

WOJSKOWA AKADEMIA TECHNICZNA
im. Jarosława Dąbrowskiego
WYDZIAŁ ELEKTRONIKI



PRACA DYPLOMOWA
MULTIMODALNY SYSTEM DETEKCJI EMOCJI

.....
(temat pracy dyplomowej)

inż. Piotr WISKI, s. Wiesława

.....
(tytuł zawodowy, imiona i nazwisko, imię ojca dyplomanta)

ELEKTRONIKA I TELEKOMUNIKACJA

.....
(kierunek studiów)

OGÓLNOAKADEMICKI

.....
(profil studiów)

INŻYNIERIA SYSTEMÓW BEZPIECZEŃSTWA

.....
(w zakresie*)

STACJONARNE STUDIA DRUGIEGO STOPNIA - MAGISTERSKIE

.....
(forma i rodzaj studiów**)

dr inż. Ewelina Majda-Zdancewicz

.....
(stopień wojskowy, tytuł i stopień naukowy, imię i nazwisko promotora pracy dyplomowej)

WARSZAWA 2023

strona celowo pozostawiona pusta

strona celowo pozostawiona pusta

strona celowo pozostawiona pusta

Spis treści

Wstęp.....	7
1. Przegląd zagadnień związanych z metodami wykrywania emocji na podstawie wybranych sygnałów fizjologicznych.....	9
1.1. Biosygnal jako źródło informacji o układzie nerwowym	11
1.1.1. Respiracja	11
1.1.2. Elektrokardiografia.....	13
1.1.3. Elektrodermografia	16
1.2. Wpływ emocji na wybrane biosygnały	20
1.3. Przegląd używanych metod estymacji emocji na podstawie wybranych sygnałów biomedycznych.....	24
2. Zdefiniowanie i pomiar sygnałów wejściowych dla systemu detekcji emocji.....	25
2.1 Przegląd użytych modułów do akwizycji biosygnałów	25
2.1.1 Grove - GSR sensor.....	25
2.1.2 SparkFun Single Lead Heart Rate Monitor - AD8232	27
2.1.3 Termistorowy czujnik przepływu powietrza	28
2.2 Założenia projektowe oraz opracowanie oprogramowania dla wybranego mikrokontrolera	31
2.2.1 Opracowanie kodu dla mikrokontrolera	33
2.2.2 Wizualizacja i zapis danych	41
2.2.3 Akwizycja danych oraz wykonanie i przebieg badań	43
3. Opracowanie algorytmu ekstrakcji cech	49
3.1 System kontroli wersji GitHub.....	49
3.2 Konwersja oraz filtracja zarejestrowanych sygnałów	50
3.3 Ekstrakcja deskryptorów	57
3.3.1 Sygnal aktywności elektrodermalnej	57
3.3.2 Sygnal elektrokardiograficzny	60
3.3.3 Sygnal respiracji	63
4. Eksperymenty i wyniki otrzymanych rezultatów	65
4.1 Ewaluacja skuteczności aplikacji <i>Classification Learner</i>	65
4.2 Ewaluacja skuteczności aplikacji <i>Neural Net Pattern Recognition</i>	70
5. Opracowanie docelowej konstrukcji systemu multimodalnego	74

5.1	Konfiguracja sprzętowa docelowego urządzenia	74
5.2	Konfiguracja programowa docelowego urządzenia.....	77
5.3	Schemat docelowego urządzenia	79
Podsumowanie		81

Wstęp

Wraz z postępem technologicznym nasza kompetencja w zakresie rozumienia i analizy ludzkich emocji jest coraz bardziej kompletna. Emocje w życiu codziennym odgrywają znaczącą rolę wpływając m. in. na nasze zachowanie oraz podejmowane decyzje, interakcje społeczne, zdrowie psychiczne i fizyczne, kreatywność czy efektywność pracy. Z tego powodu istnieje rosnące zapotrzebowanie na narzędzia i systemy, które umożliwią detekcję i analizę emocji w ramach różnych kontekstów życia codziennego. Odpowiednie wykorzystanie pozyskanej wiedzy o emocjach pozwoliłoby na wprowadzenie rewolucyjnych zmian w wielu dziedzinach życia, poprawiając jakość naszego bytu oraz otwierając nowe możliwości w rozwoju technologii, medycyny, psychologii, nauk społecznych i rozrywki.

Celem niniejszej pracy jest zbadanie możliwości detekcji emocji przy wykorzystaniu różnego rodzaju sygnałów generowanych przez ludzki organizm. Multimodalny system rozpoznawania emocji bazuje w głównej mierze na sygnałach generowanych przez autonomiczny układ nerwowy takich jak sygnał elektrokardiograficzny (EKG), sygnał elektrodermalny (EDA) i sygnał respiracyjny. W ramach pracy wykorzystano techniki przetwarzania sygnałów i uczenia maszynowego oraz przeprowadzono testy eksploatacyjne zaproponowanej koncepcji zaprojektowanego systemu dla różnych scenariuszy. Następnie na podstawie przeprowadzonych badań i ich analiz zaproponowano szczegółowy projekt docelowej konstrukcji multimodalnego systemu detekcji emocji.

W rozdziale pierwszym dokonano syntetycznego przeglądu zagadnień związanych z potencjalnymi rodzajami sygnałów możliwych do wykorzystania w projektowanym systemie uzupełnionymi o metody ekstrakcji cech dystynktywnych. Dodatkowo rozdział ten został wzbogacony o przegląd aktualnych badań nad metodami wykrywania oraz klasyfikacji emocji. W rozdziale drugim opisano proces tworzenia oprogramowania na mikrokontroler STM32 oraz opis modułów użytych do akwizycji wybranych sygnałów. Rozdział ten obejmuje także opis etapu tworzenia aplikacji, która ma za zadanie wizualizację i zapis wartości pomiarowych. Ponadto zawarto w nim opis metodologii i proces przeprowadzania badań. Rozdział trzeci przedstawia część programową zrealizowaną w oprogramowaniu MATLAB przy wsparciu systemu

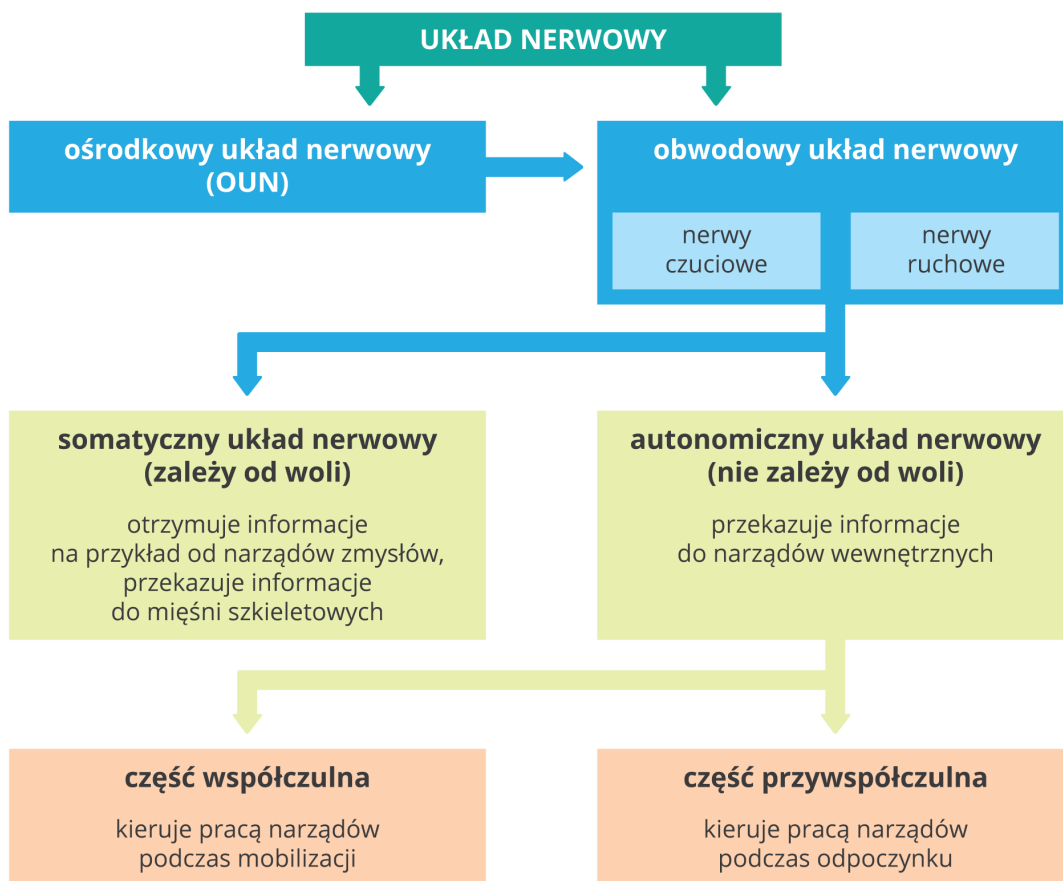
GitHub definiującą etap filtracji sygnału oraz ekstrakcję deskryptorów. W rozdziale czwartym zdefiniowano metody uczenia sieci neuronowych i rejestracji dodatkowych danych. Rozdział piąty prezentuje kompleksowy opis docelowej konstrukcji systemu multimodalnego oraz kierunek dalszych badań. Praca zakończona jest podsumowaniem oraz wnioskami.

1. Przegląd zagadnień związanych z metodami wykrywania emocji na podstawie wybranych sygnałów fizjologicznych

Emocje definiuje się jako silne uczucie wywołane określoną sytuacją [1]. Próby detekcji i klasyfikacji emocji są w ostatnim czasie bardzo często poruszanym zagadnieniem ze względu na mnogość ich zastosowania w edukacji, bezpieczeństwie czy medycynie [2, 3, 4, 5]. Aktualna wiedza oraz postęp technologiczny zbliża nas z każdym krokiem do urzeczywistnienia się koncepcji Web 5.0 [6]. Głównym celem tej wizji jest umożliwienie komunikacji za pomocą interfejsu człowiek-maszyna, który pozwoli na zrozumienie naszych myśli przez komputer lub drugą osobę na poziomie podstawowym. W tym celu korzysta się z emocji jako źródła informacji, które niesie ze sobą ogromny ładunek informacyjny. Aktualnie jednak nie są one w pełni wykorzystywane na szeroką skalę. Ich wykorzystanie pozwolić może na usprawnienie personalizacji reklam, pomoc w zrozumieniu emocji autystycznym osobom, leczenie problemów psychicznych, profilaktykę pro zdrowotną z wykorzystaniem telemedycyny, zwiększenie bezpieczeństwa poprzez wykrywanie potencjalnych zamachów terrorystycznych czy generowanie rozrywki w czasie rzeczywistym m.in.: filmów dopasowanych do naszego aktualnego nastroju, charakteru czy też preferencji użytkownika. Na podstawie przedstawionych przykładów stwierdzić można, że dogłębna analiza sygnałów skorelowanych z emocjami może pozwolić na oszczędność czasu i pieniędzy, zwiększenie efektywności, a także stworzyć nowe dziedziny czy zawody.

Detekcja emocji to złożony proces wymagający interdyscyplinarnej wiedzy oraz niosący za sobą analizę szeregu zmian psychofizycznych w ludzkim organizmie objawiających się na wiele sposobów [7]. Co ważne, zdecydowana większość ludzi wykazuje podobne reakcje na identyczne bodźce emocjonalne, dzięki czemu możliwa jest ekstrakcja pewnych cech i klasyfikacja poszczególnych emocji. Za pomocą odpowiednich urządzeń można pozyskać wybrane parametry fizjologiczne, będące deskryptorami dla poszukiwanych ludzkich emocji. Najbardziej wiarygodnymi i niezawodnymi sygnałami wydają się być te, za powstawanie których w głównym stopniu odpowiedzialny jest *autonomiczny układ nerwowy* [8], głównie ze względu na fakt, iż powstają one w znacznym stopniu niezależnie od ludzkiej woli, co powoduje iż występuje tu dużo większa trudność w ich ukryciu i zafałszowaniu. Zgoła odmienna

reakcja występuje w przypadku emocji, za które odpowiedzialny jest, *układ somatyczny*, gdzie świadomość ma bardzo duży wpływ na odpowiedź układu spowodowaną emocjonującą sytuacją. Uproszczony model omawianego układu nerwowego przedstawiono na rysunku. 1.1



Rys. 1.1 Uproszczony model układu nerwowego [8]

Na przedstawionym schemacie widoczny jest podział układu nerwowego na podukłady. Ośrodkowy układ nerwowy (OUN) jest jednostką kontrolującą, analizującą oraz przetwarzającą sygnały zewnętrzne i wewnętrzne, a następnie kieruje odpowiedzi do narządów wykonawczych. Obwodowy układ nerwowy odbiera te zadania, a następnie jego dwa podukłady somatyczny oraz autonomiczny kierują pracą poszczególnych narządów. Układ somatyczny otrzymuje sygnały ze środowiska zewnętrznego natomiast układ autonomiczny od środowiska wewnętrznego ludzkiego organizmu. Układ autonomiczny zawiera w sobie część współczulną oraz przywspółczulną. Obie części mogą wpływać na ten sam narząd, lecz ich działanie jest przeciwstawne. Część współczulna aktywuje pewne mechanizmy odpowiednich

narządów i przygotowuje organizm do reakcji w sytuacji stresowej, natomiast układ przywspółczulny działa na nie odwrotnie, prowadząc do obniżenia poziomu pobudzenia fizjologicznego organizmu. Przykładowo, układ współczulny powoduje rozszerzenie źrenic w gałce ocznej, a układ przywspółczulny powoduje ich zwężenie [8, 9].

W ramach rozdziału pierwszego pracy przedstawiono opis wybranych biosygnatów, które generowane są przy współpracy autonomicznego układu nerwowego.

1.1. Biosygnal jako źródło informacji o układzie nerwowym

1.1.1. Respiracja

Respiracja jest zespołem procesów fizjologicznych polegających na wymianie gazów między organizmem, a środowiskiem zewnętrznym. Głównym efektem tych działań jest termoregulacja ciała, dostarczenie tlenu do komórek organizmu i usunięcie z niego dwutlenku węgla [10]. Procesy te zachodzą w cyklach wdechowo-oddechowych polegających na wdechach i wydechach powietrza. Ilość cykli wykonywanych w ciągu jednej minuty określana jest jako *częstość oddechu* i określana skrótem RR (ang. *Respiratory Rate*). Zależy ona od wielu czynników takich jak: wiek, płeć, stan układu krążenia, stopień aktywności fizycznej czy stan zdrowotny lub emocjonalny. W zależności od wieku zakładając, że dana osoba jest zdrowa oraz znajduje się w stanie spoczynku wyróżnić możemy różne zakresy średniej częstości oddechów. Normy przedstawiono zbiorczo w tabeli 1.1.

Tab. 1.1 Normy częstości oddechu spoczynkowego w zależności od wieku [11]

Wiek	Liczba oddechów na minutę
Od urodzenia do 6 tygodni	30-40
6 miesięcy	25-40
3 lata	20-30
6 lat	18-25
10 lat	17-23
Dorośli	15-18
50 lat	18-25
Osoby w podeszłym wieku ≥ 65 lat	12-28
Osoby w podeszłym wieku ≥ 80 lat	10-30

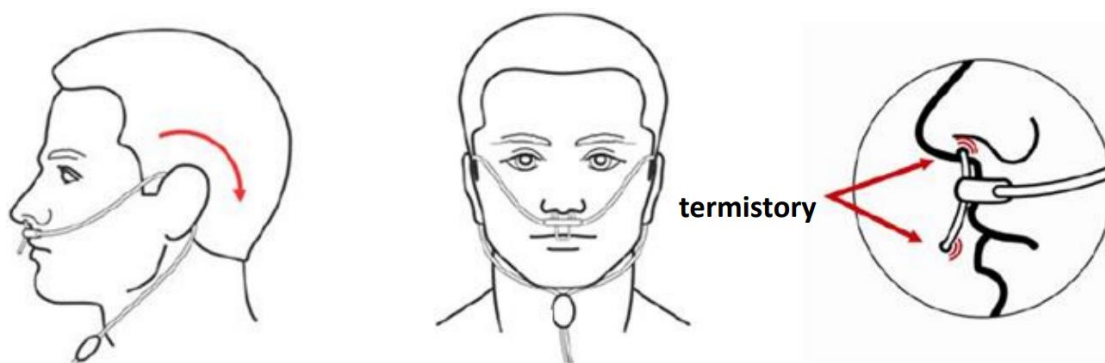
U przeciętnego zdrowego dorosłego człowieka w stanie spoczynku szacuje się, że liczba oddechów powinna wynosić między 12 a 20 oddechów na minutę niezależnie od

płci. Normy te różnią się w pewnym stopniu między sobą w zależności od wybranej literatury [12].

Analiza parametrów czasowych takich jak *czas wdechu* (T_i) oraz *czas wydechu* (T_e) pozwala na pozyskanie informacji, na podstawie których można dokonać estymacji stanu zdrowotnego czy emocjonalnego badanej osoby. Zaznaczyć należy, że czas wydechu jest nieco dłuższy niż czas wdechu w stanie spoczynku [13]. Nieprawidłowe częstotliwości oddechów można podzielić na trzy kategorie: Tachypnea (wysoka częstota oddechów), Bradypnea (niska częstota oddechów) lub Bezdech (zatrzymanie oddychania) [14, 15, 16]. Sondowań częstotliwości oddechu można dokonywać za pomocą wielu technik pomiaru kontaktowych oraz bezkontaktowych takich jak:

- dźwiękowe monitorowanie oddechu,
- analiza ruchów klatki piersiowej,
- wykrywanie przewodnictwa w płucach,
- analiza sygnału fotopletyzmo graficznego i elektrokardiograficznego,
- radarowe monitorowanie oddechu,
- analiza obrazów wideo z kamer czy
- monitorowanie przepływu powietrza [17, 18].

Ostatnia z wymienionych technik pomiaru realizowana może być na kilka sposobów, takich jak na przykład pomiar temperaturowy. Ten sposób pomiarowy jest jednym z najczęściej stosowanych ze względu na prostą i niezawodną budowę urządzeń, łatwość wykonania badania oraz miarodajne wyniki. Pomiar ten polega na wykorzystaniu procesu termoregulacji, gdzie wdychane przez człowieka zimniejsze powietrze jest ogrzewane przez ciało podczas czynności oddechowych, a następnie wydychane jako cieplejsze naprzemiennie chłodząc oraz nagrzewając czujnik termiczny. Różnica temperatur czujnika termicznego może osiągać wartość nawet 15°C [19]. Korzystając z tej zależności istnieje możliwość pomiaru częstotliwości oddechu oraz czasów wdechu i wydechu na przykład za pomocą czujników termistorowych zamieszczonych pod nosem jak przedstawiony na rysunku 1.2.

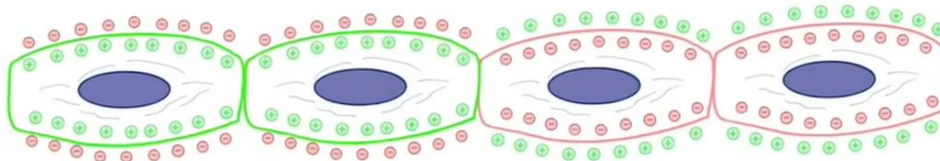


Rys. 1.2 Oprzyrządowanie do pomiaru częstości oddechów [20, 21]

Widoczny układ pomiarowy wykorzystuje właściwości termiczne termistorów, które zmieniają swoją rezystancję w zależności od temperatury otoczenia. W ten sposób odpowiednio przetwarzając i analizując uzyskane zmiany sygnału elektrycznego możliwe jest określenie parametrów oddechowych badanej osoby.

1.1.2. Elektrokardiografia

Kolejnym sygnałem opisującym zmiany wywołane zmianą stanu emocjonalnego może być *sygnał elektrokardiograficzny (EKG lub ECG)* [5]. Jest to zapis elektrycznej aktywności serca w czasie. Powstaje w wyniku cyklicznego skurczu i rozkurczu mięśnia sercowego spowodowanych odpowiednio depolaryzacją i repolaryzacją tego mięśnia. Nośnikami ładunków elektrycznych w ludzkim ciele są jony rozmieszczone wzdłuż błon komórkowych, które zostały zobrazowane na rysunku 1.3 [22, 23, 24].



Rys. 1.3 Uproszczony model błon komórkowych [24]

Błony komórkowe mięśnia sercowego są błonami półprzepuszczalnymi dla odpowiednich jonów pozwalając na przepływ ładunków elektrycznych między wnętrzem, a zewnątrz komórek. W spoczynku wewnątrzkomórkowy ładunek elektryczny jest ujemny, a zewnętrzny dodatni, co tworzy tzw. *potencjał spoczynkowy* (w przybliżeniu -90 mV) [25, 26]. Depolaryzacja następuje, gdy wewnętrzna

powierzchnia błony komórkowej staje się mniej ujemnie naładowana w stosunku do warstwy zewnętrznej (uznać można, że od -40mV do $+50\text{mV}$), co odpowiada *potencjałowi czynnościowemu* [25, 27]. Repolaryzacja jest zjawiskiem odwrotnym do depolaryzacji, a jej zadaniem jest przywrócenie potencjału spoczynkowego. Na rysunku 1.3 błony z prawej strony posiadają potencjał spoczynkowy, natomiast błony zobrazowane po stronie lewej strony potencjał czynnościowy, w wyniku czego nastąpi depolaryzacja lub repolaryzacja w zależności od wielu czynników takich jak różnica potencjałów elektrycznych błon komórkowych czy procesów zachodzących wewnątrz komórek [27]. Dla uproszczenia w przedstawionym przykładzie układ dwóch komórek posiadających różnoimienny ładunek można potraktować jako dipol elektryczny [24, 28, 29].

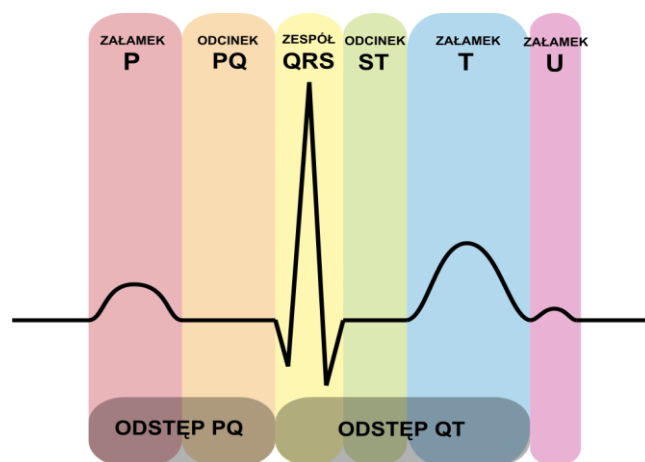
Badanie EKG polega na pomiarach różnicy potencjałów elektrycznych między podpiętymi na ciele elektrodami. Mierzone między nimi różnice napięć generowane są podczas rozchodzących się skurczy mięśnia sercowego składającego się z błon komórkowych. Poszczególne „kombinacje” elektrod nazywa się odprowadzeniami z czego wyszczególnić można następujących 12:

- 3 dwubiegunowe kończynowe Einthovena (I, II, III),
 - 3 jednobiegunowe kończynowe wzmocnione Goldbergera (aVR, aVL, aVF),
 - 6 jednobiegunowych przedsercowych Wilsona (V1, V2, V3, V4, V5, V6)
- [30, 31].

Zarejestrowane za pomocą elektrod pobudzenia elektryczne na elektrokardiogramie interpretować można jako wektorowa suma potencjałów dipoli elektrycznych zarejestrowanych między elektrodami, które opisać można wzorem [32] :

$$\vec{p} = q \cdot \vec{d} \quad (1.1)$$

Gdzie q jest ładunkiem elektrycznym, a d odległością między środkami ładunków elektrycznych. Wektor ten zawsze jest skierowany od ładunku ujemnego do dodatniego, w efekcie czego, gdy ruch ładunków elektrycznych rejestrowany jest od elektrody ujemnej do elektrody dodatniej na wykresie odnotować można wzrost wartości, a spadek w sytuacji przeciwnej, gdy ruch ładunków elektrycznych obserwuje się od elektrody dodatniej do ujemnej. Dzięki tym zależnościom zespół impulsów elektrycznych występujących podczas pracy serca tworzy przebieg zbliżony do przedstawionego na rysunku 1.4, który reprezentuje jeden pełny okres pracy serca.



Rys. 1.4 Sygnał EKG z naniesionymi oznaczeniami [33]

Pozioma linia nazywana *linią izoelektryczną* i stanowi punkt odniesienia, gdzie nie są rejestrowane żadne pobudzenia, które jako punkty charakterystyczne na przebiegu są nazywane załamekami odpowiednio P, Q, R, S, T oraz U. Czasy trwania linii izoelektrycznej między tymi punktami jest nazywany *odcinkiem*, a *odstęp* jest czasem trwania odcinków i sąsiadującego załamka lub załameków. Prawidłowe wartości oraz czasy poszczególnych załameków w czasie wykonywania spoczynkowego EKG dla dorosłej zdrowej osoby powinny mieścić się w granicach przedstawionych w tabeli 1.2 [34].

Tab. 1.2 Normy parametrów spoczynkowego EKG. [34, 35, 36, 37]

Parametr	Czas trwania [s]	Amplituda [mV]	Opis
P	~0,04	~ 0,1–0,25	Ta fala jest wynikiem skurczu prążkowie (lub depolaryzacji). Załamek P przekraczający typowe wartości może wskazywać na przerost przedsionków.
PR	0,12–0,20	–	Odstęp PR mierzony od początku załamka P do początku załamka Q. Reprezentuje czas trwania depolaryzacji (skurczu) przedsionków.
Zespół QRS	0,08–0,12	–	Zespół QRS mierzony od początku załamka Q do końca załamka S. Reprezentuje czas trwania depolaryzacji (skurczu) komór. Jeśli czas trwania jest dłuższy, może to wskazywać na obecność bloków odnogi pęczka Hisa.
QT/QTc	~0,41	–	Mierzy się go od początku załamka Q do końca załamka T. Odstęp QT reprezentuje czas trwania skurczu i rozkurczu komór. Czas trwania QT/QTc zmienia się odwrotnie proporcjonalnie do częstości akcji serca.
Odcinek PQ	~0,5	–	Ukazuje przejście depolaryzacji przez węzeł przedsionkowo – komorowy i pęczek przedsionkowo – komorowy.
Odcinek ST	Powinien trwać 120 ms	–	Przedstawia wolną repolaryzację komór mięśnia sercowego

Odcinek TP	–	–	Odpowiada okresowi kiedy komory i przedsionki znajdują się w rozkurczu.
Odstęp PQ	Powinien trwać 150 ms.	–	Odpowiada za przewodzenie depolaryzacji od węzła zatokowo przedsionkowego do komór mięśnia sercowego.
Odstęp QT	nie powinien przekraczać 0,40 s.	–	Odpowiada czasowi trwania depolaryzacji i repolaryzacji komór mięśnia sercowego. Jest zależny między innymi od częstości rytmu serca (im jest ona większa, tym odstęg QT jest krótszy)
Odstęp RR	–	–	Jest to odległość pomiędzy wierzchołkami dwóch kolejnych załameków R. Służy do obliczenia czynności rytmu serca, ponieważ jest wyrazem czasu trwania jednej ewolucji serca.
Odstęp PP	–	–	Podobnie jak w przypadku odstępu R-R, to odległość między wierzchołkami dwóch kolejnych załameków P. Jeżeli rytm serca jest miarowy, odstęg PP jest równy odstępowi RR.

Analizując zapis elektrokardiograficzny oraz znając odpowiednie normy określić można częstość pracy serca (**HR** - ang. *Heart Rate*), dominujący rytm serca, występowanie częstych zaburzeń rytmu serca, niektóre wrodzone wady serca, cechy pogrubienia mięśnia sercowego lub powiększenia przedsionków, zaburzenia poziomu elektrolitów, cechy niedokrwienia mięśnia sercowego, przebytego zawału czy dokonywać estymacji stanu emocjonalnego badanej osoby [5].

