

CZASOPISMO KOGNITYWISTYCZNE

# kognireceptor

1/2025



Śląskie Koło Kognitywistyczne  
"Do góry mózgi"

Uniwersytet Śląski  
w Katowicach

# Kognireceptor

czasopismo kognitywistyczne

Numer 3

Kognireceptor – czasopismo kognitywistyczne zostało stworzone przez Do Góry Mózgami – Śląskie Koło Kognitywistyczne działające na Uniwersytecie Śląskim w Katowicach.

Redakcję czasopisma stanowią: Redaktor Naczelny: Maksymilian Polewka

Zastępca Redaktora Naczelnego: Zuzanna Michura, Magdalena Oleszkiewicz

Edytor: Kaja Gałuszka

Zespół Redakcyjny w składzie: Marta Biela, Kaja Gałuszka, Mateusz Hadamczyk, Sylwia Korzon, Adam Małota, Witold Marzęda, Zuzanna Michura, Paulina Mieszczak, Magdalena Oleszkiewicz, Maksymilian Polewka, Adam Tkaczyk.

Czasopismo publikowane jest na stronie: [kognireceptor.us.edu.pl](http://kognireceptor.us.edu.pl)

1/2025

Kontakt do czasopisma: [kognireceptor.czasopismo@gmail.com](mailto:kognireceptor.czasopismo@gmail.com)

Kontakt do Śląskiego Koła Kognitywistycznego: [dogorymозgami@us.edu.pl](mailto:dogorymозgami@us.edu.pl)

## Spis treści

Od redakcji.....	5
Aleksandra Górka – <i>Być jak człowiek: rozpoznawanie cech antropomorficznych</i> .....	6
Szymon Pluta – <i>Wykrywanie i klasyfikacja choroby Alzheimera za pomocą metod uczenia maszynowego na podstawie obrazów MRI</i> .....	27
Ewelina Bąk – <i>Badanie pilotażowe semantycznej załamki N400 w zdaniach zakończonych homofonami w postaci ikon</i> .....	66
Paweł Ficek – <i>Czy posiadanie dziecięcej twarzy sprawia, że wydajemy się zabawniejsi? Raport z projektu</i>	95

## Od redakcji

Oddajemy w Państwa ręce trzeci numer „Kognireceptora” – czasopisma, które od początku swojego istnienia stara się łączyć różnorodne perspektywy badawcze w obrębie nauk o poznaniu. Cieszy nas, że projekt Śląskiego Koła Kognitywistycznego „Do Góry Mózgami” nie tylko utrzymuje się na mapie studenckich czasopism naukowych, ale z każdym numerem zyskuje nowe, inspirujące głosy.

Wydanie otwiera tekst Aleksandry Górki, który opisuje mechanizmy sprawiające, że zaczynamy dostrzegać ludzkie właściwości w obiektach niebędących ludźmi. To refleksja niezwykle ważna w dobie coraz śmielszych interakcji człowieka z maszynami. Szymon Pluta w swoim artykule pokazuje, jak narzędzia wywodzące się z badań nad sztuczną inteligencją mogą służyć diagnozie i lepszemu zrozumieniu schorzeń neurodegeneracyjnych, sięgając do danych, których ludzkie oko nie jest w stanie precyzyjnie odczytać. Z kolei Ewelina Bąk zagłębia się w fascynujący mechanizm przetwarzania języka, badając, jak nasz mózg radzi sobie z bodźcami graficznymi zastępującymi słowa. Numer zamyka Paweł Ficek, który przygląda się zjawisku *babyface* i sprawdza, czy efekt ten przekłada się także na ocenę poczucia humoru, łącząc tym samym psychologię społeczną z badaniami nad pierwszym wrażeniem i ewolucyjnymi uwarunkowaniami atrakcyjności.

Wszystkie prezentowane teksty łączy wysoki poziom merytoryczny, a ich autorzy wykorzystują szeroką gamę narzędzi: od eksperymentów behawioralnych, przez analizę danych neuroobrazowych, po badania potencjałów wywoływanych. Dziękujemy autorom i autorkom za zaufanie, jakim nas obdarzyli, oraz za trud włożony w przygotowanie artykułów. Mamy nadzieję, że lektura trzeciego numeru „Kognireceptora” będzie dla Państwa nie tylko okazją do poszerzenia wiedzy, ale także inspiracją do własnych poszukiwań.

Zapraszamy do kontaktu i życzymy miłej lektury!

*Zespół redakcyjny*

## **Być jak człowiek: rozpoznawanie cech antropomorficznych**

### **Abstrakt:**

Rozpoznawanie twarzy i ludzkiej sylwetki jest jednym z kluczowych procesów poznawczych umożliwiających funkcjonowanie w kontekście społecznym. Badania neurokognitywne wskazują, że mózg dysponuje wyspecjalizowanymi mechanizmami odpowiedzialnymi za przetwarzanie tych bodźców szczególnie w obrębie struktur kory wzrokowej, takich jak obszary wrzecionowate. Odpowiadają one za szybkie, zautomatyzowane rozpoznawanie twarzy oraz ciał, co pozwala na natychmiastowe reagowanie na sygnały społeczne i emocjonalne. Zjawiska takie jak pareidolia czy dolina niesamowitości wskazują, że system percepcyjny człowieka jest wyjątkowo wrażliwy na wszelkie podobieństwa i odstępstwa od typowej formy ludzkiej. Rozpoznawanie sylwetki i ruchu człowieka także opiera się na specyficznych wzorcach percepcyjnych, umożliwiających identyfikację nawet przy minimalnych wskazówkach wizualnych. Zdolność do dostrzegania cech ludzkich ma charakter uprzywilejowany i jest zakorzeniona w strukturach neuronalnych. Zrozumienie tych mechanizmów ma znaczenie nie tylko dla psychologii percepcji i neuronauki poznawczej, ale także dla badań nad robotyką oraz estetyką wizualną: animacją czy projektowaniem – dziedzin, które eksplorują granice między tym, co ludzkie, a tym, co tylko ludzkie przypomina.

**Słowa kluczowe:** pareidolia, dolina niesamowitości, rozpoznawanie twarzy, rozpoznawanie ludzkiej sylwetki, FFA, FBA, neuroestetyka

### **Abstract:**

The recognition of faces and the human body is one of the key cognitive processes enabling effective functioning in social contexts. Neurocognitive research indicates that the brain possesses specialized mechanisms responsible for processing these stimuli, particularly within the structures of the visual cortex, such as the fusiform areas. These regions support rapid, automated identification of faces and bodies, allowing for immediate responses to social and emotional cues. Phenomena such as pareidolia and the uncanny valley suggest that the human perceptual system is exceptionally sensitive to both similarities and deviations from the typical human form. The recognition of body shape and movement is also based on specific perceptual patterns that enable identification even with minimal visual information. The ability to detect human-like features is privileged and deeply rooted in neural structures. Understanding these mechanisms is significant not only for the psychology of perception and cognitive neuroscience but also for research in robotics and visual aesthetics – fields such as animation and design that explore the boundaries between what is human and what merely resembles the human.

**Keywords:** pareidolia, uncanny valley, face recognition, body recognition, fusiform face area, fusiform body area, neuroaesthetics

## Wprowadzenie

Twarz i ludzkie ciało to bodźce istotne społecznie i poznawczo. Człowiek posiada wyjątkową zdolność rozpoznawania cech antropomorficznych w otaczającym świecie. Ta umiejętność jest kluczowa w codziennym życiu na wielu poziomach. W artykule „Być jak człowiek: rozpoznawanie cech antropomorficznych” analizuję to zjawisko z perspektywy interdyscyplinarnej – obejmującej neurobiologię, psychologię percepcji i analizę konkretnych przykładów projektowania użytkowego. Celem tej pracy jest przedstawienie zjawisk poznawczych dotyczących analizy i rozpoznawania ludzkiej twarzy i sylwetki.

W pierwszej części skupiam się na mechanizmach rozpoznawania twarzy z punktu widzenia neurobiologii oraz towarzyszących im zjawisk, takich jak pareidolia czy efekt doliny niesamowitości, które wskazują na silną poznawczą potrzebę identyfikowania twarzy. Druga część poświęcona jest percepcji ciała i ruchu – od rozpoznawania ludzkiej sylwetki po analizę cielesności, niejednokrotnie balansującej na granicy między tym, co znajome, a tym, co obce.

## 1. TWARZ

### 1.1. Rozpoznawanie twarzy neurobiologicznie

Rozpoznawanie twarzy nie jest w pełni wyjaśnionym procesem neurologicznym. Twarz jako identyfikator, naturalnie przypisywany do osoby wraz z komunikatem niewerbalnym w postaci mimiki, pełni ogromną rolę w budowaniu więzi społecznych. Do zrozumienia przebiegu procesu rozpoznawania twarzy przybliży nas badanie zaburzeń i deficytów. Jednym z nich jest prozopagnozja. Jest to deficyt neurologiczny charakteryzujący się niezdolnością do rozpoznawania twarzy, pomimo prawidłowych funkcji poznawczych, intelektualnych i przetwarzania wzrokowego, skorelowany z uszkodzeniami w obszarach korowych<sup>1</sup>. Osoby dotknięte prozopagnozją nie odróżniają twarzy nowych od tych wcześniej im znanych. Prozopagnozja stała się punktem wyjścia pytania: które części mózgu są odpowiedzialne za rozpoznawanie i percepcję twarzy. Za pomocą obrazowania funkcjonalnego Nancy Kanwisher i współpracownikom<sup>2</sup> udało się wyodrębnić takie części obszarów korowych, które pozostają szczególnie aktywne podczas prezentowania bodźców związanych z twarzami, w porównaniu do prezentowania innych bodźców, np. obiektów. Ten obszar został nazwany *Fusiform Face Area* (FFA) i jest strukturą należącą do zakrętu wrzecionowatego, kory dolno-skroniowej<sup>3</sup>. Obszar FFA jest niewielki, a u większości badanych jest większy w prawej półkuli mózgu<sup>4</sup>. Potwierdza to także fakt, że to właśnie uszkodzenia prawej półkuli nierzadko prowadzą do prozopagnozji<sup>5</sup>.

Pojawiły się także hipotezy, podważające teorię iż twarze miałyby być jedynym bodźcem rozpoznawanym przez FFA. Merim Bilalic pisze: „Nie można zaprzeczać istotności modułu FFA w

---

<sup>1</sup> Zob. J. K. J. K. E. Steeves et al., „The Fusiform Face Area Is Not Sufficient for Face Recognition: Evidence from a Patient with Dense Prosopagnosia and No Occipital Face Area”, *Neuropsychologia* 44, nr 4 (2006): 3.

<sup>2</sup> Zob. N. Kanwisher et al., „The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception”, *Journal Neuroscience* 17, nr 11 (1997): 4304.

<sup>3</sup> Tamże, 4302.

<sup>4</sup> Tamże, 4306.

<sup>5</sup> Tamże, 4302.

rozpoznawaniu twarzy. Uszkodzenia modułu i struktur mu bliskich powodują niezdolność rozpoznawania twarzy, a twarze aktywują moduł FFA dwa razy częściej niż inne bodźce”<sup>6</sup>. Twarze są nie tylko bodźcem kluczowym, ale przede wszystkim bodźcem bardzo często przez nas trenowanym. Trudno jest więc ocenić, czy struktura ta jest ewolucyjnie przystosowana do kategorii bodźca jakim są twarze, czy może mieć także inne funkcje. Bilalić wysnuwa tezę, że FFA nie jest obszarem odpowiedzialnym jedynie za rozpoznawanie twarzy, ale, że bierze udział w rozróżnianiu kategorii wszelkich obiektów<sup>7</sup>. Przeprowadza eksperyment, w którym eksperci szachowi i nowicjusze przyglądają się konfiguracjom figur szachowych. Eksperyment potwierdził, iż aktywność obszaru FFA podczas przyglądania się szachownicy była wyższa wśród ekspertów szachowych niż u nowicjuszy<sup>8</sup>. Obszar FFA może więc być odpowiedzialny nie tylko za rozpoznawanie twarzy, ale także za rozpoznawanie dobrze nam znanych i często doświadczanych bodźców. Szachownica, tak jak twarz, składa się z indywidualnych części, które możemy postrzegać w odosobnieniu, ale ich kontekst zmienia się w relacji do innych figur. W tym sensie postrzeganie szachownicy przez szachowych ekspertów przypomina postrzeganie twarzy. Obserwowanie gońca można porównać do obserwowania jednego oka, a układ figur na szachownicy do oglądania twarzy. „Moduł FFA nie jest modułem odpowiedzialnym jedynie za rozpoznawanie twarzy, a raczej bardziej uniwersalnym mechanizmem mózgu, który poprzez kontakt z konkretnym bodźcem szkoli się i działa poprzez scalanie ze sobą indywidualnych części bodźca w całość. FFA nie jest odpowiedzialny za indywiduację, a jego celem jest szybkie i efektywne przetwarzanie złożonych bodźców”<sup>9</sup>. Bilalić traktuje twarz jako bodziec, którego rozpoznawanie można wytrenować dokładnie tak, jak każdy inny.

Jednymi z słynniejszych badań, które dowodzą mają wrodzonym umiejętnościom rozpoznawania twarzy są badania na noworodkach, przedstawione po raz pierwszy przez Carolyn C. Goren, Merrill Sarty i Paula Y. Wu w 1975 roku. Badania wykazały, że już kilka minut po porodzie noworodki chętniej przyglądały się twarzom harmonijnym niż tym zniekształconym i kształtom, które nie przypominały twarzy<sup>10</sup>. W badaniu replikacyjnym przeprowadzonym w 1991 przez Marka Johnsona i współpracowników potwierdzono, że noworodki w pierwszej godzinie życia chętniej przyglądają się twarzom o prawidłowym rozmieszczeniu części składowych<sup>11</sup>. Znaczna różnica występowała jednak pomiędzy kształtami pustymi, czyli bodźcem „mniej interesującym”, a pomiędzy prawidłowym kształtem przypominającym twarz, niż pomiędzy twarzą standardową, a tą z poprzestawianymi częściami składowymi. W związku z tym, we wnioskach autorzy podkreślili, że warto badaniu poddać taki bodziec, w którym zamiast standardowych części składowych twarzy, takich jak oczy czy nos, w tym samym rozmieszczeniu zawiera inne obiekty niezwiązane z twarzą<sup>12</sup>. Badania potwierdzają, że twarz jest bezsprzecznie bodźcem interesującym dla noworodków – w badaniu częściej obracały głowę w stronę wzroku za twarzą, niż za nie-twarzą. John Morton i Mark

---

<sup>6</sup> M. Bilalić, „Revisiting the Role of the Fusiform Face Area in Expertise”, *Journal Cognitive Neuroscience* 28 (2016): 1345.

<sup>7</sup> Tamże, 1345.

<sup>8</sup> Tamże, 1349.

<sup>9</sup> Tamże, 1356.

<sup>10</sup> Zob. C. C. Goren et al., „Visual Following and Pattern Discrimination of Face-like Stimuli by Newborn Infants”, *Pediatrics* 56, nr 4 (1975): 544.

<sup>11</sup> Zob. M. H. Johnson et al., „Newborns’ Preferential Tracking of Face like Stimuli and Its Subsequent Decline”, *Cognition* 40, nr 1–2 (1991): 7.

<sup>12</sup> Tamże, 10.

H. Johnson odwołują się do kontekstu etologicznego, przywołując przykład kurcząt, które po wykluciu podążają za przedmiotem, który w jakiś sposób przypomina im kształt głowy dorosłego osobnika, tak jak w słynnym przypadku Konrada Lorenza i jego butów wojskowych<sup>13</sup>. Według autorów w procesie wdrukowania (*imprintingu*) noworodki, analogicznie do kurcząt, które podążają za bodźcem przypominającym im matkę, zaczynają preferować widok twarzy.

