2. Rozkład danych pobudzenia był przesunięty w kierunku niższych wartości a rozkład danych dominacji był silnie przesunięty w kierunku wartości wyższych.

Najczęściej występującą kombinacją wartościowości, pobudzenia i dominacji były wartości {5, 1, 9}. Pojawiły się aż 119 razy. Słowne emocje należące do zbioru {spokój, nic, brak, neutralna emocja} pojawiły się 212 razy. Pokazuje to, że w zbiorze danych przeważały neutralne emocje.

Emocje słownie wpisane przez badanych w wielu przypadkach były niemożliwe do przypisania tylko do jednej grupy emocji. System starający się opisać emocje tylko na podstawie sześciu rodzin podstawowych emocji mógłby za mało trafnie przybliżać rzeczywistość.

Te same obrazy wywoływały różne emocje u różnych badanych. Ze słownych odpowiedzi wynika, że reakcja uczestnika zależy od skojarzeń i indywidualnych doświadczeń.

Danymi wejściowymi do sieci były tabele składające się z synchronizowanych danych pulsu, punktów charakterystycznych twarzy, ścieżki wzrokowej, wielkości żrenicy oraz stopnia otwarcia oka badanego.

Rozważono sieć dokonującą predykcji wartości wartościowości, pobudzenia i dominacji, jako problem regresji (dane ciągłe od 1 do 9) lub klasyfikacji (9 klas, ponieważ badany nie może zaznaczyć wartości ułamkowej w ankiecie SAM). Kolejnym dylematem było czy jedna sieć powinna przewidywać wszystkie parametry, czy lepiej wytrenować trzy sieci przewidujące tylko jeden z nich. Z tego powodu wytrenowano 20 architektur sieci regresyjnej przewidującej wszystkie parametry oraz 20 architektur trzech sieci klasyfikujących, połączonych jedną warstwą wejściową i dalej całkiem rozdzielonych.

Trenowanie sieci było trudne. Optymalizator łatwo osiągał minimum lokalne funkcji kosztu, dlatego niewiele epok polepszało rezultaty sieci dla zbioru walidacyjnego. Wykorzystano dwie

strategie trenowania sieci. W jednej dla 50 epok ustawiano funkcję wczesnego zatrzymywania uczenia sieci, jeśli od 3 epok nie poprawiła rezultatów. Każdy model trenowano 5 razy i zapisywano ten, który osiągał najlepsze wyniki dla zbioru testowego. W drugiej uruchamiano uczenie sieci dla 50 epok bez wczesnego zatrzymywania uczenia. Wykonywano ten proces tylko raz.

Dla modeli regresyjnych pierwsze podejście dawało marne rezultaty. W drugim było lepiej, udało się uzyskać wartość predykcyjną wynoszącą 24%. Dla modeli klasyfikacyjnych oba podejścia były porównywalne. Najlepszy z nich osiągnął dokładność 27%.

Dokonano augmentacji danych treningowych dodając szum gaussowski o średniej w punkcie 0 i odchyleniu standardowym 0,07. Szum był generowany w postaci tego samego wektora dla wszystkich wierszy tabeli, ale różnych wektorów między próbami. Dodatkowo podjęto próbę sztucznego zbalansowania klas, poprzez dokonanie augmentacji tylko dla tych próbek, których etykiety były niedoreprezentowane w podstawowej wersji zbioru.

Wzięto pod uwagę model regresyjny i klasyfikujący o architektach, które przynosiły najlepsze rezultaty w poprzednim treningu. Wyuczono te architektury na danych z augmentacją oraz tych sztucznie zbalansowanych, jednak przyniosły gorsze rezultaty.

W przyszłych próbach przeprowadzenia eksperymentu zbierającego dane biometryczne podczas ekspresji emocji należałoby inaczej dobrać bazę bodźców. Zamiast skupiać się na całej skali emocji, lepiej wybrać w większości obrazki wywołujące silne emocje, ponieważ nawet bodźce skrajne mogą nie wywoływać większych reakcji u części badanych. Można też zwiększyć liczbę zmysłów, na które działają bodźce i skorzystać z bazy dźwięków lub zapachów.

Szansa losowej predykcyjnej wartości z ankiety SAM to  $1/729$ . Przygotowane modele sieci osiągały znacznie wyższe wyniki. Mimo to nie przyniosły spodziewanych rezultatów.

Żeby poprawić wyniki sieci neuronowej należałoby zebrać znacznie większą bazę reakcji emocjonalnych. Można również rozważyć użycie już wytrenowanego klasyfikatora na dużym zbiorze danych i wykonanie fine-tuningu. Przydatnym może się też okazać zrobienie innej formy augmentacji danych np. usuwanie części wierszy tabeli.

Dodatkowo przy dalszych próbach rozwoju badania można wykonać inną formę synchronizacji danych mimiki, pulsu i ścieżki wzrokowej. W tej pracy dane były przymierzane do znaczników czasowych klatek filmu i wiersze tabeli pulsu były powielane a wiersze tabeli ścieżki wzrokowej usuwane. Zamiast tego można przymierzać dane do tabeli ścieżki wzrokowej i powieść wszystkie inne modalności. Można też pozostawić przymierzanie do danych mimiki, ale zamiast usuwać wiersze, to połączyć je w dłuższy wektor.

Korzystanie z biblioteki *tensorflow* i *keras* było bardzo frustrujące i przysporzyło wielu problemów. Łatwo uczyć z nimi proste sieci neuronowe, ale jeśli projekt jest nieco bardziej skomplikowany, to lepiej skorzystać z innych narzędzi.

## 8. Dodatki

### 8.1. Wyniki ankiet

Tabela 8.1. Wyniki ankiet.