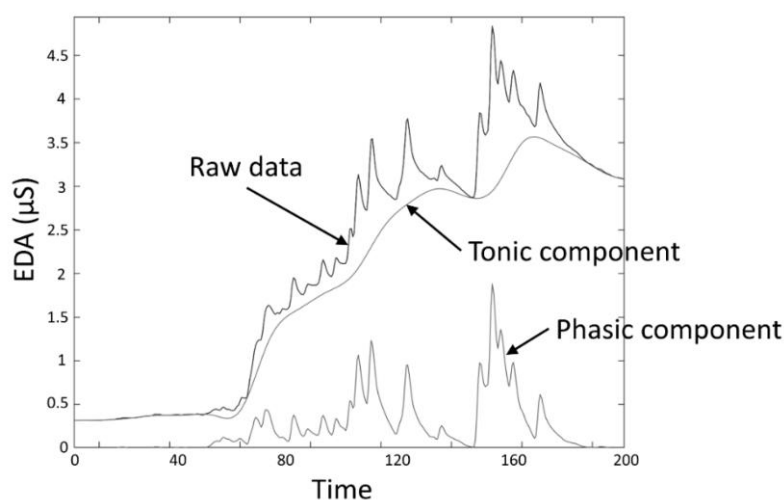
1.1.3. Elektrodermografia

Ostatnim biomarkerem jest sygnał *aktywności elektrodermalnej* (**EDA** – ang. *Electrodermal Activity*) odzwierciedlającej ogólne stany pobudzenia regulowane jedynie przez współczulną część układu nerwowego (brak wpływu części przywspółczulnej – „hamującej”), które objawiają się zmianami w przewodnictwie elektrycznym skóry. Zmiany te spowodowane są wzmożoną aktywnością gruczołów potowych odpowiedzialnych za poziom potliwości, która powoduje wzrost przewodnictwa elektrycznego skóry, gdyż pot składa się szacunkowo w 98% wody, a pozostałe około 2% to związki takie jak chlorek sodu czy kwas mlekowy i mocznik. Aktywność gruczołów potowych w głównej mierze jest odpowiedzialna za termoregulację, jednakże ważnym zagadnieniem na drodze ewolucji okazała się wydajność chwytu pozwalająca na lepsze przystosowanie się człowieka do środowiska zewnętrznego, która jest realizowana dzięki zwilżeniu skóry. Z tego powodu, ich zagęszczenie jest większe między innymi po wewnętrznej stronie stóp i dłoni. Gruczoły potowe są bardziej podatne na bodźce psychologiczne niż bodźce fizyczne, dzięki czemu uważane są za dobre wskaźniki pomocnicze estymacji stanu emocjonalnego.

Dzięki ich dużemu zagęszczeniu w wymienionych obszarach ułatwione jest wykrycie potencjalnych zmian aktywności elektrodermalnej umieszczając elektrody pomiarowe na palcach dłoni [38, 39, 40, 41].

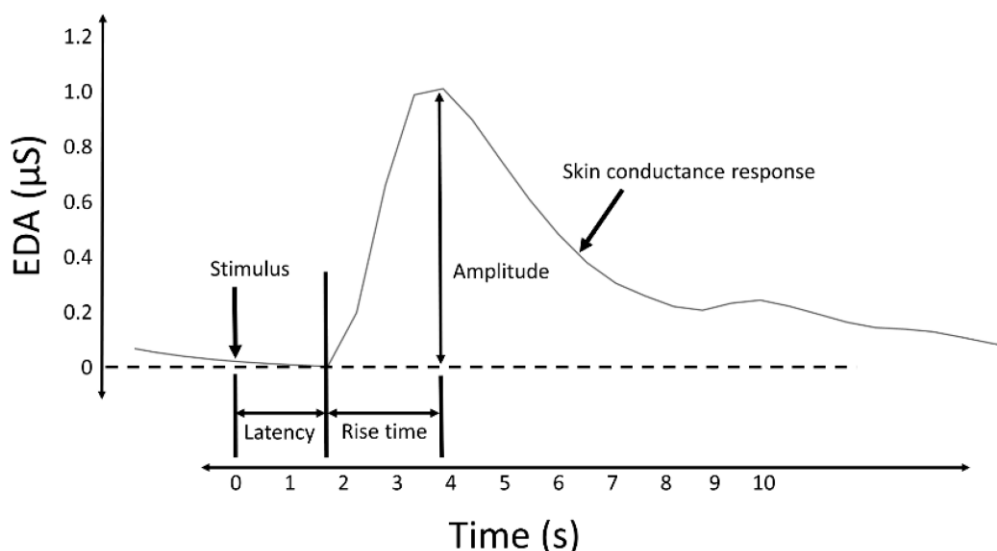
Pot przechodzi przez pewną liczbę kanałów w gruczołach potowych na różnych poziomach, w zależności od poziomu pobudzenia. Zobrazować można to jako zespół połączonych ze sobą równolegle rezystorów o zmiennej rezystancji. Im więcej potu jest wydzielane tym więcej kanałów jest wypełnionych, w konsekwencji czego następuje obniżenie rezystancji [38].

Sygnał EDA rozłożyć można na dwie składowe, gdzie pierwsza jest określana jako *składowa toniczna* poziomu przewodnictwa skórno (SCL – ang. *Skin Conductance Level*). Odnosi się do niskoczęstotliwościowych względnie długotrwałych powolnych fluktuacji sygnału EDA charakteryzujących się zmianami poziomu przewodnictwa skórno. Druga składowa określana jako *składowa fazowa* reakcji przewodnictwa skórno (SCR – ang. *Skin Conductance Response*) odnotowywana jako wysokoczęstotliwościowa, krótkotrwała zmiana amplitudy od wartości poziomu podstawowego do szczytu odpowiedzi. Obie te składowe uważa się, że niosą informacje o stanie emocjonalnym danej osoby [40]. Przykładowy przebieg sygnału EDA w czasie zobrazowano na rysunku 1.5, gdzie dokonano dekonstrukcji sygnału na wymienione składowe. Sygnał pierwotny oznaczony jest jako „Raw data”, składowa SCL jako „Tonic component”, a SCR jako „Phasic component”.



Rys. 1.5 Wykres przedstawiający typową odpowiedź przewodnictwa skóry wraz z odseparowanymi składowymi sygnału. [38]

Podstawowe miary ilościowe SCR określa się jako amplitudę, czas narastania oraz opóźnienie sygnału będące czasem pomiędzy wystąpieniem bodźca, a odpowiedzią układu. Parametry te zilustrowane oraz opisane zostały na rysunku 1.6.



Rys. 1.6 Przebieg sygnału elektrodermalnego z naniesionymi punktami charakterystycznymi [38]

Sygnał EDA najczęściej mierzony jest za pomocą dwóch metod: metody *egzosomatycznej* oraz metody *endosomatycznej*. Ponieważ skóra jak i inne narządy generuje własne pole elektryczne w metodzie endosomatycznej przedmiotem badań jest różnica potencjałów między dwoma określonymi punktami na powierzchni skóry [39, 40]. Z tego powodu pomiar EDA metodą endosomatyczną polega na umieszczeniu jednej elektrody w miejscu o dużym nagromadzeniu gruczołów potowych (wewnętrznej stronie dłoni), a drugą w miejscu względnie nieaktywnym takim jak przedramię. Tak wyznaczony potencjał jest zazwyczaj ujemny na dłoni, dzięki czemu reakcje EDA są widoczne w zapisach. W praktyce pomiar ten polega na wyznaczeniu różnicy potencjałów między wybranym punktem na ciele będącym pewnego rodzaju źródłem siły elektromotorycznej oraz potencjałem odniesienia, którego wartość powinna wynosić zero oraz być stabilna w czasie. Ten rodzaj pomiaru rzadko jest stosowany ze względu na trudności z jego wykonaniem oraz analizą wyników. Trudności te wynikają między innymi z obecności wielu niepożądanych składowych oraz problemów w uzyskaniu stabilnego poziomu odniesienia. Bardziej popularna metoda egzosomatyczna wykonywana jest na dwa sposoby, a mianowicie z wykorzystaniem prądu stałego (DC)

oraz przemiennego (AC). Pomiar z wykorzystaniem prądu stałego (DC) najczęściej polega na podaniu niewielkiej wartości napięcia stałego (np. 0,5V) pomiędzy dwie elektrody umieszczone na ciele, do których dołączono szeregowo rezystor o niewielkiej wartości od 200 do 1000 Ω . Ze względu na rzędowną wartość rezystancji skóry wynoszącą 100 k Ω , można uznać, że rezystancja dodatkowo dołączona do obwodu jest pomijalnie mała i nie ma wpływu na wartość prądu elektrycznego przepływającego przez ten obwód jednak uwzględnia się niewielki spadek napięcia. Za pomocą odpowiedniego układu dokonuje się pomiaru prądu przepływającego przez rezystor oraz zgodnie z prawem Ohma obliczana jest wartość rezystancji czy też częściej konduktancji będącej odwrotnością rezystancji, co zapisać można jako:

$$G = \frac{I}{U} \quad (1.2)$$

Gdzie I jest prądem płynącym w danym punkcie, a U jest wartością napięcia. Wyznaczona w ten sposób konduktancja odzwierciedla zmiany elektrodermalne.

Elektrody z wykorzystaniem tej metody dla uzyskania najlepszych wyników powinno umieszczać się na opuszkach palca wskazującego oraz środkowego lewej dłoni. Umieszczenie to podyktowane jest znaczną liczbą gruczołów potowych występujących w tym miejscu oraz tym, że ręka lewa u większości ludzi nie jest dominującą, dzięki czemu osoba badana jest w bardziej komfortowej pozycji, co jest ważne podczas przeprowadzania eksperymentu, aby nie wywoływać podświadomych niepożądanych reakcji stresowych.

Pomiar metodą egzosomatyczną z wykorzystaniem prądu przemiennego (ang. AC - *alternating current*) polega nie tylko na zmierzeniu wartości rezystancji, lecz także reaktancji co w nomenklaturze naukowej określane jest mianem *bioimpedancji elektrycznej*. Bioimpedancja jest obliczana w podobny sposób jak z wykorzystaniem prądu stałego, gdzie wykorzystując prawo Ohma oraz znając wartość przepływającego prądu w danym punkcie wyznaczana jest wartość napięcia lub znając wartość napięcia wyznaczana jest wartość prądu. Następnie znając wartość prądu oraz napięcia obliczana jest wartość impedancji czy też susceptancji. Wykonując badania tą metodą możliwe jest wykrycie w szerokim zakresie częstotliwości właściwości elektrycznych tkanek biologicznych, gdzie część rzeczywista jest związana z rezystancją tkanek, natomiast część urojona z pojemnością elektryczną tkanek [40]. Pomiar taki wykorzystywany jest między innymi przez inteligentne wagi łazienkowe, które dokonują estymacji między

innymi proporcji tkanki mięśniowej, tłuszczowej i kostnej w organizmie za pomocą elektrod umieszczonych w odpowiednich miejscach, gdzie stawiane są stopy.

1.2. Wpływ emocji na wybrane biosygnały

Naukowcom na całym świecie w dużym stopniu zauważają korelacje między sygnałami fizjologicznymi człowieka regulowanymi przez autonomiczny układ nerwowy (AUN), a stanem emocjonalnym danej osoby. Odzwierciedla to liczba artykułów i prac naukowych zespołów badawczych z całego świata. W ramach publikacji „*Autonomic nervous system activity in emotion: A review*” autorstwa Sylvia D. Kreibig 134 badania naukowe z tego zakresu w syntetyczny sposób podsumowano [13]. Publikacje te skupiają się na monitorowaniu stanu fizjologicznego ludzi z różnych grup wiekowych, przy użyciu wielu różnych czujników biometrycznych umieszczonych na ciele badanych oraz badania ich wpływu w połączeniu z określonymi emocjami, bądź grupami emocji. Najczęstszymi metodami wywoływania emocji były odpowiednio dobrane sygnały audiowizualne w postaci filmów, zdjęć czy muzyki z dedykowanych do tego celu baz danych jak i życia codziennego. Z uwagi na wielostopniową złożoność określenia danej emocji starając się być przy tym jak najbardziej obiektywnym przyjmowano, że określony sygnał audiowizualny wywoływał jedną określoną emocję. Wyszczególniono w jego ramach łącznie 15 emocji, z podziałem na emocje pozytywne (*przywiązanie, rozrywka, zadowolenie, szczęście, radość, przyjemność, duma i ulga.*) oraz negatywne (*gniew, lęk, wstręt, zakłopotanie, strach i smutek*). Podanie dokładnych wartości liczbowych rejestrowanych zmian fizjologicznych byłoby trudne do zdefiniowania, ponieważ tego rodzaju informacja diagnostyczna jest nieobiektywna z uwagi na zróżnicowanie grup wiekowych badanych osób, niejednorodność bodźców wywołujących emocje oraz sposobu przeprowadzania eksperymentu. Z tego powodu w tabeli 1.3 i 1.4 podano jedynie informacje o wzroście (↑), spadku (↓), wynikach niejednoznacznych (↕) lub braku zmian danego parametru (---).

Jednym z analizowanych parametrów była wartość tętna oznaczana jako HR (ang. *heart rate*). Tętno jest rytmicznym pulsowaniem krwi w naczyniach tętniczych odzwierciedlającym częstotliwości skurczów i rozkurczów pracującego serca. Uważa się, że u przeciętnej dorosłej zdrowej osoby spoczynkowa wartość tętna powinna mieścić się w granicach od 60 do 100 uderzeń na minutę. Wartość tętna można obliczyć

na wiele sposobów, a jednym z nich może być analiza elektrokardiogramu z wykorzystaniem poniższego wzoru:

$$HR = \frac{t}{RR_{sr}} \quad (1.3)$$

Gdzie t jest czasem trwania analizowanego przebiegu, a RR_{sr} jest średnią odległością między załawkami R analizowanego przebiegu EKG [42].

Z uwagi to, że tętno jest zależne od czynników zewnętrznych jak i wewnętrznych w tym odczuwanych emocji może być ważnym parametrem w procesie detekcji emocji. Ważnym parametrem jest także zmienność rytmu serca w czasie, co opisuje kolejny parametr oznaczany jako HRV (ang. *Heart Rate Variability*) [43]. Parametr ten jest ogólnym stwierdzeniem dotyczącym zmian rytmu serca, który jest określany na podstawie wielu wskaźników takich jak:

- Pierwiastek kwadratowy ze średniej kwadratowej różnic między kolejnymi odstępami RR (RMSSD – ang. *Root Mean Square of Successive Differences*) wyrażanej za pomocą wzoru:

$$RMSSD = \sqrt{\frac{\sum_{i=1}^{N-1} (RR_{i+1} - RR_i)^2}{N-1}} \quad (1.4)$$

gdzie RR_i oznacza numer zarejestrowanego odstępu między załawkami R w czasie, a N liczbę wszystkich odstępów R-R [44, 45].

- Odchylenie standardowe odstępów między kolejnymi normalnymi skurczami (SDNN – ang. *Standard Deviation of NN (Normal-to-Normal) Intervals*), które wyraża wzór:

$$SDNN = \sqrt{\frac{\sum_{i=1}^N (RR_i - \overline{RR})^2}{N}} \quad (1.5)$$

gdzie \overline{RR} jest wartością średnią czasów odstępów między załawkami RR [45].

- Procent kolejnych normalnych sinusowych odstępów R-R o długości większej niż 50 ms (pNN50), co przedstawia wzór:

$$pRR50 = \frac{\text{Liczba}(RR > 50ms)}{N-1} \cdot 100 \quad (1.6)$$

- Udział składowych wysokoczęstotliwościowych (HF – ang. *High Frequency*) oraz niskoczęstotliwościowych (LF – ang. *Low Frequency*) w widmie sygnału EKG [45].

Uwzględniając wyżej wymienione parametry sygnału EKG można ocenić poziom HRV.

Tab. 1.3 Wskaźniki emocjonalne dla emocji negatywnych [13]

Parametr	Rodzaj emocji						
	Gniew	Niepokój	Wstręt do zanieczyszczeń	Wstręt do okaleczeń	Zakłopotanie	Strach	Smutek
Tętno (HR)	↑	↑	↑/(---)	↓	↑	↑	↓↑
Zmienność rytmu serca (HRV)	↓	↓	↑	(---)	↓	↓	↓↑
Odpowiedź przewodnictwa skórniego (SCR)	↑	↑	↑	↑	brak danych	↑	↓
Szybkość reakcji przewodnictwa skórniego (nSRR)	↑	↑	↑	↑	brak danych	↑	↓↑
Poziom przewodnictwa skórniego (SCL)	↑	↑	↑	↑	↑	↑	↓↑
Częstotliwość oddechu (RR)	↑	↑	↑	↑	brak danych	↑	↓↑
Czas wdechu (Ti)	↓	↓	↓	(---)	brak danych	↓/ (---)	↓
Czas wydechu (Te)	↓	↓	↑	(---)	brak danych	↓	↓

Na podstawie danych zawartych w tabeli 1.3 zaobserwować można, że parametry związane z układem sercowo-naczyniowym takie jak HR w przypadku negatywnych emocji na ogół wzrasta natomiast HRV maleje. Parametry związane z EDA takie jak SCR, nSRR i CSL w zdecydowanej większości są wyższe. Parametry związane z respiracją takie jak częstotliwość oddechu osiąga wartości wyższe od referencyjnych, natomiast Ti oraz Te można uznać, że na ogół spadają w przypadku badanych emocji negatywnych. Emocją, która w znacznym stopniu odbiegała wartościami od pozostałych jest smutek. Spowodowane może być to silnie zróżnicowaną reakcją na smutek, co tłumaczyłoby niejednoznaczne wyniki większości parametrów.

Niejednoznaczność ta może być zaletą, gdyż parametry tej emocji osiągają bardzo specyficzne wartości niespotykane dla pozostałych emocji.

Tab. 1.4 Wskaźniki emocjonalne dla emocji pozytywnych [13]

	Rodzaj emocji							
Parametr	Sympatia	Rozrywka	Zadowolenie	Szczęście	Radość	Przyjemność	Duma	Ułga
Puls (HR)	↓	↕	↓	↑	↑	↕	↕	↑/(---)
Zmienność rytmu serca (HRV)	brak danych	↑	↕	↓	↑	↑	(---)	brak danych
Odpowiedź przewodnictwa skórniego (SCR)	brak danych	↑	(---)	brak danych	brak danych	↑	brak danych	↓
Szybkość reakcji przewodnictwa skórniego (nSRR)	brak danych	↑	brak danych	↑	↑	↑	↑	↓
Poziom przewodnictwa skórniego (SCL)	↑	brak danych	↓	↑/(---)	(---)	↑	brak danych	↓
Częstotliwość oddechu (RR)	brak danych	↑	↕	↑	↑	↕	brak danych	↓
Czas wdechu (Ti)	brak danych	↓	↑	↓	brak danych	↕	brak danych	↑
Czas wydechu (Te)	brak danych	brak danych	↑	↓	brak danych	↕	brak danych	↑

W oparciu o dane w tabeli 1.4 przedstawiającej wpływ pozytywnych emocji zauważyć można, że HR jest bardzo dynamicznym parametrem, którego stan nie w sposób ocenić grupowo ze względu na jego dużą zmienność dla emocji pozytywnych. Natomiast można zaobserwować trend wzrostowy jeśli chodzi o wartości HRV. Parametry EDA w przeważającej liczbie przypadków przedstawiają wzrost poza emocją ulgi, gdzie SCR, nSRR oraz SCL posiadają wartości niższe. W przypadku parametrów oddechowych ze względu na brak danych oraz mieszane wyniki trudno zdecydowanie ocenić emocje zbiorczo, jednak uznać można, że nie jest obserwowany widoczny trend dla badanych emocji.

1.3.Przegląd używanych metod estymacji emocji na podstawie wybranych sygnałów biomedycznych

Rozpoznawanie emocji to bardzo szeroki obszar badań, w którym wykorzystuje się wiele metod przetwarzania sygnałów, algorytmów czy metody uczenia maszynowego. W pracy Liu, Xin et al. wykorzystano dwie sieci SVM, których działanie opiera się na wyznaczeniu hiperpłaszczyzny klasyfikującej dane na dwie klasy maksymalizując odległości między klasami [46]. Sieć SVM₁ klasyfikowała emocje na neutralne oraz nienaturalne. Następnie emocje nienaturalne klasyfikowane były za pomocą sieci SVM₂ na pozytywne oraz negatywne. Uzyskano dokładność na poziomie 92,5%.

Inne podejście zaproponowali C. Jing, G. Liu and M. Hao w swojej pracy stosując algorytm Tabu Search (TS) w celu wybrania najlepszej kombinacji cech [47]. Algorytm rozpoczyna od początkowego rozwiązania i iteracyjnie dokonuje modyfikacji, uwzględniając listę ruchów "tabu" w celu uniknięcia powtórzeń. Ocenivszy jakość każdej modyfikacji, wybiera najbardziej optymalną rozwiązanie, co prowadzi do aktualizacji dotychczas zaproponowanego rozwiązania. Po zakończeniu działania TS użyto algorytmu Fisher-KNN, który mapuje uzyskane cechy na wielowymiarową przestrzeń, uwzględniając najbardziej znaczące cechy z najmniejszą odległością euklidesową do K najbliższych cech danej klasy, przypisuje je do odpowiednich grup (emocje pozytywne lub negatywne). W ten sposób uzyskano na grupie badanych 400 osób skuteczność rozpoznania na poziomie ok. 95%.

Podejście oparte o rozmytą sieć neuronową zaproponowali Malkawi, M., Murad, O. [48]. Rozmyta sieć neuronowa wykorzystując nieprecyzyjne dane i rozmyte relacje między parametrami określa stopniową przynależność do klas, która jest opisana przez funkcję Gaussa dla każdego parametru. Funkcje te między parametrami mogą mieć punkty wspólne lub być od siebie odległe. Ustalono zakresy wcześniej zmierzonych 14 parametrów w pewnych zakresach dla każdej z 22 emocji osobno. Wyniki okazały się bardzo obiecujące, lecz potrzebna wydaje się być większa liczba wysokiej jakości danych treningowych.

Metod klasyfikacji emocji jest bardzo wiele jednak wybór najlepszej z nich zależy od najróżniejszych czynników takich jak dostępność danych, jakość danych, charakterystyka danych, dostępna moc obliczeniowa jak i celu i kontekstu klasyfikacji (grupa emocji lub pojedyncza emocja).

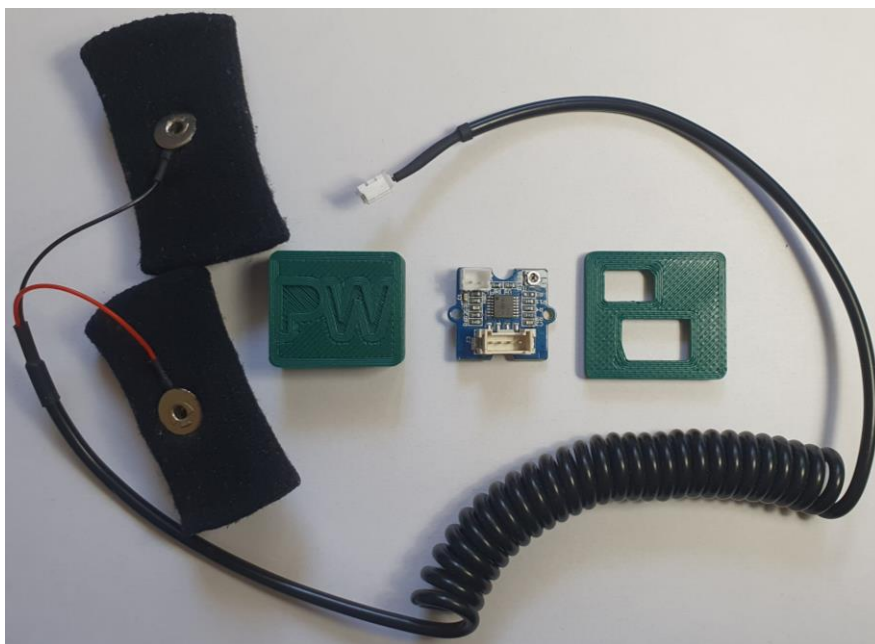
2. Zdefiniowanie i pomiar sygnałów wejściowych dla systemu detekcji emocji

Pomiar odpowiednich biosygnali jest ważnym zagadnieniem, by móc następnie wyekstrahować kluczowe determinanty prowadzące do zdefiniowania modalności behawioralnej konkretnej emocji, czy też grupy emocji, dlatego biorąc pod uwagę zagadnienia zawarte w rozdziale pierwszym postanowiono dokonać monitorowania sygnałów kontrolowanych przez autonomiczny układ nerwowy takich jak sygnał EKG, EDA oraz respiracji. Aby tego dokonać należało wywołać odpowiednie emocje na respondentach. Analizując omawiane wcześniej prace badawcze oraz powiązane z nimi zauważyć można, że większość z nich wykorzystywała specjalnie stworzone do tego celu bazy danych afektywnych takich jak DEAP [49] zawierające materiały wideo czy IAPS [50], czyli Międzynarodowym Systemem Obrazów Afektywnych, z których Autor także postanowił skorzystać jednak nie uzyskał odpowiedzi zwrotnej. Postanowiono w takim razie skorzystać z największego serwisu Internetowego zawierającego treści wideo jakim jest YouTube. Następnie wywołując odpowiedni stan afektu na respondentach poprzez prezentacje wybranego materiału wideo dokonano akwizycji danych przy użyciu przedstawionych poniżej urządzeń.

2.1 Przegląd użytych modułów do akwizycji biosygnali

2.1.1 Grove - GSR sensor

Pierwszym z użytych urządzeń był moduł „*Grove - GSR sensor*” służący do akwizycji sygnału EDA, który został zobrazowany na rysunku 2.1 [51]. Na jego płytce PCB znajduje się układ *Lm324*, który składa się z czterech niezależnych wzmacniaczy operacyjnych z kompensacją częstotliwości o dużym wzmocnieniu, co zostało zawarte w nocie katalogowej układu [52]. Oznacza to, że moduł może wzmocnić sygnały wejściowe w znacznym stopniu, a dzięki kompensacji częstotliwości jest w stanie zapewnić ich stabilność w określonym paśmie częstotliwości.



Rys. 2.1 Moduł do akwizycji sygnału EDA w zaprojektowanej obudowie wraz z elektrodami pomiarowymi

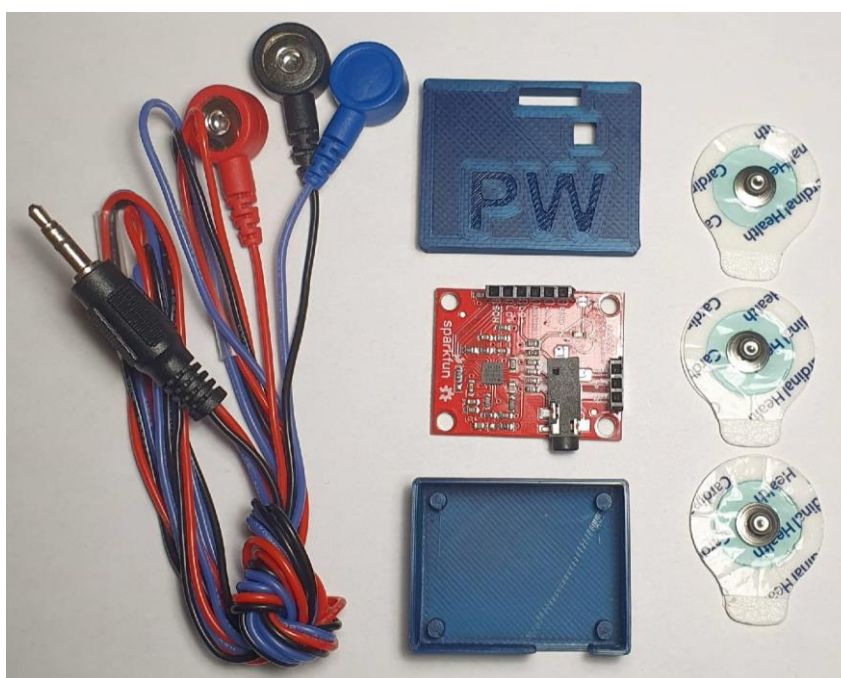
Moduł zasilany jest napięciem stałym o wartości 3,3V lub 5V. Równolegle do wejścia modułu dolutowany został kondensator ceramiczny w obudowie THT (ang. *Through-Hole Technology*) o wartości 100nC w celu filtracji potencjalnych wpływających na wyniki pomiarów wysokoczęstotliwościowych zakłóceń pochodzących z zewnętrznego układu zasilania.

Na wyjściu modułu oznaczonym jako *SIG* uzyskiwana jest analogowa wartość napięcia w zakresie od 0V do 3,3V reprezentująca zmiany sygnału EDA. Czułość sygnału na wyjściu można regulować potencjometrem oznaczonym jako *R1*. Do modułu dołączone zostały odpowiednie elektrody pomiarowe ukryte w materiałowych rękawach [53]. Producent nie informuje o rodzaju materiału z jakiego wykonane są elektrody, jednak powinny one być wykonane z materiału przewodzącego, o niskiej pojemności cieplnej, niskiej impedancji, odpowiednio dużej powierzchni oraz nie reaktywnego ze skórą i potem. Domniemywać można na podstawie oceny wizualnej oraz cech fizycznych elektrod, że użyty materiał to stal nierdzewna spełniająca w znacznej części przedstawione cechy, a także jej stosunkowo niska cena korzystna dla producenta. Moduł został zamknięty w własnoręcznie zaprojektowanej obudowie z naniesionymi inicjałami Autora w oprogramowaniu Solid Edge 2022 w wersji studenckiej oraz przekonwertowany na format zrozumiały dla drukarki 3D w oprogramowaniu UltiMaker Cura 5.4.0. Wydruk powstał z materiału typu PLA na drukarce 3D Artillery

Sidewinder X2 dzięki życzliwości firmy SiDLY. Obudowa ta została stworzona w celu ochrony modułu przed zniszczeniem oraz zapobiegnięciu wywołania przypadkowych zwarc i zakłóceń na odsłoniętej płytce PCB.