Inną ważną obserwacją, która podkreśla istotność twarzy jako bodźca u noworodków i niemowląt jest ich umiejętność mimikry wyrazów twarzy takich jak otwieranie ust i wysuwanie języka, które zostało zaobserwowane przez Andrew N. Meltzoffa i M. Keitha Moore<sup>14</sup> i nazwane aktywnym wielozmysłowym naśladowaniem (*active intermodal matching*)<sup>15</sup>. Badaniu poddawano niemowlęta w wieku kilkunastu tygodni, którym matka, bądź nieznajomy, który miał nie przypominać rodzica, pokazywał się z otwartymi ustami lub wysuniętym językiem. Znacząca większość dzieci imitowała dorosłych, z przewagą naśladowania nieznajomego. Prawidłowa obserwacja twarzy jako kluczowego bodźca w rozpoznawaniu emocji jest ważna dla rozwoju dzieci ze względu na społeczny wymiar twarzy. Mark H. Johnson na podstawie wieloletnich badań, także badań deficytów takich jak prozopagnozja, ślepowidzenie (*blindsight*) czy zespół nieuwagi stronnej<sup>16</sup>, jako miejsce rozpoznania bodźca jako twarzy uznaje drogę podkorową, czyli: wzgórek górny, poduszkę międzymózgowia i ciała migdałowe<sup>17</sup>, natomiast identyfikacja następuje w rejonie FFA (zakręt wrzecionowaty) oraz zakręcie potylicznym dolnym<sup>18</sup>. Nie wyjaśnia to natomiast, dlaczego noworodki w ciągu godziny od narodzin są bardziej wyczulone na widok twarzy.

Większość badań związanych z rozpoznawaniem twarzy skupia się na bodźcach wzrokowych. FFA był wiązany z bodźcami wzrokowymi, także ze względu na powiązanie z pierwszorzędną korą wzrokową i głównie na tym aspekcie skupiały się badania dotyczące tego obszaru. Badania osób niewidomych od urodzenia pozwoliły ustalić, że obszar FFA aktywuje się także u nich podczas haptycznego eksplorowania modelu twarzy 3D, dokładnie tak samo jak w przypadku bodźców wzrokowych<sup>19</sup>, co oznacza, że bodźce wzrokowe mogą zostać zastąpione przez inne bodźce związane z twarzą i wciąż aktywować obszar FFA. Autorzy badania potwierdzają to eksperymentem, który używa dźwięków związanych z twarzą (np. śmiech czy żucie). Porównanie obszarów aktywowanych bodźcami słuchowymi i haptycznymi, pokazało zbieżność w obszarze FFA<sup>20</sup>. Obszar zakrętu wrzecionowatego wydaje się więc modułem, który dotyczy bezpośrednio bodźców, niezależnie od zmysłu je dostarczającego, związanych z twarzą. Autorzy skłaniają się ku hipotezie, która wskazuje na sąsiedztwo obszaru FFA z strukturami wyższej kory mózgowej, co pozwala wyjaśnić dlaczego twarz, jako ważny bodziec ludzkiego życia społecznego, korzysta z osobnego modułu do jego rozpoznawania. Jedna z hipotez mówi, że ten obszar kory mózgowej staje się selektywny dla twarzy, ponieważ otrzymuje ustrukturyzowane informacje wizualne, przekazywane przez specyficzne

<sup>13</sup> Zob. M. H. Johnson i J. Morton, *Biology and Cognitive Development: The Case of Face Recognition* (Blackwell Scientific Publications, 1993), 66.

<sup>14</sup> Zob. A. N. Meltzoff i M. K. Moore, „Early Imitation within a Functional Framework: The Importance of Person Identity, Movement, and Development”, *Infant Behavior and Development* 15, nr 4 (1992): 488.

<sup>15</sup> Zob. J. Jaracz: „Neurobiologia percepcji ekspresji emocji twarzy”, *Neuropsychiatria i Neuropsychologia* 5 (2010), 11.

<sup>16</sup> Zob. M.H. Johnson: „Subcortical face processing”, *Nature Reviews Neuroscience* 6 (2005), 10: 766.

<sup>17</sup> Tamże, 766.

<sup>18</sup> Tamże, 767.

<sup>19</sup> N. A. Ratan Murty, Teng, D. Beeler i in.: „Visual experience is not necessary for the development of face-selectivity in the lateral fusiform gyrus”, *Proceedings of the National Academy of Sciences* 117 (2020), 37: 23012.

<sup>20</sup> Tamże.

połączenia z pierwszorzędowej kory wzrokowej, gdzie najczęściej występują bodźce związane z twarzą. Nie jest to jednak możliwe w przypadku osób niewidomych od urodzenia, ponieważ aktywacja tego obszaru poprzez korę wzrokową nigdy nie miała u nich miejsca<sup>21</sup>. Mimo, że udało się dowiedzieć, że obszar FFA jest aktywny także u osób niewidomych od urodzenia, do tej pory nie wiemy, czy rozpoznawanie twarzy jest wrodzone. Wymagałoby to przeprowadzenia badań na niewidomych noworodkach, które nie miały haptycznego kontaktu z twarzami, co utrudnia także fakt, że noworodki w życiu płodowym mogą mieć haptyczny kontakt z własną twarzą.

Wiemy więc, że struktury podkorowe są odpowiedzialne za rozpoznanie bodźca jako twarz, a struktura FFA za identyfikację twarzy. Nie wiemy natomiast, czy u wszystkich ludzi FFA jest w stanie przetwarzać inne kategorie obiektów, ani czy ta funkcja percepcji jest wrodzona czy w pewien sposób wyuczona.

## 1.2. Zjawisko pareidolii

Pareidolia, czyli iluzoryczne postrzeganie zmysłowe, występuje, kiedy jakiś bodziec zewnętrzny wywołuje postrzeganie bodźca nieistniejącego<sup>22</sup>. Najczęściej występującym i badanym wariantem pareidolii jest ta związana z twarzami. To zjawisko często towarzyszy nam w codziennym życiu – zauważamy ludzkie (a czasami zwierzęce) twarze na powierzchni drewna, kuchennych płytek, Księżycy, w maskach samochodów czy fasadach domów.

---

<sup>21</sup> Tamże.

<sup>22</sup> Zob. J. Liu, J. Li, L. Feng i in.: „Seeing Jesus in toast: neural and behavioral correlates of face pareidolia”, *Cortex* 53 (2014), 62.

Rys. 1. i 2. Przykłady pareidolii.



Źródło: Archiwum własne.

Pareidolia jest odmianą apofenii, czyli ludzkiej zdolności i tendencji do odnajdywania znaczących wzorców w otaczającej rzeczywistości<sup>23</sup>. Apofenia jest rodzajem błędu w postrzeganiu, który ma za zadanie uniknąć pominięcia potencjalnie niebezpiecznego bodźca. Jak pisze Robert G. Bednarik: „koszt dostrzeżenia fałszywego wzorca jest znacznie mniejszy, niż pominięcie prawdziwego wzorca”<sup>24</sup>. Z ewolucyjnego punktu widzenia znacznie lepiej jest dostrzec coś, czego nie ma, niż nie dostrzec potencjalnego zagrożenia, którym może być drapieżnik czy inny człowiek. Jak dowiodły badania na reżusach<sup>25</sup>, człowiek nie jest odosobniony w doświadczaniu zjawiska pareidolii. W badaniu analizowano wzory fiksacji wzroku małp na obrazach przedmiotów, które miały przypominać (małpie) twarze i porównywano je z obrazami przedmiotów nie przypominającymi twarzy. Analiza wykazała, iż małpy fiksowały wzrok najczęściej w obszarze „oczu” i „ust” iluzorycznych twarzy<sup>26</sup>. To badanie jest spójne z tym, co wiemy na temat układu nerwowego reżusów i jego podobieństwie do ludzkiego systemu rozpoznawania i przetwarzania twarzy. Badanie na reżusach wykazało też, że możemy spodziewać się różnic gatunkowych w mapach fiksacji na iluzorycznych twarzach, to znaczy, że przedstawiciele danego gatunku będą z większą łatwością dostrzegać bodźce posiadające cechy gatunkowe podobne do danego osobnika<sup>27</sup>. Systemy te wykazują się jednak dużą tolerancją w myśl przywołanej wcześniej tezy Bednaricka – fałszywe dostrzeżenie bodźca jest relatywnie niskokosztowe, w porównaniu do niedostrzeżenia prawdziwego bodźca.

<sup>23</sup> Zob. L. F. Zhou, M. Meng: „Do you see the «face»? Individual differences in face pareidolia”, *Journal of Pacific Rim Psychology* 14 (2020), 2: 1.

<sup>24</sup> R.G. Bednarik: „Rock art and pareidolia”, *Rock art research* 33 (2016), 2: 169.

<sup>25</sup> Zob. J. Taubert, G. Wardle, M. Flessert i in.: „Face pareidolia in the rhesus monkey”, *Current Biology* 27 (2017), 16: 2507.

<sup>26</sup> Tamże, 2506.

<sup>27</sup> Tamże, 2508.

Pareidolia może być wykorzystywana w sztuce, projektowaniu i wzornictwie. Jej implikacje są szczególnie ważne dla projektantów przedmiotów codziennego użytku. Ta forma antropomorfizacji może i polepszać i pogarszać odbiór produktu. Przedmioty, które będą imitować mimikę stanów emocjonalnych, mogą bezpośrednio wpływać na stan emocjonalny odbiorcy/użytkownika<sup>28</sup> i wywoływać sympatię lub współczucie, jak np. w przypadku oglądania smutnego budzika (Zob. Rys. 1). Różne cechy fizyczne przedmiotów mogą być utożsamiane z ich cechami „osobowości” – według Pankaja Aggarwala i Ann L. McGill, polega to na uruchomieniu konkretnego schematu, do którego przypisujemy dany obiekt<sup>29</sup>. Co więcej, przypisanie obiektu do danej kategorii zostaje wynagrodzone satysfakcją, dlatego wielu marketingowców decyduje się na produkty, które można łatwo przypisać do jakiejś kategorii. Ten efekt może jednak zostać zagłuszony – „Na przykład, kiedy produkt wydaje się być dobrym dopasowaniem do kategorii «kryminalista», konsumenci mogą ocenić go negatywnie, pomimo satysfakcji płynącej ze zgodności z aktywowanym schematem”<sup>30</sup>. Dlatego w procesie projektowania niezwykle ważne jest rozpoznanie i świadome użycie zjawiska pareidolii. Schemat „twarz” wzmacnia się, kiedy postrzegany bodziec, odbierany jako twarz, uśmiecha się, w przeciwieństwie do tego, kiedy wyraża smutek<sup>31</sup>.

Jeden z najpowszechniejszych przypadków pareidolii, to rozpoznawanie twarzy w maskach samochodów<sup>32</sup>, więc „wyraz twarzy” pojazdów jest często poddawana badaniu cechą. W jednym z badań przeprowadzonym przez Jana R. Landwehra i współpracowników porównywano różne kombinacje części składowych frontu samochodu. Manipulacji na obrazach poddawano przednie światła i maskownicę. Światła były pochylone do środka (wyrażając złość lub agresję) lub na zewnątrz (wyrażając spokój i serdeczność), natomiast maskownica miała przypominać uśmiech lub grymas. Badanie to pozwoliło wykazać, że nie tyle samo rozpoznanie twarzy jest kluczowe, do postrzegania bodźca w sposób antropomorficzny, ale także konkretny „wyraz twarzy” frontu samochodu. Najchętniej wybieraną przez badanych kombinacją, była ta, która wywoływała niezgodność – maskownica przypominająca uśmiech, w parze ze światłami przypominającymi agresywne oczy. Takie zestawienie wywoływało jednocześnie pobudzenie i pozytywny odbiór<sup>33</sup>. W badaniu Aggarwal i McGill prezentowano dwa zmanipulowane obrazy frontu samochodu – jeden z maskownicą przypominającą uśmiech, a drugi grymas. Dodatkowo, w badaniu torowano odbiorców historią dotyczącą samochodu: jednej grupie przedstawiono kontekst, w którym samochód traktowany był jako obiekt, natomiast narracja wokół samochodu dla drugiej grupy przedstawiała go w formie antropomorficznej w roli protagonisty historii. Ci badani, którym przedstawiono samochód w kontekście ludzkim, znacznie częściej traktowali samochód jako osobę. Także maskownica w formie uśmiechu, silniej skłaniała odbiorcę do personifikacji samochodu w przeciwieństwie do grymasu, zgodnie z założeniem, że uśmiech silniej wpisuje się w schemat twarzy<sup>34</sup>. Przyjazny „wyraz

<sup>28</sup> Zob. A. Wodehouse, R. Brisco, E. Broussard i in.: „Pareidolia: characterising facial anthropomorphism and its implications for product design”, *Journal of Design Research* 16 (2018), 2: 86.

<sup>29</sup> Zob. P. Aggarwal, A. L. McGill: „Is That Car Smiling at Me? Schema Congruity as a Basis for Evaluating Anthropomorphized Products”, *Journal of Consumer Research* 34 (2007), 4: 469.

<sup>30</sup> Tamże.

<sup>31</sup> Tamże, 473.

<sup>32</sup> Zob. A. Wodehouse, R. Brisco, E. Broussard i in.: „Pareidolia: characterising facial anthropomorphism and its implications for product design”, wyd. cyt., 87.

<sup>33</sup> Zob. J. R. Landwehr, A. L. McGill, A. Herrmann: „It's got the look: The effect of friendly and aggressive «facial» expressions on product liking and sales”, *Journal of marketing* 75 (2011), 3: 141.

<sup>34</sup> Zob. P. Aggarwal, A. L. McGill: „Is That Car Smiling at Me? Schema Congruity as a Basis for Evaluating Anthropomorphized Products”, wyd. cyt., 473.

twarzy” obiektu znacząco wpływa na jego pozytywny odbiór. Sam wygląd obiektu, jest ważny, natomiast w celu wzmocnienia satysfakcji płynącej z aktywacji schematu, ważna jest odpowiednia narracja związana z produktem czy obiektem. Takie wykorzystanie pareidolii i antropomorfizacji jest częstą praktyką marketingową – tworzenie maskotki marki (*brand hero*) może sprzyjać utożsamieniu jej cech z całą marką i ociepleniem jej wizerunku. Przykładami znanych na całym świecie maskotek są postaci M&M’s czy Michelin Man<sup>35</sup>.

Pareidolia jako błąd w postrzeganiu stanowi także ryzyko nadinterpretacji wzorców. Może przejawiać się to, np. widzeniem twarzy Jezusa w toście, tworzeniem teorii spiskowych czy poparciem dla myślenia magicznego<sup>36</sup>. Jednym z przykładów takich błędów w postrzeganiu, które miały wpływ na sztukę i naukę, jest historia błędnych interpretacji malowideł naskalnych przytaczana przez Bednarika. W Mongolii Wewnętrznej Zhou Yushu, rektor uczelni i badacz sztuki naskalnej zebrał setki ogromnych głazów, które rzekomo miały przedstawiać twarze i inne motywy. Podjął się transportu wielu ton kamieni i umieścił je w specjalnie utworzonym depozycie. Na podstawie swoich odkryć opracował teorię kultu trzech cesarzy sprzed 6000 lat. Zaprosił innych znanych badaczy w celu wpisania znaleziska na Listę światowego dziedzictwa UNESCO. Wkrótce okazało się, że głazy pozbawione są jakichkolwiek śladów ludzkich narzędzi, a wszystkie wzory były wynikiem naturalnych procesów geologicznych. Podczas demonstracji techniki tworzenia odbitek wzorów okazało się, że osoby je wykonujące nie odwzorowywały faktycznych śladów na skałach, lecz podążały za własnymi wyobrażeniami. Co więcej, jeden z zaproszonych profesorów po czasie także zaczął dostrzegać motyw twarzy, którego nie było tam w rzeczywistości<sup>37</sup>. Bednarik uważa ten przykład, za jeden z najbardziej spektakularnych przypadków zbiorowej pareidolii, gdzie błąd postrzegania doprowadził do daleko idących wniosków. Autor zaleca krytyczne podejście do znalezisk i stosowanie się do metod badawczych<sup>38</sup>. Ten przykład ilustruje, jak bardzo ludzie wyczuleni są na rozpoznawanie bodźców związanych z twarzą oraz jak silnie wierzą i ulegają swojej percepcji.