<b>Uczestnik</b>	<b>Nazwa obrazka</b>	<b>Emocja</b>	<b>W</b>	<b>P</b>	<b>D</b>
1	Animals 221 h	neutralne	5	1	9
1	Animals 195 h	uroczość	7	7	6
1	Animals 060 h	odraza	2	7	3
1	People 190 h	podziw	8	7	5
1	People 044 h	uroczość	8	5	6
1	Animals 006 v	neutralną	5	3	8
1	Animals 208 h	uroczość	8	8	3
1	People 198 h	obrzędzenie	2	8	3
1	Faces 227 v	radość	7	5	6
1	Faces 356 h	radość	7	6	6
1	Faces 318 v	spokój	5	3	8
1	Faces 182 h	spokój	7	7	7
1	People 037 h	smutek	3	6	4
1	Faces 115 h	radość	8	6	6
1	Faces 218 h	niepokój	4	7	3
1	Faces 149 v	obrzędzenie	3	7	4
1	Faces 037 h	irytacja	6	6	5
1	Landscapes 165 h	ekscytacja	7	6	5
1	Faces 283 h	smutek	3	7	3
1	Landscapes 120 v	spokój	7	7	3
1	Landscapes 091 h	niepokój	3	7	3
1	Objects 239 v	neutralność	5	5	5
1	Objects 128 h	obrzędzenie	3	7	4
1	Objects 136 h	niepokój	4	6	4
1	People 075 v	śmiech	4	7	2
1	Objects 093 h	obrzędzenie	4	6	8
1	People 127 h	strach	2	9	2
1	People 074 v	niepokój	4	7	4
1	Objects 115 h	niepokój	4	7	6
1	Animals 136 h	uroczość	8	2	4
2	Animals 221 h	obrzędzenie	4	3	8
2	Animals 006 v	strach	3	7	4
2	Objects 093 h	obojętność	5	1	9
2	Animals 195 h	złość	4	5	4
2	Animals 060 h	smutek	2	7	3
2	Animals 208 h	radość	7	4	9
2	Faces 227 v	obojętność	5	6	8
2	Faces 356 h	rozbawienie	8	7	6
2	Faces 318 v	współczucie, smutek	3	6	6
2	People 127 h	złość	1	8	8
2	Faces 149 v	obrzędzenie	4	6	4
2	Faces 182 h	brak	5	2	9

2	Faces 115 h	szczęście	6	3	9
2	Faces 218 h	obojętność	5	2	9
2	Faces 037 h	smutek	3	6	3
2	Faces 283 h	smutek	1	8	2
2	Landscapes_165_h	podziw	7	3	8
2	Landscapes_120_v	podziw	6	2	9
2	Landscapes_091_h	brak	5	2	9
2	Objects_239_v	złość	3	6	7
2	People_198_h	obrzydzenie	4	6	8
2	Objects_128_h	smutek	4	7	4
2	Objects_136_h	smutek	3	7	4
2	Objects_115_h	brak	5	2	9
2	People_075_v	złość	1	9	3
2	People_074_v	podziw	8	6	9
2	People_037_h	smutek	4	7	5
2	People_190_h	podziw	9	7	8
2	People_044_h	szczęście	7	6	8
2	Animals_136_h	niechęć	4	7	8
3	Animals_221_h	lekkie zdegustowanie	4	2	9
3	Animals_006_v	neutralnie	6	1	9
3	Objects_093_h	ciekawość	8	3	8
3	Animals_195_h	neutralnie	5	1	9
3	Animals_060_h	lekką przykrość	3	3	7
3	Animals_208_h	zabawnie	9	3	8
3	Faces_227_v	obojętność	6	1	9
3	Faces_356_h	zmieszanie	9	3	7
3	Faces_318_v	neutralnie	6	1	8
3	People_127_h	zdziwienie	4	3	7
3	Faces_149_v	litość, lekkie przerażenie, zdziwienie	3	4	8
3	Faces_182_h	obojętność	5	2	8
3	Faces_115_h	neutralnie	7	2	6
3	Faces_218_h	neutralnie, zaintrygowana	4	3	7
3	Faces_037_h	obojętność	5	2	9
3	Faces_283_h	lekka przykrość	3	2	8
3	Landscapes_165_h	spokój	7	1	8
3	Landscapes_120_v	neutralnie	5	1	9
3	Landscapes_091_h	znudzenie	5	1	9
3	Objects_239_v	neutralnie	5	1	9
3	People_198_h	odraza, zniszczenie	1	5	5
3	Objects_128_h	obojętność	5	2	9
3	Objects_136_h	obojętność	5	1	9
3	Objects_115_h	lekkie zaciekawienie	6	1	8
3	People_075_v	poruszenie	3	4	6
3	People_074_v	neutralnie	5	1	9
3	People_037_h	odraza	2	5	7
3	People_190_h	spokój	7	1	9
3	People_044_h	obojętność	5	1	9
3	Animals_136_h	lekki strach	5	2	7

4	Animals_006_v	brak	5	1	9
4	Animals_179_h	brak	5	1	9
4	Faces_235_v	satysfakcja	5	1	9
4	Objects_328_h	satysfakcja	5	1	9
4	Faces_135_h	zadowolenie	5	1	9
4	Landscapes_102_h	spokój	6	1	9
		brak, lekkie obrzydzenie, nienawidzę suszonych grzybów			
4	Landscapes_075_h	brak	4	1	9
4	People_025_v	brak	5	1	9
4	Landscapes_120_v	spokój	5	1	9
4	Landscapes_024_v	nostalgię	5	1	9
4	Objects_201_v	nostalgię, radość	5	1	8
4	People_198_h	brak, delikatne współczucie	5	1	9
4	Objects_311_h	nostalgię, radość	7	1	9
4	People_127_h	zażenowanie	5	1	9
4	Objects_093_h	brak	5	1	9
4	People_075_v	zażenowanie	5	1	9
4	People_217_h	współczucie, smutek	5	1	9
4	Objects_115_h	nostalgię, radość, satysfakcję	8	1	8
4	People_074_v	brak	5	1	9
4	Animals_101_v	delikatną irytację	4	2	9
5	Animals_221_h	radość, przyjemność	8	6	9
5	Animals_006_v	ciekawość, obojętność	5	4	8
5	Objects_093_h	zdezorientowanie, ciekawość	7	6	7
5	Animals_195_h	radość, obojętność	6	4	9
5	Animals_060_h	ciekawość, współczucie	6	6	8
5	Animals_208_h	radość	7	4	9
5	Faces_227_v	obojętność	5	2	9
5	Faces_356_h	radość, sympatia	8	5	8
5	Faces_318_v	podejrzliwość, niechęć, ciekawość	3	4	9
5	People_127_h	ciekawość, złość	3	6	8
5	Faces_149_v	ciekawość, wstręt, współczucie	5	6	8
5	Faces_182_h	radość, spokój	7	4	9
5	Faces_115_h	radość, sympatia	8	5	8
5	Faces_218_h	obojętność, znudzenie	5	2	9
5	Faces_037_h	smutek, współczucie, rozbawienie	4	5	8
5	Faces_283_h	smutek, współczucie	3	7	8
5	Landscapes_165_h	radość, spokój	8	5	9
5	Landscapes_120_v	radość, spokój	9	5	9
5	Landscapes_091_h	ciekawość	6	3	9
5	Objects_239_v	rozbawienie	7	4	9
5	People_198_h	wstręt, ciekawość, rozbawienie	6	6	8
5	Objects_128_h	rozczarowanie	3	3	9
5	Objects_136_h	obojętność	5	2	9
5	Objects_115_h	ciekawość	7	4	9
5	People_075_v	rozbawienie, rozczarowanie	5	4	9
5	People_074_v	spokój, zainteresowanie	7	5	9
5	People_037_h	radość, ciekawość, współczucie	7	5	8