2.1.2 SparkFun Single Lead Heart Rate Monitor - AD8232

Następnym z użytych modułów jest „*SparkFun Single Lead Heart Rate Monitor - AD8232*”, służący do akwizycji sygnału EKG przedstawiony na rysunku 2.2 [54, 55]. Na jego płytce PCB znajduje się układ AD8232 przeznaczony do wyodrębniania, wzmacniania i filtrowania niewielkich biosygnatów jak wynika z noty katalogowej [56]. Moduł zasilany jest napięciem stałym o wartości 3,3V. Do wejścia zasilania modułu również został dolutowany identyczny jak poprzednio kondensator ceramiczny o wartości 100uF w celu filtracji zakłóceń. Urządzenie posiada także analogowe wyjście sygnału EKG w zakresie od 0V do 3,3V oznaczone jako *OUT*. Dodatkowymi wejściami są *LO-* oraz *LO+* odpowiedzialnymi kolejno za wykrycie podłączenia elektrody ujemnej oraz elektrod dodatnich. Ostatnie wejście oznaczone jako *SDN* jest odpowiedzialne za wprowadzenie układu w stan czuwania po podaniu stanu wysokiego odpowiadającego napięciu 3,3V. Także tu zaprojektowano specjalną obudowę w celu ochrony modułu przed uszkodzeniem oraz zapobiegnięciu wywołania przypadkowych zwarc i zakłóceń bardzo czułego układu na odsłoniętej płytce PCB.



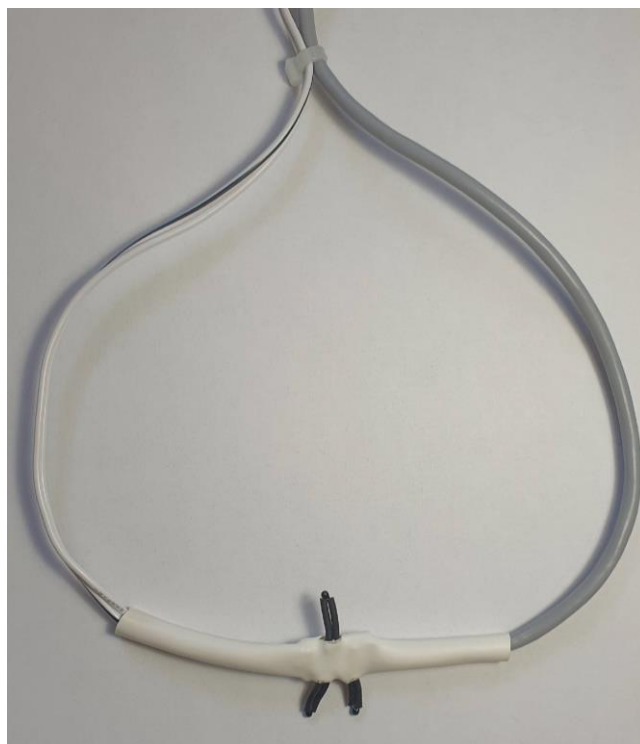
Rys. 2.2 Moduł do akwizycji sygnału EKG w zaprojektowanej obudowie wraz z przewodami oraz elektrodami pomiarowymi

Dodatkowo do układu zakupione zostały przewody do elektrod „*SparkFun CAB-12970*„ realizujące 3 biegunowe odprowadzenia [57, 58]. Przewody te podłączane są za pomocą złącza jack 4mm do znajdującego się na płycie modułu. Wadą użytych przewodów jest brak ekranu, przez co są podatne na zakłócenia, jednak po zakupie drugiej pary przewodów wielokrotnie krótszych obserwowano na wyjściu przetwornika szum na tym samym poziomie wynoszącym około 200mV na wyjściu modułu, co oznaczać może, że głównym źródłem zakłóceń jest inny czynnik.

Do przewodów podłączone zostały elektrody chlorkowo-srebrowe „*H124SG*” zawierające przewodzący adhezyjny hydrożel przyłączony do pianki polietylenowej ze złączem wykonanym ze stali nierdzewnej [59, 60].

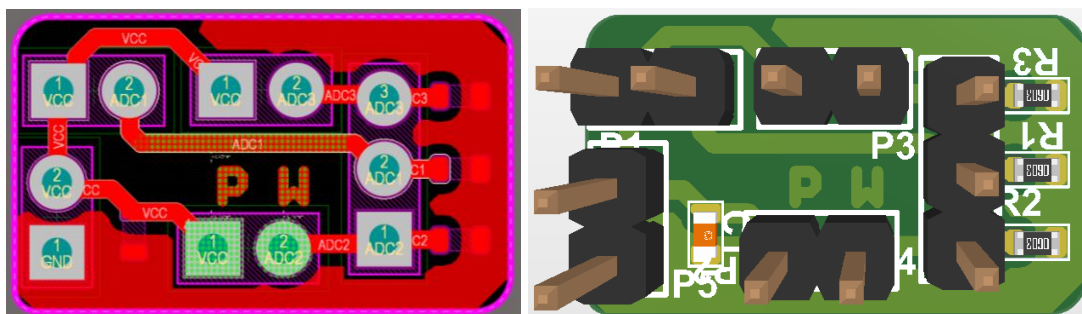
2.1.3 Termistorowy czujnik przepływu powietrza

Moduł przepływu powietrza postanowiono zaprojektować własnoręcznie ze względu na wysokie ceny urządzeń sięgających setki funtów w sklepach specjalistycznych [61]. Zdecydowano się na zbudowanie urządzenia do pomiaru częstości oddechu z wykorzystaniem termistorów jako czujników temperaturowych przepływu powietrza, którego finalna konstrukcja widoczna jest na przedstawionym poniżej rysunku 2.3.



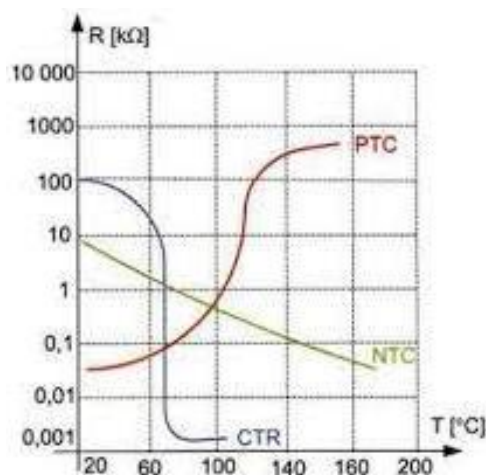
Rys. 2.3 Wykonany termistorowy czujnik przepływu powietrza

Do zaprojektowania płytki modułu wybrano oprogramowanie Altium Designer 2022 w wersji 22.6.1 używając licencji studenckiej, a zaprojektowany układ połączeń elektrycznych przedstawiono na rysunku 2.4.



Rys 2.4 Projekt wykonanego czujnika przepływu powietrza. Z lewej strony schemat połączeń elektrycznych oraz rozkład komponentów z prawej strony

Dla ułatwienia późniejszej pracy montażowej w miejscu montażu przewodów oraz termistorów postanowiono użyć szablonu montażu (ang. *footprint*) podwójnych złącz kołkowych, by następnie wywiercić dziurki w ich miejscu na płytce PCB. Zaprojektowana płytka PCB jest niewielka aby nie przeszkadzać respondentom w noszeniu oraz by mieściła się pod nosem, a jej wymiary to 1cm na 2,5cm oraz 3mm grubości nie wliczając wystających termistorów, których długość może być regulowana. Każdy z trzech czujników znajduje się na zaprojektowanej oraz wytrawionej płytce PCB. Termistory wystają poza obrys płytki tak aby dwa znajdowały się w okolicach nozdrzy oraz jeden w okolicach ust rejestrując przepływy powietrza. Wykorzystano termistory **MF52A473J3950** typu NTC (ang. *Negative Temperature Coefficient*) posiadających ujemny współczynnik temperaturowy o wartości znamionowej rezystancji 47kOhm w obudowie typu THT [62]. Charakteryzuje je bliska liniowej charakterystyka rezystancyjno-temperaturowa w szerokim przedziale w przeciwieństwie do termistorów PTC i CRT, których charakterystyki zostały wykreślone wspólnie na rysunku 2.5. Użyty termistor posiada wartość $\beta = 3950K$, co w praktyce odnosi się do stopnia nachylenia charakterystyki rezystancyjno-temperaturowej, gdzie im wyższa wartość β , tym większe zmiany rezystancji będą miały miejsce w odpowiedzi na zmiany temperatury. Uznano że wartość β używanych termistorów jest w zupełności wystarczająca dla zbudowanego urządzenia.



Rys. 2.5 Charakterystyka rezystancyjno-temperaturowa różnych typów termistorów [63]

Aby poprawnie odczytać wartość napięcia odzwierciedlającą wartość temperatury przez przetwornik analogowo-cyfrowy mikrokontrolera postanowiono umieścić termistor w układzie dzielnika napięcia obliczając uprzednio wartość rezystora za pomocą równania 2.1. Przyjmując, że układ zasilany będzie napięciem $U = 3,3V$, a wartość rezystancji znamionowej termistora $R_t = 47k\Omega$ poszukiwano wartości rezystora R_r z uwzględnieniem napięcia wyjściowego U_t .

$$U_t = U \cdot \frac{R_t}{R_t + R_r} = 3,3V \cdot \frac{47k\Omega}{47k\Omega + 15k\Omega} \approx 2,5V \quad (2.1)$$

Obliczona wartość rezystora R_r wynosi $15k\Omega$, a bazowa wartość napięcia na wyjściu wynosi $2,5V$. Jest to wartość pożądana ponieważ nie przekracza wartości $3,3V$, czyli maksymalnego zakresu odczytu przetwornika analogowo-cyfrowego (ADC – ang. *Analog-to-Digital Converter*) jednocześnie przekraczając połowę wartości. Jest to optymalna wartość, gdyż ogrzany termistor zmniejszy wartość swojej rezystancji do około $15k\Omega$ podczas przepływu powietrza z ust jednocześnie zmniejszając napięcie na wyjściu układu do około $1,65V$, co zapewni dobrą rozdzielczość odczytu pomiarów. Wartości rezystancji ogrzanego termistora oraz napięć wyjściowych zmierzono multimetrem UNI-T 139B. Identyczne układy dzielników napięcia zastosowano dla wszystkich trzech termistorów, dzięki czemu wyniki z każdego czujnika są znormalizowane. Na wejściu układu dołączono szeregowo kondensatory ceramiczne C_1 i C_2 filtrujące zakłócenia zewnętrzne o wartości odpowiednio $10\mu C$ i $100nC$. Na dodatkowy kondensator C_2 nie uwzględniony w wcześniejszych projektach zdecydowano się w celu filtracji zakłóceń z szerszej gamy częstotliwości. Następnie

dolutowane zostały 1,5 metrowe przewody zasilające z lewej strony oraz przewody sygnałowe ekranowane z prawej strony układu.

Tak skonstruowany układ został umieszczony w białej koszulce termokurczliwej aby nie podrażniał skóry, był bardziej higieniczny, a także nie powstawały zakłócenia powodowane przez kontakt skóry z ścieżkami na płytce PCB. Na nóżki wystających termistorów założone zostały czarne koszulki termokurczliwe z tych samych powodów. Termistory dzięki długim nóżkom mogą być bez przeszkód wyginane i dostosowane do każdej osoby indywidualnie, co okazało się ogromną zaletą urządzenia podczas przeprowadzania badań.

2.2 Założenia projektowe oraz opracowanie oprogramowania dla wybranego mikrokontrolera

Ponieważ wybrane oraz zbudowane czujniki na swoich wyjściach podają analogowy sygnał napięcia, który w celu późniejszej analizy i obliczeń należy przetworzyć na postać cyfrową potrzebny będzie odpowiedni przetwornik analogowo-cyfrowy (ADC) o odpowiedniej rozdzielczości. Sygnały należy także próbkować z odpowiednią częstotliwością, co najmniej odpowiadającej częstotliwości Nyquista. Sygnał przepływu powietrza jest sygnałem wolnozmiennym, którego częstotliwość chwilowa nie powinna przekraczać podczas eksperymentu wartości równej 1,5Hz odpowiadających 90 oddechów na minutę, gdzie górna granica normy w stanie spoczynku dla zdrowej osoby dorosłej to 18 - 20 oddechów na minutę, czyli około 0,3Hz. Sygnał EDA zazwyczaj jest próbkowany z częstotliwością 70Hz, jednak istnieją źródła mówiące, że niektóre informacje w ten sposób są tracone i dobrym rozwiązaniem jest częstotliwość próbkowania 1kHz, a nawet 2kHz [64]. Sygnał EKG zaleca się aby próbkować w celu analizy HRV z częstotliwością 1kHz [65]. W celu akwizycji sygnałów z pożądaną częstotliwością potrzebny będzie licznik (ang. *timer*) zliczający iteracyjnie od 0 do 999 oraz wyzwalaający przerwanie wyzwalające pojedynczą konwersję sygnału przez przetwornik ADC w pętli programu.

Uwzględniając wymienione kryteria do akwizycji sygnałów postanowiono wybrać płytkę deweloperską *STM32 Nucleo-F446RE* zobrazowaną na rysunku 2.6. Wybrana została ze względu na rozbudowaną 32-bitową strukturę mikroprocesora z rdzeniem ARM Cortex-M4, która umożliwi zarejestrowanie sygnału EDA, EKG oraz trzech

sygnałów z czujników termistorowych z pożądaną przez Autora częstotliwością próbkowania jednocześnie.



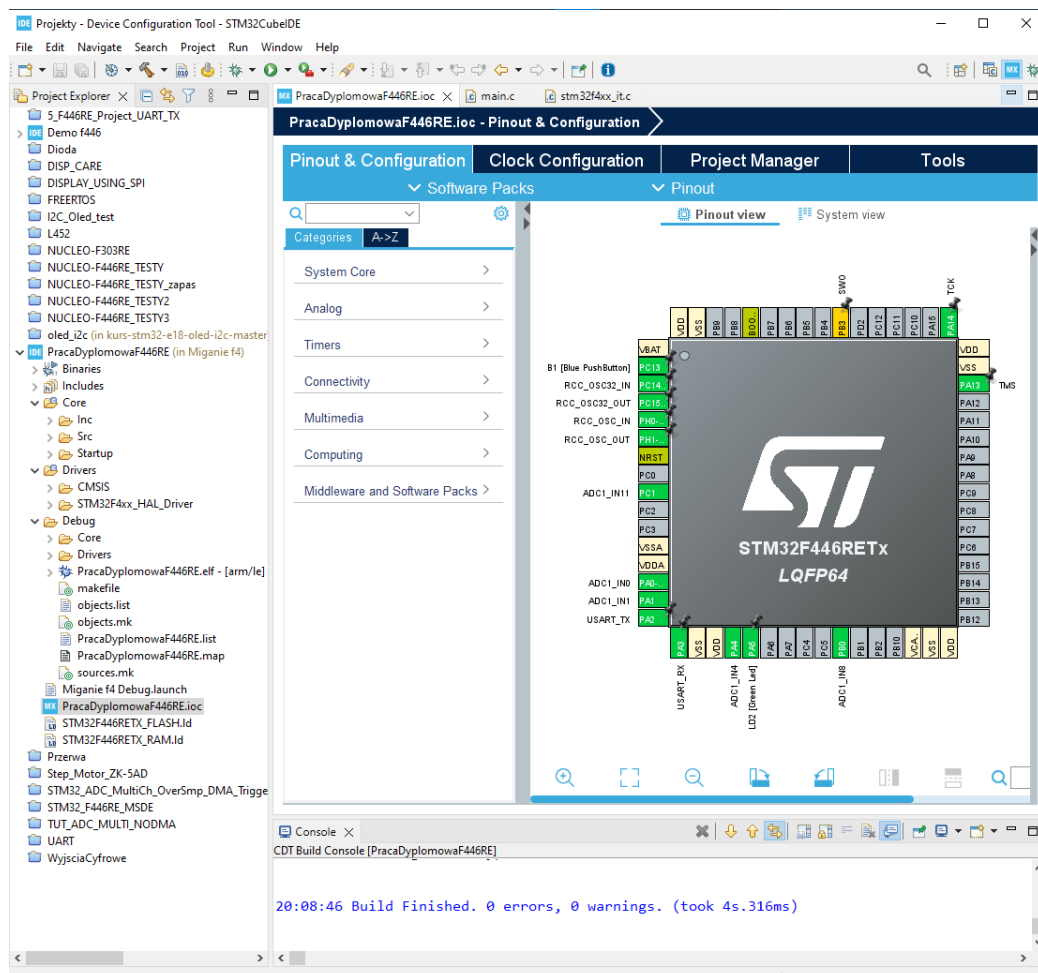
Rys. 2.6 Płytki deweloperska STM32 Nucelo-F446RE

W górnej części płytki zobrazowanej na rysunku 2.6 znajduje się debugger/programator *ST-LINK/V2-1* umożliwiający zaprogramowanie mikroprocesora oraz późniejsze debugowanie kodu, co podczas pracy z płytką okazało się nieocenioną pomocą. Dużą zaletą prezentowanej płytki deweloperskiej jest stosunkowo łatwa implementacja kodu z pomocą wspomnianego programatora oraz bibliotek HAL (ang. *Hardware Abstraction Layer*) z użyciem zintegrowanego środowiska programistycznego *STM32CubeIDE* umożliwiających szybkie generowanie kodu bez zagłębiania się w rejestry mikroprocesora. Alternatywnym podejściem jest korzystanie z bibliotek LL (ang. *Low Layer*) lub CMSIS (ang. *Cortex Microcontroller Software Interface Standard*) dających pełną kontrolę nad rejestrami procesora lecz implementacja kodu wymaga zazwyczaj pracy z obszerną dokumentacją (często ponad 2000 stron) oraz jest dużo bardziej czasochłonna i trudna.

Producent stworzył także oprogramowanie *STM32CubeMonitor* umożliwiające między innymi wizualizację oraz zapis zarejestrowanych sygnałów w popularnych formatach jak *csv*. Niewątpliwą zaletą jest aktualne wsparcie i rozwój oprogramowania przez firmę ST oraz obszerna dokumentacja oprogramowania jak i mikroprocesora, do którego implementacja kodu przedstawiona została w dalszej części.

2.2.1 Opracowanie kodu dla mikrokontrolera

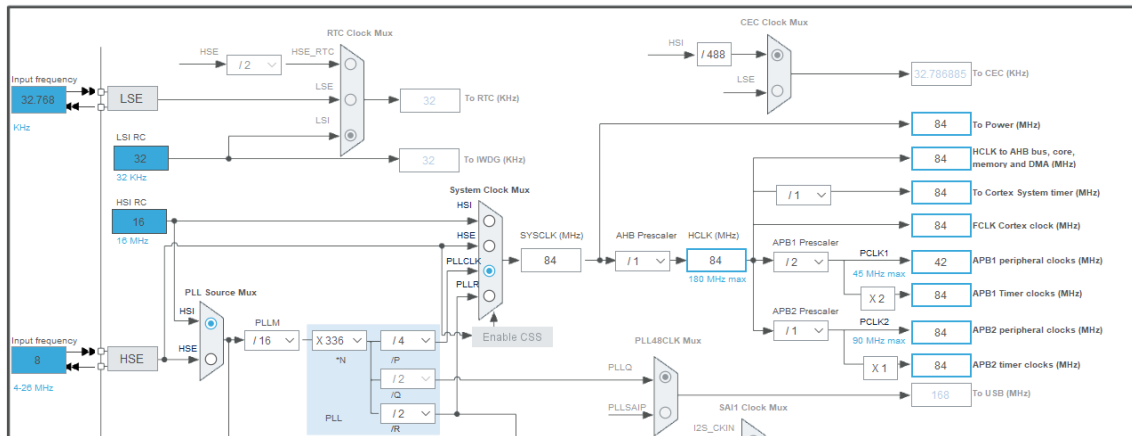
Aby zaprogramować płytkę deweloperską połączono ją z komputerem za pomocą odpowiedniego przewodu USB. W oprogramowaniu STM32CubeIDE po wybraniu odpowiedniej płytki deweloperskiej, która jest używana oraz nazwaniu projektu i wygenerowaniu projektu w trybie domyślnym ukazuje się główne okno robocze przedstawione na rysunku 2.7 zawierające graficzne interaktywne okno.



Rys. 2.7 Główne okno robocze projektu w środowisku STM32CubeIDE

W lewym panelu środowiska programistycznego widoczne jest drzewo projektów powiązanych ze środowiskiem. Znajduje się tam także nowo utworzony na potrzeby pracy dyplomowej projekt o nazwie *PracaDyplomowaF446RE* wraz ze wszystkimi plikami konfiguracyjnymi. Z prawej strony widoczne jest domyślnie uruchomione okno pliku *PracaDyplomowaF446RE.ioc* z zakładkami *Pinout & Configuration*, *Clock Configuration*, *Project Manager* oraz *Tools*. Otwarta domyślnie widoczna na

powyższym rysunku zakładka *Pinout & Configuration* z prawej strony przedstawia graficzne interaktywne przedstawienie wyprowadzeń (*pinów*) mikrokontrolera. Z lewej strony znajduje się rozwijana lista konfiguracji zasobów mikrokontrolera w takich jak *Timers* czy *Analog*, których konfiguracja zostanie przedstawiona w dalszej części. Na wstępie należało dokonać konfiguracji zegarów mikrokontrolera w zakładce *Clock Configuration* widocznej na rysunku 2.8.



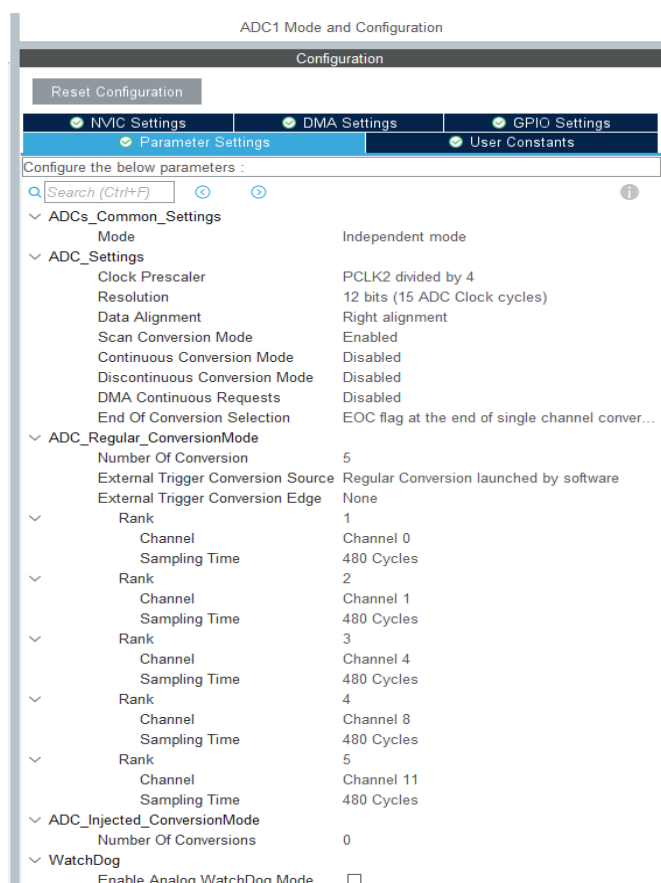
Rys. 2.8 Fragment okna konfiguracyjnego zegara mikrokontrolera STM32 Nucelo-F446RE z zakładki *Clock Configuration* w oprogramowaniu STM32CubeIDE

Przedstawiony mikrokontroler taktowany może być przez wiele rodzajów zegarów zewnętrznych i wewnętrznych o różnej częstotliwości taktowania takich jak widoczne z lewej strony na rysunku 2.8 LSE (ang. *Low Speed External*), LSI (ang. *Low Speed Internal*), HSI (ang. *High Speed Internal*) oraz HSE (ang. *High Speed External*). Zazwyczaj zegary zewnętrzne charakteryzują się większą dokładnością oraz szybkością w porównaniu z zegarami wewnętrznymi i tak jest także w tym przypadku. Główny sygnał zegarowy w projekcie pochodzi jednak z HSI taktowanego częstotliwością 16MHz oraz przechodzi przez szereg dzielników i mnożników częstotliwości, multiplekserów oraz pętlę PLL (ang. *Phase-Locked Loop*) służącej do generowania sygnałów o określonej częstotliwości i dokładności. W efekcie tego częstotliwość wyjściowa wynosi 84MHz na wszystkich szynach poza APB1 (ang. *Advanced Peripheral Bus 1*), a dokładniej PCLK1, gdzie jest to częstotliwość 42MHz, która nie będzie używana w projekcie [66, 67, 68]. Częstotliwość 84MHz została wybrana ze względu na to, że maksymalna częstotliwość na widocznych szynach zegarowych wynosiła dzięki użyciu HSI aż 90MHz będąca kłopotliwą w późniejszych obliczeniach

w przeciwieństwie do 84MHz. Te ustawienia kończą konfigurację zegarów mikrokontrolera.

Powracając do zakładki *Pinout & Configuration* oraz po rozwinięciu wcześniej wspomnianej listy *Analog* zawierającą ustawienia dotyczące przetworników *ADC1*, *ADC2*, *ADC3* i *DAC*. Postanowiono skorzystać z przetwornika *ADC1*, którego parametry konfiguracyjne zobrazowane zostały na rysunku 2.9. Ustawienia które nie zostały omówione pozostały w trybie domyślnym ze względu na rozbudowaną budowę peryferii, co tyczyć się będzie także reszty konfigurowanych komponentów.

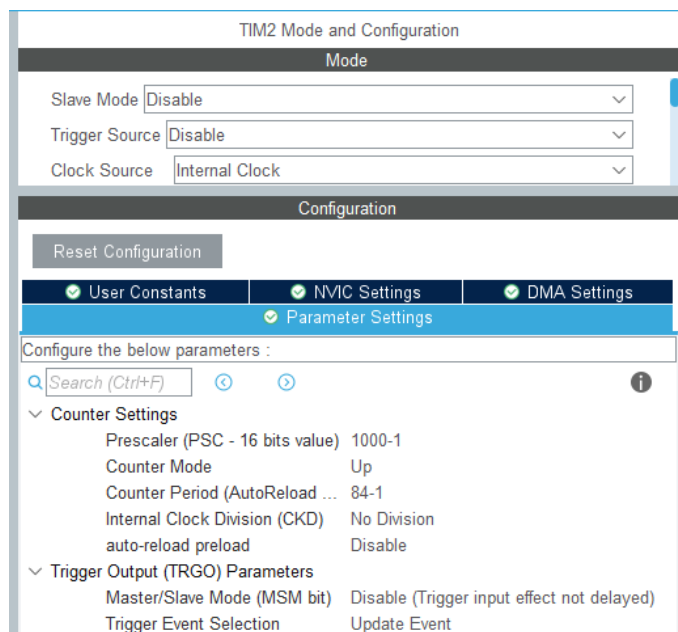
Na wstępie w dokumentacji mikrokontrolera sprawdzono źródło taktowania przetwornika *ADC1* [66]. Okazała się być nim szyna *APB2* taktowana jak wcześniej ustalono z częstotliwością 84MHz. W kolejnym kroku skonfigurowano *IN0*, *IN1*, *IN4*, *IN8* oraz *IN11* jako wejścia dla przetwornika zaznaczając odpowiednie pola.



Rys. 2.9 Okno konfiguracyjne użytego przetwornika analogowo-cyfrowego *ADC1* w programie *STM32CubeIDE*

Przetwornik działa w trybie *independent mode*, co oznacza że działa samodzielnie i niezależnie od innych przetworników. *Clock Prescaler* czyli dzielnik częstotliwości zegara ADC ustawiono na *PCLK divided by 4*, czyli podzielono częstotliwość taktowania przetwornika ($APB2 = 84\text{MHz}$) przez 4, co było minimalną dostępną wartością. Częstotliwość zegara przetwornika w wyniku tej operacji wyniosła 21MHz. *Resolution*, czyli rozdzielczość przetwornika ustawiono na maksymalną wartość, czyli 12 bitów zapewniające bardzo dobre odwzorowanie sygnału. Następnie włączono *Scan Conversion Mode*, co zapewniło automatyczną programową konwersję wartości przez przetwornik z każdego kanału. *Number Of Conversion* zostało ustawione na 5, gdyż z takiej ilości kanałów należało dokonać konwersji. Po ustawieniu tej wartości pojawiły się opcje konfiguracji od *Rank 1* do *Rank 5*. Opcji *Rank 1* przyporządkowano *Channel 0*, *Rank 2* jako *Channel 1*, *Rank 3* jako *Channel 4*, *Rank 4* jako *Channel 8* i *Rank 5* jako *Channel 11*. Nazwa kanału odpowiada przypisanym wejściom na płytce. Dla przykładu *Channel 0* jest odpowiednikiem wcześniej wymienionego wejścia *IN0*. Dla każdej z tych opcji *Sampling Time* ustawiono na maksymalną wartość równą 480 *Cycles*. Oznacza to, że sygnał będzie próbkowany przez 480 cykli zegara *PCLK2* podzielonego przez 4, a uśredniony wynik zostanie zapisany w odpowiednim rejestrze procesora. Operacja ta zapewnia dużo większą dokładność pomiaru i odporność na zakłócenia przy pomijalnym opóźnieniu, które można wyrazić obliczając stosunek ilości cykli do częstotliwości pracy przetwornika (21MHz), co wynosi w przybliżeniu 22,86μs na jedną próbkę sygnału [69]. Po skonfigurowaniu przetwornika należało dostosować ustawienia *timera*.