Zjawisko pareidolii dotyczy nas w codziennym życiu, wpływa na nasze decyzje marketingowe, postrzeganie przedmiotów codziennego użytku w sposób pozytywny, lub negatywny. Nasza zdolność do antropomorfizacji, może nieść daleko idące konsekwencje, takie jak w przypadku badacza naskalnych malowideł. Szczególnie projektantki i wzornicy powinni być świadomi tego zjawiska i używać go krytycznie.

---

<sup>35</sup> Tamże, 468.

<sup>36</sup> Zob. A.Chudzik: „Atrakcyjność internetowej magii: o myśleniu mityczno-magicznym we współczesnym świecie”, *Media i Społeczeństwo* 15, (2021), 2: 68.

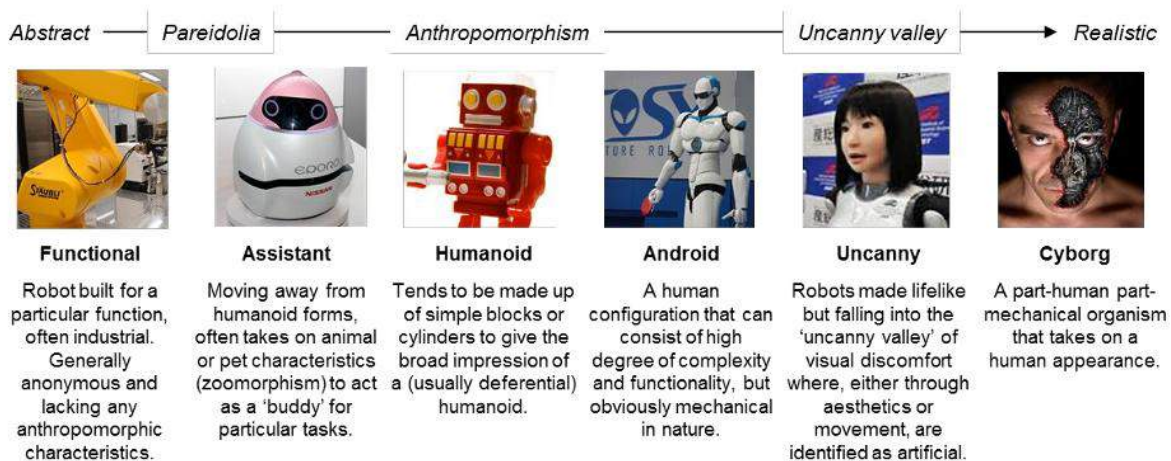
<sup>37</sup> Zob. R.G. Bednarik: „Rock art and pareidolia”, wyd. cyt., 169–171.

<sup>38</sup> Tamże, 178.

### 1.3. Dolina niesamowitości w rozpoznawaniu twarzy

Wodehouse wraz ze współpracownikami stworzyli skalę, którą nazwali Ruchomą skalą antropomorfizmu<sup>39</sup> (Zob. Rys. 3.). Jest to skala, która określa zakres w jakim obiekty, w szczególności roboty, przypominają ludzi. Początek skali stanowią formy bardzo abstrakcyjne – roboty czysto funkcjonalne: roboty przemysłowe czy odkurzacze samojezdzące, które stosunkowo rzadko bywają antropomorfizowane. Pareidolia, jako rodzaj „przypadkowej emergencji twarzy”<sup>40</sup> na skali zostały umieszczone pomiędzy czymś abstrakcyjnym a antropomorfizmem. Do kategorii antropomorfizmu, zaliczone zostały roboty asystenckie, humanoidy czy androidy, które przypominają człowieka, ale są na tyle mechaniczne, że nie mają na celu go udawać. Zjawisko doliny niesamowitości (*Uncanny Valley*), to zjawisko pojawiające się przy końcu skali. Charakteryzuje obiekty, które powodują wizualny dyskomfort, poprzez nieudolne naśladowanie prawdziwego człowieka<sup>41</sup>.

Rys. 3. Ruchoma skala antropomorfizmu



Źródło: A.Wodehouse i in.: „Pareidolia: characterising facial anthropomorphism and its implications for product design”, *Journal of Design Research*, 16(2), 2018, 88.

Ta skala nawiązuje do słynnej pracy robotyka Masahiro Moriego, który w 1970 roku stworzył pojęcie doliny niesamowitości<sup>42</sup>. Stworzony przez niego wykres, posiada oś podobieństwa i oś akceptacji, której wartości negatywne odnoszą się do uczucia niepokoju, upiorności (Mori w swojej oryginalnej pracy używa do określenia osi Y japońskiego neologizmu *shinwakan*, co w konsultacji z autorem zostało przetłumaczone na język angielski jako *affinity*<sup>43</sup>, co z kolei na język polski

<sup>39</sup> Zob. A.Wodehouse, R. Brisco, E. Broussard i in.: „Pareidolia: characterising facial anthropomorphism and its implications for product design”, wyd. cyt., 87–88.

<sup>40</sup> Tamże, 87.

<sup>41</sup> Tamże, 88.

<sup>42</sup> Zob. M. Mori, K. F. MacDorman, N. Kageki: „The uncanny valley [from the field]”, *IEEE Robotics & automation magazine* 19 (2012), 2: 98.

<sup>43</sup> Tamże, 98.

dosłownie tłumaczy się na pokrewieństwo/bliskość/uznanie<sup>44</sup>). Doliną niesamowitości określa się załamanie w wykresie, które pojawia się, kiedy coś, co miało przypominać człowieka widocznie nie spełnia oczekiwań w odbiorze, a tym samym nie jest akceptowane. Mori jako przykład podaje prostetyczną rękę, która wyglądem nie odbiega od prawdziwej – w dotyku natomiast budzi „upiorne” wrażenie. Innym przykładem przywoływanym przez Moriego są japońskie kukły *bunraku*, które choć podczas spektaklu mogą nie wydawać się straszne, wyrwane z artystycznego kontekstu wywołują efekt doliny niesamowitości<sup>45</sup>.

To zjawisko stało się przedmiotem wielu badań empirycznych i przeniknęło do kultury popularnej. Środowisko akademickie jest podzielone co do istnienia efektu, choć wiele przeprowadzonych badań je potwierdza<sup>46</sup>. Christoph Bartneck ze współpracownikami w swoim badaniu nie potwierdzili tez Moriego i zwrócili szczególną uwagę na wielowymiarowość antropomorfizmu i sympatii wobec robotów. Ich badanie charakteryzowała jednak mała i niereprezentatywna próba (32 osoby, wszystkie studiujące na uniwersytecie)<sup>47</sup>. W 2012 roku Moore zaprezentował model oparty na bayesowskim, który wyjaśnia, dlaczego pewne obiekty, zwłaszcza te na granicy kategorii, mogą wywoływać niepokój<sup>48</sup>. Autor podkreśla, że samo nazwa osi Y w artykułach Moriego, ze względu na to, że jest neologizmem, co utrudnia rozumienie i tłumaczenie, może być przyczyną nieporozumień i sprzecznych wyników badań empirycznych<sup>49</sup>. Jego model zakłada, iż efekt doliny spowodowany jest napięciem percepcyjnym (*perceptual tension*) oraz nierównościami w poziomach zgodności z kategorią<sup>50</sup>. Konflikt percepcyjny może pojawić się, na przykład, gdy humanoidalny robot ma twarz niezwykle przypominającą ludzką, ale jego oczy wydają się nienaturalne i mało realistyczne.

W jednym z nowszych badań uczestnikom poddano ocenie aż 80 różnych wersji twarzy robotycznych ze spektrum od bardzo mechanicznych do tych najbardziej realistycznych<sup>51</sup>. Oceniano podobieństwo do człowieka, sympatyczność oraz poziom zaufania, które badano na podstawie gry inwestycyjnej, opartej na teorii gier<sup>52</sup>. Efekt doliny niesamowitości był wyraźnie widoczny w obu ocenach. Twarze o umiarkowanym stopniu podobieństwa do ludzi były postrzegane jako najmniej sympatyczne<sup>53</sup> i najmniej godne zaufania<sup>54</sup>. Te badania są zgodne także z badaniami nad neurobiologicznym rozpoznawaniem twarzy. James ze współpracownikami<sup>55</sup> badając mózg za pomocą neuroobrazowania prezentowali badanym twarze ludzi i zwierząt, realistyczne i animowane.

---

<sup>44</sup> *Affinity definition* | Cambridge dictionary. [dictionary.cambridge.org/us/dictionary/english-polish/affinity](https://dictionary.cambridge.org/us/dictionary/english-polish/affinity) [dostęp: 22.03.2025].

<sup>45</sup> M. Mori, K. F. MacDorman, N. Kageki: „The uncanny valley [from the field]”, wyd. cyt., 99.

<sup>46</sup> M. B. Mathur, D. B. Reichling: „Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley”, *Cognition* 146 (2016), 22.

<sup>47</sup> Zob. C. Bartneck, T. Kanda, H. Ishiguro i in.: „My robotic doppelgänger-A critical look at the uncanny valley”, In *RO-MAN 2009-The 18th IEEE international symposium on robot and human interactive communication* (2009), 269–276.

<sup>48</sup> R.K. Moore: „A Bayesian explanation of the ‘Uncanny Valley’ effect and related psychological phenomena”, *Scientific reports* 2 (2012), 1: 1.

<sup>49</sup> Tamże, 1.

<sup>50</sup> Tamże, 3.

<sup>51</sup> M. B. Mathur, D. B. Reichling: „Navigating a social world with robot partners”, wyd. cyt., 23.

<sup>52</sup> Tamże, 29.

<sup>53</sup> Tamże, 28.

<sup>54</sup> Tamże, 29.

<sup>55</sup> Zob. T.W. James, R.F. Potter, Lee i in.: „How realistic should avatars be?”, *Journal of Media Psychology* 27 (2015), 3: 113–114.

Wyniki ich badań sugerują, że moduły rozpoznawania twarzy, są dużo bardziej wyczułone na realizm w twarzach nie-ludzkich niż ludzkich<sup>56</sup>.

Istnieje wiele pomysłów na wyjaśnienie źródła efektu doliny niesamowitości. Podobnie jak w przypadku pareidolii, na przedstawicielach gatunku makaków zaobserwowano mechanizm bardzo podobny do doliny niesamowitości<sup>57</sup>, co potwierdza ewolucyjną historię zjawiska. Na wykresie zaprezentowanym przez samego Moriego w najniższym punkcie doliny znajdują się zwłoki, a nieco wyżej zombie<sup>58</sup>, co sugeruje, że choroba, zakażenia, śmierć i strach przed nią, mogą być źródłem powstawania tego zjawiska. Dziś dolina niesamowitości stanowi wyzwanie przede wszystkim dla projektantek wizerunku robotów, ale także dla projektantów zabawek czy animatorów. Pragmatycznym dowodem na istnienie doliny niesamowitości są filmy animowane dużych wytwórni, które spotkały się z finansową porażką i słabym odbiorem publiczności w związku z błędami popełnionymi na etapie projektowania twarzy postaci. Jednym z filmów, które stały się utożsamiane w przestrzeni publicznej z terminem doliny niesamowitości jest animowany film „The Polar Express” z 2004 roku.

Rys. 4. Bohaterowie filmu „The Polar Express”



Źródło: Klatka z filmu „The Polar Express” (2004), Warner Bros. Pictures.

Tytuł jednej z recenzji tego filmu nosił tytuł „Chłodne spojrzenie, ciepłe serce” (ang. „Cold eyes, warm heart”), co podkreśla znaczenie oczu w odbiorze postaci i występowaniu zjawiska doliny niesamowitości<sup>59</sup>. Wyniki badań przeprowadzone przez Christine E. Looser i Thalię Wheatley potwierdziły, iż oczy odgrywają największą rolę w rozpoznaniu żywotności twarzy, choć jedna cecha twarzy w odosobnieniu nie jest jeszcze uznawana za żywą lub nie. Sprawdzano wizerunki ludzi i manekinów, które płynnie w siebie przechodząc, tworzyły spektrum. Podczas badania stwierdzono

<sup>56</sup> Tamże, 115.

<sup>57</sup> Zob. A. Steckenfinger, A. A. Ghazanfar: „Monkey visual behavior falls into the uncanny valley”, *Proceedings of the National Academy of Sciences* 106 (2009), 43: 18363.

<sup>58</sup> Zob. M. Mori, K. F. MacDorman, N. Kageki: The uncanny valley [from the field], wyd. cyt., 99.

<sup>59</sup> Zob. T. Geller: „Overcoming the Uncanny Valley”, *IEEE Computer Graphics and Applications* 28 (2008), 4: 15.

też, że ocena żywotności jest kategorięczna<sup>60</sup> – dlatego projekt postaci w filmie może przeważać całkowicie o jego sukcesie.

Dolina niesamowitości stanowi wyzwanie dla projektantów robotów, twórczyn animacji i realistycznych symulacji postaci. Mimo licznych badań wciąż nie ma jednoznacznych odpowiedzi na pytanie, dlaczego niektóre sztuczne twarze wywołują u ludzi niepokój. Wiadomo jednak, że im bardziej sztuczna postać przypomina człowieka, tym większe ryzyko, że odbiorcy uznają ją za dziwną lub nienaturalną, jeśli nie osiągnie ona pełnego realizmu. Istnieje wiele teorii próbujących wyjaśnić to zjawisko, lecz dolina niesamowitości wciąż wymaga dalszych badań i analiz, zwłaszcza w kontekście tworzenia postaci, tak, aby były odbierane jako przekonujące i akceptowalne wizualnie.

## 2. CIAŁO I RUCH

### 2.1. Rozpoznawanie ludzkiego ciała

Ludzkie ciało, podobnie jak ludzka twarz, jest bodźcem uprzywilejowanym w postrzeganiu. Aktywuje inne systemy niż przyglądanie się obiektom. Potwierdzają to między innymi badania nad efektem inwersji. Słynna iluzja ze zdjęcia Margaret Thatcher<sup>61</sup>, znana jako *Thatcher effect* sfabrykowana przez Petera Thompsona w 1980 roku mówi nam wiele o twarzy jako bodźcu przetwarzanym w całości. Zdjęcie twarzy prezentowane jest odwrócone, natomiast usta i oczy osoby na zdjęciu znajdują się w prawidłowym położeniu – na pierwszy rzut oka nie zauważamy błędu, czytamy twarz jako całość, dopisujemy sobie prawidłowe położenie oczu i ust – po obróceniu zdjęcia do prawidłowej pozycji widzimy twarz dziwną i zniekształconą. Efekt ten doprowadził do przyjęcia twierdzenia, iż twarz jako wyjątkowy bodziec przetwarzana jest konfiguracyjnie (*configural processing*) – czyli na podstawie odległości między poszczególnymi składowymi twarzą, uruchomione jest przetwarzanie holistyczne<sup>62</sup>. Według metaanalizy stworzonej przez Jasona W. Griffina i Florę Oswald ciała także są przetwarzane konfiguracyjnie<sup>63</sup>. Efekt inwersji w przypadku ciała jest słabszy niż w przypadku twarzy, co można tłumaczyć ograniczoną ekspozycją na bodźce związane z ciałem w okresie niemowlęctwa (w związku z ograniczonym polem widzenia). Jest on jednak znacznie silniejszy niż w przypadku obserwacji odwróconych obiektów<sup>64</sup>. Jak piszą autorzy metaanalizy badania przeprowadzone przy pomocy ERP także potwierdzają istnienie efektu inwersji ciała (BIE – *body inversion effect*): „Twarze ludzkie wywołują negatywny potencjał związany ze zdarzeniem, osiągający szczyt około 170 ms po rozpoczęciu bodźca (określany jako N170). Latencja N170 jest dłuższa, a amplitudy większe w przypadku twarzy odwróconych w porównaniu do twarzy przedstawionych prawidłowo. Te efekty nie obejmują obiektów, co sugeruje wyspecjalizowane

---

<sup>60</sup> Zob. C. E. Looser, T. Wheatley: „The Tipping Point of Animacy: How, When, and Where We Perceive Life in a Face”, *Psychology Science* 21 (2010), 12: 1857–58.