5	People 190 h	radość, ciekawość	8	5	9
5	People 044 h	radość, sympatia	8	4	9
5	Animals 136 h	rozbawienie	7	5	9
6	Animals 006 v	strach	3	3	3
6	Animals 179 h	spokój	5	1	1
6	Faces 235 v	radość	7	7	4
6	Objects 328 h	obrzydzenie	5	8	7
6	Faces 135 h	radość	7	1	1
6	Landscapes 102 h	spokój	7	1	1
6	Landscapes 075 h	szok	5	5	2
6	People 025 v	strach	3	7	5
6	Landscapes 120 v	radość	9	2	1
6	Landscapes 024 v	niepokój	5	2	1
6	Objects 201 v	spokój	7	1	1
6	People 198 h	obrzydzenie	1	9	7
6	Objects 311 h	neutralność	5	1	1
6	People 127 h	przerażenie	2	7	7
6	Objects 093 h	zaciekawienie	6	4	1
6	People 075 v	złość	5	5	5
6	People 217 h	obrzydzenie	1	9	9
6	Objects 115 h	zaciekawienie	6	1	1
6	People 074 v	neutralność, nic	5	1	1
6	Animals 101 v	radość	7	4	1
7	Animals 221 h	spokój	7	1	9
7	Animals 006 v	zaniepokojenie	4	2	7
7	Objects 093 h	opanowanie	8	1	9
7	Animals 195 h	szczęście	8	6	9
7	Animals 060 h	smutek	3	5	7
7	Animals 208 h	zaciekawienie, nostalgia	7	4	6
7	Faces 227 v	szczęście	8	5	6
7	Faces 356 h	zaciekawienie	6	3	8
7	Faces 318 v	zazdrość	4	3	8
7	People 127 h	tress, zażenowanie	2	5	7
7	Faces 149 v	obrzydzenie, zaskoczenie	3	6	8
7	Faces 182 h	nostalgia	4	5	6
7	Faces 115 h	szczęście, zmartwienie	6	3	8
7	Faces 218 h	nostalgia, szczęście, spokój	7	3	9
7	Faces 037 h	nostalgia, szczęście	7	4	9
7	Faces 283 h	tress, zaskoczenie	3	1	9
7	Landscapes 165 h	szczęście, odkrywczość, podniecenie	8	4	8
7	Landscapes 120 v	nostalgia, szczęście, tęsknota	8	5	9
7	Landscapes 091 h	zdziwienie, dezorientacja	5	3	7
7	Objects 239 v	tesyknota	3	4	5
7	People 198 h	obrzydzenie	3	6	7
8	Animals 006 v	niepokój	4	4	7
8	Animals 179 h	swoboda	9	1	9
8	Faces 235 v	neutralnie	6	2	9
8	Objects 328 h	głód	7	4	9

8	Faces_135_h	spokój	9	1	9
8	Landscapes_102_h	spokój	6	1	9
8	Landscapes_075_h	obojętność	6	1	9
8	People_025_v	lekki niepokój	4	2	9
8	Landscapes_120_v	spokój, szczęście	9	1	9
8	Landscapes_024_v	obojętność	6	1	9
8	Objects_201_v	nostalgia	9	1	9
8	People_198_h	obrzydzenie	2	6	7
8	Objects_311_h	obojętność	5	1	9
8	People_127_h	lekki niepokój	3	5	7
8	Objects_093_h	obojętność	5	1	9
8	People_075_v	zaskoczenie	5	4	7
8	People_217_h	niepokój, lekki stres	3	4	7
8	Objects_115_h	obojętność	5	1	9
8	People_074_v	spokój, bezpieczeństwo	9	1	9
8	Animals_101_v	optimizm	9	1	9
9	Animals_006_v	neutralne	5	1	9
9	Animals_179_h	radość	7	2	5
9	Faces_235_v	neutralne	5	1	9
9	Objects_328_h	neutralny	5	1	9
9	Faces_135_h	neutralny	6	1	9
9	Landscapes_102_h	zachwyt	6	2	9
9	Landscapes_075_h	neutralny	5	1	9
9	People_025_v	smutek	4	2	7
9	Landscapes_120_v	neutralne	5	1	9
9	Landscapes_024_v	neutralny	5	1	9
9	Objects_201_v	neutralny	5	1	9
9	People_198_h	niepokój	4	2	8
9	Objects_311_h	neutralny	5	1	9
9	People_127_h	niepokój	4	2	8
9	Objects_093_h	zachwyt	6	2	9
9	People_075_v	neutralny	5	1	9
9	People_217_h	wstręt	4	2	9
9	Objects_115_h	neutralny	5	1	9
9	People_074_v	neutralny	5	1	9
9	Animals_101_v	radość	7	3	8
10	Animals_006_v	dyskomfort	5	2	3
10	Animals_179_h	radość	9	4	9
10	Faces_235_v	radość	7	3	9
10	Objects_328_h	śmiech, zdziwienie	8	3	8
10	Faces_135_h	radość, spokój	7	2	9
10	Landscapes_102_h	brak	5	1	9
10	Landscapes_075_h	ciekawość	6	2	8
10	People_025_v	obrzydzenie	4	2	5
10	Landscapes_120_v	tęsknota, spokój	6	2	7
10	Landscapes_024_v	neutralne	5	1	9
10	Objects_201_v	tęsknota	7	2	9
10	People_198_h	obrzydzenie	3	3	4