Timer jest modułem zliczającym cykle zegarowe zmniejszając lub zwiększając wartość wewnętrznego licznika, by w momencie osiągnięcia pewnej wartości wywołać daną czynność oraz rozpocząć kolejną iterację zliczania. Aby uzyskać pożądaną częstotliwość pracy *timera* należało dokonać jego konfiguracji. W tym celu wybrano *timer* oznaczony jako *TIM2* ze względu na jego bardziej rozbudowaną budowę, a jego okno konfiguracyjne przedstawione zostało na rysunku 2.10.



Rys. 2.10 Okno konfiguracyjne timera TIM2

Jako źródło taktowania czyli *Clock Source* wybrano *Internal Clock*, będący wewnętrznym zegarem. Po sprawdzeniu w dokumentacji mikroprocesora odczytano, że szyną sygnału zegarowego jest *APB1* o częstotliwości 84MHz widoczną także na rysunku 2.8. Producent podaje w dokumentacji wzór na obliczenie częstotliwości zdarzeń timera, który wygląda w następujący sposób:

$$Update_event = \frac{TIM_CLK}{((PSC + 1) \cdot (ARR + 1) \cdot (RCR + 1))} \quad (2.2)$$

Gdzie *TIM_CLK* jest częstotliwością sygnału zegarowego timera, *PSC* jest wartością rejestru preskalera, *ARR* jest wartość rejestru Auto-Reload oraz *RCR* jest wartością rejestru Repeat Counter [70, 71].

Pierwszym z dostępnych ustawień był *Preescaler* odpowiadający wartości rejestrów *PSC* o rozdzielczości 16 bitów, co daje maksymalną wartość dzielnika równą 65536. Postanowiono ustawić jego wartość na 1000 oraz odejmując 1, ponieważ rejestry indeksowane są od 0. *Counter Period* odpowiadający wartości rejestrów *RCR* ustawiono na wartość 84-1. *Auto-reload preload* odpowiadające rejestrowi *ARR* pozostawiono wyłączony, co w domyśle oznacza wartość 0. Podstawiając do równania wszystkie wartości oraz pamiętając, że *TIM_CLK* będący częstotliwością szyny sygnału *APB1* równej 84MHz otrzymano oczekiwaną wartość **1000Hz**. Na koniec ustawiono *Trigger Event Section* jako *Update event*, co oznacza, że, dana operacja lub funkcja zostanie wyzwolona po osiągnięciu zdarzenia *Update Event* w *timerze*. Na koniec

w zakładce *NVIC Setings* ustawiono wyzwolenie globalnego przerwania w momencie osiągnięcia zdarzenia (*Update_event*) zaznaczając *TIM2 global interrupt* [66, 67, 70, 71]. Po skonfigurowaniu wszystkich opcji użyto skrótu klawiszowego *ctrl+B* służącego do generacji kodu.

Po wygenerowaniu kodu dokonano jego edycji. Zaczęto od wybrania z drzewa projektu pliku *main.c* zawartego w podfolderze *Src* folderu *Core*. Opracowanie kodu zostało wykonane w języku programowania C będącym najpopularniejszym językiem programowania dla mikrokontrolerów głównie ze względu na niewielki rozmiar kodu oraz szybkość działania. Aby nie przedstawiać obszernego kodu wygenerowanego automatycznie podczas wcześniejszych operacji postanowiono przedstawić jedynie jego fragmenty. Pierwszy fragment został przedstawiony na listingu 2.1. Ponieważ środowisko wytycza specjalne fragmenty celu łatwiejszej edycji kodu oraz spójności z trybem graficznym przedstawionym uprzednio edycji kodu można dokonać jedynie w wytyczonych fragmentach, które jak w listingu rozpoczynają się od komentarza */* USER CODE BEGIN 0 */* oraz kończy po komentarzu */* USER CODE END 0 */*.

Listing 2.1.

```
/* USER CODE BEGIN 0 */

uint16_t ADC_VAL[5];

void ADC_Select_CH0(void)
{
    ADC_ChannelConfTypeDef sConfig = { 0 };
    sConfig.Channel = ADC_CHANNEL_0;
    sConfig.Rank = 1;
    sConfig.SamplingTime = ADC_SAMPLETIME_480CYCLES;

    if (HAL_ADC_ConfigChannel(&hadc1, &sConfig) != HAL_OK)
    {
        Error_Handler();
    }
}

.
.
.

/* USER CODE END 0 */
```

Powyższy kod obrazuje zadeklarowanie zmiennej tablicowej typu `uint16_t` do przechowywania danych z przetwornika oraz stworzenie funkcji **ADC_Select_CH0** typu `void` nie zwracającej żadnej wartości. Wewnątrz funkcji deklarowana jest zmienna 'sConfig' typu 'ADC_ChannelConfTypeDef', która jest strukturą używaną do konfiguracji kanałów ADC. W strukturze *sConfig* ustawiane są parametry konfiguracyjne dla konkretnego kanału, którym jest *ADC_CHANNEL_0*. Następnie, w warunku *if*, wołowana zostaje funkcja *HAL_ADC_ConfigChannel*, która konfiguruje wybrany przetwornik ADC na podstawie ustawień zawartych w strukturze *sConfig*. Jeśli funkcja *HAL_ADC_ConfigChannel* zwróci wartość różną od *HAL_OK*, oznaczać to będzie, że konfiguracja nie powiodła się, oraz wywołana zostaje funkcja *Error_Handler()*, która obsługuje błąd. Konfiguracja ta jest powielona dla pozostałych kanałów, co symbolizują trzy kropki umieszczone w kodzie. Zmienione zostały dla pozostałych kanałów jedynie nazwy wywoływanych funkcji. W następnym fragmencie listingu 2.2 przedstawionym poniżej zainicjowano start skonfigurowanego *timera* w trybie przerwania.

Listing 2.2.

```
/* USER CODE BEGIN 2 */

// Start timer
HAL_TIM_Base_Start_IT(&htim2);

/* USER CODE END 2 */
```

W następnym listingu 2.3 przedstawiony poniżej widoczna jest implementacja funkcji zwrotnej **HAL_TIM_PeriodElapsedCallback**, która jest wywoływana, gdy następuje przerwanie odliczania wskazanego *timera* (*timer2*). Funkcja zwrotna (*callback*) jest to funkcja, która jest przekazywana jako argument do innej funkcji lub modułu, aby została wywołana w odpowiednim momencie lub w odpowiedzi na określone zdarzenie.

W przypadku przedstawionego kodu **HAL_TIM_PeriodElapsedCallback** jest funkcją zwrotną, która jest automatycznie wywoływana przez bibliotekę *Cube HAL* w momencie, gdy następuje przepełnienie licznika oraz wywoływane jest z tego

powodu wcześniej skonfigurowane przerwanie. Funkcja przyjmuje jako argument wskaźnik na strukturę `TIM_HandleTypeDef` przechowującą konfigurację timera.

Listing 2.3.

```
/* USER CODE BEGIN 4 */
void HAL_TIM_PeriodElapsedCallback(TIM_HandleTypeDef *htim)
{
    if (htim == &htim2)
    {
        ADC_Select_CH0();
        HAL_ADC_Start(&hadc1);
        HAL_ADC_PollForConversion(&hadc1, 1000);
        ADC_VAL[0] = HAL_ADC_GetValue(&hadc1);
        HAL_ADC_Stop(&hadc1);

        ADC_Select_CH1();
        HAL_ADC_Start(&hadc1);
        HAL_ADC_PollForConversion(&hadc1, 1000);
        ADC_VAL[1] = HAL_ADC_GetValue(&hadc1);
        HAL_ADC_Stop(&hadc1);

        ADC_Select_CH4();
        HAL_ADC_Start(&hadc1);
        HAL_ADC_PollForConversion(&hadc1, 1000);
        ADC_VAL[2] = HAL_ADC_GetValue(&hadc1);
        HAL_ADC_Stop(&hadc1);

        ADC_Select_CH8();
        HAL_ADC_Start(&hadc1);
        HAL_ADC_PollForConversion(&hadc1, 1000);
        ADC_VAL[3] = HAL_ADC_GetValue(&hadc1);
        HAL_ADC_Stop(&hadc1);

        ADC_Select_CH11();
        HAL_ADC_Start(&hadc1);
        HAL_ADC_PollForConversion(&hadc1, 1000);
        ADC_VAL[4] = HAL_ADC_GetValue(&hadc1);
        HAL_ADC_Stop(&hadc1);

        HAL_GPIO_TogglePin(LD2_GPIO_Port, LD2_Pin);
    }
}
/* USER CODE END 4 */
```

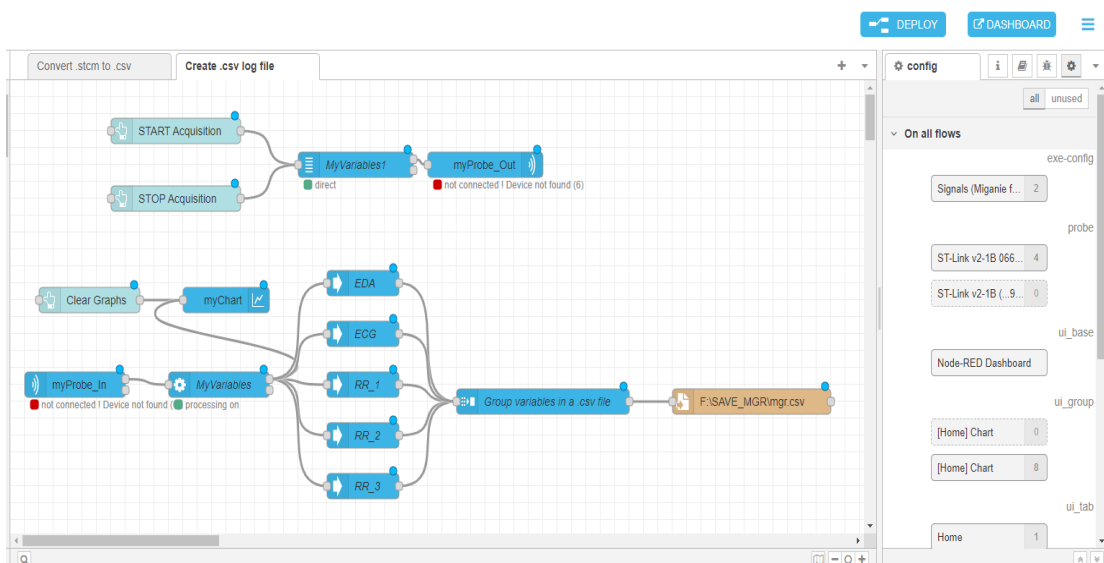
Wewnątrz funkcji znajduje się warunek `if` sprawdzający czy przerwanie pochodzi z `timera2`. Jeśli warunek jest spełniony wywoływana zostaje funkcja

ADC_Select_CH0() napisana wcześniej, która konfiguruje przetwornik ADC dla kanału CH0. Następnie rozpoczyna się konwersja analogowo-cyfrowa dla wybranego kanału przy użyciu funkcji HAL_ADC_Start(&hadc1). Kolejna linia kodu wywołuje funkcję oczekującą na zakończenie konwersji z ograniczeniem czasowym równym 1000ms. Następnie zapisywany zostaje wynik do zmiennej tablicowej używając funkcji HAL_ADC_GetValue(&hadc1). Na koniec wywoływana jest funkcja kończąca konwersję dla wybranego kanału. Sekwencja tych działań następnie jest powielana dla każdego kanału. Na końcu przedstawionego kodu znajduje się funkcja powodująca miganie diodą zieloną z częstotliwością wywoływania przerwania przez timer.

Tak napisany kod został skompilowany, zapisany oraz wgrany do mikroprocesora płytki deweloperskiej. Po wykonaniu tych działań do wejścia na płytce oznaczonego jako A0 podpięte zostało wyjście sygnałowe modułu „Grove - GSR sensor”, do wejścia A1 wyjście modułu „SparkFun Single Lead Heart Rate Monitor - AD8232” oraz do pozostałych skonfigurowanych wejść podłączono moduł przepływu powietrza. Wejście A2 jako wyjście czujnika nozdrza lewego, A4 jako wejście nozdrza prawego i A3 jako wejście czujnika z okolic ust [66].

2.2.2 Wizualizacja i zapis danych

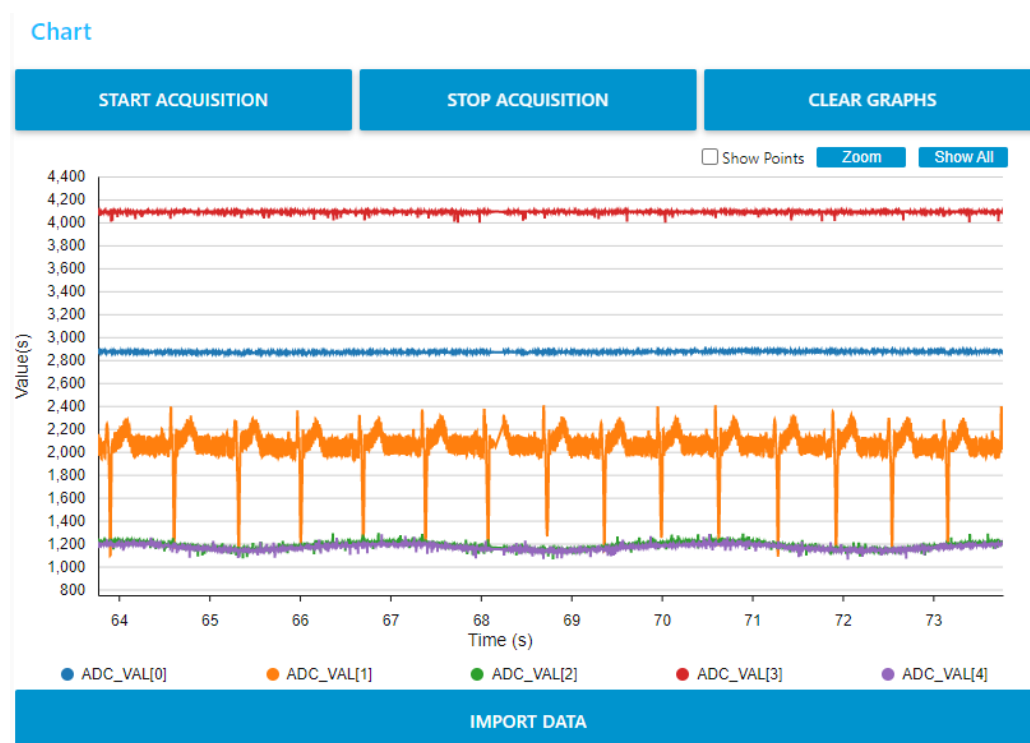
Aby zwizualizować wyniki akwizycji oraz je zapisać postanowiono użyć oprogramowania STM32CubeMonitor widocznego na rysunku 2.11.



Rys. 2.11 Schemat blokowy do wizualizacji oraz zapisu danych w oprogramowaniu STM32CubeMonitor

Zrealizowano to umieszczając odpowiednie konfigurowalne bloki w polu roboczym. Oprogramowanie umożliwia także zaimportowanie kodu zewnętrznego. Producent na swojej oficjalnej stronie udostępnia instrukcje do utworzenia przykładowych schematów oraz kod do zaimportowania programów, z czego skorzystano. Wyniki zaimportowanego oraz skonfigurowanego schematu widoczne są przedstawionym powyżej rysunku [72].

W górnej części rysunku 2.11 znajdujące się bloki służą do akwizycji sygnału wejściowego. W bloku *MyVariables1* podano ścieżkę do pliku wykonywalnego o rozszerzeniu .elf (ang. *Executable and Linkable Format*). Plik o rozszerzeniu .elf zawiera skompilowany kod maszynowy z napisany wcześniej w *STM32CubeIDE*. W bloku tym także wybrano zmienne *ADC_VAL[0]*, *ADC_VAL[1]*, *ADC_VAL[2]*, *ADC_VAL[3]* oraz *ADC_VAL[4]*, które będą obserwowane na wykresie oraz zapisywane do pliku. Blok *myProbe_Out* jest odpowiedzialny za otwarcie i zamknięcie połączenia jak i wysyła polecenia do wybranej sondy jaką jest w tym przypadku *ST-Link v2-1B*. Poniżej znajduje się schemat realizujący zapis odczytywanych wartości do pliku o rozszerzeniu .csv. Schemat ten rozpoczyna się od bloku *myProbe_In* służącego do odbioru danych z wybranej sondy *ST-Link v2-1B*. Następnie w bloku *MyVariables* wybrano z listy zmienne które następnie będą wyświetlane oraz obserwowane czyli *ADC_VAL*. Następnie kolejne pięć bloków nazwano adekwatnie do rodzaju rejestrowanego sygnału oraz przyporządkowano odpowiadające im zmienne. W kolejnym bloku o nazwie „*Group variables in a .csv file*” zadeklarowano liczbę sygnałów które będą zapisywane, czyli 5. W ostatnim bloku o nazwie „*F:\SAVE_MGR\mgr.csv*” podano absolutną ścieżkę do zapisu danych z czujników. Po zadeklarowaniu wszystkich wymienionych wartości oraz kliknięciu w prawym górnym rogu przycisku **DEPLOY** wszystkie okręgi obok bloków powinny zmienić kolor z czerwonego na zielony jeśli wszystkie urządzenia są ze sobą podłączone, a mikrokontroler jest podpięty do komputera. Po wykonaniu wszystkich wymienionych kroków oraz wybraniu w górnej prawej części ekranu opcji **DASHBOARD** ukazuje się okno przedstawione na rysunku 2.12.



Rys. 2.12 Okno programu STM32CubeMonitor wizualizujące i zapisujące wyniki pomiarów

W oknie tym po wciśnięciu START ACQUISITION rozpocznie się rysowanie przebiegów z wszystkich dotychczas skonfigurowanych czujników oraz jednocześnie ich zapis pliku w formacie csv. Przycisk STOP ACQUISITION kończy rejestrację oraz obrazowanie sygnałów. Przycisk CLEAR GRAPHICS służy do wyczyszczenia ekranu roboczego z pomiarów.

2.2.3 Akwizycja danych oraz wykonanie i przebieg badań

Zgodnie z założeniami projektowymi ze względu na brak dostępu do dedykowanej bazy afektywnej do badań wykorzystano treści wideo zamieszczone na platformie YouTube. Podyktowane to było ogromną bazą materiałów wysokiej rozdzielczości oraz jakości, a także ich łatwym dostępem w przeciwieństwie do specjalistycznych baz afektywnych. Ponadto warto zauważyć, iż wspomniane bazy danych emocji także bardzo często wykorzystywały materiały zamieszczone w serwisie YouTube [49, 50].

Po przeprowadzeniu przez Autora wywiadu z osobami pochodzącymi z różnych środowisk, uwzględniając komentarze pod materiałami w serwisie YouTube oraz bazując na własnych odczuciach i percepcji, do badań wybrano 6 materiałów wideo,

które uznano za wyraźnie i powszechnie wywołujące określone emocje. Jednak aby określić zmiany zachodzące podczas indukcji danej emocji należało w pierwszej kolejności zdefiniować tzw. stan odniesienia. Stanem tym jest stan neutralny z uwagi na to, że towarzyszy on ludziom przez większość czasu. Ponadto charakteryzuje się łatowścią wywołania i dość prostym opisem. Do wywołania tego stanu posłużono się wyświetleniem neutralnego czarnego obrazu na pełnym ekranie monitora. Aby zapewnić różróżnialność wyników postanowiono wybrać oprócz emocji neutralnej sześć innych jakimi są: *radość*, *rozbawienie*, *smutek*, *złość*, *wstręt do okaleczeń* oraz *ulga* w efekcie czego łącznie zdefiniowano siedem stanów emocjonalnych.

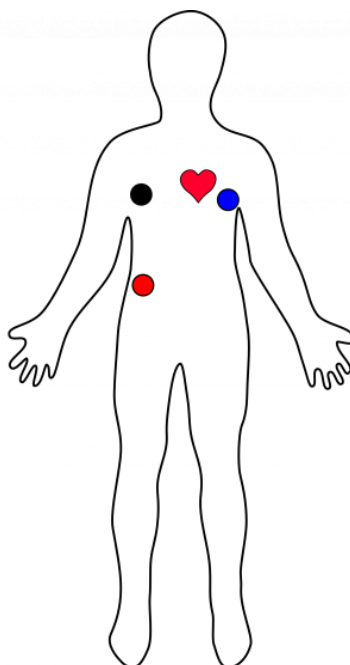
Emocję typu „radość” wywoływano przy użyciu utworu wraz z teledyskiem autorstwa Pharela Williamsa pt. „*Happy*” [73]. Utwór ten został wybrany ze względu na treść piosenki manifestującej stan radości. Słowa: „*because I’m happy*”, mogą wpływać na podświadomość i pozytywny odbiór treści („*fake it till you make it*”). Następną emocją było rozbawienie. W tym celu wybrano wideo przedstawiające śmieszne sytuacje związane z kotami [74]. Fragmenty te przedstawiały je najczęściej w rutynowych sytuacjach z życia codziennego. To powodowało najczęściej szeroki uśmiech na twarzy respondentów oraz spontaniczne wydechy powietrza, co świadczy o poprawnym doborze materiału. Kolejną emocją był smutek. Wywoływany był poprzez wyświetlenie fragmentu filmu animowanego „*Król lew*” przedstawiającego śmierć ojca głównego bohatera filmu spowodowaną przez jego brata, którą widział syn uśmierconego lwa [75]. Fragment ten przez wiele towarzyszących mu elementów jest powszechnie uznawany za jeden ze smutniejszych kadrów w historii kinematografii. Kolejną emocją była złość wywoływana poprzez przedstawienie wiadomości politycznych dotyczących „fit for 55” w programie informacyjnym telewizji rządowej TVP INFO [76]. Tematyka programu dotyczyła mniejszości seksualnych, decyzji ekonomicznych dotyczących ograniczenia emisji CO₂, działań opozycji i partii rządzącej oraz wielu treści pobocznych. Materiał ten poruszał szeroką gamę tematów kontrowersyjnych przy czym założono, że, z pośród wymienionych choć jeden z elementów może wywołać złość u badanego. Kolejną z emocji był wstręt do okaleczeń, co zostało wywołane poprzez przedstawienie fragmentów krwawych i brutalnych walk w ringu z użyciem szklanych świetłówek, krzesel oraz innych ostrych narzędzi [77]. Fragment ten wywoływał u respondentów widoczne kaprysy na twarzy oraz nerwowe ruchy prawą ręką. Ostatnią z emocji była *ulga* wynikająca z zatrzymania prezentowania wcześniejszego materiału przysparzającego dużą dawkę stresu, a spowodowana spokojną, kojącą i rytmiczną

muzyką instrumentalną lo-fi z obrazem piszącej młodej dziewczyny w przytulnym pokoju z kotem machającym ogonem [78]. Materiał ten był wyświetlany nagle przerywając brutalne sceny. Emocje wywoływane były w wymienionej kolejności aby zachować pewną przyczynowość emocji oraz spójność, która pozwalała na zachowanie jak największej dozy immersji badanych.

Każdy pomiar trwał około 2 minuty, a całość badania trwała 30 minut, o czym informowano respondentów przed rozpoczęciem badań, w momencie założenia pełnej aparatury. Sugerowane było obranie komfortowej pozycji, ze względu na długość trwania badania oraz faktu, iż, w trakcie pomiarów należało zachować od wskazanego momentu ciszę oraz dodatkowo ograniczyć do minimum ruch, który zakłócał akwizycję sygnałów i w efekcie przebieg badania. W następnej kolejności respondenci byli informowani, że oglądać będą materiały wideo mające na celu wywołanie u nich określonych emocji. Każdą z emocji oprócz wstrętu do okaleczeń oraz ulgi zaczynało rejestrować przynajmniej po minucie od przedstawienia materiału wideo/obrazu aby zaindukować w nich pożądane emocje. Wyjątek stanowi zarejestrowanie emocji neutralnej, gdzie pomiar nastąpił po 3 minutach, z uwagi na konieczność kalibracji modułów pomiarowych. Rejestracja rozpoczynała się w miarę możliwości u każdego badanego w tym samym fragmencie wyświetlanego materiału. Ewentualny wpływ na opóźnienie rejestracji miał moment kliknięcia przycisku START ACQUISITION w programie *STMCubeMonitor*, odpowiedzialnym za rejestrację badań. Wszystkie materiały były wyświetlane na laptopie *MSI GV72 8RC* z ekranem o przekątnej 17,3 cala oraz wbudowanymi głośnikami, co umożliwiło komfortowy odbiór danych przez respondentów oraz jednoczesną akwizycję danych na dysk typu SSD. Badania były realizowane zazwyczaj w mieszkaniach respondentów. Z tego względu aby jak najbardziej uniezależnić rejestrację sygnałów od warunków zewnętrznych usuwano wszystkie rozpraszające elementy z pola widzenia badania, tak aby pozostał jedynie laptop. Zaletą takiego rozwiązania jest fakt, iż respondenci nie byli dodatkowo zestresowani zmianą środowiska pozostając w dobrze im znanym miejscu..

W pomieszczeniu oprócz respondenta oraz osoby przeprowadzającej badania nie znajdowała się żadna inna osoba. Przyklejenie elektrod do badania EKG było wykonywane przez osobę nadzorującą badanie, natomiast ułożenie elektrod do pomiaru sygnału EDA oraz respiracji było ściśle uzgadniane z respondentem. Elektrody do akwizycji sygnału EDA po przeprowadzonej dogłębnej analizie w rozdziale poprzednim umieszczano na lewej ręce środkowego oraz wskazującego palca. Dodatkowo

poinstruowano każdego z respondentów, aby metalowe elektrody wyczuwalne w materiałowych rękawach były umieszczone na środku opuszków palców. Przed przymocowaniem elektrod do badania EKG przemywano dokładnie izopropanolem dane fragmenty skóry, do którego podłączono elektrody, aby usunąć obumarły naskórek pogarszający właściwości przewodzące. Producent modułu *SprakFun* na swojej stronie przedstawia schemat podłączeń elektrod do ciała zobrazowany na rysunku 2.13 [79]. Przedstawia on realizację *trójkąta Einthovena*. Ułożenie to zapewnia obserwację najważniejszych zależności między pracującymi komorami serca. Dodatkowym atutem tego ułożenia jest bliskość elektrod od serca, co skutkuje wyraźniejszym zobrazowaniem, które jest bardziej odporne na przekłamanie i szumy.



Rys. 2.13 Schemat podłączeń elektrod do ciała [80]

Elektrody oznaczone kolorem **czarnym** oraz **niebieskim** przymocowane zostały w wgłębieniu poniżej obojczyka precyzyjnie nazywanym jako region pachwowy lewy oraz prawy [80]. Elektrode w kolorze **czerwonym** umieszczano na ostatnim z żeber w linii prostej poniżej elektrody w kolorze czarnym. Dodatkowo proszono osoby badane o zdjęcie metalowych łańcuszków z szyi oraz wszelkiego rodzaju biżuterii, która zakłócać mogła przebieg badania. Moduł akwizycji przepływu powietrza umieszczano pod nosem, a przewody zasilające i sygnałowe poprowadzono opierając je za uszami respondenta podobnie jak noszone są okulary. Następnie regulowano docisk

modułu do ciała w celu ustawienia urządzenia w poprawnej, stałej i komfortowej dla respondenta pozycji. Termistory pomiarowe modułu ze względu na swoją budowę i rodzaj obudowy umożliwiały dostosowanie się do twarzy badanego. Instruowano respondentów, aby termistory umieszczone były pod każdym z nozdrzy tak aby nie miały bezpośredniego kontaktu z ciałem, co mogłoby powodować zafałszowanie wyników. Tak wyregulowane termistory utrzymywały się w niemal niezmienniej pozycji podczas trwania badania. Podczas przeprowadzania badania Autor pozostawał za respondentem w odległości nie przekraczającej 3 metrów w celu nie wzbudzania dodatkowych emocji. Jego obecność ograniczała się do rejestracji pomiarów oraz zmian treści video. Na rysunku 2.14 przedstawiono respondenta w pełnej aparaturze pomiarowej podczas badania.



Rys. 2.14 Respondent z założoną aparaturą pomiarową podczas próbnego badania

Respondent ma umieszczone elektrody EKG pod bluzką, a fotografię tę wykonano podczas badania próbnego, kiedy testowano działanie modułów pomiarowych, poprawność zapisu danych, kolejność wywoływania emocji, ułożenie elementów oraz przebieg badania. Powyższy rysunek przedstawia ostateczną konfigurację badania.