<sup>61</sup> P. Thompson: „Margaret Thatcher: A new illusion”, *Perception* 9 (1980), 4: 483–484.

<sup>62</sup> D. Maurer, R. Le Grand, C. J. Mondloch: „The many faces of configural processing”, *Trends in cognitive sciences* 6 (2002), 6: 255.

<sup>63</sup> J. W. Griffin, F. Oswald: „A multilevel Bayesian meta-analysis of the body inversion effect: Evaluating controversies over headless and sexualized bodies”, *Psychonomic Bulletin & Review* 29 (2022), 5: 1559.

<sup>64</sup> Tamże.

konfiguracyjne przetwarzanie twarzy. Wykazano podobny potencjał, który przejawia się w bodźcach przedstawiających ciała – określane jako N190<sup>65</sup>.

Podobieństwa między rozpoznawaniem twarzy i ciała sięgają także konkretnych struktur mózgu. Ciało ludzkie, jak piszą autorzy artykułu „Functional MRI analysis of body representation and body parts in the extrastriate and fusiform body area”, jest przetwarzane głównie przez dwa moduły w mózgu EBA (*extrastriate body area*) i FBA (*fusiform body area*). EBA odpowiada za przetwarzanie części ciała, a FBA uaktywnia się stopniowo – silniej przy obserwacji torsów i bezgłowych ciał, najsilniej przy obserwowaniu kompletnej sylwetki człowieka<sup>66</sup>, co przypomina przetwarzanie twarzy i aktywność modułu FFA. Obserwacja ciał neurologicznie przypomina więc obserwację twarzy. Jak piszą Andrea Orlandi i Matteo Candidi: „Wiele badań przeprowadzonych od wczesnych lat 2000 potwierdza wzmożone reakcje hemodynamiczne w EBA i FBA w reakcji na obrazy przedstawiające prawdziwe ciała lub części ciała, a także: sylwetki, linearne rysunki i «ludziki» w porównaniu z twarzami, zwierzętami i przedmiotami”<sup>67</sup>.

Podobnie jak w przypadku badań nad rozpoznawaniem twarzy, rozpoznawanie ludzkich sylwetek zostało przebadane na niemowlętach. Prezentowano im prawidłowe figurki postaci ludzkich, a także figurki ciał rozczłonkowanych lub nieprawidłowych – z nogami w miejscu rąk itd. Wykazano, że 18-miesięczne i 15-miesięczne dzieci dłużej przyglądały się nieprawidłowym figurkom, co oznacza, że są w stanie odróżniać prawidłowe i nieprawidłowe ciała. 12-miesięczne dzieci nie prezentowały podobnych zdolności – przypatrywały się wszystkim figurkom tak samo długo, co sugeruje, że ich reprezentacja prawdziwego ciała nie jest wystarczająco szczegółowa, żeby odróżnić ciała standardowe od tych niestandardowych<sup>68</sup>.

Reprezentacja ciała w pamięci jest także zależna od naszych oczekiwań i przewidywań dotyczących możliwości i ograniczeń biomechanicznych ciała oraz fizycznych możliwości rzeczywistości, np. grawitacji. W badaniu Qiu Han wraz ze współpracownikami prezentowali badanym rysunki ciał z nienaturalnym położeniem rąk nad głową. Okazało się, że badani zapamiętywali ciała z mniejszym nachyleniem – takim jakie byłoby możliwe bez uszkodzeń dla ciała oraz zgodnie z przewidywanymi prawami grawitacji. Efekt nie został zaobserwowany gdy ciała były odwrócone, co dodatkowo potwierdza efekt BIE. Obraz ciała w pamięci jest zdominowany przez nasze wcześniejsze doświadczenia i stara się dorównywać optymalnym oczekiwaniom – zgodnie z modelem bayesowskim<sup>69</sup>. Ciało wiąże się więc z szeregiem założeń na jego temat, dzięki którym oglądane obrazy mogą być modyfikowane w pamięci tak, aby sprzyjały oczekiwaniom.

Przetwarzanie ciał jako bodźców jest istotnie społecznie, dotyczy naszego codziennego życia od jego pierwszych chwil. Postrzeganie ludzkiej sylwetki może wiązać się bezpośrednio z aktywacją

---

<sup>65</sup> Tamże, 1560.

<sup>66</sup> Zob. J. C. Taylor, A. J. Wiggett, P. E. Downing: „Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas”, *Journal of neurophysiology* 98 (2007), 3: 1632.

<sup>67</sup> A. Orlandi, M. Candidi: „Towards a Neuroaesthetics of Interactions: Insights from Dance on the Aesthetics of Individual and Interacting Bodies”, *iScience* 28 (2025), 5: 2.

<sup>68</sup> Zob. M. Heron, V. Slaughter: „Infants’ responses to real humans and representations of humans”, *International Journal of Behavioral Development* 34 (2010), 1: 43.

<sup>69</sup> Zob. Q. Han, M. Gandolfo, M. V. Peelen: „Prior knowledge biases the visual memory of body postures”, *iScience* 27 (2024), 4: 6.

konkretnego schematu należącego do antropomorficznej kategorii. W badaniu Aggarwal i McGill<sup>70</sup> (podobnym do tego, w którym badano grymas i uśmiech przodu samochodu) prezentowano zestaw napojów w butelkach. Badano pozytywny wpływ aktywowania schematu rodziny lub bliźniąt na ocenę produktu. Butelki w kategorii rodzina dopasowywano tak, aby różniły się wielkością: dwie mniejsze i dwie większe butelki, i porównywano z zestawem, który składał się z czterech takich samych butelek. Szklana butelka jest bodźcem, który może być łatwo odczytywany jako przypominający sylwetkę człowieka<sup>71</sup>. Uruchamiano schemat kategorii poprzez przedstawianie zestawów butelek jako „rodzina produktów” lub „linia produktów”. Badani, którym aktywowano schemat poprzez nazwanie butelek „rodziną produktów” łatwo rozpoznawali w butelkach „rodziny” – komentowali np. „wyglądają jak rodzina, ojciec i matka z brzegu, dziecko i młodsze dziecko po środku” lub „przypominają mi rodzinę”. Była to znaczna różnica w porównaniu z grupą, której przedstawiono „linię produktów”<sup>72</sup>. To badanie potwierdza silny wpływ narracji na rozpoznawanie schematów, ale także na rozpoznawanie ludzkiej sylwetki w przedmiotach codziennego użytku. Może to oznaczać, że rozpoznawanie ludzkiej sylwetki jest zależne od kontekstu i spodziewanego przebiegu wydarzeń.

Postrzeżenie i przetwarzanie ludzkiej sylwetki przypomina postrzeżenie twarzy. Tak jak widząc twarz, analizujemy odległość i prawidłowość ułożenia ust, oczu i nosa, podobnie przyglądając się ciału oczekujemy rąk, nóg, głowy w odpowiednich miejscach. Postrzeżenie ciała nie poddaje się tak silnej schematyzacji co twarz, co uzasadnia się krótszą ekspozycją na bodźce ciała w kluczowych momentach nauki postrzegania wizualnego – niemowlęctwie. Ciało ludzkie postrzegane jest także za pomocą silnych kategorii i oczekiwań wobec jego zdolności fizycznych, a także kategorii antropomorficznych kojarzących się bezpośrednio z ludzką sylwetką, np. rodzina czy bliźniaki.

## 2.2. Rola ruchu w rozpoznawaniu ludzkiej sylwetki

We wszelkich próbach odwzorowania ludzkiego ciała, poza wyglądem i stopniem podobieństwa, pojawia się kwestia ruchu. Na rozpoznawanie ludzkiej sylwetki i stopień jej realizmu ogromny wpływ ma ruch, o czym także pisze Mori<sup>73</sup>. O ile sam wygląd zdaje się nietrudny do odwzorowania, ruch ludzki jest znacznie bardziej skomplikowanym zadaniem – także ze względu na trudność w odwzorowaniu organicznej materii: mięśni i ścięgien, które są elastyczne i tworzą kompleksowość ludzkiego aparatu motorycznego<sup>74</sup>.

W 1975 Gunnar Johansson, psycholog poznawczy, wykonał eksperyment, który udowodnił, że użycie zaledwie 12 świetlnych punktów, przyczepionych do poruszającego się w ciemności człowieka, pozwala innym ludziom na rozpoznanie wykonywanej przez aktora czynności. „Początkowo badani wydają się skonfundowani, ponieważ jedyne co widzą to losowa konstelacja światła, ale jak tylko aktor wstaje i zaczyna się poruszać, obserwatorzy zaczynają postrzegać światła

<sup>70</sup> Zob. P. Aggarwal, A. L. McGill: „Is That Car Smiling at Me? Schema Congruity as a Basis for Evaluating Anthropomorphized Products”, wyd. cyt., 473.

<sup>71</sup> Tamże, 470.

<sup>72</sup> Tamże, 475.

<sup>73</sup> Zob. M. Mori, K. F. MacDorman, N. Kageki: „The uncanny valley [from the field]”, wyd. cyt., 99.

<sup>74</sup> Zob. A. Sciutti, C. Ansuini, C. Becchio i in.: „Investigating the ability to read others' intentions using humanoid robots”, *Frontiers in psychology* 6 (2015), 1362: 4.

jako przymocowane do wcześniej niewidocznej postaci ludzkiej. Nie tylko bez wahania rozróżniają bieg od chodzenia, ale zauważają niewielkie anomalie takie jak utykanie. W innym eksperymencie nagraliśmy tańczącą żywy taniec folkowy parę, a obserwatorzy od razu zobaczyli, że 24 mieniące się punkty reprezentują tańczącą parę”<sup>75</sup>. Ten eksperyment dowodzi, że ruch człowieka jest charakterystyczny i ludzie z łatwością rozpoznają sylwetkę, a także konkretną czynność, na podstawie jedynie kilku punktów świetlnych przymocowanych do aktora. Badani nie rozpoznawali w punktach świetlnych postaci, dopóki nie zaczęła się ona poruszać.

Podobny eksperyment ponad 40 lat później przeprowadza Agnieszka Wykowska wraz z współpracownikami. Tym razem nie starają się jedynie udowodnić, że ludzie są szczególnie wyczuleni na ludzką sylwetkę, ale sprawdzić jak małe ruchy wzorowane na ruchach człowieka wpływają na postrzeganie robotów humanoidalnych. Robot umieszczony był w sytuacji wskazywania na ekran. W jednym z wariantów miał on reagować na pojawiający się na ekranie bodziec z stałym, zaprogramowanym opóźnieniem, a w drugim opóźnienie generowane było przez eksperymentatora-człowieka siedzącego w innym pomieszczeniu. Wyniki jednoznacznie wskazały na ponadprzeciętną wrażliwość badanych na zachowanie ludzkie, choć nie byli oni w stanie wyjaśnić, gdzie znajduje się źródło ich przekonania<sup>76</sup>. Wprowadzenie niewielkich, subtelnych różnic, które zbliżyłyby zachowanie humanoidalnych robotów do ludzkiego zachowania, sprawiłoby, że roboty te będą postrzegane bardziej jako, jak nazywa je Wykowska, „naturalne” podmioty<sup>77</sup>. To badanie dowodzi temu, jak bardzo ludzie wyczuleni są na ruch specyficzny dla człowieka. W związku z tym, robotyka poza projektowaniem wyglądu robotów, bierze pod uwagę także naturalność ruchu, co sprowadza się do analizy skomplikowanych modeli ruchu ludzkiego. Ocena ruchu przyczynia się także do wykrycia intencji, dlatego jest ważna ze względu na przewidywanie rozwoju sytuacji, w której się znajdujemy.

W jednym z badań „An object for an action, the same object for other actions: effects on hand shaping” dotyczącym intencji czynności, mierzono kąty chwytu i nachylenia rąk w zależności od czynności, którą badani mieli wykonać z użyciem butelki. Części z badanych zalecono po prostu chwycić butelkę, innym chwycić i rzucić, nalać płynu, położyć w wyznaczonym miejscu lub przekazać innej osobie. Okazało się, trajektoria rąk znacznie różni się w zależności od intencji – powodu dla którego chwytamy butelkę<sup>78</sup>. Bazując na tym badaniu, przeprowadzono inne, które miało na celu sprawdzenie, czy ruchy inicjacyjne są odróżnialne i czy intencja z jaką podejmowane jest działanie może zostać odczytana jedynie na ich podstawie. Zarejestrowano cztery kategorie ruchów sięgania, które określono kooperatywnymi lub kopetytywnymi, z naturalną lub przyspieszoną prędkością. Badani obserwowali nagrania wideo, które nie zawierały fragmentu z inicjowanym działaniem, a jedynie ruch ręki wiodący do niego. Na tej podstawie badani byli w stanie ocenić, jaka intencja kryła się za ruchem ręki z wysoką dokładnością. W drugiej części badania, zamiast ramienia prezentowana była twarz aktora, co pozwoliło na rozróżnienie przypadków, w których twarz jest

---

<sup>75</sup> G. Johansson: „Visual motion perception”, *Scientific American* 232 (1975), 6: 87.

<sup>76</sup> Zob. A. Wykowska, T. Chaminade, G. Cheng: „Embodied artificial agents for understanding human social cognition”, *Philosophical Transactions of the Royal Society B: Biological Sciences* 371 (2016), 1693: 6.

<sup>77</sup> Tamże.

<sup>78</sup> Zob. C. Ansuini, L. Giosa, L. Turella i in.: „An object for an action, the same object for other actions: effects on hand shaping”, *Experimental Brain Research* 185 (2008), 118.

bodźcem istotniejszym, a w których mniej istotnym w ocenie intencji ruchu. Wyniki badania sugerują, że i ruch i twarz odgrywają istotną rolę w określeniu intencji<sup>79</sup>.

Ruch staje się zapisem intencji, może działać jako ostrzeżenie przed niebezpieczeństwem czy zaproszenie do kooperacji, podobnie jak interpretacja wyrazu twarzy. Pisząc na temat ruchu i intencji, trudno nie wspomnieć o słynnym eksperymencie Fritza Heidera i Marianne Simmel, który przedstawiał figury geometryczne poruszające się w różnych konfiguracjach. Figury nie miały żadnych fizycznych cech antropomorficznych – kwadrat był kwadratem, trójkąt trójkątem, a badani, którzy zostali poproszeni o opisanie sytuacji przedstawionej w animacji, łatwo przypisywali im cechy takie jak: agresywny, słaby, odważny, rozsądny, a nawet z łatwością nazywali czynności zachodzące między figurami: gonią się, pocałowały się, chowa się przed<sup>80</sup>. Ruch figur na animacji, bezpośrednio wpływał na ocenę ich zachowania i nadawał im ludzkich cech. Pozwalał także na ocenę ich intencji.