10	Objects 311 h	radość	7	3	9
10	People 127 h	złość, smutek	3	2	7
10	Objects 093 h	neutralne	5	1	9
10	People 075 v	śmiech	7	3	9
10	People 217 h	ból, strach	2	3	5
10	Objects 115 h	neutralne	5	1	9
10	People 074 v	dyskomfort	5	2	6
10	Animals 101 v	radość	8	3	9
11	Animals 221 h	obrzydzenie	4	2	9
11	Animals 006 v	zaskoczenie, niepokoj	2	4	6
11	Objects 093 h	obrzydzenie	3	3	7
11	Animals 195 h	raczej neutralne zaskoczenie	5	3	5
11	Animals 060 h	trwoga, tress, zaniepokojenie	6	5	5
11	Animals 208 h	neutralny	5	1	9
11	Faces 227 v	entuzjazm	7	4	7
11	Faces 356 h	ciekawość, zdziwienie	7	4	7
11	Faces 318 v	zdziwienie, zaniepokojenie	5	2	5
11	People 127 h	wrogość	3	5	5
11	Faces 149 v	obrzydzenie	1	6	3
11	Faces 182 h	neutralne	5	1	6
11	Faces 115 h	wesołość	6	2	8
11	Faces 218 h	brak sympatii	3	1	6
11	Faces 037 h	niepokój	4	2	5
11	Faces 283 h	lekki niepokój	4	1	6
11	Landscapes 165 h	wolność, beztroska, spokój	7	6	9
11	Landscapes 120 v	spokój	6	4	6
11	Landscapes 091 h	zdziwienie, neutralne	5	1	8
11	Objects 239 v	neutralne	4	2	7
11	People 198 h	obrzydzenie	1	8	3
11	Objects 128 h	przytłoczenie, poczucie biedy	3	3	5
11	Objects 136 h	neutralność, ignorancja	4	2	5
11	Objects 115 h	neutralność, smutek, niepewność	5	1	6
11	People 075 v	zażenowanie	4	2	9
11	People 074 v	zaniepokojenie, niepokój	5	4	6
11	People 037 h	obrzydzenie	4	3	4
11	People 190 h	ciekawość, lekkie zaniepokojenie	5	4	4
11	People 044 h	neutralność, spokój	5	1	9
11	Animals 136 h	spokój	5	2	9
12	Animals 006 v	Zaniepokojony	4	2	8
12	Animals 179 h	Szczęśliwy	7	5	7
12	Faces 235 v	Spokojny	5	1	9
12	Objects 328 h	Głodny	3	7	7
12	Faces 135 h	Nostalgia	6	3	4
12	Landscapes 102 h	Zdziwienie	4	6	6
12	Landscapes 075 h	Zadowolony	8	7	6
12	People 025 v	Zaniepokojony	2	7	7
12	Landscapes 120 v	Spokojny	8	3	9
12	Landscapes 024 v	Nostalgia	7	3	9

12	Objects_201_v	Ból	1	8	3
12	People_198_h	Obrzydzenie	8	9	9
12	Objects_311_h	Spokój	5	1	9
12	People_127_h	Obrzydzony, zaniepokojony	3	6	3
12	Objects_093_h	Nostalgia	7	3	6
12	People_075_v	Obrzydzenie	2	8	8
12	People_217_h	Zaniepokojony	3	7	6
12	Objects_115_h	Zadowolony	9	7	8
12	People_074_v	Zaciekawiony	6	8	5
12	Animals_101_v	Szczęśliwy	9	7	4
13	Animals_221_h	neutralnie, brak	5	6	9
13	Animals_195_h	odczucie głodu	8	1	7
13	Animals_060_h	brak	5	2	9
13	People_190_h	nostalgia, tęsknoty	8	1	8
13	People_044_h	brak	5	1	9
13	Animals_006_v	brak	5	2	8
13	Animals_208_h	brak albo delkatne ciepło na sercu	6	2	8
13	People_198_h	obrzędzenie	2	7	5
13	Faces_227_v	brak	5	3	8
13	Faces_356_h	brak, neutralna	5	2	8
13	Faces_318_v	brak	5	2	8
13	Faces_182_h	brak	5	2	8
13	People_037_h	brak	5	2	8
13	Faces_115_h	brak	5	2	8
13	Faces_218_h	brak	5	2	8
13	Faces_149_v	brak, lekki smutek	4	2	8
13	Faces_037_h	brak	5	3	8
13	Landscapes_165_h	nostalgia	7	1	8
13	Faces_283_h	lekki smutek	3	2	8
13	Landscapes_120_v	nostalgia	7	1	8
13	Landscapes_091_h	brak	5	2	8
13	Objects_239_v	brak	5	2	8
13	Objects_128_h	brak	5	2	8
13	Objects_136_h	brak	5	2	8
13	People_075_v	rozbawienie absurdem, lekko ziryutowany idiotyzmem	5	3	6
13	Objects_093_h	brak	5	2	8
13	People_127_h	ziryutowany	3	3	6
13	People_074_v	zdenerwowanie przez ryzyko poboru	3	4	6
13	Objects_115_h	brak	5	2	8
13	Animals_136_h	brak	5	2	8
14	Animals_221_h	Obrzydzenie	2	5	3
14	Animals_195_h	Neutralną	5	1	9
14	Animals_060_h	Obrzydzenie, smutek	3	5	4
14	People_190_h	Neutralną	5	5	9
14	People_044_h	Neutralną	5	1	9
14	Animals_006_v	Strach, niepokój	2	7	3
14	Animals_208_h	Zadowolenie	9	5	9

14	People 198 h	Obrzydzenie, odraza	1	9	4
14	Faces 227 v	Neutralną	5	1	9
14	Faces 356 h	Radość	9	6	9
14	Faces 318 v	Neutralną	5	1	9
14	Faces 182 h	Neutralną	5	1	9
14	People 037 h	Strach, obrzydzenie	3	6	8
14	Faces 115 h	Zadowolenie	6	1	9
14	Faces 218 h	Neutralną	5	1	9
14	Faces 149 v	Współczucie, obrzydzenie, przerażenie	2	6	7
14	Faces 037 h	Neutralną	5	1	9
14	Landscapes 165 h	Neutralną	5	1	9
14	Faces 283 h	Współczucie, smutek	3	6	5
14	Landscapes 120 v	Neutralną	5	1	9
14	Landscapes 091 h	Neutralną	5	1	9
14	Objects 239 v	Neutralną	5	1	9
14	Objects 128 h	Neutralną	5	1	9
14	Objects 136 h	Niepewność	5	1	9
14	People 075 v	Zażenowanie	1	6	7
14	Objects 093 h	Neutralną	5	1	9
14	People 127 h	Zażenowanie, głupota, ciekawość	2	7	7
14	People 074 v	Neutralną	5	1	9
14	Objects 115 h	Neutralną	5	1	9
14	Animals 136 h	Zauroczenie	9	6	9
15	Animals 221 h	lekle rozbawiony	7	2	8
15	Animals 195 h	neutralnie	5	1	9
15	Animals 060 h	Współczucie, niezadowolenie	1	4	5
15	People 190 h	neutralne	5	1	9
15	People 044 h	lekle rozbawiony	6	1	8
15	Animals 006 v	neutralne	5	1	9
15	Animals 208 h	radość	9	3	4
15	People 198 h	współczucie bólu	3	2	7
15	Faces 227 v	neutralne	5	1	9
15	Faces 356 h	pozytywnie, lekkie rozbawienie	7	2	8
15	Faces 318 v	neutralność	5	1	9
15	Faces 182 h	neutralność	5	1	9
15	People 037 h	neutralność	5	1	9
15	Faces 115 h	neutralność	6	1	9
15	Faces 218 h	neutralność	5	1	8
15	Faces 149 v	współczucie	3	3	7
15	Faces 037 h	neutralność	5	2	8
15	Landscapes 165 h	spokój	7	1	9
15	Faces 283 h	lekki niepokój, irytacja	3	5	5
15	Landscapes 120 v	pominięte	1	1	1
15	Landscapes 091 h	neutralność	5	1	9
15	Objects 239 v	neutralność	5	1	9
15	Objects 128 h	niepokój	4	3	7
15	Objects 136 h	neutralne	5	1	9
15	People 075 v	lekkie zobojętnienie	5	2	8