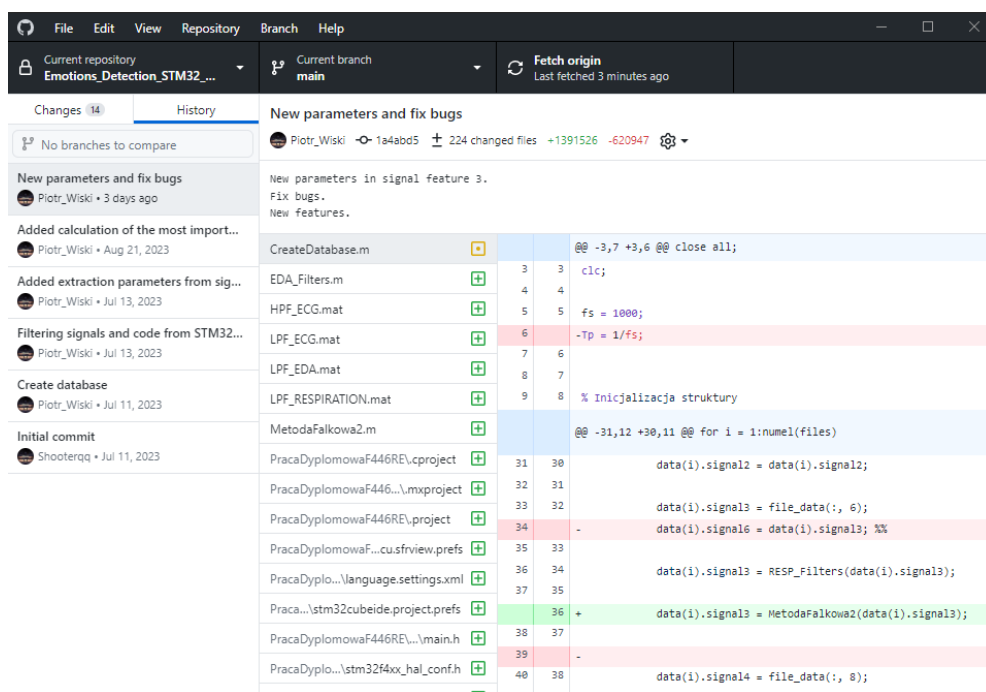
Najbardziej znaczącą różnicą podczas przeprowadzania badań było używanie w miarę możliwości krzeseł z podłokietnikami dzięki czemu ręce respondentów mniej się męczyły. Badanie zostało przeprowadzone na 20 osobach w wieku od 18 do 60 roku życia ze zdecydowaną przewagą osób w wieku 23 – 25 lat. Wszystkie osoby były niepalące. Spośród badanych siedem osób to kobiety, a trzymastu to mężczyźni. Każdy z zarejestrowanych sygnałów zapisywano w oddzielnym pliku o unikatowej nazwie z uzupełnianymi w późniejszym czasie metadanymi obejmującymi rodzaj emocji, reprezentowaną liczbą od 0 do 6 odpowiadającym kolejności ich wywoływania. Dodatkowo każda osoba otrzymała indywidualny numer identyfikacyjny (*ID*) oraz akronim płci (1 - mężczyzna, 0 kobieta). Pierwotnie wprowadzono również etykiety wagi i wzrostu, jednak podejście to zostało odrzucone ze względu na zbyt małą ilość danych do takiej analizy oraz potencjalnie niewielką korelację tych parametrów ze stanem emocjonalnym. Wszystkie badania wykonano w okresie czasu od 01.06.2023 r. do 31.08.2023 r. Po wstępnej analizie zarejestrowanych sygnałów stwierdzono, iż nie wszystkie z nich mogą zostać wykorzystane do dalszych eksperymentów, głównie ze względu na fakt, iż w trakcie badań dochodziło m.in. do niekontrolowanych uszkodzeń końcówek przewodów modułu przepływu powietrza, słabego dolegania elektrod EKG do ciała (prawdopodobnie spowodowanego przez pot), ruchu badanych czy zbyt wysokiej temperatury otoczenia spowodowanej uwcześnie panującymi upałami, co powodowało nadmierne pocenie się respondentów. To z kolei powodowało że elektrody do pomiaru reakcji elektrodermalnej nie rejestrowały pożądaných wartości po pewnym czasie badania, z powodu nadmiernej wilgoci. Tego rodzaju zdarzenia były powodem wykluczenia pomiarów 6 badanych w efekcie czego pozostało 14 unikalnych badań eksperymentalnych. Błędy te zostały zidentyfikowane dopiero na etapie wstępnej analizy sygnałów i ekstrakcji cech, ponieważ w czasie przeprowadzania badania nie było możliwości sprawdzenia jakości sygnałów przez wyświetlany na całym ekranie materiał wywołujący emocje. Potencjalnym rozwiązaniem tego problemu mógłby być drugi monitor, gdzie możliwa byłaby wizualna analiza zarejestrowanych sygnałów. Wszystkie zarejestrowane dane zajmują na dysku łącznie 1,08GB, a średni rozmiar jednego pliku to 7,89MB. Tak zdefiniowana obszerna baza danych jest podstawą do analizy opisane w rozdziale trzecim pracy.

3. Opracowanie algorytmu ekstrakcji cech

Analizę danych w celu identyfikacji emocji przeprowadzono przy użyciu oprogramowania MATLAB w wersji R2021a. Wybór ten podyktowany był dość prostą składnią programowania, wieloma wbudowanymi aplikacjami i modułami do filtrowania danych, udziałem społeczności oprogramowania, aplikacjami do uczenia i implementacji sieci neuronowych, łatwością wizualizacji danych i szeroką kompatybilnością polegająca między innymi na możliwości konwersji kodu napisanego w oprogramowaniu na inne języki programowania, w tym C.

3.1 System kontroli wersji GitHub

Aby usprawnić wymianę informacji oraz postępów w kodzie między *Promotorem* i *Autorem* postanowiono wykorzystać system kontroli wersji GitHub oparty na systemie rozproszonym Git [81]. Platforma ta umożliwia tworzenie repozytoriów zawierających przykładowo kod źródłowy programu jak zrobiono w tym przypadku. Na rysunku 3.1 przedstawiono okno robocze oficjalnej aplikacji GitHub na komputer.



Rys. 3.1 Okno aplikacji GitHub z widoczną historią zmian kodu projektu

Powodem wykorzystania platformy była możliwość tworzenia różnych wersji kodu na przestrzeni rozwijania oprogramowania, do których można powrócić w niemal dowolnym momencie bez konsekwencji, a zmiany widoczne są dla zespołu zaraz po zamieszczeniu kodu w repozytorium tworząc i wypychając zmiany. Osoby w zespole niezależnie od miejsca mogą pracować nad tym samym kodem równocześnie na swoich komputerach ładując kod z repozytorium. Dzięki takiemu działaniu możliwa jest do prześledzenia pełna historia zmian kodu wraz z opisami, a to pozwala łatwiej kontrolować jego rozwój. Kolejną z ważnych zalet tego rodzaju rozwiązania jest możliwość pracy na tzw. gałęziach (*branches*). Gałęzie to osobne miejsca gdzie można pracować nad kodem niezależnie od innych miejsc zawierających kod czyli gałęzi niemniej jednak nadal w ramach jednego projektu. W ramach jednego projektu gałęzie mogą się w pewnym momencie połączyć. Przykładowo można w osobnej gałęzi stworzyć nową funkcjonalność, którą następnie dołącza się do głównej gałęzi oprogramowania kiedy jest gotowa i przetestowana, dzięki czemu do kodu zostają wdrożone nowe elementy unikając wielu kłopotów jak kolizje czy błędy oprogramowania czy bezpieczeństwa. Dzięki użyciu przedstawionych funkcjonalności w rozwoju oprogramowania zyskano wiele korzyści wynikających z uporządkowania kodu w efekcie czego znacznie przyspieszono pracę.

3.2 Konwersja oraz filtracja zarejestrowanych sygnałów

Pliki zawierające wyniki badań zapisane w formacie *csv* wraz z metadanymi postanowiono początkowo przekonwertować na format wygodny do analizy. W tym celu stworzono program o nazwie *CreateDatabase* tworzący strukturę danych o nazwie *data*, do której zapisywano wszystkie odfiltrowane sygnały pochodzące z czujników biometrycznych. Istotny fragment kodu programu *CreateDatabase* przedstawiono poniżej w listingu 3.1.

Listing 3.1.

```
% Filtracja i zapis sygnału EDA
data(i).signal1 = file_data(:, 2);
data(i).signal1 = EDA_Filters(data(i).signal1);

% Filtracja i zapis sygnału EKG
data(i).signal2 = file_data(:, 4);

% Filtracja i zapis sygnału respiracji
data(i).signal3 = file_data(:, 6);
```

```

data(i).signal3 = RESP_Filters(data(i).signal3);
data(i).signal3 = Wavelet_Filters(data(i).signal3);

% Filtracja i zapis sygnału respiracji
data(i).signal4 = file_data(:, 8);

% Filtracja i zapis sygnału respiracji
data(i).signal5 = file_data(:, 10);

% Zapis metadanych dotyczących sygnału
data(i).age = file_data(1, 11);
data(i).height = file_data(1, 12);
data(i).weight = file_data(1, 13);
data(i).gender = file_data(1, 14);
data(i).fs = file_data(1, 15);
data(i).emotion = file_data(1, 16);
data(i).id = file_data(1, 17);

```

Podczas jednej iteracji zwiększającej wartość zmiennej *i* wewnątrz struktury *data* tworzono jeden wiersz odpowiadający zarejestrowanej jednej emocji zapisanej w pojedynczym pliku csv. W pierwszej kolumnie struktury o nazwie *signal1* danego wiersza zapisywano sygnał aktywności elektrodermalnej. Następnie nadpisywano dane te wywołując funkcję *EDA_Filters*, której argumentem była zmienna *signal1*. Funkcja ta stworzona została do filtracji tego rodzaju sygnałów. Wewnątrz omawianej funkcji zawiera się kod przedstawiony w listingu 3.2.

Listing 3.2.

```

function EDA_F = EDA_Filters(edaf)

    load('LPF_EDA.mat');

    outputEDA = filter(LPF_EDA, 1, edaf);
    outputEDA = outputEDA(50:end);

    filtered_eda = movmean(outputEDA, 15);

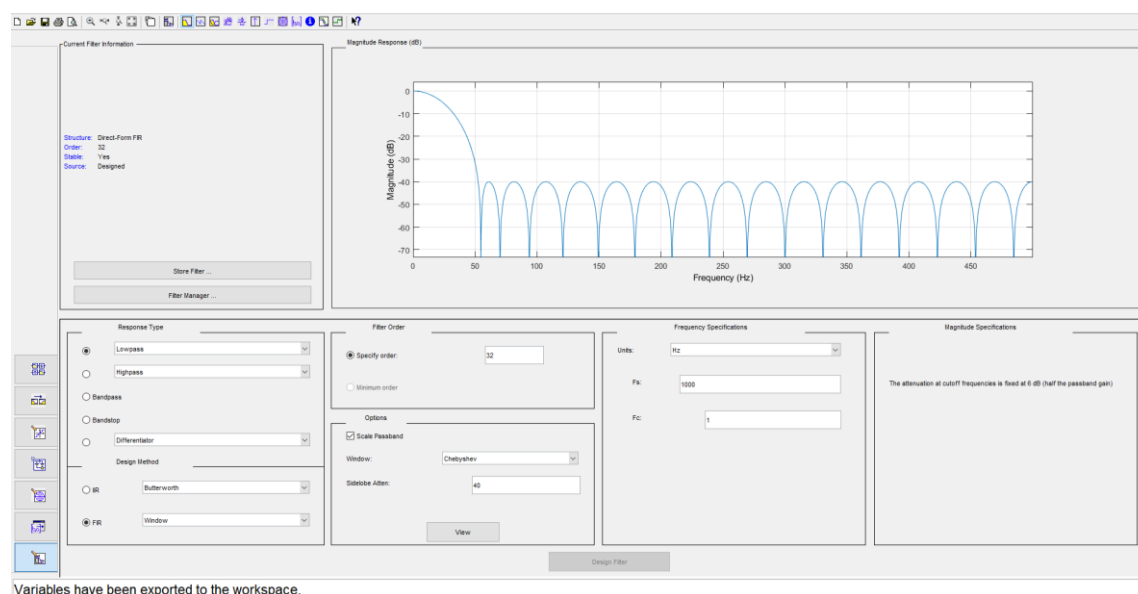
    EDA_F = filtered_eda;

end

```

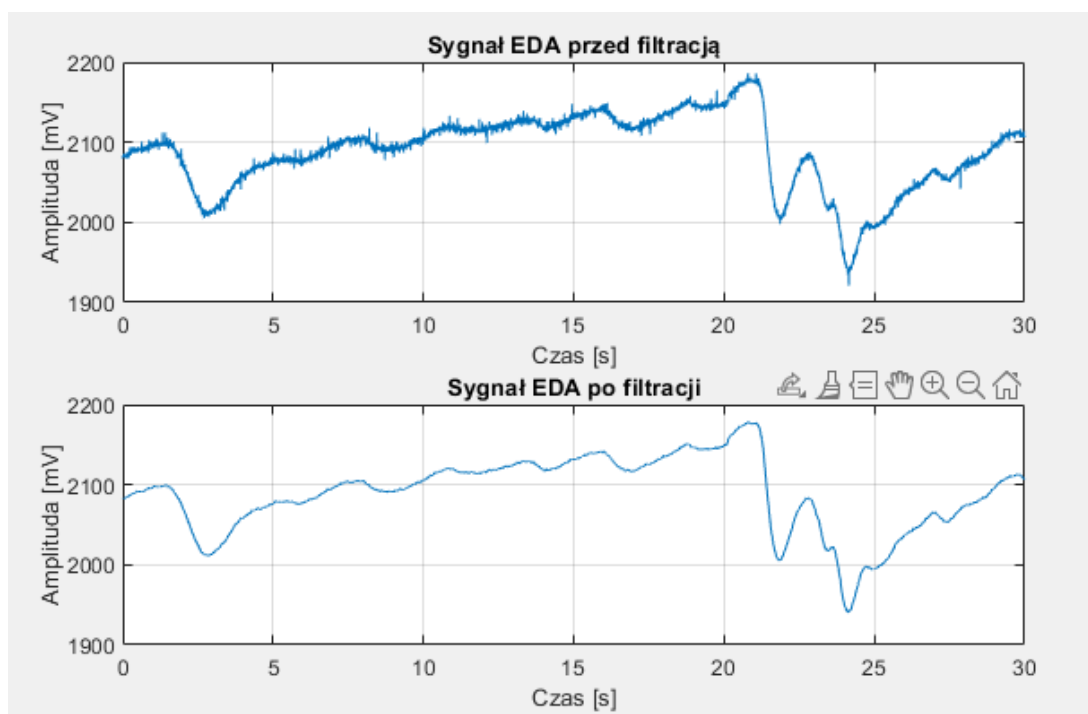
W pierwszej kolejności ładowany do przestrzeni zmiennych roboczych był wektor o nazwie *LPF_EDA.mat*. Wektor ten zawierał współczynniki zaprojektowanego uprzednio filtru za pomocą narzędzia o nazwie *Filter Designer* dostępnego w oprogramowaniu MATLAB. Na rysunku 3.2 przedstawiono okno robocze programu z ustalonymi parametrami filtru. Ponieważ sygnał wejściowy był niskiej jakości przez

szum kwantowania oraz niską rozdzielczość sygnału wyjściowego (spowodowaną rozdzielczością modułu przetwornika *Grove*) postanowiono zaprojektować filtr dolnoprzepustowy (LP – ang. *Lowpass filter*), który wyeliminuje wymienione niedoskonałości. W tym celu postanowiono wybrać filtr typu FIR (ang. *Finite Impulse Response*). Charakteryzuje się on skończoną odpowiedzią impulsową, liniowym przesunięciem fazowym, opóźnieniem równym ilości współczynników i gwarantowaną stabilnością. Głównymi wadami tego rodzaju filtru są duża złożoność obliczeniowa i stosunkowo szerokie pasmo przejściowe [82, 83].



Rys. 3.2 Zaprojektowany filtr dolnoprzepustowy FIR dla sygnału EDA

Zaprojektowano filtr 32 rzędu, co przekłada się na 33 współczynniki. Dodatkowo aby pozbyć się zafałowań charakterystyki częstotliwościowej sygnału wyjściowego spowodowanych filtracją użyto okna *Chebysheva* z parametrem *Sidelobe Atten* (tłumienie bocznych pasm) równym 40 odnoszącym się do poziomu tłumienia sygnału poza pasmem przepustowym filtru. W polu F_s (ang. *Frequency sampling*) podano częstotliwość sygnału filtrowanego, natomiast w polu F_c (ang. *Frequency cutoff*) częstotliwość odcięcia ustawiono na 1 Hz. Po przeprowadzeniu filtracji za pomocą funkcji `filter` dokonano dodatkowej filtracji za pomocą funkcji `movmean`, która działa jako filtr uśredniający o szerokości okna równego 15 próbek. Ostatecznie nadpisywano sygnał usuwając pierwsze 50 próbek sygnału ze względu na opóźnienie wprowadzone przez filtr FIR równe 33 próbki. Efekt filtracji widoczny jest na rysunku 3.3, gdzie zobrazowano fragment sygnału przed i po operacji filtrowania.

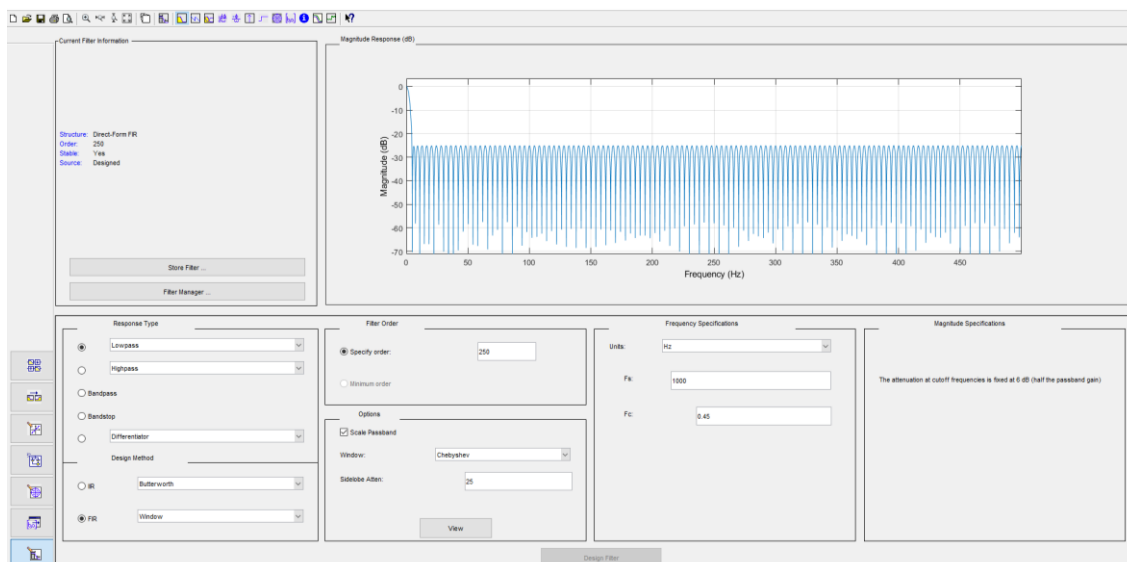


Rys. 3.3 Porównanie sygnału EDA przed oraz po operacji filtracji

Widoczne jest usunięcie szumu wysokoczęstotliwościowego zachowując pożądane właściwości sygnału EDA.

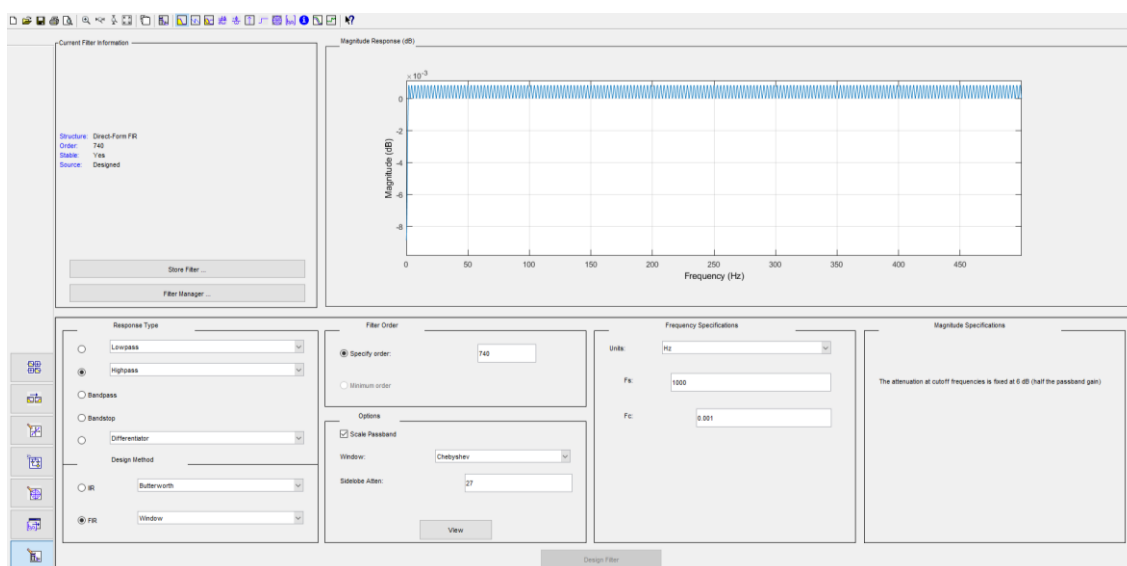
Następnym krokiem w pliku *CreateDatabase* było zapisanie sygnału EKG do wiersza drugiego w strukturze *data* o nazwie *signal2* bez filtracji, ponieważ w dalszej części pracy użyto funkcji, która filtruje sygnał lub funkcji nie wymagającej filtracji do przeprowadzenia analizy i parametryzacji sygnału.

Następnie w wierszu trzecim struktury *data* o nazwie *signal3* zapisano sygnał respiracji oraz poddano go filtracji używając filtru o parametrach przedstawionych na rysunku 3.4.



Rys. 3.4 Zaprojektowany filtr dolnoprzepustowy FIR dla sygnału respiracji

Dodatkowo zaprojektowano filtr górnoprzepustowy z parametrami podanymi na rysunku 3.5. Filtry te zastosowano w funkcji `RESP_Filters` na sygnale respiracyjnym aby usunąć wolnozmiennne i szybkozmiennne fluktuacje wynikające głównie z zakłóceń z układu zasilania.



Rys. 3.5 Zaprojektowany filtr górnoprzepustowy FIR dla sygnału respiracji

Po filtracji dolnoprzepustowej oraz górnoprzepustowej wewnątrz funkcji dokonywano usunięcia pierwszej sekundy pomiaru z powodu opóźnienia wynikającego z filtracji. Kolejno dokonywano filtracji uśredniającej za pomocą funkcji `movmean` o wielkości

okna równego 250 próbek. Na koniec wyznaczano wartość średnią sygnału, która zostawała od niego odjęta, aby końcowy sygnał oscylował w okolicy zera.

Tak wyfiltrowany wstępnie sygnał zostawał poddawany filtracji właściwej wewnątrz funkcji `Wavelet_Filters`, której kod przedstawiono na listingu 3.3.

Listing 3.3.

```
function filtered_signal = Wavelet_Filters(in_sig)

    wavelet_type = 'sym4'; % Typ falki
    level = 12;           % Ilość detali
    det_cofs = [];

    [cofs, levels] = wavedec(in_sig, level, wavelet_type);
    for k = 1:level
        det_cofs(:, k) = wrcoef('d', cofs, levels, wavelet_type, k);
    end

    % Suma detali użytych do rekonstrukcji sygnału
    det_sum = sum(det_cofs(:, 8:12), 2);

    % Zastosowanie filtru usredniającego
    filtered_det_sum = movmean(det_sum, 100);

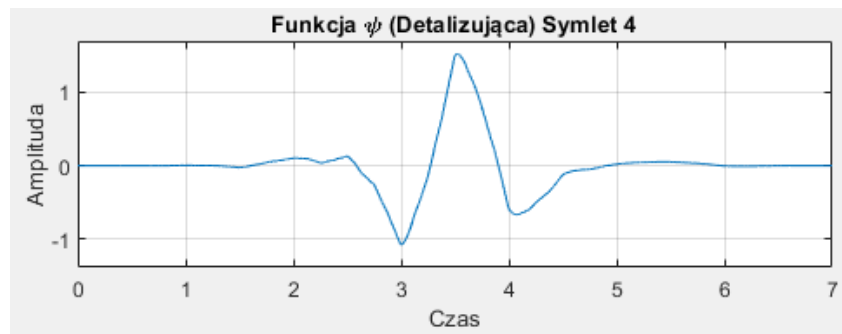
    filtered_signal = filtered_det_sum;
end
```

Wewnątrz przedstawionej funkcji zastosowano filtrację przy użyciu dyskretnej transformaty falkowej (DWT – ang. *Discrete Wavelet Transform*). Zasada jej działania jest podobna do transformaty Fouriera, lecz jądrem nie są funkcje sinusoidalne, a funkcje nazywane falkami, oznaczane symbolem Ψ . Falki to funkcje o określonych kształtach, które są skalowalne i przesuwane wzdłuż sygnału poddanego transformacji falkowej. Podczas tej operacji sygnał zostaje rozłożony na składowe częstotliwościowe i czasowe, które odpowiadają użytej falce w różnych skalach i położeniach czasowych. Dekompozycja sygnału jest dokonywana na podstawie korelacji kształtu sygnału z falką. Aby uzyskać wyodrębnienie cech charakterystycznych powiązanych z kształtem użytej falki z sygnału w różnych kontekstach czasowych i częstotliwościowych na przestrzeni całego sygnału dokonuje się odwrotnej transformaty falkowej (IDWT – ang. *Inverse Discrete Wavelet Transform*). W efekcie tych operacji możliwe jest wydobycie pożądanych właściwości sygnału usuwając zniekształcenia i szумы. Ogólny wzór dyskretnej transformaty falkowej przedstawiono poniżej:

$$\text{DWT}(f, \Psi) = \frac{1}{\sqrt{a}} \sum_n f(n) \cdot \Psi\left(\frac{n-b}{a}\right) \quad (3.1)$$

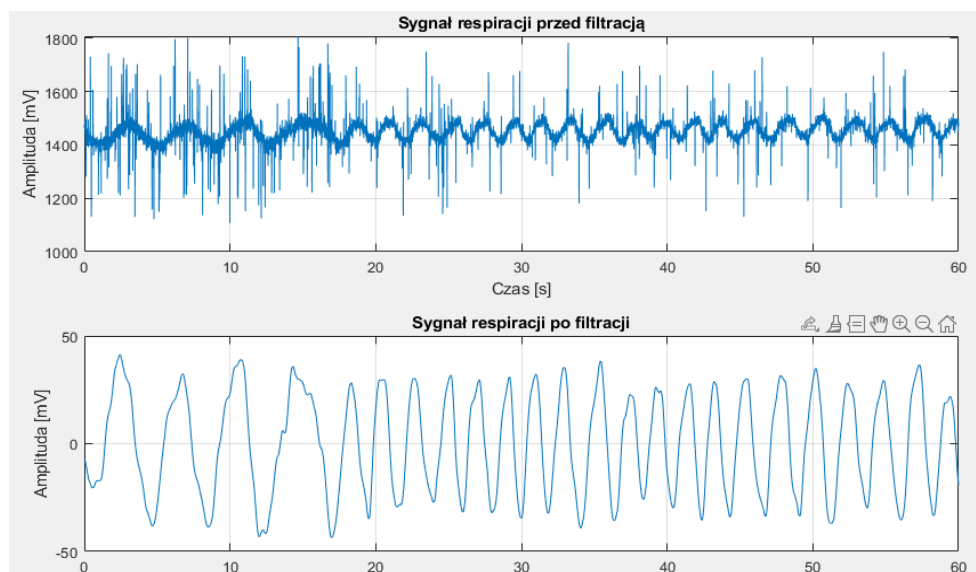
Gdzie a jest parametrem skali, b jest parametrem przesunięcia w dziedzinie czasu, a $f(n)$ analizowanym sygnałem [84, 85, 86, 87, 88].

Na potrzeby funkcji `Wavelet_Filters` wybrano falkę *Symlet 4*, której ogólny kształt detalu zilustrowano na rysunku 3.6. Użyto jej z uwagi na kształt przypominający szukany sygnał respiracji, gdy przeskalowano ją do detali o numerze od 8 do 12 aby dokonać rekonstrukcji sygnału.



Rys. 3.6 Zilustrowanie ogólnego kształtu detalu falki Symlet 4

Na koniec sygnał lekko wygładzono stosując filtr uśredniający o szerokości okna równej 100 próbek. Tak wyfiltrowany fragment sygnału został zobrazowany na rysunku 3.7.