Ruch może służyć nam także do identyfikacji. W jednym z badań zatytułowanym „Recognizing people from their movement”<sup>81</sup>, którego celem było sprawdzenie co jest ważniejsze podczas postrzegania ruchu: własne doświadczenie motoryczne, czy nabyta zdolność do wizualnej analizy ruchu innych ludzi, badanym przedstawiano (podobnie jak w eksperymencie Johanssona) postaci złożone z kilkunastu świetlnych punktów. Nagrania przedstawiały samych badanych, ich przyjaciół i nieznajomych w ruchu, a zadaniem była próba identyfikacji. Badani najlepiej radzili sobie z rozpoznawaniem własnej sylwetki, gorzej z rozpoznawaniem przyjaciół, ale wciąż lepiej niż nieznajomych, co do których trafienia były wynikiem przypadku. Co więcej, lepsze wyniki w zgadywaniu osiągnięto, kiedy bohaterowie wykonywali czynności bardziej charakterystyczne dla konkretnej osoby, np. taniec, niż w przypadku bardziej uniwersalnych takich jak bieg czy chodzenie. Wyniki porównano z rozpoznawaniem statycznych punktów świetlnych przedstawiających te same postaci – tutaj wszystkie prawidłowe odpowiedzi były w granicach przypadku<sup>82</sup>, co oznacza, że ruch stanowi istotny czynnik przy identyfikacji osób znajomych i nieznajomych.

Rozpoznawanie ruchu służy nam do oceny intencji, identyfikacji i oceny poziomu zagrożenia. Szybki, niespodziewany ruch, którego doświadczamy na co dzień, np. odgłos w bliskiej odległości podczas spaceru, wywołuje w nas strach i wzbudza reakcję obronną. Rozpoznawanie ruchu typowego dla ludzi używa innych systemów niż rozpoznawanie ruchu obiektów – „jedynie ludzki ruch jest powiązany z następstwem aktywności w prawym płacie skroniowym, co sugeruje, że układ wzrokowy rozróżnia ludzki ruch od ruchu obiektów”<sup>83</sup>. Ludzkie ciało jest bodźcem szczególnym, także ze względu na działanie neuronów lustrzanych, które aktywują się podczas patrzenia na czynność wykonywaną przez innych. W badaniu przeprowadzonym przez Arietę Chouhourelou uczestnikom pokazywano nagrania w których człowiek poruszał się w pustym pomieszczeniu, obok innego człowieka lub obok pudełka. Nagranie zostało sfabrykowane w taki sposób, że na każdym samo nagranie poruszającej się osoby było takie samo. Losowo przydzielono badanym nagranie i

---

<sup>79</sup> Zob. L. Sartori, C. Becchio, U. Castiello: „Cues to intention: the role of movement information”, *Cognition* 119 (2011), 2: 250.

<sup>80</sup> Zob. F. Heider, M. Simmel: „An experimental study of apparent behavior”, *American Journal of Psychology* 57 (1944), 243–259.

<sup>81</sup> Zob. F. Loula, Prasad, K. Harber i in. : „Recognizing people from their movement”, *Journal of Experimental Psychology: Human Perception and Performance* 31 (2005), 1: 210.

<sup>82</sup> Tamże, 217.

<sup>83</sup> M. Shiffrar, M. D. Kaiser, A. Chouhourelou: „Seeing human movement as inherently social”, *The science of social vision* (2011), 266.

polecono ocenić intensywność ruchu na wideo. Okazało się, że badani, którym pokazano nagranie, na którym znajdowała się inna osoba, ocenili intensywność ruchu znacznie wyżej, niż ci, którym prezentowano to samo nagranie, ale z pudełkiem lub w pustym pomieszczeniu<sup>84</sup>. Oznacza to, że kontekst sytuacji społecznej może bezpośrednio wpływać na odbiór ruchu innych ludzi.

Realistyczne odwzorowanie ruchu stanowi duże wyzwanie i jest znaczące w odbiorze podmiotów naśladowujących ludzi, co jest ważne, na przykład w projektowaniu robotów społecznych. Postrzeganie społecznych robotów jako towarzyszy, a więc podmiotów akceptowalnych i nie wzbudzających niepokoju, jest ważne podczas wykonywania przez nie ich funkcji. Jak piszą autorzy artykułu „Robots As Intentional Agents: Using Neuroscientific Methods to Make Robots Appear More Social” szczególnie w dziedzinach nauczania, rozrywki czy opieki medycznej, roboty przypominające ludzi mogą mieć pozytywny wpływ na samopoczucie, wyrażanie emocji przez osoby starsze, czy wykazywać efekty terapeutyczne w kontakcie z dziećmi ze spektrum autyzmu. Ludzki wygląd robotów społecznych oddziałuje pozytywnie na ich odbiór i sprawia, że lepiej wykonują swoje zadania, ale prawdopodobnie to właśnie ich zachowanie jest cechą nawet bardziej krytyczną w postrzeganiu ich przez ludzi niż humanoidalny wygląd<sup>85</sup>. Wspomniany wcześniej eksperyment Wykowskiej utwierdza w przekonaniu, że ludzkie mikroruchy także mają wpływ bezpośredni wpływ na odbiór zachowania robota w sytuacji społecznej.

Ruch pełni ogromną rolę w rozpoznawaniu sylwetki człowieka, a także identyfikacji i odbioru intencji. Ruch i zachowanie robotów humanoidalnych bezpośrednio wpływa na ich odbiór jako podmiotów społecznych, co za tym idzie ich projektowanie powinno być szczególnie wrażliwe na realistyczne modelowanie ich ruchu. Ocena ruchu wiąże się bezpośrednio z oceną sytuacji społecznej i wpływa na późniejsze zachowanie, np. ucieczkę w chwili zagrożenia lub wręcz przeciwnie, wzrost zaufania do zbliżającej się postaci.

### 2.3. Ciało inne i horror

Dolina niesamowitości nie dotyczy jedynie twarzy – nie ma też konkretnych reguł, które mogłyby zapewnić jej wystąpienie czy też nie. Nie ulega jednak wątpliwości, że istnieje dziwność (*uncanniness*) towarzysząca patrzeniu na ciała nie do końca ludzkie, zdeformowane, chore, martwe – inne niż te do, których przywykliśmy. Poza tym, że wywołuje uczucie niepokoju, sprawia, że nie chcemy (nie możemy) oderwać od nich wzroku. Píše o tym Denis Dutton rozważając, dlaczego ludzie lubią przyglądać się tragedii. Przywołuje fragment „Państwa” Platona, gdzie Leoncjusz widząc martwe zwłoki pod murem walczył z pragnieniem spojrzenia na nie i przegrał. Zwrócił się wtedy do własnych oczu słowami „Patrzcie na własny użytek, wy niegodziwi nikczemnicy, cieszcicie się pięknym widokiem!”, krytykując naturę własnej duszy, która chce za wszelką cenę oglądać katastrofę. Jak pisze Dutton: „niektórzy ludzie będący świadkami wypadku drogowego wiedzą, o czym mówi Platon”<sup>86</sup>. Platon w tym fragmencie pragnie skrytykować teatralną tragedię, ale przypadkiem mówi nam też o sile z jaką przyciągają nas ciała (nie)ludzkie – martwe, okaleczone,

<sup>84</sup> Tamże, 269–270.

<sup>85</sup> E. Wiese, G. Metta, A. Wykowska: „Robots as intentional agents: using neuroscientific methods to make robots appear more social”, *Frontiers in psychology* 8 (2017), 2–4.

<sup>86</sup> D. Dutton, „Instynkt sztuki. Piękno, zachwyty i ewolucja człowieka”, tłum. J. Luty, (Kraków: Copernicus Center Press 2019), 57.

dziwne. Podkreśla to makabra w słynnym obrazie „Saturn pożerający jedno ze swoich dzieci” Francisco Goyi czy znaczna popularność horrorów, w tym w szczególności gatunku *body horror*.

Potworność często opiera się na figurze ludzkiej – potwór Frankenstein z powieści Mary Shelley to człowiek pozlepiany z kawałków, ożywiony za pomocą impulsów elektrycznych, Slenderman to ludzka postać bez twarzy z nienaturalnie wydłużonymi kończynami funkcjonująca w legendach miejskich, Terminator to robot stworzony na wzór człowieka, który zamierza przejąć kontrolę nad światem, a Pale Man z filmu *Labirynt Fauna*, wygląda jak człowiek, ale porusza się w niepokojący sposób, a jego oczy znajdują się po wewnętrznej stronie dłoni. Ta skażona ludzkość budzi w widzach niepokój.

Fenomen horrorów i (nie)ludzkiej potworności może mieć swoje korzenie w obszarach mózgu odpowiedzialnych za przetwarzanie ludzkiej sylwetki. W badaniu przeprowadzonym przez Emily Cross i współpracowników przedstawiono badaniem zdjęcia postaci ludzkiej w standardowych pozach oraz takich gdzie ciało było nienaturalnie wygięte. Okazało się, że EBA i FBA silniej reagowały na nietypowe postawy ciała niż na postawy zwyczajne. W odróżnieniu od EBA i FBA, inne obszary mózgu – takie jak dolny zakręt czołowy, dolny płacik ciemieniowy, tylna bruzda skroniowa górna oraz fragment zakrętu wrzecionowatego – wykazywały zmniejszoną aktywność podczas ponownego oglądania tej samej postawy, niezależnie od kąta, z którego była przedstawiona. Oznacza to, że obszary EBA i FBA nie tyle reagują na samą formę postawy, co raczej przetwarzają jej strukturę w sposób niezależny od perspektywy, tworząc bardziej abstrakcyjną reprezentację ciała i nie ulegając habituacji<sup>87</sup>. Co więcej, obszary związane z systemem neuronów lustrzanych, nie różnicowały postaw zwykłych i nienaturalnych, choć wcześniejsze badania sugerowały, że są one bardziej aktywne przy obserwacji znajomych ruchów. Te wyniki mogą sugerować, że obszary EBA i FBA mogą kodować postawy ciała niezależnie od doświadczenia danej osoby. Obserwując nienaturalne wygięcia ciała w horrorach, możemy być skłonni do utożsamiania się z bohaterami poprzez aktywację neuronów lustrzanych, angażując nasze ciała do silniejszego przeżywania.

W swojej rozprawie doktorskiej Sarah Downes analizuje gatunek horroru pod względem przeżyć neuroestetycznych. Jak pisze: „horror jest bezsprzecznie najbardziej ucieleśnionym gatunkiem kina z perspektywy zawartości filmowej, a także cielesnej reakcji odbiorcy”<sup>88</sup>. Szczególne miejsce w tym procesie zajmują postaci, które przypominają człowieka, ale są od niego różne „albo w sposób psychologiczny albo fizyczny”<sup>89</sup>. Świetnie sprawdza się w tym miejscu wykorzystanie zjawiska doliny niesamowitości, którego „najbardziej prawdopodobną przyczyną jest naturalna awersja do ciał lub ludzi, których wygląd poddany został zmianom w wyniku zakażenia. W horrorach czy legendach miejskich o potworach czy żywych trupach mają one zdolność do zakażenia. Tworzenie więc postaci wpadających w kategorię doliny niesamowitości jest silnym narzędziem do tworzenia postaci, które są przerażające i zapadają w pamięć, pod warunkiem, że zostanie to wykonane precyzyjnie, ponieważ gdy postać będzie za bardzo przypominać człowieka, odbiorcy mogą zacząć jej współczuć, zamiast się jej bać”<sup>90</sup>.

<sup>87</sup> Zob. E. Cross, E. C. Mackie, G. Wolford i in.: „Contorted and ordinary body postures in the human brain”, *Exp Brain Res* 204 (2010), 406.

<sup>88</sup> Downes: „Bodily sensation in contemporary extreme horror film”, (Loughborough University: Thesis 2014), 4.

<sup>89</sup> Tamże.

<sup>90</sup> L. Nummenmaa: „Psychology and neurobiology of horror movies”, *Projections* 18 (2024), 2: 46.

Wszelkie ingerencje w integralność czy spójność ludzkiego ciała, wywołują w odbiorcach niepokój, ale także poznawczą ciekawość. Przyczynia się to do wykorzystania efektów takich jak dolina niesamowitości czy nienaturalny ruch w filmach grozy. Dziwni, zakażeni, nieproporcjonalni ludzie są projektowani tak, aby wywołać emocjonalną reakcję u odbiorców.

### **Zakończenie**

Człowiek, w szczególności jego ciało, sylwetka i ruch są istotnymi i uprzywilejowanymi bodźcami w postrzeganiu oraz przedstawienie zjawisk poznawczych dotyczących analizy i rozpoznawania ludzkiej twarzy i sylwetki w perspektywie interdyscyplinarnej.

W części „Twarz” przybliżyłam mechanizm rozpoznawania twarzy, a także konkretne zjawiska jego dotyczące, ze szczególnym uwzględnieniem pareidolii i doliny niesamowitości. Zwróciłam szczególną uwagę na wykorzystanie wyżej wymienionych zjawisk w projektowaniu produktów i marketingu, a także na błędy poznawcze i produkcyjne. W części „Ciało i ruch” skupiłam się na sylwetce człowieka, a także ruchu, jako sposobie identyfikacji, ocenie intencji i niebezpieczeństwa. Podkreślałam w nim także, jak modyfikacje ciała wpływają na jego odbiór – powodując niepokój, ale też poznawczą ciekawość, co wymaga dalszych analiz szczególnie w kontekście tworzenia i odbioru sztuki.

## Bibliografia

- „Affinity”, Cambridge Dictionary, dostęp 22 marca 2025, [dictionary.cambridge.org/us/dictionary/english-polish/affinity](https://dictionary.cambridge.org/us/dictionary/english-polish/affinity).
- Aggarwal, P., i A. L. McGill. „Is That Car Smiling at Me? Schema Congruity as a Basis for Evaluating Anthropomorphized Products”. *Journal of Consumer Research* 34 (2007): 468–79.
- Ansuini, C., L. Giosa, L. Turella, G. Altoè, i U. Castiello. „An Object for an Action, the Same Object for Other Actions: Effects on Hand Shaping”. *Experimental Brain Research* 185 (2008): 111–19.
- Bednarik, R. G. „Rock art and pareidolia”. *Rock Art Research: The Journal of the Australian Rock Art Research Association (AURA)* 33 (2016): 167–81.
- Bilalić, M. „Revisiting the Role of the Fusiform Face Area in Expertise”. *Journal Cognitive Neuroscience* 28 (2016): 1345–57.
- Chudzik, A. „Atrakcyjność internetowej magii. O myśleniu mityczno-magicznym we współczesnym świecie”. *Media i Społeczeństwo* 15 (2021): 58–77.
- Cross, E. S., E. C. Mackie, G. Wolford, i A. Hamilton. „Contorted and Ordinary Body Postures in the Human Brain”. *Exp Brain Res* 204 (2010): 397–407.
- Downes, S. „Bodily Sensation in Contemporary Extreme Horror Film”. Loughborough University, 2014.
- Geller, T. „Overcoming the Uncanny Valley”. *IEEE Computer Graphics and Applications* 28, nr 4 (2008): 11–17.
- Goren, C. C., M. Sarty, i P. Y. Wu. „Visual Following and Pattern Discrimination of Face-like Stimuli by Newborn Infants”. *Pediatrics* 56, nr 4 (1975): 544–49.
- Griffin, J. W., i F. Oswald. „A Multilevel Bayesian Meta-Analysis of the Body Inversion Effect: Evaluating Controversies over Headless and Sexualized Bodies”. *Psychonomic Bulletin & Review* 29, nr 5 (2022): 1558–93.
- Han, Q., M. Gandolfo, i M. V. Peelen. „Prior Knowledge Biases the Visual Memory of Body Postures”. *Iscience* 27, nr 4 (2024).
- Heider, F., i M. Simmel. „An Experimental Study of Apparent Behavior”. *American Journal of Psychology* 57 (1944): 243–59.
- Heron, M., i V. Slaughter. „Infants’ Responses to Real Humans and Representations of Humans”. *International Journal of Behavioral Development* 34, nr 1 (2010): 34–45.
- James, T. W., R. F. Potter, S. Lee, S. Kim, R. A. Stevenson, i A. Lang. „How Realistic Should Avatars Be?”. *Journal of Media Psychology* 27, nr 3 (2015): 109–17.
- Jaracz, J. „Neurobiologia percepcji ekspresji emocji twarzy”. *Neuropsychiatria i Neuropsychologia* 5 (2010): 3–4.
- Johansson, G. „Visual motion perception”. *Scientific American* 232, nr 6 (1975): 76–89.
- Johnson, M. H., S. Dziurawiec, H. Ellis, i J. Morton. „Newborns’ Preferential Tracking of Face like Stimuli and Its Subsequent Decline”. *Cognition* 40, nr 1–2 (1991): 1–19.
- Johnson, M. H., i J. Morton. *Biology and Cognitive Development: The Case of Face Recognition*. Blackwell Scientific Publications, 1993.
- Johnson, M. H., i J. Morton. „Subcortical Face Processing”. *Nature Reviews Neuroscience* 6, nr 10 (2005): 766–74.
- Kanwisher, N., J. McDermott, i M. M. Chun. „The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception”. *Journal Neuroscience* 17, nr 11 (1997): 4302–11.
- Landwehr, J. R., A. L. McGill, i A. Herrmann. „It’s Got the Look: The Effect of Friendly and Aggressive «facial» Expressions on Product Liking and Sales”. *Journal of Marketing* 75, nr 3 (2011): 132–46.
- Liu, J., J. Li, L. Feng, L. Li, J. Tian, i K. Lee. „Seeing Jesus in Toast: Neural and Behavioral Correlates of Face Pareidolia”. *Cortex* 53 (2013): 60–77.
- Looser, C. E., i T. Wheatley. „The Tipping Point of Animacy: How, When, and Where We Perceive Life in a Face”. *Psychological Science* 21 (2010): 1854–62.
- Loula, F., S. Prasad, K. Harber, i M. Shiffrar. „Recognizing People from Their Movement”. *Journal of Experimental Psychology: Human Perception and Performance* 31, nr 1 (2005): 210–20.