15	Objects 093 h	neutralność	5	1	9
15	People 127 h	złość, irytacja	1	5	5
15	People 074 v	spokój	6	2	8
15	Objects 115 h	neutralność	5	1	9
15	Animals 136 h	radość	8	3	7
16	Animals 221 h	spokój	5	1	9
16	Animals 195 h	spokój	5	1	9
16	Animals 060 h	zaskoczenie	4	2	9
16	People 190 h	spokój	5	1	9
16	People 044 h	spokój	5	1	9
16	Animals 006 v	spokój	5	1	9
16	Animals 208 h	spokój	5	1	9
16	People 198 h	obrzydzenie	5	2	9
16	Faces 227 v	spokój	5	1	9
16	Faces 356 h	spokój	5	1	9
16	Faces 318 v	spokój	5	1	9
16	Faces 182 h	spokój	5	1	9
16	People 037 h	spokój	5	1	9
16	Faces 115 h	spokój	5	1	9
16	Faces 218 h	spokój	5	1	9
16	Faces 149 v	zaskoczenie	4	3	8
16	Faces 037 h	spokój	5	2	9
16	Landscapes 165 h	spokój	5	1	9
16	Faces 283 h	smutek	4	2	9
16	Landscapes 120 v	spokój	5	1	9
16	Landscapes 091 h	spokój	5	1	9
16	Objects 239 v	spokój	5	1	9
16	Objects 128 h	spokój	5	1	9
16	Objects 136 h	spokój	5	1	9
16	People 075 v	zaskoczenie	5	2	9
16	Objects 093 h	spokój	5	1	9
16	People 127 h	smutek	4	1	9
16	People 074 v	spokój	5	1	9
16	Objects 115 h	spokój	5	2	9
16	Animals 136 h	spokój	5	1	9
17	Animals 221 h	Neutralna	5	1	9
17	Animals 195 h	nic	5	1	9
17	Animals 060 h	neutralne	5	1	9
17	People 190 h	neutralna	6	1	9
17	People 044 h	nic	5	1	9
17	Animals 006 v	nic	5	1	9
17	Animals 208 h	nic	5	1	9
17	People 198 h	skupienie	5	1	9
17	Faces 227 v	nic	5	1	9
17	Faces 356 h	zdziwienie	5	1	9
17	Faces 318 v	nic	5	1	9
17	Faces 182 h	nic	5	1	9
17	People 037 h	nic	5	1	9

17	Faces 115_h	nic	5	1	9
17	Faces 218_h	nic	5	1	9
17	Faces 149_v	nic	5	1	9
17	Faces 037_h	nic	5	1	9
17	Landscapes_165_h	spokój	7	1	9
17	Faces 283_h	nic	4	1	9
17	Landscapes_120_v	spokój	6	1	9
17	Landscapes_091_h	nic	5	1	9
17	Objects_239_v	nic	5	1	9
17	Objects_128_h	nic	5	1	9
17	Objects_136_h	nic	5	1	9
17	People_075_v	ziryutowanie	5	2	9
17	Objects_093_h	nic	5	1	9
17	People_127_h	ziryutowanie	4	1	9
17	People_074_v	skupienie	5	1	9
		dezorientacja, co to jest miecz czy skarpeta?			
17	Objects_115_h		5	2	9
17	Animals_136_h	nic	5	1	9

## 8.2. Nazwy kolumn tabeli wejściowej do sieci

Tabela 8.2. Nazwy kolumn zsynchronizowanych danych oraz ich modalność.

Numer	Modalność	Nazwa kolumny
1	Mimika - nagranie	frameNr
2	Mimika - punkty charakterystyczne	point0_x
3		point1_x
4		point2_x
5		point3_x
6		point4_x
7		point5_x
8		point6_x
9		point7_x
10		point8_x
11		point9_x
12		point10_x
13		point11_x
14		point12_x
15		point13_x
16		point14_x
17		point15_x
18		point16_x
19		point17_x
20		point18_x
21		point19_x
22		point20_x

23	point21_x
24	point22_x
25	point23_x
26	point24_x
27	point25_x
28	point26_x
29	point27_x
30	point28_x
31	point29_x
32	point30_x
33	point31_x
34	point32_x
35	point33_x
36	point34_x
37	point35_x
38	point36_x
39	point37_x
40	point38_x
41	point39_x
42	point40_x
43	point41_x
44	point42_x
45	point43_x
46	point44_x
47	point45_x
48	point46_x
49	point47_x
50	point48_x
51	point49_x
52	point50_x
53	point51_x
54	point52_x
55	point53_x
56	point54_x
57	point55_x
58	point56_x
59	point57_x
60	point58_x
61	point59_x
62	point60_x
63	point61_x
64	point62_x
65	point63_x

66	point64_x
67	point65_x
68	point66_x
69	point67_x
70	point0_y
71	point1_y
72	point2_y
73	point3_y
74	point4_y
75	point5_y
76	point6_y
77	point7_y
78	point8_y
79	point9_y
80	point10_y
81	point11_y
82	point12_y
83	point13_y
84	point14_y
85	point15_y
86	point16_y
87	point17_y
88	point18_y
89	point19_y
90	point20_y
91	point21_y
92	point22_y
93	point23_y
94	point24_y
95	point25_y
96	point26_y
97	point27_y
98	point28_y
99	point29_y
100	point30_y
101	point31_y
102	point32_y
103	point33_y
104	point34_y
105	point35_y
106	point36_y
107	point37_y
108	point38_y

109		point39_y
110		point40_y
111		point41_y
112		point42_y
113		point43_y
114		point44_y
115		point45_y
116		point46_y
117		point47_y
118		point48_y
119		point49_y
120		point50_y
121		point51_y
122		point52_y
123		point53_y
124		point54_y
125		point55_y
126		point56_y
127		point57_y
128		point58_y
129		point59_y
130		point60_y
131		point61_y
132		point62_y
133		point63_y
134		point64_y
135		point65_y
136		point66_y
137		point67_y
138	Puls	hr
139	Ścieżka wzrokowa	Sensor
140		Average calibration accuracy (mm)
141		Average calibration precision SD (mm)
142		Average calibration precision RMS (mm)
143		Average calibration accuracy (degrees)
144		Average calibration precision SD (degrees)
145		Average calibration precision RMS (degrees)
146		Average calibration accuracy (pixels)
147		Average calibration precision SD (pixels)
148		Average calibration precision RMS (pixels)
149		Average validation accuracy (mm)
150		Average validation precision SD (mm)
151		Average validation precision RMS (mm)