Rys. 3.7 Porównanie sygnału respiracji przed oraz po operacji filtracji

Dodatkowo sygnał nie posiada składowej stałej, co na etapie ekstrakcji cech dystynktywnych okazało się bardzo istotne. Widoczne jest bardzo dobre odwzorowanie sygnału dzięki dobraniu odpowiednio dopasowanej falki oraz detali dla odwrotnej transformaty falkowej. Takie podejście umożliwia precyzyjnie wskazanie maksimum sygnału. Zobrazowany sygnał zapisywany pod nazwą *signal3* w strukturze *data* pochodzi z termistora umieszczonego pod jednym z nozdrzy. Ponieważ sygnał ten dostarczał wystarczającej ilości informacji sygnał z drugiego nozdrza potraktowano jako sygnał redundantny oraz referencyjny dla sygnału z pierwszego nozdrza. Sygnały te posiadały niemal identyczne właściwości, dlatego postanowiono drugiego z nich nie używać w trakcie późniejszej analizy. W pojedynczych przypadkach z racji rejestrowania przez jeden z czujników mniej precyzyjnej wartości sygnału lub uszkodzonych danych dokonano wyboru tego sygnału, który w ocenie Autora był wyższej jakości. Termistor umieszczony w okolicy ust w późniejszym czasie usunięto z uwagi na trudność w dopasowaniu wszystkich trzech czujników do twarzy respondentów oraz niską użyteczność. Nie rejestrował on żadnych wartości ponieważ każdy z respondentów oddychał jedynie przez nos. Sygnał ten nadal był każdorazowo rejestrowany dla zachowania spójności plików.

W celu ewentualnej przyszłej analizy tych sygnałów są one także zapisywane do struktury jako *signal4* oraz *signal5*.

Ostatnim krokiem był zapis metadanych takich jak wiek do struktury pod zmienną o nazwie *age*, wzrost jako *height*, waga jako *weight*, płeć jako *gender*, częstotliwość próbkowania sygnałów jako *fs*, rodzaj reprezentowanej emocji jako *emotion* oraz numer identyfikacyjny danej osoby jako *id*.

3.3 Ekstrakcja deskryptorów

W celu ekstrakcji cech dystynktywnych stworzono trzy funkcje odpowiadające za ekstrakcję deskryptorów. Każda z funkcji odpowiada za analizę i ekstrakcję cech bazując na jednym sygnale (Sygnał EDA, sygnał EKG, sygnał respiracji) .

3.3.1 Sygnał aktywności elektrodermalnej

Ekstrakcja cech z sygnału aktywności elektrodermalnej zdefiniowana jest za pomocą funkcji *SignalOneFeatures*. Sygnał ten rozłożyć można na dwie składowe: składową

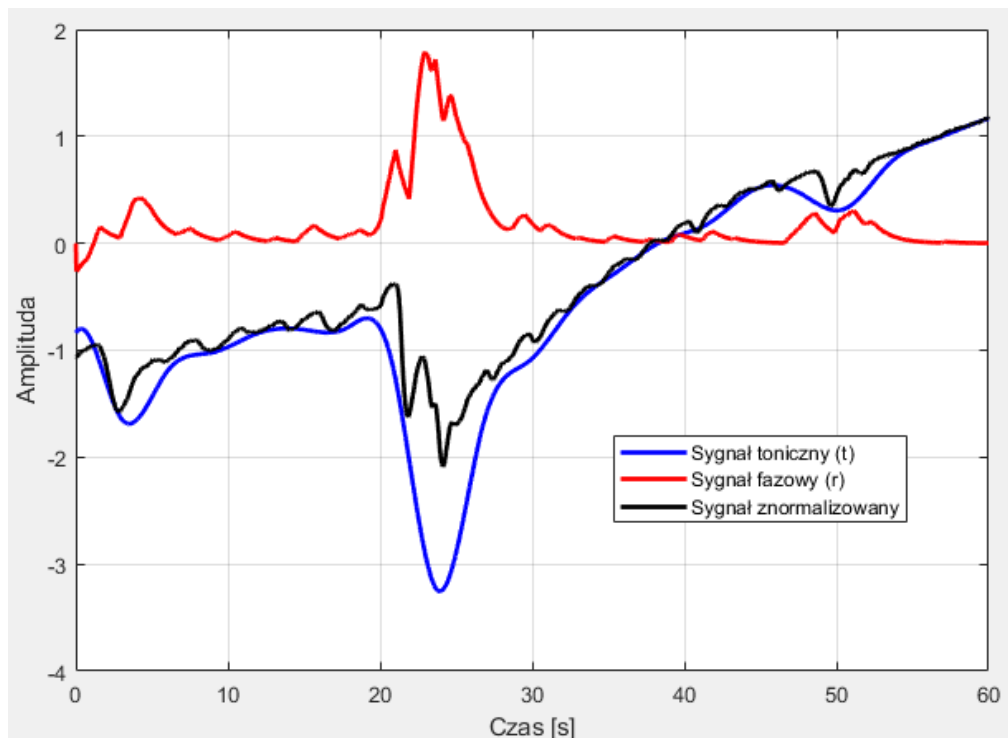
toniczną (SCL) niosącą informację o ogólnym poziomie pobudzenia oraz składową fazową (SCR) reprezentującą głównie poziom pobudzenia emocjonalnego spowodowanego działaniem układu współczulnego. Składowa SCR świadczyć może o poziomie odczuwalnych emocji jak i ich rodzaju. W tym celu dokonano dekompozycji sygnału za pomocą stworzonej przez społeczność oprogramowania MATLAB funkcji o nazwie `cvxEDA` [89]. Jej działanie w ogólności polega na znalezieniu za pomocą metod optymalizacyjnych funkcji celu dokonującej dekompozycji sygnału na wymienione składowe. Funkcja ta przyjmuje łącznie 8 argumentów. Pierwszy z nich to sygnał wejściowy poddawany dekompozycji. Drugim jest interwał próbkowania, czyli odwrotność częstotliwości próbkowania. Pozostałe 6 parametrów określono eksperymentalnie przeprowadzając obszerne testy na analizowanych sygnałach w efekcie czego przyjęto wartości przedstawione w listingu 3.4.

Listing 3.4.

```
% Ustalenie parametrów dla funkcji cvxEDA
tau0 = 1.2;           % slow time constant of the Bateman function
tau1 = 0.01;          % fast time constant of the Bateman function
delta_knot = 2.97;    % time between knots of the tonic spline function
alpha = 0.005;        % penalization for the sparse SMNA driver
gamma = 0.1;          % penalization for the tonic spline coefficients
solver = 'quadprog'; % sparse QP solver to be used, 'quadprog' or
'sedumi'

% Wywołanie funkcji cvxEDA do analizy sygnału
[r, p, t, l, d, e, obj] = cvxEDA(normalized_signal, 1/fs, tau0,...
    tau1, delta_knot, alpha, gamma, solver);
```

Parametry te uznano za dobry kompromis pomiędzy dekompozycją na składową fazową reprezentowaną końcowo jako zmienna r oraz składową toniczną reprezentowaną jako zmienna t . Na rysunku 3.6 zostały wspólnie wykreślone składowe sygnału oraz sam sygnał, który przed tą operacją został znormalizowany.



Rys. 3.6 Zilustrowanie sygnału EDA przed oraz po operacji dekompozycji na składowe

Wspomniany kompromis polegał na dostosowaniu szybkości dopasowywania się składowej tonicznej do sygnału. Ponieważ wartości składowej tonicznej oraz fazowej są w pewnym sensie odwrotnie proporcjonalne względem siebie, starano się, aby sygnał fazowy wyraźnie odzwierciedlał znaczące różnice potencjałów oraz nakreślał subtelniejsze zmiany, które, choć mniej istotne nie wynikały z błędów w metodzie pomiarowej, miały jednak wpływ na decyzje dotyczące prognozowania emocji. Ponieważ przedstawiony przebieg dotyczy emocji szczęścia nie zawiera on dużej wartości składowej fazowej i jest zgodny z teorią.

Po tej operacji możliwa była właściwa analiza sygnałów EDA. Pierwszym krokiem było znalezienie maksimów za pomocą funkcji `findpeaks` z parametrem `'MinPeakWidth'` wynoszącym 0,2 odnoszącym się do znalezienia maksimum funkcji o minimalnej szerokości szczytu równym 0,2. Ponieważ analizowane sygnały są różnej długości wszystkie obliczane parametry są obliczane w odniesieniu do czasu trwania sygnału w minutach. W efekcie tego obliczono parametry takie jak: liczba maksimów sygnału zarejestrowana w ciągu minuty, ilość szczytów powyżej amplitudy równej 1 w ciągu minuty, średnia amplituda znalezionej maksimum w ciągu minuty, średnia wartość sygnału fazowego w ciągu minuty.

3.3.2 Sygnał elektrokardiograficzny

Ekstrakcję cech z sygnału EKG postanowiono przeprowadzić głównie na podstawie tętna (HR) oraz zmienności rytmu serca (HRV). Podyktowane było to dość słabą jakością sygnałem oraz koniecznością użycia zaawansowanych algorytmów w celu ekstrakcji szerszej gamy cech. W tym celu zdefiniowano autorską funkcję `SignalTwoFeatures` iteracyjnie analizującą sygnały EKG, której fragment kodu przedstawiono w listingu 3.5.

Listing 3.5.

```
wt = modwt(ecgsig, 4, 'sym4'); % DWT z wykorzystaniem falki sym4
wtrec = zeros(size(wt)); % Inicjalizacja wektora rekonstrukcji
falkowej
wtrec(3:4, :) = wt(3:4, :); % Zachowanie określonej skali falkowej

% Rekonstrukcja sygnału po DWT
y = imodwt(wtrec, 'sym4'); % Odwrotna Transformacja Falkowa
y = abs(y).^2; % Kwadrat wartości amplitudy
avg = mean(y); % Średnia wartość kwadratu

if avg < 0.35

    [Rpeaks, Rlocs] = findpeaks(y, t, 'MinPeakHeight',...
        4 * avg, 'MinPeakDistance', 300); % Wyszukanie pików R
    nohb = length(Rlocs);
    timelimit = length(ecgsig)/(fs * 60); % Czas trwania sygnału w
minutach
    QRS_per_min = length(Rlocs) / timelimit; % Liczba pików R na
minutę

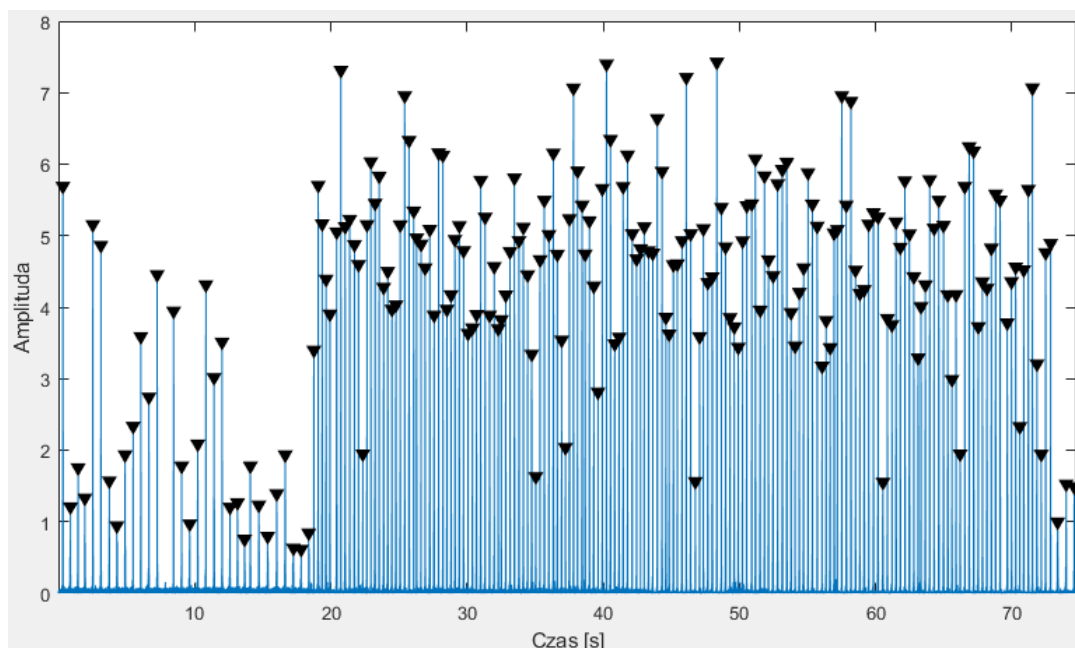
else

    %% --- Wyznaczanie załamek R z wykorzystaniem algorytmu Pan-
Tompkins --- %%
    [Rpeaks, Rlocs, delay] = pan_tompkin(signal.signal2, fs, 0);
    timelimit = length(signal.signal2)/(fs * 60); % Obliczenie
czasu sygnału w minutach
    QRS_per_min = length(Rpeaks)/timelimit; % Liczba pików R na
minutę

end
```

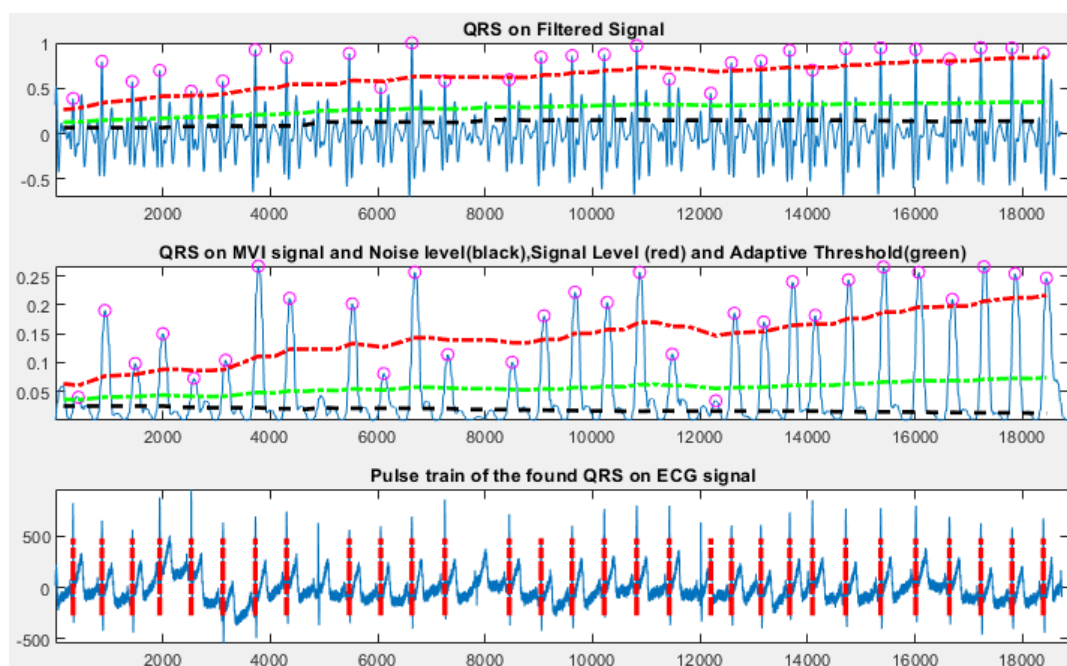
Funkcja wykorzystuje transformatę falkową, która ma za zadanie poszukiwanie kompleksu QRS w sygnale EKG, a następnie na podstawie ich ilości oraz położenia obliczone zostały HR i HRV. W tym celu ponownie użyto falki *Symlet 4* z powodu korelacji kształtu jej funkcji przeskalowanego do poziomu 3 i 4 do kształtu fal zespołu QRS. Tym razem użyto funkcji `modwt`, ponieważ jest to bardziej zaawansowana funkcja

pozwalająca na jednoczesną analizę sygnału na wielu poziomach skalowych. Dzięki jej użyciu możliwe jest uwzględnienie różnych aspektów sygnału na różnych skalach. W tym przypadku efekt działania obu funkcji można przyjąć że jest taki sam. Funkcja ta rozkłada sygnał na falki Symlet 4 o skali od 1 do 4, a następnie dokonywana jest odwrotna transformacja falkowa jedynie z użyciem detali o numerze 3 oraz 4 przy użyciu funkcji `imodwt`. Kolejno obliczano wartość średnią uzyskanego sygnału. Po wstępnych badaniach zauważono, że algorytm ten działa bardzo dobrze nawet na sygnałach zakłóconych ruchem badanego. Ważne jest jednak, aby sygnał charakteryzował się niskim poziomem szumu. Po dogłębnej analizie przyjęto, że sygnał posiadający wartość średnią powyżej **0,35** powinien być przetwarzany przy użyciu innego rodzaju algorytmu. W tym celu użyto funkcji stworzonej przez społeczność oprogramowania MATLAB o nazwie `pan_tompkin` [90]. Istotne jest to, że funkcja ta zawiera trzy adaptacyjne progi do określania poziomów szumu i sygnału. Domyślnie nie jest ona używana ze względu na skłonność do wykrywania załamków P zamiast kompleksu QRS, co w pewnym stopniu jest akceptowalne. Niemniej jednak, nie zawsze jest w stanie dokładnie zlokalizować poszukiwane załamki, nawet po edycji kodu jej źródłowego przez Autora. Procedura działania tej funkcji polega na odsumieniu sygnału EKG przez różne filtry, w tym filtry IIR, a następnie wzmocnieniu wartości załamków R przy pomocy filtru pochodnej. Widoczne na rys. 3.7 piki odpowiadają położeniu załamków R po dokonaniu dekompozycji falkowej. Znalezienie widocznych maksimów nie sprawia trudności, dzięki czemu załamki są prawidłowo zliczane przy użyciu funkcji `findpeaks`. Wadą tego rodzaju przetwarzania sygnału oprócz wcześniej wymienionej niskiej odporności na szumy wysokoczęstotliwościowe jest obecność wstęg bocznych pików widocznych w przybliżeniu. Nie jest to problemem przy wykorzystaniu funkcji `findpeaks` z zadeklarowaną minimalną odległością między maksimami.



Rys. 3.7 Efekt działania filtracji falkowej z naniesionymi maksimami

Na rysunku 3.8 przedstawiono wbudowane w funkcję `pan_tompkin` okno ukazujące wyniki działania algorytmu. Zauważyć można, że wszystkie załamki zostały prawidłowo zlokalizowane poza załamkiem w okolicach 8000 próbek, który nie został prawidłowo zarejestrowany oraz błędnie zarejestrowany załamek P w okolicach 12000 próbek zamiast załamka R.



Rys. 3.8 Efekt działania funkcji „pan_tompkin” ze znalezionymi załamkami R sygnału

Oba algorytmy na przestrzeni dłuższego czasu trwania sygnału powinny przedstawiać podobne wyniki wartości HR oraz nieco mniej zbliżone wyniki analizy HRV. Dzięki znalezionym wartościom załamków obliczone zostały deskryptory HR, SDDN, RMSSD i pRR50 zgodnie z wzorami 1.3-1.6 przedstawionymi w rozdziale pierwszym pracy. Dodatkowo zdefiniowano cechę: średnie bezwzględne odchylenie (MAD ang. *Mean Absolute Deviation*) amplitud od szczytu do szczytu (PPA ang. *Peak-to-Peak Amplitude*) opisaną poniższym wzorem:

$$PPAMAD = \frac{1}{n} \sum_{i=1}^n |A_i - \hat{A}| \quad (3.2)$$

Gdzie:

- A to i -ta amplituda od szczytu do szczytu,
- \hat{A} to mediana wszystkich amplitud od szczytu do szczytu.

Parametr ten wybrano ze względu na odporność na wartości skrajne oraz wiele wnoszące informacje o sygnale [91].

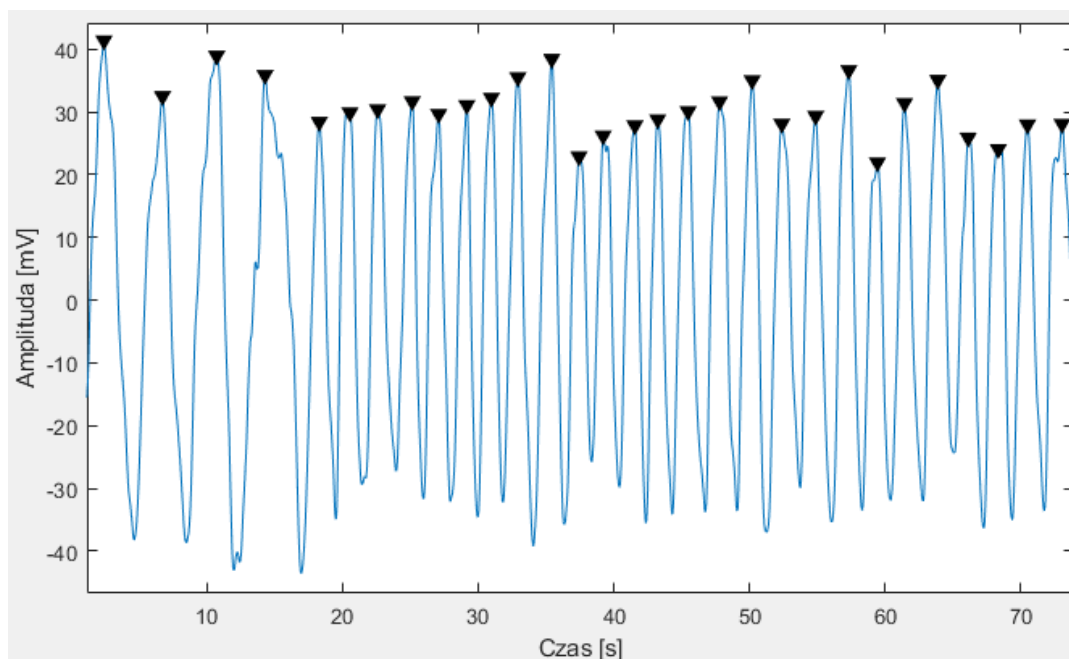
3.3.3 Sygnał respiracji

Działanie ostatniej z funkcji o nazwie `SignalThreeFeatures` również odbywa się iteracyjnie na sygnale respiracji. Ponieważ sygnał ten został uprzednio przefiltrowany transformatą falkową nie wymagał dodatkowej zaawansowanej analizy. Zdecydowano zatem skoncentrować się na wyznaczeniu częstości oddechów oraz parametrów z nią związanych. W tym celu użyto funkcji `findpeaks` z parametrami przedstawionymi w listingu 3.6.

Listing 3.6.

```
[peaks, locs] = findpeaks(signal.signal3, fs, 'MinPeakWidth',
0.495, 'MinPeakDistance', 0.95, 'MinPeakHeight', mean(signal.signal3));
```

Parametry minimalnej wysokości sygnału, minimalnej szerokości szczytu oraz minimalnej odległości między szczytami dobrano eksperymentalnie. Na rysunku 3.9 przedstawiono znalezione maksima za pomocą powyższej funkcji. Widoczne jest uchwycenie lokalnych maksimów sygnału o niższej częstotliwości chwilowej jak i wyższej.



Rys. 3.9 Wykres sygnału respiracji z wykrytymi i naniesionymi maksimami

W wyniku różnych wzorców oddechowych niektórych respondentów zaobserwowano sytuacje, w których w trakcie wydychania powietrza przerywali na chwilę proces, by po chwili kontynuować. W rezultacie takich zachowań zarejestrowane były dwa bardzo bliskie sobie lokalne maksima sygnału oddechowego, co było interpretowane jako dwa pełne cykle oddechowe. Rozwiązaniem mogłoby okazać się użycie bardziej zaawansowanego algorytmu. Ponieważ błąd występował bardzo rzadko nie miał znaczącego wpływu na wyniki końcowe.

Następnie na podstawie wyznaczonych maksimów wyznaczono poniższe deskryptory: ilość oddechów na minutę, SDNN, średnia wysokość szczytu oraz PPA MAD.

Tak zdefiniowane macierze parametrów były zapisywane w plikach o rozszerzeniu .mat aby nie wykonywać ponownie bardzo długotrwałych obliczeń przez stację roboczą. Omawiany kod programu starano się pisać w sposób zrozumiały umieszczając dużą ilość komentarzy oraz z przyjętymi dobrymi praktykami by ewentualna późniejsza edycja w celu dodania czy zmiany ustawień parametrów nie sprawiała kłopotów.

4. Eksperymenty i wyniki otrzymanych rezultatów

W wyniku wstępnych analiz i eksperymentów zdefiniowano trzy macierze cech będące potencjalnymi deskryptorami stanów emocjonalnych. Na tym etapie dzięki możliwości wizualizacji oraz analizy wyników usunięto 6 uszkodzonych pomiarów z badań respondentów. W efekcie tego pozostało 14 pomiarów z badań, które posłużyły do kolejnego etapu eksperymentów, czyli procesu klasyfikacji. Do tego celu użyto aplikacji do uczenia maszynowego dostępnych w oprogramowaniu MATLAB. Analizując uzyskane dane oraz analizowany przypadek klasyfikacji Autor pracy postanowił wykorzystać dwa narzędzia, które wydały się spełniać założenia projektu. Pierwszym z nich była aplikacja *Classification Learner*, natomiast drugim aplikacja *Neural Net Pattern Recognition*.

4.1 Ewaluacja skuteczności aplikacji *Classification Learner*

Classification Learner to narzędzie do uczenia maszynowego umożliwiające tworzenie i ocenę modeli klasyfikacyjnych [92]. Działa na zasadzie uczenia nadzorowanego (z nauczycielem) przy użyciu wielu modeli klasyfikujących. Proces ten opiera się na dostarczeniu danych wejściowych wraz z odpowiadającymi im etykietami klasy. Dostarczone dane są automatycznie odpowiednio od ich rodzaju normalizowane. Aplikacja umożliwia zrealizowanie procesu uczenia, walidacji i klasyfikacji danego modelu. Wykorzystywane modele to między innymi drzewa decyzyjne, analiza dyskryminacyjna, maszyny wektorów nośnych, regresja logistyczna, algorytm najbliższych sąsiadów, naiwny klasyfikator Bayesa, aproksymacja jądra, klasyfikacja zespołowa czy sieci neuronowe. Kiedy modele są wytrenowane, *Classification Learner* ocenia ich wydajność na podstawie dostarczonych danych. W efekcie użytkownik może wybrać najbardziej odpowiadający model i dokonać jego optymalizacji w zależności od potrzeb. Gotowy i wytrenowany model może być wyeksportowany jako funkcja programu MATLAB, tak aby możliwe było klasyfikowanie nowych danych. Istotnym funkcjonalnością aplikacji jest intuicyjny graficzny interfejs przedstawiony na rysunku 4.1. Zobrazowane okno pojawia się po kliknięciu przycisku *New Session* odpowiadającego za dodanie danych do uczenia modeli w aplikacji *Classification Learner*. W oknie *Data Set Variable* wybrane zostały dane wejściowe *data_table*, które zdefiniowano na etapie ekstrakcji cech. Zmienna *data_table* zawiera ID osoby i rodzaj

emocji jako zmienne katagoryczne oraz 14 wcześniej obliczonych parametrów jako zmienne typu *double*. Wewnątrz omawianej zmiennej znajdują się jedynie parametry dotyczące emocji neutralnej oraz wstrętu do okaleczeń, ponieważ ich wartości są najmniej zbliżone do siebie, co daje większą szansę na prawidłową klasyfikację danych. W oknie *Response* wybrano zmienną *Var2* zawierającą odpowiedź etykiety klasy. Zmienna ta zawiera liczbę 0 lub 5 interpretowaną jako kategoria. Kategoria 0 odpowiada stanowi neutralnemu natomiast 5 wstrętowi do okaleczeń.

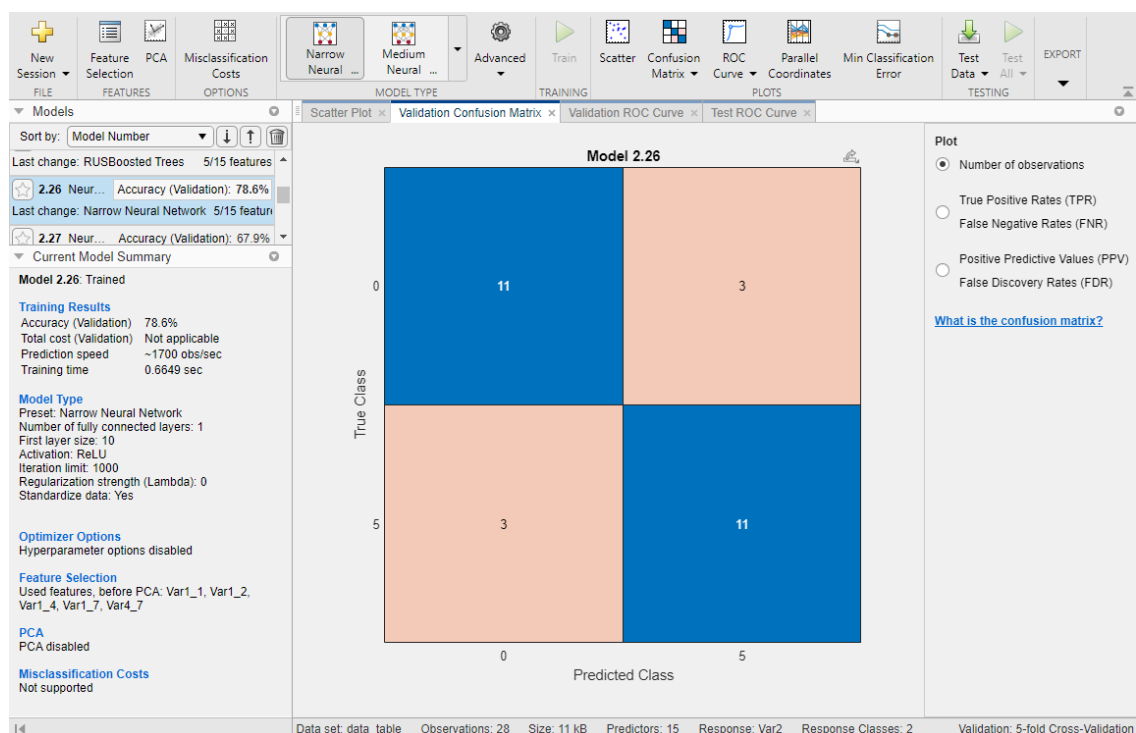
The screenshot shows the 'Data set' and 'Validation' tabs of the Classification Learner application. In the 'Data set' tab, 'data_table' is selected as the Data Set Variable (28x4 table). Under 'Response', 'From data set variable' is selected, and 'Var2' is chosen (categorical, 2 unique). A list of predictors is shown with checkboxes for Var1_1 through Var1_7, all of which are selected. Below the list are 'Add All' and 'Remove All' buttons, and a link to 'How to prepare data'. The 'Validation' tab has three options: 'Cross-Validation' (selected), 'Holdout Validation', and 'Resubstitution Validation'. 'Cross-Validation' is described as protecting against overfitting by partitioning the data set into folds. The 'Cross-validation folds' are set to 5. 'Holdout Validation' is recommended for large data sets, with 'Percent held out' set to 25. 'Resubstitution Validation' offers no protection against overfitting. At the bottom, there are 'Start Session' and 'Cancel' buttons, and a link to 'Read about validation'.