- Mathur, M. B., i D. B. Reichling. „Navigating a Social World with Robot Partners: A Quantitative Cartography of the Uncanny Valley”. *Cognition* 146 (2016): 22–32.
- Maurer, D., R. Le Grand, i C. J. Mondloch. „The Many Faces of Configural Processing”. *Trends in Cognitive Sciences* 6, nr 6 (2002): 255–60.
- Meltzoff, A. N., i M. K. Moore. „Early Imitation within a Functional Framework: The Importance of Person Identity, Movement, and Development”. *Infant Behavior and Development* 15, nr 4 (1992): 479–505.
- Moore, R. K. „A Bayesian Explanation of the ‘Uncanny Valley’ Effect and Related Psychological Phenomena”. *Scientific Reports* 2, nr 1 (2012): 864.
- Mori, M., K. F. MacDorman, i N. Kageki. „The Uncanny Valley [from the Field]”. *IEEE Robotics & Automation Magazine* 19, nr 2 (2012): 98–100.
- Nummenmaa, L. „Psychology and Neurobiology of Horror Movies”. *Projections* 18, nr 2 (2024): 37–64.
- Orlandi, A., i M. Candidi. „Towards a Neuroaesthetics of Interactions: Insights from Dance on the Aesthetics of Individual and Interacting Bodies”. *iScience* 28, nr 5 (2025): 1–31.
- Ratan Murty, N. A., S. Teng, D. Beeler, A. Mynick, A. Oliva, i N. Kanwisher. „Visual Experience Is Not Necessary for the Development of Face-Selectivity in the Lateral Fusiform Gyrus”. *Proceedings of the National Academy of Sciences* 117, nr 37 (2020): 23011–20. <https://doi.org/10.1073/pnas.2004607117>.
- Royle, N. *The uncanny*. University Press, 2003.
- Sartori, L., C. Becchio, i U. Castiello. „Cues to Intention: The Role of Movement Information”. *Cognition* 119, nr 2 (2011): 242–52.
- Sciutti, A., C. Ansuini, C. Becchio, i G. Sandini. „Investigating the Ability to Read Others’ Intentions Using Humanoid Robots”. *Frontiers in Psychology* 6 (wrzesień 2015). <https://doi.org/10.3389/fpsyg.2015.01362>.
- Shiffrar, M., M. D. Kaiser, i Chouhourelou. „Seeing Human Movement as Inherently Social”. *The Science of Social Vision*, 2011, 262–77.
- Steckenfinger, A., i A. A. Ghazanfar. „Monkey Visual Behavior Falls into the Uncanny Valley”. *Proceedings of the National Academy of Sciences* 106, nr 43 (2009): 18362–66.
- Steeves, J. K. E., J. C. Culham, B. C. Duchaine, et al. „The Fusiform Face Area Is Not Sufficient for Face Recognition: Evidence from a Patient with Dense Prosopagnosia and No Occipital Face Area”. *Neuropsychologia* 44, nr 4 (2006): 594–609. <https://doi.org/10.1016/j.neuropsychologia.2005.06.013>.
- Taubert, J., S. G. Wardle, M. Flessert, D. A. Leopold, i L. G. Ungerleider. „Face Pareidolia in the Rhesus Monkey”. *Current Biology* 27, nr 16 (2017): 2505–9.
- Taylor, J. C., A. J. Wiggett, i P. E. Downing. „Functional MRI Analysis of Body and Body Part Representations in the Extrastriate and Fusiform Body Areas”. *Journal of Neurophysiology* 98, nr 3 (2007): 1626–33.
- Thompson, P. „Margaret Thatcher: A New Illusion”. *Perception* 9, nr 4 (1980): 483–84.
- Wiese, E., G. Metta, i A. Wykowska. „Robots as Intentional Agents: Using Neuroscientific Methods to Make Robots Appear More Social”. *Frontiers in Psychology* 8, nr 1663 (2017).
- Wodehouse, A., R. Brisco, E. Broussard, i A. Duffy. „Pareidolia: Characterising Facial Anthropomorphism and Its Implications for Product Design”. *Journal of Design Research* 16, nr 2 (2018): 83–98.
- Wykowska, A., T. Chaminade, i G. Cheng. „Embodied Artificial Agents for Understanding Human Social Cognition”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371, nr 1693 (2016): 1–9.
- Zhou, L. F., i M. Meng. „Do You See the «face»? Individual Differences in Face Pareidolia”. *Journal of Pacific Rim Psychology* 14, nr 2 (2020): 1–8.

## **Wykrywanie i klasyfikacja choroby Alzheimera za pomocą metod uczenia maszynowego na podstawie obrazów MRI**

### **Abstrakt**

Choroby otępienne, takie jak choroba Alzheimera, stanowią coraz większe wyzwanie dla systemów ochrony zdrowia. Wczesne i dokładne rozpoznanie jest kluczowe dla efektywnej terapii, jednak standardowe metody diagnostyczne często cechują się ograniczoną skutecznością. Wykorzystanie algorytmów sztucznej inteligencji w procesie diagnozowania owych chorób może okazać się nieocenionym wsparciem, a nawet zapewnić wyższą precyzję w porównaniu do istniejących narzędzi. W tej pracy przedstawiono użycie metod uczenia maszynowego, zwłaszcza konwolucyjnych sieci neuronowych (CNN), do analizy obrazów mózgu z rezonansu magnetycznego (MRI) w celu wykrywania i klasyfikacji choroby Alzheimera.

**Słowa kluczowe:** choroba Alzheimera, uczenie maszynowe, konwolucyjne sieci neuronowe, obrazowanie metodą rezonansu magnetycznego

### **Abstract**

Dementia diseases, such as Alzheimer's disease, are a growing challenge for health care systems. Early and accurate diagnosis is crucial for effective therapy, but standard diagnostic methods often have limited effectiveness. The use of artificial intelligence algorithms in the process of diagnosing these diseases can prove to be an invaluable support and even provide higher precision compared to existing tools. This work presents the use of machine learning methods, particularly convolutional neural networks (CNNs), to analyze magnetic resonance imaging (MRI) brain scans for the detection and classification of Alzheimer's disease.

**Keywords:** Alzheimer's disease, machine learning, convolutional neural networks, magnetic resonance imaging

## Wstęp i cel pracy

Choroba Alzheimera jest najczęstszą przyczyną otępienia na świecie i stanowi poważne wyzwanie zdrowotne, społeczne oraz ekonomiczne. Szacuje się, że odpowiada za około 60% wszystkich przypadków demencji<sup>1</sup>, a liczba chorych rośnie wraz ze starzeniem się społeczeństw. Wczesne i trafne rozpoznanie tej jednostki chorobowej ma kluczowe znaczenie dla wdrożenia działań terapeutycznych, spowolnienia postępu choroby oraz poprawy jakości życia pacjentów i ich rodzin. Tradycyjne metody diagnostyczne opierają się na ocenie klinicznej, testach neuropsychologicznych oraz obrazowaniu medycznym, w szczególności metodą rezonansu magnetycznego (MRI). Obrazy MRI pozwalają uwidocznic zmiany strukturalne mózgu charakterystyczne dla choroby Alzheimera, takie jak zanik hipokampa czy uogólniona atrofia kory mózgowej. Analiza tych obrazów jest jednak czasochłonna, wymaga dużego doświadczenia klinicznego i obarczona jest subiektywnością.

W ostatnich latach rośnie zainteresowanie wykorzystaniem metod sztucznej inteligencji<sup>2</sup>, a zwłaszcza głębokich sieci neuronowych (deep learning), do wspomagania procesów diagnostycznych. Konwolucyjne sieci neuronowe (CNN) znalazły szczególne zastosowanie w analizie danych obrazowych<sup>3</sup>, dzięki zdolności automatycznego wykrywania złożonych wzorców i hierarchicznych reprezentacji cech. W kontekście diagnostyki choroby Alzheimera CNN umożliwiają rozpoznawanie subtelnych różnic w obrazach MRI, które mogą umykać klasycznej ocenie specjalisty. Mimo dużego potencjału tej technologii, wciąż istnieją trudności związane z jej praktycznym zastosowaniem: ograniczona dostępność dużych i zbalansowanych zbiorów danych, zróżnicowanie obrazów MRI pod względem jakości i parametrów technicznych, a także ryzyko przeuczenia modeli.

Celem niniejszej pracy jest zaprojektowanie i implementacja własnego modelu opartego na konwolucyjnej sieci neuronowej, przeznaczonego do automatycznej klasyfikacji stopnia zaawansowania choroby Alzheimera na podstawie obrazów MRI. Praca skupia się na ocenie skuteczności modelu w rozróżnianiu czterech klas: brak otępienia, bardzo łagodne otępienie, łagodne otępienie, umiarkowane otępienie. Zaproponowana architektura zostanie porównana z wybranymi rozwiązaniami opisanymi w literaturze, a jej działanie zostanie ocenione przy użyciu szeregu metryk jakościowych, takich jak dokładność (accuracy), czułość (recall), swoistość (specificity), miara F1, a także ROC, AUC oraz macierz pomyłek. Najistotniejszą z nich jest dokładność.

W pracy podjęto próbę odpowiedzi na pytanie, czy stosunkowo prosta i lekka architektura CNN, zaprojektowana od podstaw i dopasowana do charakteru danych, może zapewnić wyniki porównywalne lub lepsze od rozbudowanych modeli opisywanych w literaturze.

Struktura pracy została podzielona na sześć głównych rozdziałów. W rozdziale drugim przedstawiono tło teoretyczne, obejmujące charakterystykę choroby Alzheimera, rolę obrazowania MRI, podstawy sztucznych sieci neuronowych oraz stosowane metryki ewaluacyjne. Trzeci rozdział opisuje zastosowane technologie, w tym środowisko programistyczne, biblioteki oraz sprzęt

---

<sup>1</sup> Rzecznik Praw Obywatelskich, *Sytuacja osób chorych na chorobę Alzheimera*, wyd. 2 (Warszawa: Biuro RPO, 2016), s. 12, <https://www.rpo.gov.pl/pl/content/sytuacja-osob-chorych-na-chorobe-alzheimera>.

<sup>2</sup> Geert Litjens i in., „A Survey on Deep Learning in Medical Image Analysis,” *Medical Image Analysis* 42 (2017): 60–88, <https://doi.org/10.1016/j.media.2017.07.005>, s. 77.

<sup>3</sup> Ryosuke Yamashita, Masahiro Nishio, Richard K. G. Do i Kuni O. Togashi, „Convolutional Neural Networks: An Overview and Application in Radiology,” *Insights into Imaging* 9, nr 4 (2018): 611–629, <https://doi.org/10.1007/s13244-018-0639-9>, s. 615.

obliczeniowy. Metodologia badań, obejmująca zbiór danych, architekturę modelu, proces treningu oraz metryki, została zaprezentowana w rozdziale czwartym. Rozdział piąty zawiera wyniki eksperymentów wraz z wizualizacjami oraz porównaniem z literaturą. Całość pracy kończy rozdział szósty, w którym zawarto podsumowanie, omówiono ograniczenia przeprowadzonych badań oraz wskazano potencjalne kierunki dalszych prac.

## **Wprowadzenie teoretyczne**

### **Choroba Alzheimera i diagnostyka**

Choroba Alzheimera to postępujące schorzenie neurodegeneracyjne, którego przyczyny wciąż nie są do końca poznane. W jej przebiegu dochodzi do stopniowego zaniku niektórych struktur mózgu oraz zaburzeń przewodzenia impulsów nerwowych. Powoduje to utratę tzw. wyższych funkcji poznawczych, takich jak pamięć, mowa, zdolność rozumienia i oceny sytuacji czy wykonywania obliczeń. Z czasem chory traci samodzielność, a w zaawansowanym stadium jest całkowicie zależny od pomocy opiekunów lub rodziny<sup>4</sup>.

Na chwilę obecną choroba pozostaje nieuleczalna, jednak stosowane terapie – farmakologiczne i psychospołeczne – mogą złagodzić objawy oraz spowolnić jej rozwój, co przekłada się na lepszą jakość życia pacjentów. Dlatego szybkie rozpoznanie choroby ma ogromne znaczenie i daje szansę na skuteczniejsze wsparcie osoby chorej.

Według szacunków IHME (Institute for Health Metrics and Evaluation), w 2019 roku w Polsce na chorobę Alzheimera i schorzenia pokrewne cierpiało około 585 000 osób, co stanowiło około 1,5% całej populacji. Na tle globalnym, w 2022 roku szacowano, że z chorobą Alzheimera żyje około 39 milionów ludzi na świecie, a według prognoz liczba ta może wzrosnąć do 65 milionów w 2030 roku oraz przekroczyć 100 milionów w 2050 roku. W samej Unii Europejskiej, w tym samym 2022 roku, z demencją typu alzheimerowskiego zmagало się około 9 milionów osób, co czyni to schorzenie jednym z najwiękzych wyzwań zdrowotnych starzejącego się społeczeństwa<sup>5</sup>.

Według raportu RPO z 2016 roku, około 60% cierpiących z powodu otępienia, chorowało na Alzheimera. Niestety, jedynie 15-20% chorych na otępienia otrzymało diagnozę i podjęło się leczenia. Dodatkowo, w ciągu następnych 20-25 lat liczba chorych podwoi się, podobnie, jak w innych krajach<sup>6</sup>.

Nie istnieje jedno badanie, które jednoznacznie potwierdziłoby chorobę Alzheimera, jednak diagnoza kliniczna może osiągnąć nawet około 90% trafności<sup>7</sup>. Proces diagnostyczny obejmuje konsultacje z kilkoma specjalistami oraz badania obrazowe:

- Neuropsycholog – ocenia ogólną sprawność intelektualną pacjenta i przebieg procesów poznawczych.

---

<sup>4</sup> Najwyższa Izba Kontroli, „NIK: o opiece nad chorymi na Alzheimera,” 16 maja 2017, <https://www.nik.gov.pl/najnowsze-informacje-o-wynikach-kontroli/nik-o-opiece-nad-chorymi-na-alzheimera.html>.