152	Average validation accuracy (degrees)
153	Average validation precision SD (degrees)
154	Average validation precision RMS (degrees)
155	Average validation accuracy (pixels)
156	Average validation precision SD (pixels)
157	Average validation precision RMS (pixels)
158	Gaze point X
159	Gaze point Y
160	Gaze point left X
161	Gaze point left Y
162	Gaze point right X
163	Gaze point right Y
164	Gaze direction left X
165	Gaze direction left Y
166	Gaze direction left Z
167	Gaze direction right X
168	Gaze direction right Y
169	Gaze direction right Z
170	Pupil diameter left
171	Pupil diameter right
172	Pupil diameter filtered
173	Validity left
174	Validity right
175	Eye position left X (DACSmm)
176	Eye position left Y (DACSmm)
177	Eye position left Z (DACSmm)
178	Eye position right X (DACSmm)
179	Eye position right Y (DACSmm)
180	Eye position right Z (DACSmm)
181	Gaze point left X (DACSmm)
182	Gaze point left Y (DACSmm)
183	Gaze point right X (DACSmm)
184	Gaze point right Y (DACSmm)
185	Gaze point X (MCSnorm)
186	Gaze point Y (MCSnorm)
187	Gaze point left X (MCSnorm)
188	Gaze point left Y (MCSnorm)
189	Gaze point right X (MCSnorm)
190	Gaze point right Y (MCSnorm)
191	Eye movement type
192	Gaze event duration
193	Eye movement type index
194	Fixation point X

195	Fixation point Y
196	Fixation point X (MCSSnorm)
197	Fixation point Y (MCSSnorm)
198	Ungrouped
199	Client area position X (DACSpx)
200	Client area position Y (DACSpx)
201	Viewport position X
202	Viewport position Y
203	Viewport width
204	Viewport height
205	Full page width
206	Full page height
207	Mouse position X
208	Mouse position Y

### 8.3. Przykładowe wyniki, które dawały sieci neuronowe

Tabela 8.3. Przykładowe wyniki sieci regresyjnej. Każdy wiersz tabeli to predykcja dla kolejnej próbki danych.

Wartościowość	Pobudzenie	Dominacja
4,86495	3,48659	7,37249
4,73138	3,78889	7,25509
4,64859	3,31127	6,55800
4,73912	3,10993	7,00636
4,62289	2,81403	7,63916
5,08336	3,40774	7,26850
5,05906	3,14252	7,55113
5,16253	3,40906	7,26324
5,03957	3,21892	7,54146
4,91032	3,36489	7,46816
4,86495	3,48659	7,37249
5,38056	4,06210	5,61521
4,87055	2,82304	7,40913
4,92116	3,84435	7,10786
4,97260	3,40830	7,46103
4,93940	3,53178	7,38363
4,91435	3,00472	7,52075
4,94281	3,41104	7,41523
4,99537	4,01182	7,08091
4,98167	3,58252	7,39987
5,03000	2,76663	7,56063
5,21841	3,48804	6,99531
5,12789	3,11651	7,40297
4,91494	3,60527	7,25516
4,95012	2,86914	7,37753
5,02751	3,21701	7,66077
5,05788	3,34446	7,34976

4,97881	3,17594	7,44936
4,33450	4,12364	6,98515
4,90971	2,92334	7,57122
5,01059	2,78353	7,58589
4,89642	3,33030	7,45594
4,63633	3,61673	7,27092
4,89812	2,91098	7,65375
4,73138	3,78889	7,25509
5,03161	4,08709	5,80515
4,96828	3,77764	7,19553
5,02205	2,85423	7,13996
4,87631	3,60770	7,46498
4,93233	3,19512	7,64312
4,90940	3,42962	7,27393
4,70004	3,21667	7,51009
4,88444	3,31463	7,43887
4,63568	3,39774	7,36620
5,06288	2,74461	7,54188

Tabela 8.4. Przykładowe wyniki sieci klasyfikującej dla wartościowości. Każdy wiersz tabeli to predykcja dla kolejnej próbki danych. Pełne wyniki składały się z trzech, każda w takim formacie, po jednej dla wartościowości, pobudzenia i dominacji.

Nr	Prawdopodobieństwo, że dane przedstawiają jedną z wartości wartościowości								
	1	2	3	4	5	6	7	8	9
1	33,5%	33,4%	47,5%	62,2%	90,9%	63,8%	47,7%	27,4%	51,2%
2	28,4%	26,7%	51,1%	58,0%	92,2%	47,5%	72,2%	25,3%	53,3%
3	19,5%	26,7%	50,9%	65,0%	91,5%	47,7%	68,6%	31,7%	58,8%
4	45,4%	43,2%	42,5%	53,4%	84,0%	55,1%	38,9%	52,2%	38,2%
5	12,1%	33,0%	54,9%	74,9%	94,1%	48,3%	53,3%	37,0%	57,3%
6	21,5%	28,1%	57,3%	69,7%	90,7%	65,0%	44,3%	37,5%	45,5%
7	43,4%	34,9%	32,1%	65,4%	94,3%	80,9%	37,1%	32,4%	45,7%
8	35,2%	31,9%	39,1%	61,5%	94,4%	72,4%	40,5%	30,1%	48,5%
9	19,5%	26,7%	50,9%	65,0%	91,5%	47,7%	68,6%	31,7%	58,8%
10	34,9%	36,4%	41,7%	59,3%	94,6%	63,8%	51,1%	26,0%	51,9%
11	39,8%	29,8%	45,3%	69,4%	92,2%	51,3%	62,9%	25,3%	47,2%
12	41,0%	29,6%	34,5%	60,2%	95,3%	69,2%	44,0%	22,8%	47,9%
13	28,8%	28,5%	74,9%	74,8%	57,9%	39,6%	57,0%	36,4%	42,7%
14	32,3%	28,9%	33,3%	81,5%	95,3%	76,5%	41,1%	26,7%	57,5%
15	29,5%	30,3%	72,6%	73,2%	60,4%	43,6%	54,4%	37,8%	41,2%
16	32,0%	26,0%	63,5%	66,5%	77,6%	41,1%	67,5%	29,3%	50,9%
17	24,4%	27,2%	63,5%	67,0%	82,1%	43,9%	71,2%	30,7%	60,1%
18	44,4%	38,1%	39,7%	56,5%	89,8%	67,8%	34,8%	44,1%	40,0%
19	25,3%	28,3%	72,9%	74,8%	67,8%	43,4%	52,9%	35,7%	42,0%
20	28,6%	32,1%	55,6%	61,3%	89,2%	55,5%	52,7%	26,9%	54,2%
21	29,8%	30,7%	74,8%	73,8%	57,0%	38,5%	58,7%	35,2%	42,6%
22	45,4%	43,2%	42,5%	53,4%	84,0%	55,1%	38,9%	52,2%	38,2%
23	28,4%	26,7%	51,1%	58,0%	92,2%	47,5%	72,2%	25,3%	53,3%