	Name	Type	Range
<input checked="" type="checkbox"/>	Var1_1	double	11.3267 .. 38.5151
<input checked="" type="checkbox"/>	Var1_2	double	65.3103 .. 134.429
<input checked="" type="checkbox"/>	Var1_3	double	65.7637 .. 172.288
<input checked="" type="checkbox"/>	Var1_4	double	7.97447 .. 77.7367
<input checked="" type="checkbox"/>	Var1_5	double	0 .. 27.0176
<input checked="" type="checkbox"/>	Var1_6	double	0.0173245 .. 1.19387
<input checked="" type="checkbox"/>	Var1_7	double	0.00519128 .. 1.02256

Rys. 4.1 Okno wyboru danych aplikacji *Classification Learner*

W oknie opcji *Validation* wybrano *Cross-Validation*, czyli walidację krzyżową [93]. Użyta została w celu oceny wydajności modelu oraz zminimalizowania ryzyka przeuczenia lub niedoszacowania sieci neuronowej. Polega na podziale danych na 5 podzbiorów, gdzie jeden z nich jest używany jako zestaw testowy, a pozostałe do trenowania modelu. Proces jest powtarzany w tym przypadku 5 razy, a każdy podzbiór jest używany jako zestaw testowy dokładnie raz. Po zadeklarowaniu przedstawionych wartości zatwierdzono je przyciskiem *Start Session*. W efekcie tej operacji ukazało się okno główne programu *Classification Learner* widoczne na rysunku 4.2. W panelu tym dokonano serii eksperymentów realizowanych w różnych konfiguracjach metod klasyfikacyjnych, ilości podzbiorów kroswalidacji oraz możliwości zastosowanie

metody PCA (ang. *Principal Component Analysis*) [93]. Ostatecznie najwyższy wynik identyfikacji jaki uzyskano wynosił **78,6%** na danych walidacyjnych przy użyciu modelu *Narrow Neural Network*. *Narrow Neural Network*, inaczej nazywana wąską siecią neuronową, to rodzaj płaskiej sieci neuronowej, która zazwyczaj zawiera jedną lub kilka warstw ukrytych. Głównym wyróżnikiem tego rodzaju metody jest jej ograniczona liczba neuronów czyniącą ją efektywną w przypadku niewielkiej liczby parametrów wejściowych. To z kolei przekłada się na szybszą pracę sieci, ponieważ wymaga ona wykonania mniejszej liczby obliczeń. Kolejną istotną cechą jest jej zdolność do zapobiegania zjawisku przeuczenia (ang. *overfitting*), czyli nadmiernego dopasowania się danych treningowych do modelu, co oznacza, że może skutecznie wykonywać zadania przy umiarkowanej złożoności sieci neuronowej. Sieć ta sprawdza się wszędzie dobrze tam, gdzie wprowadzane dane są stosunkowo proste lub dostępna ilość danych jest ograniczona. Właściwości te skłoniły Autora do wykorzystania w badaniach eksperymentalnych właśnie tej metody.



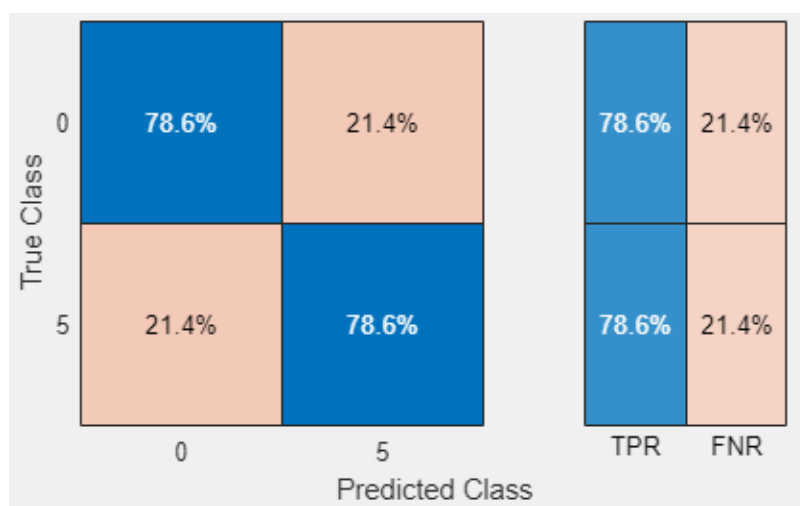
Rys. 4.2 Główne okno programu *Classification Learner* z wynikami klasyfikacji dla modelu *Narrow Neural Network*

Dokładność klasyfikacji na poziomie 78,6% uzyskano przy użyciu jedynie **5** deskryptorów:

- Ilość oddechów na minutę,
- Tętno
- Ilość szczytów składowej fazowej o zadanej szerokości na minutę
- Średnia wartość składowej fazowej na minutę
- PPA MAD znalezionych załamków R.

Dobór tych parametrów dokonano techniką selekcji w przód (ang. *Forward Selection*) [93]. Metoda ta polega na rozpoczęciu od pustego zbioru cech, a następnie dodawaniu ich iteracyjnie, jedna po drugiej. Po dodaniu każdej cechy model był trenowany, a jego skuteczność była oceniana. Jeśli dodanie konkretnej cechy prowadziło do poprawy wyników modelu, była ona zachowywana. W przypadku, gdy dodanie cechy nie przynosiło korzyści lub pogarszało wyniki, była ona odrzucana, a proces kontynuowany był z pozostałymi cechami. Wadą tej metody jest jej czasochłonność oraz wysokie wymagania obliczeniowe.

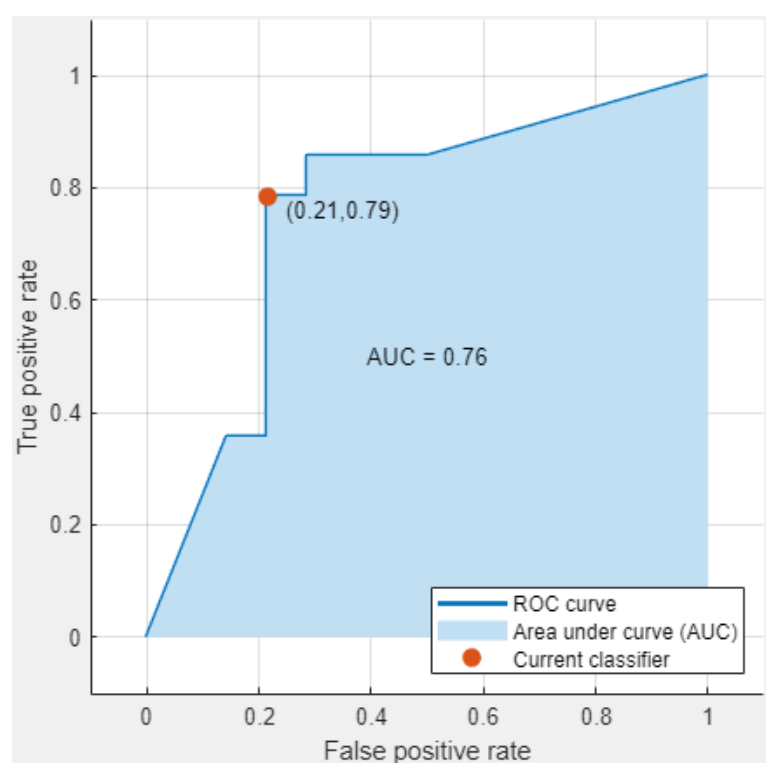
Na podstawie przedstawionego wyniku klasyfikacji stwierdzić można, że klasy 5 i 0 zostały prawidłowo sklasyfikowane w 11 przypadkach, w 3 przypadkach natomiast wystąpiła błędne rozpoznanie. Dokonując bardziej szczegółowej analizy można wyznaczyć współczynnik prawdziwie pozytywnych rozpoznań TPR (ang. *True Positive Rate*) oraz współczynnik fałszywie negatywnie zaklasyfikowanych FNR (ang. *False Negative Rate*) danych, co przedstawia wizualnie rysunek 4.3.



Rys. 4.3 Współczynniki TPR i FNR dla wybranego modelu klasyfikacji

Przedstawione wyniki można uznać za satysfakcjonujące, ze względu na trudność w zaklasyfikowaniu emocji tak wielu osób o różnej odpowiedzi układu nerwowego na emocjonujące sytuacje oraz uwzględnieniu w umiarkowanym stopniu indywidualnego podejścia do badanej osoby.

Ostatnim zagadnieniem analizy tego rozwiązania jest analiza rysunku 4.4 przedstawiającego krzywą ROC (ang. *Receiver Operating Characteristic*) [93]. Krzywa ROC to graficzne przedstawienie zależności między TPR (wrażliwością), a FPR (specyficznością) w zależności od różnych progów decyzyjnych. Im wyżej na krzywej znajduje się punkt dla danego modelu, tym lepsza jest jego wydajność. Idealny klasyfikator miałby krzywą ROC przechodzącą przez punkt (0,1), co oznaczałoby pełną wrażliwość bez żadnych fałszywych pozytywów. Czerwony punkt o współrzędnych (0.21, 0.79), oznaczony jako "*Current classifier*" na krzywej ROC, jest aktualnym punktem pracy przedstawionego klasyfikatora przy progu decyzyjnym równym 0,21. Współrzędna pozioma 0,21 odpowiada FPR, czyli wskaźnikowi fałszywych akceptacji. Współrzędna pionowa 0,79 odpowiada TPR, czyli wskaźnikowi prawdziwych akceptacji. Zmieniając punkt pracy klasyfikatora zmieni się także próg decyzyjny klasyfikatora, co wpłynie na stosunek TPR do FPR.



Rys. 4.4 Krzywa ROC wybranego modelu

AUC (ang. *Area Under the Curve*) to miara obszaru pod krzywą ROC. Im większe AUC, tym lepsza jest wydajność zastosowanego modelu. Dla idealnego klasyfikatora AUC wynosi 1, a dla zupełnie losowego klasyfikatora AUC wynosi 0,5. Aktualny próg decyzyjny klasyfikatora jest wyznaczony tak aby uzyskać jak największe pole pod krzywą ROC zachowując przy tym optymalny stosunek TPR do FPR [93].

Przedstawiony model wyeksportowano oraz zapisano jako funkcja o nazwie *NarrowNetworkModel*. Dzięki temu zabiegowi możliwa jest klasyfikacja nowych danych przy użyciu powstałego modelu.

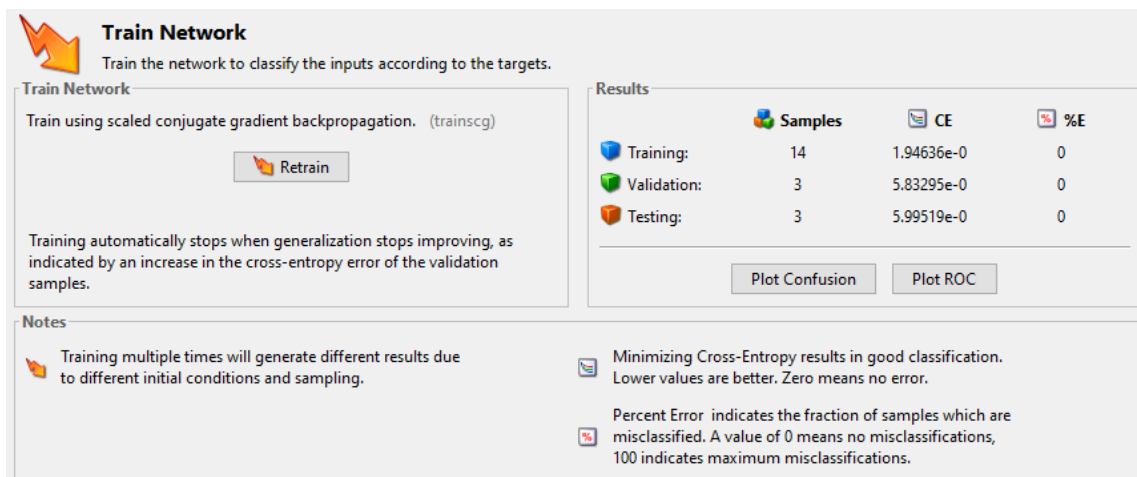
Podsumowując stworzono model dostosowany do analizowanego przypadku, odporny na przeuczenie, wydajny obliczeniowo, możliwy w implementacji także na płycie używanego mikrokontrolera STM32 oraz o satysfakcjonującej dla Autora poprawności rozpoznania. Zdefiniowany model wytworzono w oparciu o dane stosunkowo dużej grupy respondentów, dzięki czemu istnieje wysokie prawdopodobieństwo, że zaproponowane rozwiązanie będzie efektywnie funkcjonować w warunkach rzeczywistych. Ponieważ model był trenowany jedynie na jednym pomiarze danej emocji jednego respondenta domniemywać można, że model nie jest dobrze dostosowany do wykrywania emocji jednej osoby w efekcie czego uzyskanie powtarzalności wyników może być problematyczne. Problem ten wynika z nieuwzględnienia jeszcze bardziej indywidualnego podejścia do osoby badanej. Rozwiązaniem może być użycie innego modelu klasyfikatora, dostarczenie dodatkowej puli danych uczących oraz ponownego nauczania analizowanej sieci czy też stworzenie modelu sieci skoncentrowanej na wykryciu stanu emocjonalnego konkretnej osoby, co postanowiono zrealizować.

4.2 Ewaluacja skuteczności aplikacji Neural Net *Pattern Recognition*

W celu stworzenia modelu możliwie maksymalnie dostosowanego do badanej osoby zdefiniowano klasyfikator, który nauczono większą ilością danych wyższej jakości, niż w przypadku poprzedniego modelu. W tym celu postanowiono dokonać 10 pomiarów stanu neutralnego oraz 10 pomiarów emocji wywołanej wstrętem do okaleczeń na jednej osobie, która brała także udział w poprzednich badaniach. Aby dane dostarczały jeszcze większej ilości wysokiej jakości informacji postanowiono przeprowadzać pomiary trwające minimum 3 minuty, a nierzadko pojedynczy pomiar trwał ponad 4

minuty. Pomiary obu stanów emocjonalnych uzyskano przeprowadzając badania w ciągu jednego dnia z modyfikacją wstępnych założeń projektowych. Dokonano zatem podzielenia pomiarów na sesje oraz przerwy wynoszące 2 minuty między pomiarami. Ponadto pomiary dotyczące stanu neutralnego uzyskano podczas dwóch sesji, dla których rejestracji dokonano sobie odstępów około 30 minut. Kolejną różnicą było przedstawienie innej bazy materiałów afektywnych respondentowi, w celu wywołania emocji wstrętu do okaleczeń. Użyta baza zawierała sceny z przebiegu różnego rodzaju operacji medycznych przeprowadzanych przez wyspecjalizowany personel medyczny mających charakter edukacyjny (operacje kończyn czy operacje z widocznym bijącym ludzkim sercem). Aby utrzymać osobę badaną w stanie afektu zrezygnowano z przerw między pomiarami, a pomiary podzielono na 3 sesje oddzielone od siebie o około 30 minut. Tak przygotowane dane poddano konwersji oraz filtracji w programie *CreateDatabase*, a następnie parametryzacji wywołując kod programu *main*. W trakcie wykonywania kodu zawartego w pliku *main* dodatkowo tworzona była macierz parametrów wejściowych o nazwie *pattern_data_matrix_in* oraz macierz etykiet klasyfikacji danych o nazwie *pattern_data_matrix_out*. Macierz parametrów wejściowych zaprojektowano niemal identycznie jak w przypadku danych użytych w aplikacji *Classification Learner*. Macierz zawierała jedynie 5 parametrów, które najlepiej sprawdziły się w przypadku uczenia poprzedniego modelu. Tak przygotowane dane poddano analizie w aplikacji *Neural Net Pattern Recognition* w oprogramowaniu MATLAB. *Neural Net Pattern Recognition* to aplikacja umożliwiająca trening dwuwarstwowej sieci neuronowej ze sprzężeniem zwrotnym. Wewnątrz warstw ukrytych sieć zawiera funkcje sigmoidalne oraz wyjście typu softmax. Sieć jest trenowana przy użyciu skalowanej propagacji wstecznej gradientu sprzężonego [94]. W pierwszej kolejności przyporządkowano *pattern_data_matrix_in* jako dane wejściowe „*Inputs*” oraz dane zapisane wewnątrz *pattern_data_matrix_out* jako odpowiedzi klasyfikacji „*Targets*” wcześniej wprowadzonych danych. Ponieważ dane jednej emocji zapisywano w wierszu wybrano także opcję „*Matrix rows*”. W kolejnym oknie konfiguracyjnym wybrano domyślny stosunek danych wykorzystywanych do zdefiniowania modelu sieci: 70% danych testowych (14 próbek), 15% danych walidacyjnych (3 próbki) i 15% danych testowych (3 próbki). W następnym oknie wybrano ostatecznie po wykonaniu wielu testów na modelu 5 neuronów w warstwach ukrytych, gdyż ta liczba skutkowała uzyskaniem najlepszych wyników przez model. Następne okno zobrazowane na rysunku 4.5 ukazuje wyniki treningu sieci neuronowej.

Wartość %E oznaczająca procent danych zaklasyfikowanych błędnie przez model wynosi w każdym przypadku danych 0, co świadczy o poprawnej klasyfikacji wszystkich przypadków danych. Wartość CE, czyli krosentropia, jest miarą, która mierzy, jak bardzo prawdopodobieństwa przewidywane przez model różnią się od rzeczywistych etykiet w danych treningowych [93].

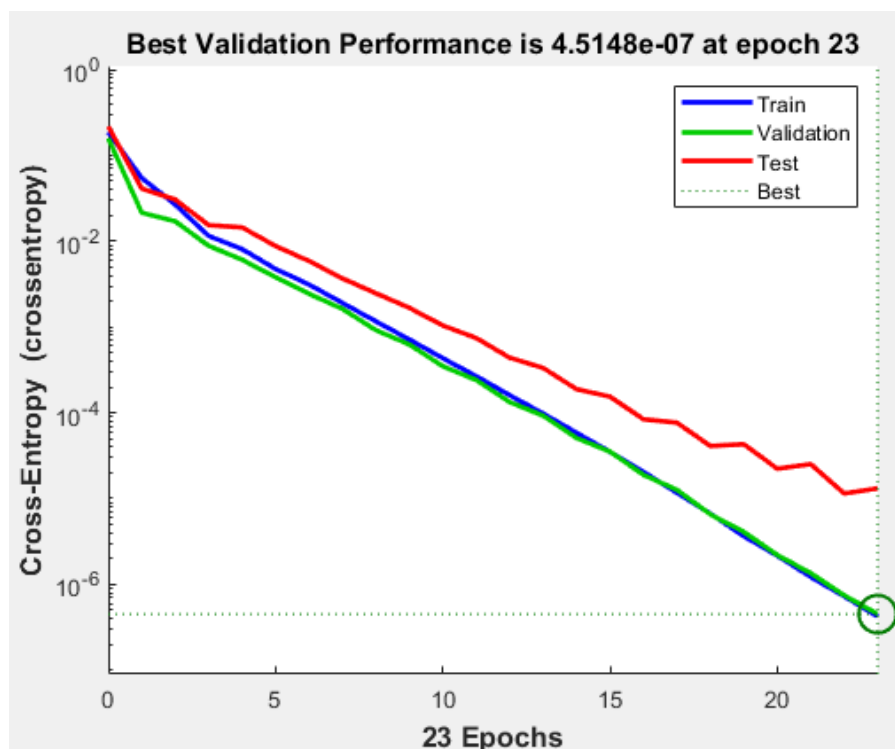


Rys. 4.5 Okno aplikacji Neural Net Pattern Recognition z wynikami treningu modelu sieci neuronowej

Oznacza to, że wartość CE informuje nas o tym, w jakim stopniu model jest pewny swoich prognoz i jak dobrze te prognozy pasują do rzeczywistych danych. Niższa wartość CE wskazuje na lepszą zgodność modelu z rzeczywistymi danymi. Uzyskane wartości CE można uznać za bardzo dobre biorąc pod uwagę brak popełnionych błędów w klasyfikacji. Z racji, że wszystkie przypadki zostały poprawnie zaklasyfikowane pole pod krzywymi ROC wynosi 1.

Rysunek 4.6 przedstawia jak w procesie uczenia sieci neuronowej wraz z każdą z 23 iteracji wzrastała wydajność modelu dla trzech rodzajów danych. Im uzyskana wartość w przypadku funkcji kosztu (ang. *Cross-Entropy*) jest mniejsza, tym lepiej świadczy o zdolności do poprawnej klasyfikacji modelu. Wartość ta nie maleje w nieskończoność aby uniknąć zbytniego dopasowania się do danych, dlatego algorytm zakończył działanie w 23 iteracji. Uzyskana wartość $4,5 \cdot 10^{-7}$ świadczy o bardzo dobrej wydajności modelu na zbiorze walidacyjnym. Wyniki uczenia modelu ukazują zbliżone wartości wydajności dla danych treningowych i walidacyjnych, co sugeruje, że model nie jest ani nadmiernie dopasowany, ani zbyt prosty. Natomiast wydajność modelu jest nieco niższa na danych testowych, ale różnica ta nie jest zbyt duża. To świadczy o tym,

iż model dobrze radzi sobie z ogólną generalizacją wzorców w danych. Istnieje także zauważalna korelacja między wynikami na danych walidacyjnych i treningowych, co oznacza, że model dobrze uchwycił te wzorce [93].



Rys. 4.6 Wykres wydajności wybranego modelu podczas uczenia

Tak nauczony model sieci neuronowej wyeksportowano oraz zapisano jako funkcja o nazwie *myNeuralNetworkFunction1Pearon* w przestrzeni roboczej oprogramowania MATLAB.

Zaproponowany oraz stworzony model sieci neuronowej cechuje się znacznie wyższą wydajnością klasyfikacji emocji niż przedstawiony w ramach aplikacji *Classification Learner*. Jest to spowodowane głównie położeniem nacisku na zindywidualizowane podejście do respondenta, wyeliminowanie wielu błędów wynikających z metodyki przeprowadzania badań m.in. dość krótki czas pomiaru, niewielka ilość pomiarów czy przeprowadzenie badania w jednym pomieszczeniu tego samego dnia. Wadą modelu jest znikome prawdopodobieństwo poprawnej klasyfikacji emocji zarejestrowanych u innej osoby, ze względu na nie uwzględnienie jej biomarkerów w procesie uczenia, które mogą być kompletnie inne przy stanie afektywnym, przez co model sieci może je błędnie klasyfikować.

5. Opracowanie docelowej konstrukcji systemu multimodalnego

Realizując koncepcję końcowego urządzenia postanowiono skoncentrować się na realizacji założeń jakimi są pogłębianie wiedzy na temat powstawania i wpływu odczuwanych w ciągu dnia emocji przez ludzi. Docelowe urządzenie na podstawie zebranych informacji generowałoby codzienne raporty o zdarzeniach z poprzedniej doby. Raporty zawierałyby informacje o potencjalnych skutkach odczuwanych emocji oraz sugestie dotyczące poprawy parametrów zdrowotnych i zachowań.

Na podstawie działań przeprowadzonych podczas tworzenia przedstawionej pracy i wyciągniętych wniosków postanowiono przedstawić koncepcję urządzenia, które będzie w stanie klasyfikować emocje wyrażane przez użytkownika. Przedstawiony poniżej podział implementacji sprzętowej oraz programowej jest bardzo płynny i umowny, ponieważ oba zagadnienia są bardzo blisko ze sobą związane w efekcie tworząc nierozrwalną całość.

5.1 Konfiguracja sprzętowa docelowego urządzenia

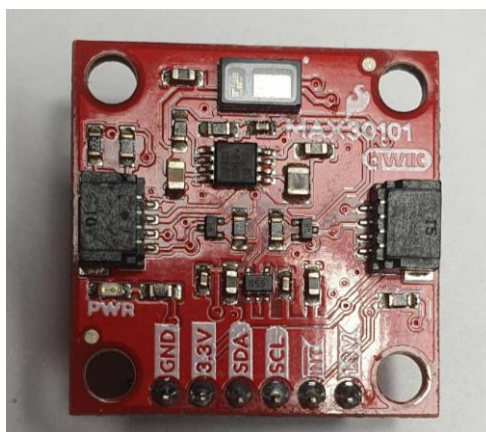
Proponowana docelowa konstrukcja multimodalnego systemu detekcji emocji może funkcjonować jako zestaw modułów działających osobno tworząc pewien ekosystem urządzeń wzajemnie się uzupełniających. Pierwszy moduł może mieć formę opaski czy też smartwatcha. Wybór zbudowania tego rodzaju urządzenia jest podyktowany znikomą ingerencją w komfort użytkownika, zapewniając przy tym bardzo dużą funkcjonalność.

Urządzenie mogłoby przeprowadzać pomiar sygnału EDA przy użyciu elektrod umieszczonych na zapięciu smartwatcha, podobnie jak w przypadku opaski Empatica E4, która jest powszechnie wykorzystywana w badaniach naukowych [96, 96]. Taki sposób pomiaru pozwoliłby na swobodne wykonywanie codziennych czynności bez zakłóceń. Jednak wada tego rozwiązania polega na przeniesieniu miejsca pomiaru z palców dłoni na nieco mniej reaktywne okolice nadgarstka.

Analiza sygnału EKG w przypadku tego typu urządzenia mogłaby przebiegać analogicznie jak w przypadku badania Holter EKG, gdzie sygnał EKG jest rejestrowany cały czas. Możliwe byłoby również rozwiązanie polegające na akwizycji sygnału w pewnych zadanych interwałach [97, 98]. Wartość dodaną stanowiłoby również

uwzględnienie dodatkowych odprowadzeń dla sygnału EKG, co zwiększyłoby odporność na zakłócenia i zapewniło szerszy zakres analizy sygnału. Analizę można również przeprowadzić na bardziej zaawansowanym poziomie po konsultacjach z ekspertami, w tym kardiologami, neurologami i psychologami. Przewody od poszczególnych elektrod poprowadzone od urządzenia na nadgarstku do klatki piersiowej powinny być ukryte blisko ciała pod swojego rodzaju materiałową siatką dociskającą przewody do ciała, tak jak w przypadku przeprowadzania wysiłkowego badania EKG lub w inny sposób ograniczający ich ruch, aby przewody nie przeszkadzały w codziennym użytkowaniu urządzenia.

Z racji wyboru urządzenia typu smartwatch intuicyjnie nasuwa się pomysł umieszczenia czujnika fotopletyzmoграфicznego (PPG – ang. *Photoplethysmogram*), który dostarczyłby razem z analizą sygnału elektrokardiograficznego informacji o respiracji po przeprowadzonej analizie dwóch sygnałów polegającej między innymi na synchronizacji obu sygnałów, analizie interwałów między załamkami R-R i amplitudy sygnału PPG [99]. Tym sposobem możliwe byłoby wyeliminowanie ingerującego w komfort respondentów czujnika częstości oddechów. Moduł z czujnikiem fotopletyzmoграфicznym przedstawiony na rysunku 5.1 w pierwszym podejściu Autor zamierzał wykorzystać w ramach realizowanej pracy. Niestety, ze względu na bardzo obszerne zagadnienie poruszane w ramach tej tematyki, trudność w integracji z mikrokontrolerem, dużą ilości zbieranych danych oraz ograniczoną ilość czasu idea ta ostatecznie została odrzucona.



Rys. 5.1 Czujnik fotopletyzmoграфiczny zaproponowany do urządzenia docelowego

Przedstawiony moduł z układem MAX30101 firmy SparkFun zawiera diody z zakresu światła podczerwonego (880nm), czerwonego (660nm) i zielonego (527nm) oraz

fotodetektor do którego trafiają dobite od ciała zakresy fali. Posiada interfejs I2C, który cechuje się dość dużą szybkością transmisji danych oraz wykorzystaniem jedynie dwóch wejść mikrokontrolera nie wliczając w to linii zasilania [100]. Dzięki diodom z zakresu długości fal czerwonych oraz podczerwonych możliwe jest przeprowadzenie analizy saturacji krwi, co potencjalnie dostarczyłoby dodatkowych informacji o stanie emocjonalnym [101]. Przedstawiony moduł zawiera także zintegrowany czujnik temperatury, który mógłby dostarczyć dodatkowych istotnych informacji o stanie emocjonalnym [48].