<sup>5</sup> Rzecznik Praw Obywatelskich, *Sytuacja osób chorych na chorobę Alzheimera*, s. 2-3

<sup>6</sup> Tamże, s. 12

<sup>7</sup> Centrum e-Zdrowia, „11 sygnałów, które wskazują na Alzheimera,” *Pacjent.gov.pl – Serwis Rzeczypospolitej Polskiej*, 21 września 2022, <https://pacjent.gov.pl/aktualnosc/11-sygnalow-ktore-wskazuja-na-alzheimera>.

- Neurolog – identyfikuje przyczyny zaburzeń funkcji poznawczych, wyklucza inne rodzaje otępień i prowadzi dalszą opiekę medyczną, często w poradni zaburzeń pamięci lub chorób neurodegeneracyjnych.
- Badania neuroobrazowe (TK lub MRI) – pozwalają wykluczyć inne przyczyny pogorszenia funkcji poznawczych (np. guzy mózgu) oraz określić lokalizację zmian patologicznych.
- Psychiatra – dokonuje oceny stanu chorego na wczesnym etapie choroby, zajmuje się leczeniem zaburzeń zachowania, udziela wskazówek opiekunom oraz pomaga w organizacji formalnej strony opieki (zaświadczenia, skierowania, recepty, informacje o dostępnych formach wsparcia).

## **Rola MRI w diagnostyce choroby Alzheimerera**

### **Czym jest MRI i dlaczego się go używa**

Magnetic Resonance Imaging (MRI) to metoda obrazowania medycznego wykorzystująca silne pole magnetyczne i fale radiowe, pozwalająca na pozyskanie szczegółowych obrazów struktur wewnętrznych mózgu bez stosowania promieniowania jonizującego. W diagnostyce chorób neurodegeneracyjnych, takich jak Alzheimer, MRI umożliwia ocenę strukturalnych zmian mózgu, w tym stopnia atrofii mózgu i zmian w objętości hipokampu<sup>8</sup>

### **Charakterystyka Alzheimerera na obrazach MRI**

Typowymi cechami otępienia alzheimerowskiego są: asymetryczna atrofia mózgu, zmniejszenie objętości hipokampu, pogłębienie bruzd mózgu i zwiększenie obecności płynu mózgowo-rdzeniowego (CSF)<sup>9</sup>.

Jak podkreślają Vemuri i Jack<sup>10</sup>, atrofia widoczna w obrazach MRI dobrze koreluje z gęstością splątków neurofibrylarnych, co uzasadnia jej rolę jako biomarkera choroby Alzheimerera. Wczesne zmiany widoczne na obrazach są skupione głównie w obrębie zakrętu przyhipokampowego i przyśrodkowej części płata skroniowego.

---

<sup>8</sup> *Medscape*, „Alzheimer Disease Imaging: Practice Essentials, Computed Tomography, Magnetic Resonance Imaging,” *Medscape*, 28 grudnia 2022, <https://emedicine.medscape.com/article/336281-overview>.

<sup>9</sup> Tamże.

<sup>10</sup> Prashanthi Vemuri i Clifford R. Jack Jr., „Role of Structural MRI in Alzheimer’s Disease,” *Alzheimer’s Research & Therapy* 2, nr 4 (2010): 23, <https://doi.org/10.1186/alzrt47>, s. 2.

## **Specyficzność obrazowania T1-weighted MRI**

W celu wykrywania Alzheimerera, często używane są obrazy T1-weighted reprezentowane jako „slices”. Są to pojedyncze przekroje warstwowe mózgu, które mogą być dalej przetwarzane jako obrazy 2D. Umożliwiają doskonałą wizualizację anatomii mózgu – tkanka nerwowa (istota szara) jest jasna, płyn mózgowo-rdzeniowy (CSF) ciemny, co ułatwia ocenę morfologiczną struktur mózgu. Pozwala to na dokładne mierzenie objętości hipokampu i innych regionów. Badania pokazują, że redukcja objętości hipokampu skutecznie różnicuje pacjentów z Alzheimerem od zdrowych osób<sup>11</sup>, dzięki czemu T1-weighted MRI uznawane są za praktyczny biomarker strukturalny tej choroby<sup>12</sup>.

## **Istotność MRI w Alzheimerze**

MRI pozwala nie tylko wykluczyć inne przyczyny otępienia – takie jak guzy, zmiany naczyniowe lub wodogłowie – ale także zidentyfikować wczesne objawy AD i monitorować postęp choroby w czasie<sup>13</sup>.

## **Sztuczna sieć neuronowa**

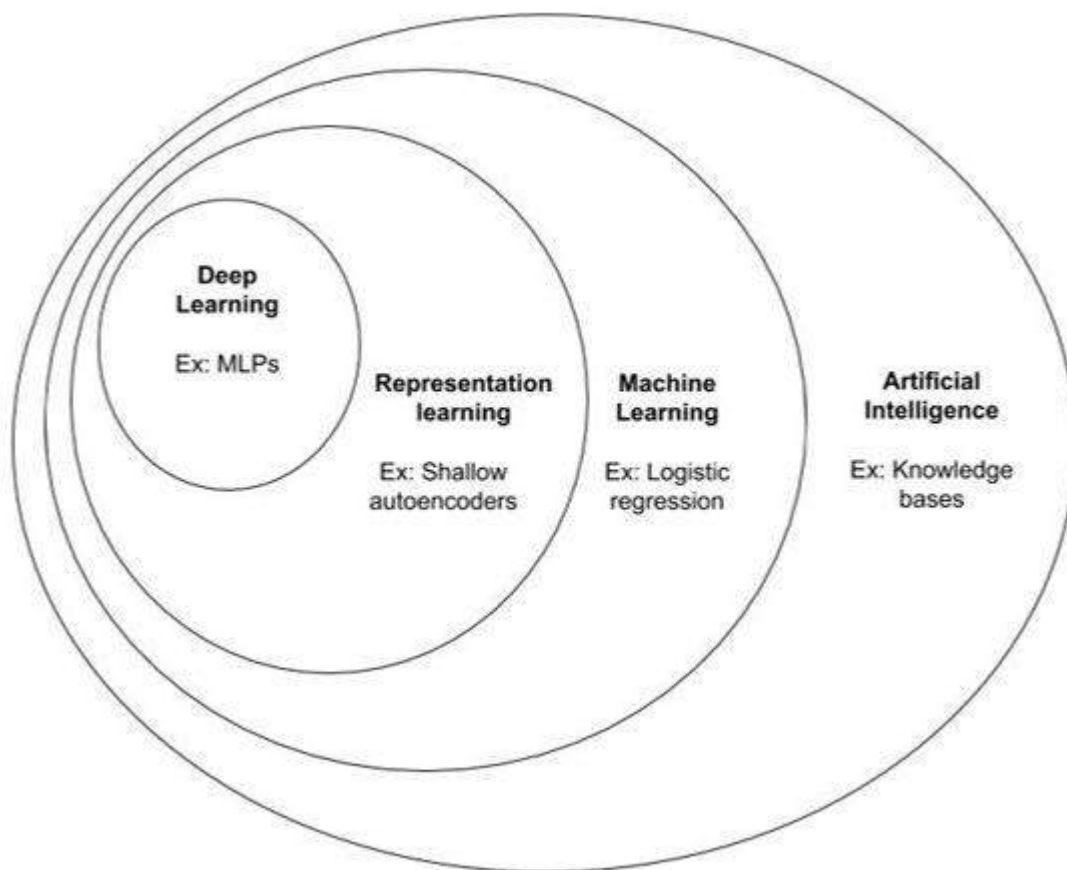
Sztuczna inteligencja (AI) to szeroka dziedzina, która obejmuje wiele dyscyplin i technik budowania inteligentnych systemów. Umożliwiają wykonywanie zadań wymagających ludzkiej inteligencji, takie jak rozumienie języka, rozpoznawanie obrazów, podejmowanie decyzji i rozwiązywanie problemów. Wszystkie opisane tutaj subdyscypliny zostały przedstawione na Rysunku 1.

---

<sup>11</sup> Tamże, s. 2-4.

<sup>12</sup> Aleksandra Marcisz i Joanna Polańska, „Can T1-Weighted Magnetic Resonance Imaging Significantly Improve Mini-Mental State Examination-Based Distinguishing Between Mild Cognitive Impairment and Early-Stage Alzheimer’s Disease?” *Journal of Alzheimer’s Disease* 92, nr 3 (2023): 941–957, <https://doi.org/10.3233/JAD-220806>.

<sup>13</sup> Frederik Barkhof, Marieke Hazewinkel, Max Binnewijzend i Robin Smithuis, „Dementia – Role of MRI,” *Radiology Assistant*, 2012/2022, <https://radiologyassistant.nl/neuroradiology/dementia/role-of-mri>.



Rysunek 1. Zależności pomiędzy różnymi dyscyplinami sztucznej inteligencji (AI) Źródło: Laraba i in., 2020, s. 12.

Uczenie maszynowe, będące subdyscypliną sztucznej inteligencji, koncentruje się na tworzeniu algorytmów i modeli statystycznych zdolnych do uczenia się na podstawie danych oraz do podejmowania decyzji i prognoz. W ramach tej dziedziny wyróżnia się różne podejścia, takie jak uczenie nadzorowane, nienadzorowane, częściowo nadzorowane oraz uczenie ze wzmocnieniem, z których każde opiera się na odmiennym sposobie wykorzystania danych. Uczenie maszynowe jest wykorzystywane w wielu aplikacjach, w tym do rozpoznawania obrazów<sup>14</sup> i mowy<sup>15</sup>, przetwarzania języka naturalnego (NLP)<sup>16</sup> i modelowania predykcyjnego<sup>17</sup> (Bishop, 2006).

Uczenie się reprezentacji (ang. *representation learning* lub *feature learning*) stanowi poddziedzinę uczenia maszynowego, skupioną na automatycznym odkrywaniu użytecznych cech danych, które mogą być następnie wykorzystane w innych zadaniach. Proces ten najczęściej przebiega w trybie nienadzorowanym, czyli bez potrzeby stosowania etykiet czy adnotacji – model

<sup>14</sup> Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever i Ruslan Salakhutdinov, „Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research* 15 (2014): 1929–1958, <http://jmlr.org/papers/v15/srivastava14a.html>.

<sup>15</sup> Geoffrey Hinton i in., „Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups,” *IEEE Signal Processing Magazine* 29, nr 6 (2012): 82–97, <https://doi.org/10.1109/MSP.2012.2205597>.

<sup>16</sup> Jacob Devlin, Ming-Wei Chang, Kenton Lee i Kristina Toutanova, „BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding,” w: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)* (Association for Computational Linguistics, 2019), s. 4171–4186, <https://arxiv.org/abs/1810.04805>.

<sup>17</sup> Christopher M. Bishop, *Pattern Recognition and Machine Learning* (Nowy Jork: Springer, 2006).

uczy się wyłącznie na podstawie dostępnych danych. Dzięki temu możliwe jest pozyskiwanie zwartych i znaczących reprezentacji, które wspierają późniejsze zadania analityczne, takie jak klasyfikacja, grupowanie czy przewidywanie.

Uczenie głębokie to subdyscyplina uczenia maszynowego inspirowana strukturą i funkcją mózgu, w szczególności sieciami neuronowymi, które tworzą mózg. Obejmuje ono szkolenie sztucznych sieci neuronowych na dużym zbiorze danych, umożliwiając sieci uczenie się i podejmowanie inteligentnych decyzji<sup>18</sup>.

W następnych sekcjach przedstawię kluczowe elementy sieci neuronowych, które mają największe znaczenie w kontekście tej pracy.

## Warstwy sieci neuronowej

### Warstwa liniowa

Warstwy liniowe są powszechnie stosowane w sieciach neuronowych. Określa się je mianem „liniowych”, ponieważ dokonują liniowej transformacji danych wejściowych.

Z matematycznego punktu widzenia, warstwa liniowa może być reprezentowana przez mnożenie macierzy, po którym następuje dodawanie elementów. Przykładowo, jeśli dane wejściowe do warstwy liniowej to macierz  $X$  o wymiarach (rozmiar partii, rozmiar danych wejściowych), a zakładamy rozmiar wyjściowy równy  $N$ , to wagi tej warstwy liniowej będą reprezentowane przez macierz  $W$  o wymiarach (rozmiar danych wejściowych, rozmiar danych wyjściowych). Bias warstwy liniowej będzie reprezentowany przez wektor  $b$  o długości  $N$ . Ogólny zapis działania warstwy liniowej przyjmuje wówczas postać:

$$\text{output} = XW + b$$

Warstwy liniowe często służą jako krytyczny element "główicy klasyfikacyjnej". Ta część sieci neuronowej pobiera dane wyjściowe z wcześniejszych warstw (np. konwolucyjnych), czyli wyodrębnione cechy, i wykorzystuje je do przeprowadzenia klasyfikacji.

### Warstwy konwolucyjne

Podczas wyodrębniania cech, warstwy konwolucyjne są często pomocne w takich celach. Początkowo wprowadzone w 1999<sup>19</sup>, obecnie często stanowią rdzeń DNN (głębokich sieci neuronowych). Warstwy te są biegłe w wyodrębnianiu cech przestrzennych danych wejściowych (głównie obrazów), co jest wyzwaniem przy użyciu tylko warstw liniowych. Dokonują tego poprzez zastosowanie stosunkowo małych (w odniesieniu do rozmiaru danych wejściowych) jąder (kerneli),

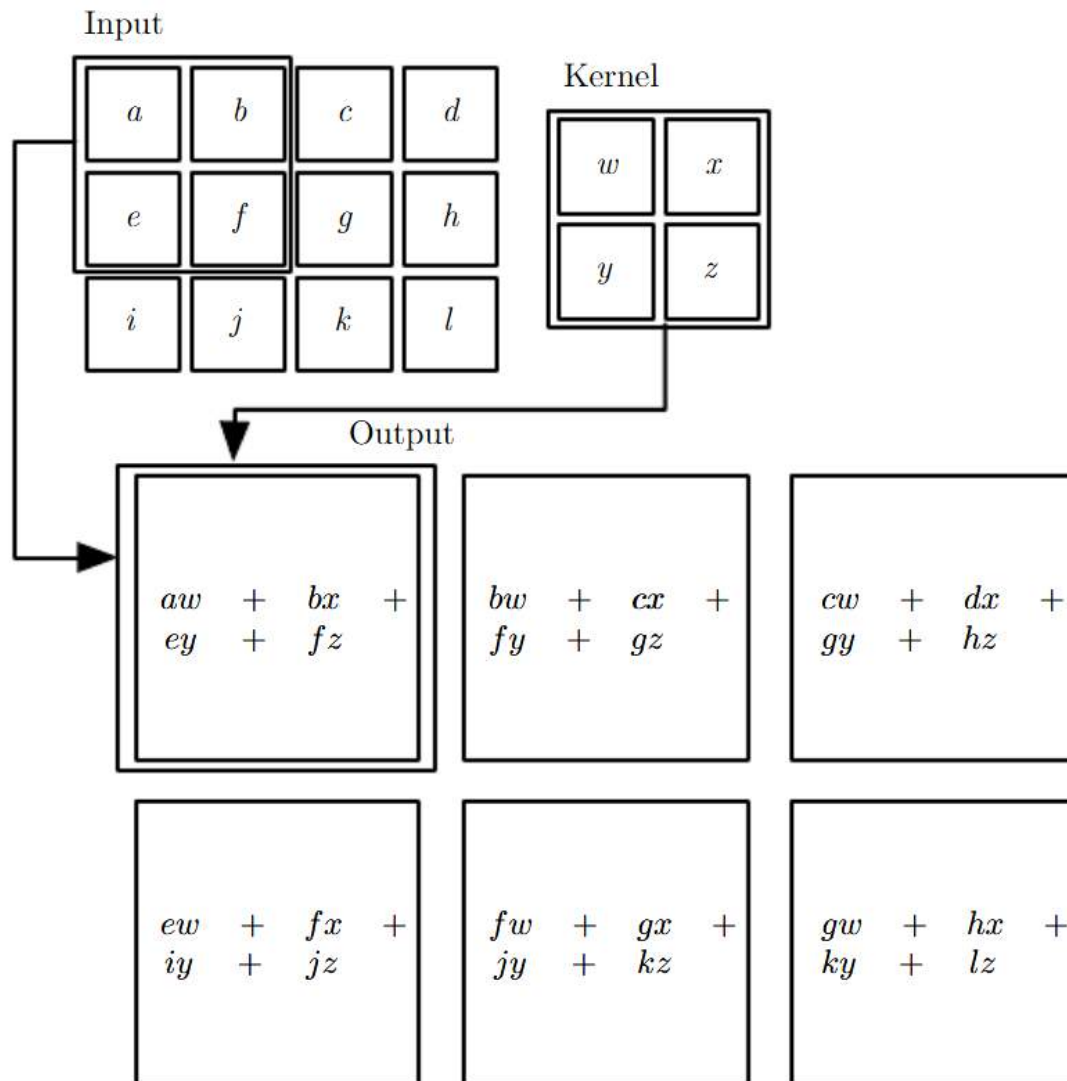
---

<sup>18</sup> Ian Goodfellow, Yoshua Bengio i Aaron Courville, *Deep Learning* (Cambridge, MA: MIT Press, 2016), <http://www.deeplearningbook.org>.