24	24,4%	27,2%	63,5%	67,0%	82,1%	43,9%	71,2%	30,7%	60,1%
25	35,1%	34,4%	33,2%	76,1%	90,1%	68,9%	45,6%	34,4%	59,2%
26	29,8%	30,7%	74,8%	73,8%	57,0%	38,5%	58,7%	35,2%	42,6%
27	52,6%	38,0%	30,4%	51,5%	91,5%	65,5%	36,8%	41,7%	38,4%
28	34,9%	36,4%	41,7%	59,3%	94,6%	63,8%	51,1%	26,0%	51,9%
29	34,4%	34,3%	46,8%	56,0%	91,3%	60,4%	46,7%	26,4%	49,0%
30	28,6%	32,1%	55,6%	61,3%	89,2%	55,5%	52,7%	26,9%	54,2%
31	21,5%	28,1%	57,3%	69,7%	90,7%	65,0%	44,3%	37,5%	45,5%
32	32,0%	26,0%	63,5%	66,5%	77,6%	41,1%	67,5%	29,3%	50,9%
33	25,1%	28,7%	59,8%	67,5%	89,0%	65,0%	38,4%	40,2%	44,8%
34	32,3%	28,9%	33,3%	81,5%	95,3%	76,5%	41,1%	26,7%	57,5%
35	29,5%	30,3%	72,6%	73,2%	60,4%	43,6%	54,4%	37,8%	41,2%
36	13,8%	28,3%	51,1%	67,4%	95,3%	52,9%	67,6%	29,3%	65,4%
37	33,1%	33,3%	43,9%	62,6%	92,6%	65,2%	44,4%	24,4%	51,3%
38	21,9%	25,5%	51,8%	71,5%	93,9%	74,0%	37,2%	34,9%	44,1%
39	21,5%	26,3%	62,4%	70,2%	85,7%	59,5%	44,5%	35,4%	41,6%
40	42,0%	30,4%	29,2%	63,1%	97,0%	81,3%	36,4%	25,8%	46,9%
41	25,3%	28,3%	72,9%	74,8%	67,8%	43,4%	52,9%	35,7%	42,0%
42	43,1%	35,8%	32,7%	64,2%	95,9%	75,9%	29,9%	32,3%	41,2%
43	24,5%	28,1%	70,6%	73,8%	72,9%	46,9%	51,1%	35,1%	42,4%
44	25,7%	26,3%	71,8%	78,8%	67,6%	40,8%	56,0%	33,6%	46,5%
45	26,5%	28,5%	69,7%	72,4%	69,8%	48,9%	50,6%	37,2%	44,3%

## 9. Bibliografia

- [1] American Psychological Association, „APA Dictionary of Psychology,” 19 kwietnia 2018. [Online]. Available: <https://dictionary.apa.org/reactivity>. [Data uzyskania dostępu: 9 marca 2024].
- [2] R. Gerrig i P. Zimbardo, Psychologia i życie wydanie nowe, 3 red., M. Materska, Red., Warszawa: Wydawnictwo naukowe PWN, 2009.
- [3] M. M. Bradley i P. J. Lang, „Measuring emotion: the self-assessment manikin and the semantic differential,” *Journal of behaviortherapy and experimental psychiatry*, pp. 49-59, 1994.
- [4] A. Soares, A. Pinheiro, A. Costa, S. Frade, M. Comesaña i R. Pureza, „Affective auditory stimuli: Adaptation of the International Affective Digitized Sounds (IADS-2) for European Portuguese,” *Behavior research methods*, tom 45, Marzec 2013.
- [5] P. Ekman, „An argument for basic emotions,” San Francisco, 1992.
- [6] European Parliament, „Artificial Intelligence Act,” 2024.
- [7] K. Vemou, A. Horvath i T. Zerdick, „Facial Emotion Recognition,” *TechDispatch*, nr 1, 2021.
- [8] LinkedIn, „What are the best strategies to reduce participant reactivity in field research studies?,” 22 Luty 2024. [Online]. Available: <https://www.linkedin.com/advice/0/what-best-strategies-reduce-participant-reactivity-field-wicgc>. [Data uzyskania dostępu: 2024 Kiwecień 2024].
- [9] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 3 red., O'Reilly Media, 2022.
- [10] M. Mazurek, „Simple neural network,” ImpiCode, 20 styczeń 2020. [Online]. Available: <https://impicode.com/blog/simple-neural-network/>. [Data uzyskania dostępu: 12 października 2024].
- [11] J. Jordan, „Neural networks: representation,” 28 czerwiec 2017. [Online]. Available: <https://www.jeremyjordan.me/intro-to-neural-networks/>. [Data uzyskania dostępu: 12 października 2024].
- [12] B. Boser, I. Guyon i V. Vapnik, „A Training Algorithm for Optimal Margin Classifier,” *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, tom 5, Sierpień 1996.
- [13] Y. Shi, Y. Tian, G. Kou i Y. Peng, Advanced Information and Knowledge Processing, London: Springer, 2011.