Dobrym rozwiązaniem wydaje się być także analiza sygnału fonicznego skupiająca się na analizie cepstralnej konkretnych głosek mowy oraz odczytanych formantów, nad czym Autor pracował podczas swojej pracy inżynierskiej [102]. Wyniki przeprowadzonych wówczas badań były bardzo obiecujące pod kątem wykrywania stresu przy pomocy tego rodzaju sygnału. W celu przeprowadzenia tej analizy wymagane jest umieszczenie mikrofonu wewnątrz proponowanego urządzenia. Wartym przetestowania pomysłem wydaje się być wykorzystanie w tym celu mikrofonu w telefonie, który stał się na przestrzeni ostatnich lat nieodłącznym narzędziem ogromnej grupy ludzi na świecie.

Następnym elementem docelowej konstrukcji mogłaby być analiza obrazu za pomocą kamery przy użyciu interfejsu DCMI (ang. *Digital Camera Interface*) dostępnego w używanym mikroprocesorze STM32F446RE [66]. Analiza obrazu mogłaby się opierać o eksplorację zdjęcia twarzy wykonywanego podczas sprawdzenia godziny, co zapewniałoby pewną dozę dyskrecji podczas użytkowania. Aby zachęcić użytkownika do poprawnego wykonania zdjęcia godzina mogłaby pokazywać się po poprawnie wykonanym zdjęciu, co można byłoby zdefiniować w ustawieniach urządzenia. Analiza obrazu lub ukazanie godziny mogłoby być wyzwalane przez żyroskop wykrywający zmianę płaszczyzny, w której się znajduje, a przy okazji możliwe byłoby analizowanie gestykulacji ruchów danej ręki czy ilość nerwowych mikro ruchów. Dodatkowo dzięki żyroskopowi możliwa byłaby do przeprowadzenia analiza liczby wykonanych kroków czy detekcja upadków i potencjalnych zdarzeń niebezpiecznych.

Potencjalnie warto rozważyć możliwość dodania funkcji sfigmomanometru poprzez stworzenie zapięcia urządzenia w formie mankietu do pomiaru ciśnienia krwi. Taki mankiet mógłby być wykorzystywany we współpracy ze smartwatchem w odpowiednich momentach. Pomiar ciśnienia skurczowego i rozkurczowego krwi oraz jego właściwe wykorzystanie stanowiłoby znaczące uzupełnienie funkcji docelowego

urządzenia. Przykład implementacji sfigmomanometru w tym kontekście można znaleźć w popularnym smartwatchu Huawei Watch D dostępnym na rynku [103, 104].

5.2 Konfiguracja programowa docelowego urządzenia

Implementacja oprogramowania obsługującego wymienione moduły powinna być wykonana na mikroprocesorze obsługującym minimum protokoły Bluetooth 5 oraz Wi-Fi jak mikroprocesory firmy Qualcomm [105]. Alternatywnym podejściem jest wykorzystanie dwóch mikroprocesorów komunikujących się między sobą, gdzie jednym z nich mógłby być używany w ramach projektu STM32F446RE lub lepszy. Drugim sugerowanym mikroprocesorem mógłby być w tej konfiguracji mikroprocesor z rodziny ESP32 firmy ESPRESSIF [106]. Układy z rodziny ESP32 znajdują zastosowanie w sprzętach IoT (ang. *Internet of Things*) ze względu na obsługę wewnętrznych modułów bezprzewodowych takich jak Wi-Fi czy Bluetooth w wersji 5 (LE – ang. *Low Energy*). Dodatkowo zawierają 32 bitową strukturę mikroprocesora. Układ ten po uzyskaniu dostępu do sieci mógłby dokonywać aktualizacji oprogramowania oraz wysyłać próbki danych do dalszego rozwoju modelu sieci neuronowej. Dzięki temu rozwiązaniu możliwe byłoby zapewnienie długotrwałego wsparcia programowego czy zabezpieczeń bezpieczeństwa wrażliwych danych biomedycznych.

Dane z pomiarów z racji na swoją objętość musiałyby być zapisywane na zewnętrznej karcie pamięci lub pamięci EEPROM z niższą częstotliwością próbkowania kosztem stosunkowo niewielkiej utraty jakości rejestrowanych sygnałów. Przy tych założeniach częstotliwość próbkowania sygnału fotopletyzmograficznego mogłaby wynosić 24 Hz, natomiast sygnał elektrodermalny mógłby być rejestrowany z częstotliwością 15 Hz. Sygnał EKG mógłby być próbkowany z częstotliwością 500 Hz, gdyż wartość ta pozwala na dość precyzyjne odwzorowanie sygnału i analizę między innymi poziomu HRV [97, 98]. Ponieważ pełen potencjał analizy sygnału EKG był niewykorzystany ze względu na trudność w analizie, słabą jakość zarejestrowanego sygnału oraz konieczność, stworzenia zaawansowanych algorytmów, należy przeprowadzić w tym kierunku dalsze badania.

Dzięki możliwości konwersji kodu napisanego w oprogramowaniu MATLAB na kod napisany w języku C przy użyciu dedykowanych aplikacji oprogramowania możliwa jest stosunkowo szybka migracja projektu na wybrany mikrokontroler. Po odpowiedniej

adaptacji, program implementowany na mikrokontrolerze powinien być pisany bezpośrednio przy użyciu rejestrów mikrokontrolera (bare-metal programming). Dzięki temu działanie urządzenia będzie szybsze, stabilniejsze i bardziej bezpieczne, a kod będzie zajmował mniejszą ilość miejsca w pamięci urządzenia. Te korzyści wynikają z faktu, że programista ma praktycznie pełną kontrolę nad oprogramowaniem mikrokontrolera i ma możliwość ingerowania w wiele niskopoziomowych aspektów działania urządzenia.

Filtracja sygnałów po przeprowadzeniu odpowiednich badań sprawdzających stosunek jakości sygnału do wykorzystywanej w tym celu mocy obliczeniowej powinna być poddana optymalizacji. Z racji na niewielką objętość zajmowanej pamięci FLASH mikroprocesora, łatwą implementację kodu i stosunkowo niewielką wymaganą moc obliczeniową preferowane byłoby używanie transformaty falkowej zamiast filtrów FIR. Na przestrzeni realizacji projektu stwierdzono, że dekompozycja sygnału EDA przy użyciu funkcji *cwxEDA* jest dużo bardziej obciążająca pod względem konieczności wykonywania obliczeń. W efekcie w docelowej konstrukcji należy zatem zastąpić algorytm metodą mniej złożoną. Przetwarzanie sygnałów realizowane w oprogramowaniu MATLAB obejmujące ekstrakcję cech dla 14 respondentów trwało niemal 30 minut z wykorzystaniem jednostki roboczej *MSI GV72 8RC* wyposażonej między innymi procesorem Intel CORE i5 8 generacji i 24GB pamięci RAM. Blisko 90% czasu trwania programu trwała dekompozycja sygnałów EDA na składowe oraz ich analiza.

Zaproponowane w poprzednim rozdziale dwa odmienne podejścia do klasyfikacji emocji nie wykluczają się wzajemnie, a docelowa konstrukcja systemu multimodalnego może zawierać zaimplementowane dwa przedstawione modele sieci neuronowych do klasyfikacji emocji. Wybór konkretnego modelu klasyfikatora mógłby się odbywać poprzez aplikację mobilną zainstalowaną na smartfonie, która łączyłaby się z urządzeniem przy pomocy transmisji Bluetooth. Początkowo użytkownik mógłby korzystać z modelu ogólnego nauczonego na danych wielu użytkowników (jak przedstawiony model w rozdziale 4.1), który byłby optymalizowany do użytkującego go badacza po przeprowadzonym wywiadzie. Zebrane dane z okresu kilku dni mogłyby następnie być wysyłane z użyciem sieci internetowej do administratora, gdzie model byłby uczony oraz dostosowywany do indywidualnych preferencji użytkownika.

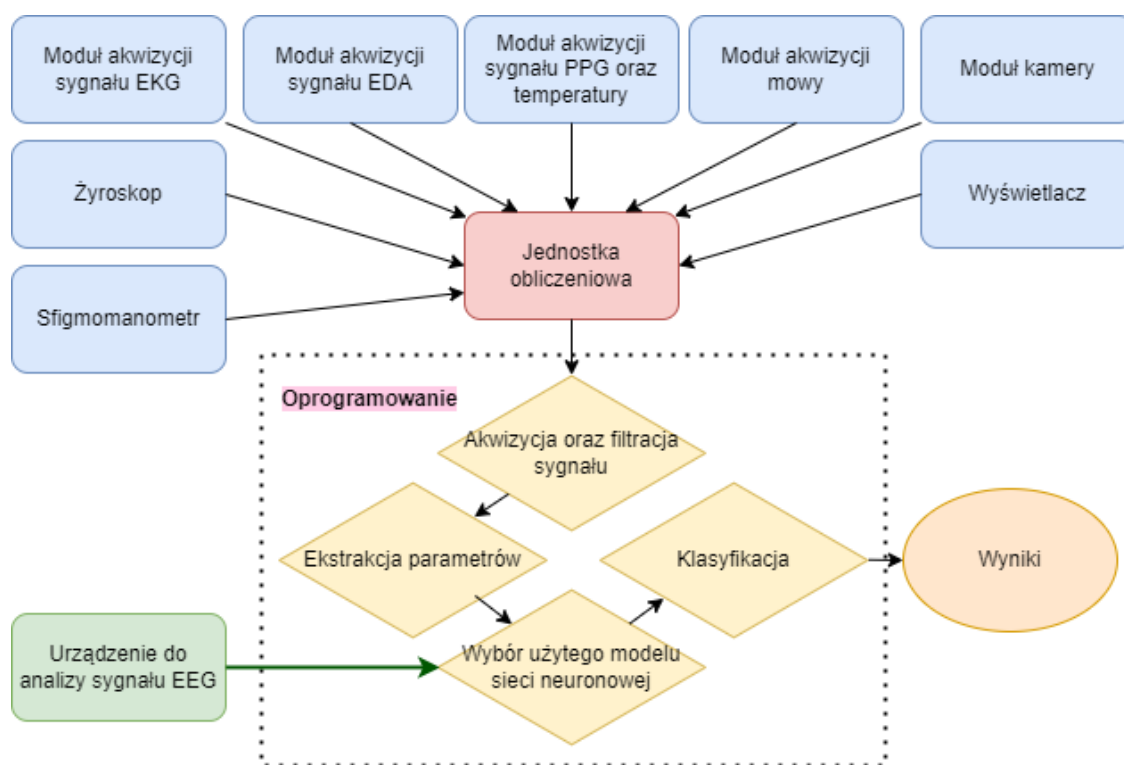
Wygenerowany na koniec każdego dnia raport przedstawiający godzinowy wykres odczuwanych emocji mógłby być edytowany i zatwierdzany przez użytkownika. Dzięki

temu dane wykorzystywane do optymalizacji sieci neuronowej byłyby sklasyfikowane, a raport po dokonaniu poprawek przedstawiał mniej błędów.

Z uwagi na fakt zebrania po pewnym czasie dużej liczby unikalnych danych o użytkowniku możliwe byłoby stworzenie na podstawie przedstawionego urządzenia ubieralnego biometrycznego systemu kontroli dostępu do wybranych stref, gdzie urządzenie łącząc się automatycznie z odpowiednio umieszczonymi w obiekcie routerami Wi-Fi śledziłoby ruch danej osoby oraz umożliwiałoby lub odmawiało wstępu.

5.3 Schemat docelowego urządzenia

Zaproponowany schemat blokowy docelowego urządzenia typu smartwatch z modułami i funkcjonalnościami przedstawiono na rysunku 5.2.



Rys. 5.2 Zaproponowany schemat blokowy docelowego urządzenia

Główna jednostka obliczeniowa, czyli mikroprocesor w proponowanej konfiguracji docelowego urządzenia pełniłaby główną rolę. Dokonywałaby akwizycji poszczególnych sygnałów w określonej kolejności aby uchwycić korelację między stanem emocjonalnym użytkownika, a zarejestrowanymi sygnałami. Następnie dokonywana byłaby filtracja sygnałów przy użyciu optymalnych metod, ekstrakcja

deskryptorów z zarejestrowanych sygnałów oraz klasyfikacja emocji na podstawie nauczonego modelu sieci neuronowej. Końcowy wynik dostępny byłby w aplikacji na telefonie użytkownika, a po połączeniu telefonu z siecią internetową generowane byłyby aktualne zalecenia dotyczące poprawy parametrów. W celu minimalizacji urządzeń rejestrujących istotnym byłoby możliwie maksymalne wykorzystanie smartfonu w całym procesie.

Kolejne urządzenie uzupełniające działanie systemu multimodalnego mogłoby przeprowadzać analizę sygnału elektroencefalograficznego (EEG). Sygnał EEG jest wykorzystywany do reprezentowania aktywności fal mózgowych i jest rejestrowany za pomocą technik nieinwazyjnych. Proces akwizycji sygnału polega na umieszczeniu elektrod w odpowiednich obszarach głowy osoby badanej. Analiza fal EEG jest wyjątkowo wymagająca ze względu na szeroki zakres częstotliwości badanych sygnałów, który sięga od kilku Hz do kilkuset Hz, ilość uzyskanych danych czy trudność w znalezieniu korelacji sygnałów. Trudności w analizie zależą także od złożoności problemu, a klasyfikacja emocji jest nadal uznawana za problem trudny oraz wymagający. Najnowsze prace badawcze w dużej mierze skupiają się właśnie na analizie sygnałów EEG. Wyniki wydają się być bardzo obiecujące, a ciągły postęp i dalsze badania w przyszłości mogą uczynić problem detekcji emocji bardzo prostym dzięki analizie EEG oraz sygnałów skorelowanych [107].

Docelowy system mógłby znaleźć zastosowanie w najróżniejszych dziedzinach jak psychoterapia, rozrywka, edukacja czy wojsko i dostarczyć znacznych korzyści.

Podsumowanie

Celem niniejszej pracy było opracowanie koncepcji multimodalnego systemu detekcji emocji. Aby podjąć się tego zadania *Autor* w pierwszej kolejności dokonał obszernego przeglądu literatury pozwalającego na odpowiednie ukierunkowanie dalej podjętych działań. Skupiały się one na analizie reakcji organizmu ludzkiego kontrolowanych przez autonomiczny układ nerwowy, które odzwierciedlają między innymi sygnał EKG, EDA oraz sygnał respiracji. Aby dokonać akwizycji wybranych biosygnali potencjalnie świadczących o rodzaju odczuwanej emocji postanowiono zaprogramować mikrokontroler z rodziny STM32 obsługujący 3 moduły czujników. Jednym z nich był czujnik do akwizycji sygnału EKG bazujący na układzie AD8232. Kolejnym był moduł Grove z układem Lm324 służący do rejestracji aktywności elektrodermalnej. Ostatni moduł do śledzenia zmian sygnału respiracji bazujący na czujnikach termistorowych został zaprojektowany od podstaw i stworzony przez Autora. Następnie przy użyciu stworzonej platformy przeprowadzono nadzorowane badania na 20 osobowej grupie respondentów. Badania te polegały na prezentacji wybranych afektywnych wideo materiałów respondentom z założoną aparaturą pomiarową, która rejestrowała 3 wymienione sygnały życiowe, w efekcie czego zarejestrowano 7 odmiennych stanów emocjonalnych, dla których średni czas trwania sygnału wynosił 2 minuty. Zarejestrowane sygnały przy użyciu oprogramowania MATLAB poddano - odpowiedniej do rodzaju sygnału - filtracji używając między innymi transformaty falkowej oraz zaprojektowanych filtrów FIR. Odfiltrowane sygnały EDA poddano dekompozycji na składową SCL i SCR, na bazie których wyekstrahowano cechy dystynktywne. Z sygnału EKG przy użyciu odpowiednich algorytmów dokonano detekcji zespołów QRS, w wyniku czego wyznaczono między innymi HR. Natomiast analizując sygnał respiracji wyznaczono momenty wdechu i wydechu powietrza. Zdefiniowane na tej podstawie parametry *stanu neutralnego* oraz *stanu wstrętu do okaleczeń* dla 14-stu respondentów posłużyły jako dane uczące sieci neuronowej *Narrow Neural Network* wygenerowanej przy użyciu aplikacji Classification Learner, dla której uzyskano 78,6% poprawności rozpoznania. Następnie badania skupiono wybierając wyłącznie jednego respondenta i przeprowadzono dodatkowe dwadzieścia 4 minutowych pomiarów, które miały na celu wywoływanie rejestrację wspomnianych wcześniej dwóch stanów emocjonalnych. Sygnały poddano analogicznej filtracji oraz

ekstrakcji cech, a uzyskane deskryptory posłużyły tym razem jako dane uczące w aplikacji *Neural Net Pattern Recognition*, gdzie uzyskano 100% poprawnej klasyfikacji danych uczących, testowych i walidacyjnych z dużą pewnością klasyfikacji sieci. Uzyskane wysokie wyniki klasyfikacji obu sieci neuronowych pozwalają stwierdzić, że wybrane cechy dystynktywne zawierają wystarczającą ilość informacji różnicujących stan neutralny oraz stan wstrętu do okaleczeń. Cechami tymi okazały się być dla obu sieci neuronowych takie parametry jak: ilość oddechów wykonywanych w ciągu minuty, wartość tętna, ilość znalezionych szczytów składowej SCR o zadanej szerokości w ciągu minuty, średnia wartość składowej SCR w ciągu minuty oraz PPA MAD znalezionych załamków R. Aby potwierdzić zasadność przyjętych ustaleń należy zrealizować powiększony zakres prac badawczych przeprowadzonych z zastosowaniem zoptymalizowanej metodyki przeprowadzania badań. Nowo powstałe procedury powinny obejmować powiększoną bazę respondentów, wykorzystywać specjalistyczne urządzenia oraz rozpatrywać poszerzony zakres cech dystynktywnych do opisanie problematyki różnicowania emocji. Na podstawie przeprowadzonych badań oraz analizy ich wyników zaproponowano schemat docelowego urządzenia systemu multimodalnego w formie smartwatcha wraz z szczegółowym opisem konstrukcji oraz zaproponowanymi usprawnieniami sprzętowymi i programowymi. Najprawdopodobniej docelowe urządzenie dzięki uwzględnieniu wymienionych uwag mogłoby dokonywać klasyfikacji większej ilości emocji z lepszą precyzją. Zaproponowano także kierunki kolejnych badań i rozwoju urządzeń, które mogłyby dokonywać detekcji ludzkich emocji.

Bibliografia

1. <https://sjp.pwn.pl/slowniki/emocje.html>, (21.09.2023 r.)
2. <https://www.mdpi.com/2076-3417/13/4/2573>
3. <https://www.sciencedirect.com/science/article/pii/S157106611930009X>
4. <https://link.springer.com/article/10.1007/s10772-011-9125-1>
5. <https://www.mdpi.com/1424-8220/20/3/592>
6. <https://hashdork.com/pl/web-5-0-explained/>
7. <https://ieeexplore.ieee.org/document/5999653>
8. <https://zpe.gov.pl/a/uklad-nerwowy/D7nf5jx0A>
9. <https://zdrowie.gazeta.pl/Zdrowie/7,140283,25033878,autonomiczny-uklad-nerwowy-czyli-sprawca-reakcji-niezaleznych.html>
10. <https://www.ncbi.nlm.nih.gov/books/NBK507838/>
11. https://en.wikipedia.org/wiki/Respiratory_rate
12. <https://www.ncbi.nlm.nih.gov/books/NBK537306/>
13. <https://www.sciencedirect.com/science/article/abs/pii/S0301051110000827>
14. <https://www.ncbi.nlm.nih.gov/books/NBK541062/>
15. <https://my.clevelandclinic.org/health/symptoms/24832-bradypnea>
16. <https://www.mp.pl/pacjent/objawy/151334.zaburzenia-oddychania>
17. <https://iopscience.iop.org/article/10.1088/1361-6579/ab299e/meta>
18. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8769001/>
19. <https://www.mdpi.com/1424-8220/19/4/908>
20. dr hab. inż. Jacek Jakubowski Wyk. 2. Techniki pomiaru sygnałów biomedycznych (część 1) sensory elektrody aparatura do badań 2. Techniki pomiaru sygnałów biomedycznych (część 1) sensory elektrody aparatura do badań. 2023r.
21. <https://blogs.ubc.ca/conphyusedgrad/2017/11/10/thermistor-airflow-sensors/>
22. <https://chem.pg.edu.pl/documents/175361/84965897/002.pdf>
23. https://brain.fuw.edu.pl/edu/index.php/Biologia_Kom%C3%B3rki/Budowa_i_funkcje_struktur
24. https://www.youtube.com/watch?v=V1BYVdrH_vM
25. https://en.wikipedia.org/wiki/Cardiac_action_potential
26. <http://www.vhlab.umn.edu/atlas/conduction-system-tutorial/cardiac-action-potentials.shtml>
27. <https://www.ncbi.nlm.nih.gov/books/NBK537194/>
28. <http://www.nataliescasebook.com/tag/e-c-g-basics>
29. <https://www.sciencedirect.com/science/article/abs/pii/0002914964900724>

30. <https://kredos.pl/artykuly/kardiologia/aparat-ekg-zasada-dzialania-i-interpretacja-wynikow-2>
31. <https://szkolnictwo.pl/szukaj/Elektrokardiografia>
32. <https://efizyka.net.pl/dipol-elektryczny-elektryczny-moment-dipolowy>
33. <https://pl.wikipedia.org/wiki/Elektrokardiografia>
34. <https://aptekazawiszy.pl/artykuly/czy-mozna-odczytac-swoj-wynik-ekg-bez-pomocy-lekarza-sprawdz.html>
35. <https://fizjotechnologia.com/przegląd-sprzetu/jak-rozumiec-i-interpretowac-krzywa-ekg.html#>
36. https://journals.viamedica.pl/kardiologia_polska/article/download/79588/61622
37. https://www.ka.edu.pl/download/gfx/ksw/pl/defaultopisy/1054/5/1/mcr1_-_ekg_w_pigulce.pdf
38. <https://www.mdpi.com/1424-8220/20/2/479dpi.com/1424-8220/20/2/479>
39. <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1469-8986.2012.01384.x>
40. <https://www.intechopen.com/chapters/70053>
41. https://connect.tobii.com/s/article/where-to-place-gsr-electrodes?language=en_US
42. https://en.wikipedia.org/wiki/Heart_rate
43. https://en.wikipedia.org/wiki/Heart_rate_variability
44. https://onlinelibrary.wiley.com/doi/full/10.1002/mus.25573?casa_token=FRwPfZzfmOwAAAAA%3AAKG0jqLyxbzRavubuXmTh09uAiUh6oX4vN-ftD-DsomMzmVFyucFeyDRdg1tEixoIWS_zaLZOEXXbgM
45. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5624990/>
46. <https://content.iospress.com/articles/technology-and-health-care/thc174747>
47. <https://ieeexplore.ieee.org/abstract/document/5190120>
48. <https://link.springer.com/article/10.1186/2192-1962-3-3>
49. <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/>
50. <https://csea.phhp.ufl.edu/media/iapsmessage.html>
51. <https://www.seeedstudio.com/Grove-GSR-sensor-p-1614.html>
52. <https://www.onsemi.com/pdf/datasheet/lm324-d.pdf>
53. <https://botland.com.pl/grove-czujniki-medyczne/11314-grove-czujnik-gsr-oporu-elektrycznego-skory-5903351246545.html>
54. <https://www.sparkfun.com/products/12650>
55. <https://botland.com.pl/czujniki-medyczne/2654-czujnik-do-pomiaru-aktywnosci-serca-pulsometr-ad8232-sparkfun-sen-12650-5904422373436.html>
56. <https://www.analog.com/media/en/technical-documentation/data-sheets/ad8232.pdf>

57. <https://botland.com.pl/czujniki-medyczne/2605-przewody-do-elektrod-biomedycznych-sparkfun-cab-12970-5904422373405.html>
58. <https://www.sparkfun.com/products/12970>
59. <https://botland.com.pl/czujniki-medyczne/2604-elektrody-biomedyczne-10szt-sparkfun-sen-12969-5903351245159.html>
60. <https://www.sparkfun.com/products/12969>
61. <https://www.ternimed.de/Alice-5-respiratory-sensor-nasal/oral-Thermistor>
62. <https://www.infinite-electronics.is/pdf/4045804917/MF52A1473J3950.pdf>
63. http://www.zstio-elektronika.pl/pliki_t_elektronik/1%20Ta-%2025_05_2020%20Badanie%20element%C3%B3w%20uk%C5%82ad%C3%B3w%20analogowych-%20Termistor.pdf
64. <https://www.biopac.com/wp-content/uploads/EDA-SCR-Analysis.pdf>
65. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6085204/>
66. <https://www.st.com/resource/en/datasheet/stm32f446mc.pdf>
67. https://www.st.com/resource/en/reference_manual/rm0390-stm32f446xx-advanced-armbased-32bit-mcus-stmicroelectronics.pdf
68. <https://www.st.com/en/evaluation-tools/nucleo-f446re.html#documentation>
69. https://www.st.com/resource/en/application_note/an3116-stm32s-adc-modes-and-their-applications-stmicroelectronics.pdf
70. https://www.st.com/content/ccc/resource/technical/document/application_note/group0/91/01/84/3f/7c/67/41/3f/DM00236305/files/DM00236305.pdf/jcr:content/translations/en.DM00236305.pdf
71. https://www.st.com/content/ccc/resource/technical/document/application_note/54/0f/67/eb/47/34/45/40/DM00042534.pdf/files/DM00042534.pdf/jcr:content/translations/en.DM00042534.pdf
72. https://wiki.st.com/stm32mcu/wiki/STM32CubeMonitor:How_to_log_data_in_a_.csv_file
73. <https://www.youtube.com/watch?v=ZbZSe6NBXs>
74. <https://www.youtube.com/watch?v=ja1sXvNCyO0>
75. https://www.youtube.com/watch?v=_rSuk27Bako
76. <https://www.youtube.com/watch?v=CBbqYv5FmfU>
77. <https://www.youtube.com/watch?v=uKfOedcdD68>
78. <https://www.youtube.com/watch?v=jfKfPfyJRdk>
79. <https://learn.sparkfun.com/tutorials/ad8232-heart-rate-monitor-hookup-guide/all>
80. <https://www.kenhub.com/en/library/anatomy/axillary-region>
81. https://github.com/Shooterqq/Emotions_Detection_STM32_F446RE
82. https://en.wikipedia.org/wiki/Finite_impulse_response

83. <https://livesound.pl/tutoriale/4809-filtr-cyfrowe.-iir-vs.-fir>
84. https://pl.wikipedia.org/wiki/Transformacja_falkowa
85. https://en.wikipedia.org/wiki/Wavelet_transform
86. <https://en.wikipedia.org/wiki/Wavelet>
87. <https://www.mathworks.com/help/wavelet/ref/wrcoef.html>
88. <https://www.mathworks.com/help/wavelet/ref/wavedec.html>
89. <https://www.mathworks.com/matlabcentral/fileexchange/53326-cvxeda>
90. <https://www.mathworks.com/matlabcentral/fileexchange/45840-complete-pan-tompkins-implementation-ecg-qrs-detector>
91. <https://www.mdpi.com/1424-8220/22/13/4855>
92. https://www.mathworks.com/help/stats/classificationlearner-app.html#responsive_offcanvas
93. Peter Bruce, Andrew Bruce, Peter Gedeck - Statystyka praktyczna w data science. 50 kluczowych zagadnień w językach R i Python. Wydanie II. 16.06.2021 r.
94. https://www.mathworks.com/help/deeplearning/ref/neuralnetpatternrecognition-app.html#responsive_offcanvas
95. <https://e4.empatica.com/e4-wristband>
96. <https://www.mdpi.com/1424-8220/23/7/3565>
97. <https://delibra.bg.polsl.pl/dlibra/publication/30815/edition/27365/content>
98. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6085204/>
99. <https://www.mdpi.com/2306-5354/10/9/1024>
100. <https://www.sparkfun.com/products/12650>
101. <https://www.mdpi.com/1424-8220/20/14/4037>
102. Piotr Wiski – Praca dyplomowa pt. Elektroniczny wykrywacz stresu. Wojskowa Akademia Techniczna 2022 r.
103. <https://consumer.huawei.com/pl/wearables/watch-d/buy/>
104. <https://www.mdpi.com/1424-8220/22/7/2538>
105. <https://www.qualcomm.com/products/technology/processors#product-list>
106. <https://www.espressif.com/en/products/socs/esp32-h2>
107. <https://www.mdpi.com/2076-3417/12/5/2527>

Wyrażam zgodę na
udostępnianie mojej pracy w Uczelni.

mgr inż. Piotr Wiski
Warszawa, dnia 21.09.2023 r.