- [14] S. Nimmaturi, „Kernel Trick in SVM,” Medium, 28 Listopad 2019. [Online]. Available: <https://medium.com/analytics-vidhya/how-to-classify-non-linear-data-to-linear-data-bb2df1a6b781>. [Data uzyskania dostępu: 2024 Listopad 2].
- [15] IBM, „What is a decision tree?,” [Online]. Available: <https://www.ibm.com/topics/decision-trees>. [Data uzyskania dostępu: 2 Listopad 2024].
- [16] IBM, „Regression trees,” 3 Listopad 2022. [Online]. Available: <https://www.ibm.com/docs/en/db2-warehouse?topic=procedures-regression-trees>. [Data uzyskania dostępu: 2 Listopad 2024].
- [17] T. Hastie, R. Tibshirani i J. Friedman, *The Elements of Statistical Learning; Data Mining, Inference, and Prediction*, Second Edition, Springer, 2009.
- [18] Inside Learning Machines, „How to Interpret Decision Trees with 1 Simple Example,” [Online]. Available: [https://insidelearningmachines.com/interpret\\_decision\\_trees/](https://insidelearningmachines.com/interpret_decision_trees/). [Data uzyskania dostępu: 2 Listopad 2024].
- [19] Geeks for geeks, „GradientBoosting vs AdaBoost vs XGBoost vs CatBoost vs LightGBM,” 2023 Maj 2023. [Online]. Available: <https://www.geeksforgeeks.org/gradientboosting-vs-adaboost-vs-xgboost-vs-catboost-vs-lightgbm/>. [Data uzyskania dostępu: 2 Listopad 2024].
- [20] B. Soni, „Understanding Boosting in Machine Learning: A Comprehensive Guide,” Medium, 28 Kwiecień 2023. [Online]. Available: Understanding Boosting in Machine Learning: A Comprehensive Guide. [Data uzyskania dostępu: 2 Listopad 2024].
- [21] V. Kazemi i J. Sullivan, „One Millisecond Face Alignment with an Ensemble of Regression Trees,” *IEEE Xplore*, tom 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1867-1874, 2014.
- [22] matemaks, „Pochodne,” [Online]. Available: <https://www.matemaks.pl/pochodne.html>. [Data uzyskania dostępu: 4 Listopad 2024].
- [23] K. Szopa, „Metody Numeryczne, Laboratorium 5, Rozwiązywanie równań różniczkowych I rzędu,” AGH, Kraków, 2020.
- [24] S. Mallick, „Histogram of Oriented Gradients explained using OpenCV,” Learn OpenCV by Big Vision, 6 Grudzień 2016. [Online]. Available: <https://learnopencv.com/histogram-of-oriented-gradients/>. [Data uzyskania dostępu: 26 Październik 2024].
- [25] pyimagesearch customers, „Histogram of Oriented Gradients (and car logo recognition),” PyImageSearch, [Online]. Available: <https://customers.pyimagesearch.com/lesson-sample-histogram-of-oriented-gradients-and-car-logo-recognition/>. [Data uzyskania dostępu: 4 Listopad 2024].

- [26] E. Chopra, „Using Histogram of Oriented Gradients (HOG) for Object Detection,” opengenus, [Online]. Available: <https://iq.opengenus.org/object-detection-with-histogram-of-oriented-gradients-hog/>. [Data uzyskania dostępu: 4 Listopad 2024].
- [27] A. Marchewka, Ł. Żurawski, K. Jednoróg i A. Grabowska, „The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database,” *Springer*, 31 Sierpień 2013.
- [28] M. Riegel, Ł. Żurawski, M. Wierzba, A. Moslehi, Ł. Klocek, M. Horvat, A. Grabowska, J. Michałowski, K. Jednoróg i A. Marchewka, „Characterization of the Nencki Affective Picture System by discrete emotional categories (NAPS BE),” *Springer*, 24 Lipiec 2015.
- [29] Laboratory of Brain Imaging, „Nencki Affective Picture System (NAPS),” [Online]. Available: <https://lobi.nencki.edu.pl/research/8/>. [Data uzyskania dostępu: 26 Listopad 2024].
- [30] Tobii, „Tobii Pro Fusion User manual,” Grudzień 2022. [Online]. Available: <https://www.yixinkeyan.com/uploadfile/202303/74260a9f6806b91.pdf>. [Data uzyskania dostępu: 2024 Listopad 7].
- [31] Tobii AB, Tobii Pro Lab (Version v 1.241) [Computer software], Danderyd, 2024.
- [32] Polar, „Polar H10 N,” [Online]. Available: [https://support.polar.com/e\\_manuals/h10-heart-rate-sensor/polar-h10-user-manual-polski/manual.pdf](https://support.polar.com/e_manuals/h10-heart-rate-sensor/polar-h10-user-manual-polski/manual.pdf). [Data uzyskania dostępu: 2024 Listopada 7].
- [33] Intel, „Intel® RealSense™ Depth Camera D415,” [Online]. Available: <https://www.intelrealsense.com/depth-camera-d415/>. [Data uzyskania dostępu: 7 Listopad 2024].
- [34] Logitech, „Instrukcja obsługi Logitech QuickCam Pro 9000,” [Online]. Available: <https://www.instrukcjaobslugipdf.pl/logitech/quickcam-pro-9000/instrukcja>. [Data uzyskania dostępu: 2024 Listopada 7].
- [35] Google, „Formularze google,” [Online]. Available: <https://www.google.com/intl/pl/forms/about/>.
- [36] The Python Package Index, „numpy 1.26.1,” 14 Wrzesień 2023. [Online]. Available: <https://pypi.org/project/numpy/1.26.1/>. [Data uzyskania dostępu: 10 Listopad 2024].
- [37] The Python Package Index, „matplotlib 3.9.2,” 13 Sierpień 2024. [Online]. Available: <https://pypi.org/project/matplotlib/>. [Data uzyskania dostępu: 10 Listopad 2024].
- [38] pandas, „pandas documentation,” 26 Wrzesień 2023. [Online]. Available: <https://pandas.pydata.org/pandas-docs/version/2.1.2/index.html>. [Data uzyskania dostępu: 10 Listopad 2024].

- [39] The Python Package Index, „tensorflow 2.18.0,” 25 Wrzesień 2024. [Online]. Available: <https://pypi.org/project/tensorflow/>. [Data uzyskania dostępu: 10 Listopad 2024].
- [40] Keras, „Keras 3 API documentation,” [Online]. Available: <https://keras.io/api/>. [Data uzyskania dostępu: 2024 Listopad 24].
- [41] scikit-multilearn, „Multi-Label Classification in Python,” [Online]. Available: <http://scikit.ml/>. [Data uzyskania dostępu: 24 Listopad 2024].
- [42] Google, „Witamy w Colab,” [Online]. Available: <https://colab.research.google.com/>. [Data uzyskania dostępu: 2024 Listopad 24].
- [43] The Python Package Index, „opencv-python 4.10.0.84,” 17 Czerwiec 2024. [Online]. Available: <https://pypi.org/project/opencv-python/>. [Data uzyskania dostępu: 10 Listopad 2024].
- [44] Dlib C++ Library, „Home,” 8 Maj 2022. [Online]. Available: <http://dlib.net/>. [Data uzyskania dostępu: 22 październik 2024].
- [45] Libraries.io, „imutils,” [Online]. Available: <https://libraries.io/pypi/imutils>. [Data uzyskania dostępu: 10 Listopad 2024].
- [46] V. Kazemi i J. Sullivan, „One Millisecond Face Alignment with an Ensemble of Regression Trees,” *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1867-1874, 2014.