<sup>19</sup> Yann LeCun i in., „Object Recognition with Gradient-Based Learning,” w: *Shape, Contour and Grouping in Computer Vision*, red. David A. Forsyth, Joseph L. Mundy, Vincenzo Di Gesù i Roberto Cipolla (Springer, 1999), s. 319–345, [https://doi.org/10.1007/3-540-46805-6\\_19](https://doi.org/10.1007/3-540-46805-6_19).

które są macierzami z elementami (wagami) obliczonymi przez algorytm wstecznej propagacji<sup>20</sup>. Kiedy jądro jest stosowane do obrazu (zazwyczaj danych wejściowych), umieszczamy je na górze obrazu, obliczamy iloczyn punktów, a następnie przesuwamy i powtarzamy, aż jądro zostanie zastosowane do całego obrazu. Proces ten pokazano na Rysunku 2. Podobnie jak w przypadku warstw liniowych, możemy dodać bias do wyjścia splotu i zastosować funkcję aktywacji F. Ostateczne wyjście splotu można wyrazić wzorem poniżej, gdzie \* jest operatorem splotu.

$$\text{output} = F(X * W + b)$$



Rysunek 2. Przykład zastosowania operacji konwolucji do obrazu (bez biasu i funkcji aktywacji). Jądro jest tutaj macierzą o wymiarach (2×2). Źródło: Goodfellow i in., 2016.

Poniżej przedstawiam parametry, które można wykorzystać do utworzenia i wykonania operacji w warstwie konwolucyjnej:

<sup>20</sup> Jonathan Romero i in., „Strategies for Quantum Computing Molecular Energies Using the Unitary Coupled Cluster Ansatz,” *Quantum Science and Technology* 4, nr 1 (2018): 014008, <https://doi.org/10.1088/2058-9565/aad3e4>.

- Liczba filtrów – liczba filtrów, które mają być zastosowane w warstwie konwolucyjnej. Liczba ta ma bezpośredni wpływ na jeden z wymiarów danych wyjściowych.
- Rozmiar jądra (kernel size) – rozmiar macierzy jądra.
- Kroki (strides) – wartość określająca odległość przesunięcia między obliczaniem iloczynów skalarnych.
- Dylatacja (dilation) – zazwyczaj ustawiona na jeden, co oznacza, że nie pomijamy żadnych wartości z danych wejściowych (nie tworzymy luk). Gdy ustawimy ją na dwa, otrzymamy lukę o rozmiarze jeden i tak dalej.

Kolejną rzeczą, o której warto wspomnieć, jest wypełnianie (padding). Przed zastosowaniem konwolucji możemy dodać elementy wokół krawędzi danych wejściowych. Dzięki temu możemy zapewnić, że rozmiar danych wyjściowych będzie taki sam jak rozmiar danych wejściowych. Zazwyczaj dodane elementy są równe zero, ale możemy użyć na przykład wypełniania lustrzanego, gdzie dodane elementy są odbiciem wartości bliskich krawędziom danych wejściowych. Na przykład, jeśli na brzegu obrazu znajduje się piksel o wartości 150, to dodatkowy piksel poza krawędzią będzie miał również wartość 150, a kolejne będą stanowiły odbicie kolejnych wartości sąsiadujących.

## Dropout

Istnieje dobrze znany problem nadmiernego dopasowania sieci neuronowych. Nadmierne dopasowanie ma miejsce, gdy model jest zbyt dobrze wyszkolony na danych szkoleniowych, do tego stopnia, że nie może uogólniać na nieznanymi danych. Aby rozwiązać ten problem, naukowcy opracowali różne techniki regularyzacji. Jedną z tych technik nazywa się Dropout<sup>21</sup>. W implementacji TensorFlow<sup>22</sup> podczas szkolenia, w każdej iteracji warstwa dropout ustawia wartości wejściowe na zero z prawdopodobieństwem  $p$  i przeskaluje pozostałe wartości wejściowe o  $\frac{1}{1-p}$  aby zapewnić, że suma wartości wejściowych pozostaje niezmienną. Takie działanie zmusza sieć do szkolenia tylko podzbiorem wag podczas analizowania pojedynczej partii. Podejście to pomaga zapobiegać nadmiernemu dopasowaniu i poprawia uogólnienie.

## MaxPooling

W sieciach konwolucyjnych często pojawia się potrzeba redukcji wymiarów map cech, aby ograniczyć liczbę parametrów i skrócić czas uczenia. Jednym ze stosowanych rozwiązań jest warstwa MaxPooling. Jej działanie polega na dzieleniu mapy cech na niewielkie fragmenty (np. okna  $2 \times 2$ ) i wybieraniu z każdego fragmentu wartości maksymalnej. Dzięki temu sieć koncentruje się na najbardziej wyróżniających cechach obrazu, ignorując mniej istotne szczegóły. MaxPooling zmniejsza rozmiary danych przepływających przez kolejne warstwy, co redukuje złożoność obliczeniową modelu i pozwala uniknąć nadmiernego dopasowania. Takie podejście dodatkowo

<sup>21</sup> Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever i Ruslan Salakhutdinov, „Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research* 15 (2014): 1929–1958, <http://jmlr.org/papers/v15/srivastava14a.html>.

<sup>22</sup> Martín Abadi i in., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, oprogramowanie dostępne na: <https://www.tensorflow.org>.

zwiększa odporność modelu na niewielkie przesunięcia czy zniekształcenia obrazu, a jednocześnie zachowuje istotne informacje potrzebne do poprawnej klasyfikacji.

## Flatten

Warstwa Flatten stosowana jest w końcowej części sieci konwolucyjnych, aby przekształcić wielowymiarowe mapy cech w jednowymiarowy wektor. Dzięki temu możliwe jest połączenie części ekstrakcji cech (konwolucje i pooling) z klasyfikatorem opartym na warstwach w pełni połączonych. Nie zostają wprowadzane dodatkowe parametry, lecz jedynie zmieniana jest struktura danych, zachowując wszystkie informacje zakodowane w poprzednich mapach cech. Takie podejście pozwala przekazać sieci pełny zestaw wyodrębnionych cech w postaci wektora, co jest niezbędne do dalszej klasyfikacji.

## Dense

Warstwa Dense, znana również jako w pełni połączona (fully connected), stanowi podstawowy element klasyfikatora w sieciach neuronowych. Każdy neuron w tej warstwie połączony jest z wszystkimi neuronami poprzedniej warstwy, co pozwala na tworzenie złożonych zależności między cechami. Każde połączenie między neuronami posiada wagę, która jest aktualizowana podczas procesu uczenia. Funkcje aktywacji, takie jak ReLU czy softmax, nadają warstwie Dense zdolność modelowania nieliniowych zależności. Dzięki temu możliwe jest przekształcenie wyodrębnionych cech w prawdopodobieństwa przynależności do poszczególnych klas, co stanowi końcowy etap procesu klasyfikacji.

## Funkcje aktywacji sieci neuronowej

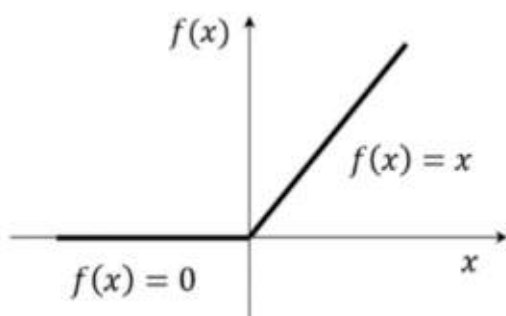
Aby uzyskać możliwość rozumienia równań nieliniowych, potrzebujemy czegoś więcej niż tylko wielu warstw liniowych ułożonych jedna na drugiej. Musimy wprowadzić nieliniowość do naszego modelu – funkcje aktywacji są jednym ze sposobów, aby to osiągnąć.

## ReLU

Obecnie jedną z najczęściej stosowanych funkcji aktywacji jest ReLU – Rectified Linear Unit<sup>23</sup> definiowana jako  $f(x) = \max(0, x)$ . Taka funkcja rozwiązała problem poprzednio stosowanej funkcji aktywacji sigmoidalnej  $\sigma(x) = \frac{1}{1+e^{-x}}$ , która była obciążona problemem z zanikającym gradientem. W przypadku sigmoidy, dla dużych wartości dodatnich i ujemnych argumentu pochodna (gradient) staje się bardzo bliska zeru, co powodowało, że podczas propagacji wstecznej (backpropagation) sygnał błędu zanikał i sieci głębokie uczyły się bardzo wolno lub wręcz wcale. Inaczej mówiąc, dla wartości  $x$  bardziej oddalonych od zera gradienty sigmoidalne są bliskie zeru, podczas gdy przy użyciu ReLU otrzymujemy stałe wartości niezerowe.

---

<sup>23</sup> Abien Fred Agarap, „Deep Learning Using Rectified Linear Units (ReLU),” *arXiv preprint arXiv:1803.08375* (2018), <http://arxiv.org/abs/1803.08375>.



*ReLU activation function*

Rysunek 3. Funkcja aktywacji ReLU. Źródło: Li i in., 2022.

## **Softmax**

Termin softmax został wprowadzony w (Bridle, 1989), o wzorze:

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

gdzie  $x_i$  jest  $i$ -tym elementem wektora wejściowego  $x$ , a  $n$  jest długością wektora. Wynikiem funkcji softmax jest wektor wartości z przedziału od 0 do 1, których suma wynosi 1. Funkcja softmax jest często stosowana w zadaniach klasyfikacji wieloklasowej w celu przekształcenia wyniku działania sieci neuronowej w rozkład prawdopodobieństwa dla możliwych klas.

## **Funkcje strat sieci neuronowej**

Podczas trenowania wag warstw sieci neuronowej chcemy zminimalizować błąd między przewidywanym wynikiem modelu a rzeczywistym wynikiem. Istnieje więcej niż jeden sposób pomiaru tego błędu, nazywane są one funkcjami strat, a ich wybór zależy od zadań uczenia maszynowego.

## Entropia krzyżowa kategoriyczna (Categorical cross entropy)

Entropia krzyżowa kategoriyczna jest obliczana przy użyciu wzoru:

$$\text{loss} = - \sum_{i=1}^n y_{\text{true},i} \log(y_{\text{pred},i})$$

gdzie  $n$  jest liczbą klas,  $y_{\text{true},i}$  jest prawdziwym prawdopodobieństwem  $i$ -tej klasy (często 0 lub 1 dla prawidłowej klasy), a  $y_{\text{pred},i}$  jest przewidywanym prawdopodobieństwem  $i$ -tej klasy.

Funkcja straty entropii krzyżowej kategoriycznej służy do szkolenia modelu w celu prawidłowego przewidywania prawdopodobieństw klas. Niska entropia krzyżowa kategoriyczna wskazuje, że model może dokładnie przewidywać prawdopodobieństwa klas. Natomiast wysoka entropia krzyżowa kategoriyczna wskazuje, że model nie może dokładnie przewidywać prawdopodobieństw klas.

## Weryfikacja procesu uczenia

Aby ocenić skuteczność sieci neuronowej, należy obserwować, jak zmieniają się wartości wybranych metryk w kolejnych epokach. Przeuczenie, to jeden z niepożądanych efektów uczenia<sup>24</sup>. Oznacza ono sytuację, w której model dobrze zapamiętuje dane treningowe, ale ma trudności z poprawnym działaniem na nowych przykładach. Żeby tego uniknąć, część danych wydziela się jako zbiór walidacyjny. Wyniki uzyskane na tym zbiorze pozwalają zauważyć, czy dana architektura uczy się zbyt dobrze na danych treningowych, a jednocześnie słabiej na walidacyjnych. W takim przypadku można przerwać uczenie wcześniej, aby zmienić architekturę modelu albo zmniejszyć liczbę epok.

## Wartość funkcji celu

Jedną z podstawowych miar postępu w uczeniu jest wartość funkcji celu. W kolejnych epokach powinna ona stopniowo maleć, co świadczy o poprawie działania modelu. Jeśli jednak spadek widoczny jest tylko dla zbioru treningowego, a na zbiorze walidacyjnym wartości pozostają bez zmian lub rosną, pojawia się ryzyko wspomnianego przed chwilą przeuczenia. Żeby temu zapobiec, stosuje się kryterium zatrzymania, które oznacza przerwanie procesu uczenia, gdy funkcja celu na danych walidacyjnych nie zmniejsza się przez określoną liczbę epok.

---

<sup>24</sup> Ryszard Tadeusiewicz i Mariusz Szaleniec, *Leksykon sieci neuronowych* (Wrocław: Projekt Nauka, Fundacja na rzecz Promocji Nauki Polskiej, 2015).

## Dokładność

Drugą miarą wykorzystaną w pracy była dokładność predykcji (accuracy). W ujęciu ogólnym określa się ją jako stosunek liczby poprawnie sklasyfikowanych przykładów do liczby wszystkich próbek w zbiorze:

$$ACC = \frac{N_T}{N_T + N_F}$$

gdzie  $N_T$  oznacza liczbę prawidłowo sklasyfikowanych przypadków, a  $N_F$  – liczbę błędnych predykcji.

W przypadku klasyfikacji wieloklasowej, obejmującej cztery kategorie odpowiadające różnym stopniom nasilenia choroby Alzheimera (brak otępienia, bardzo łagodne otępienie, łagodne otępienie, umiarkowane otępienie), dokładność wyrażana jest jako odsetek wszystkich poprawnych rozpoznań we wszystkich klasach. Należy jednak pamiętać, że podobnie jak w klasyfikacji binarnej, sama miara accuracy może być myląca, zwłaszcza gdy rozkład liczebności klas nie jest zrównoważony. W takiej sytuacji model, który częściej przewiduje klasę dominującą, może osiągać pozornie wysoką dokładność, mimo że jego zdolność do rozróżniania poszczególnych klas pozostaje niska. Dlatego przy ocenie jakości klasyfikatora wieloklasowego należy posiłkować się także innymi metrykami, takimi jak precyzja (precision), czułość (recall), miara F1 czy analiza macierzy pomyłek, które omawiam w następnych podrozdziałach.

## Precyzja

Precyzja (precision) określa, jaki odsetek próbek zaklasyfikowanych przez model do danej klasy rzeczywiście do niej należy. Innymi słowy, precyzja mierzy dokładność pozytywnych przewidywań modelu. Wzór wygląda tak:

$$\text{Precision} = \frac{TP}{TP + FP}$$

gdzie TP (*True Positives*) oznacza liczbę poprawnie zaklasyfikowanych przykładów danej klasy, a FP (*False Positives*) liczbę przykładów błędnie zaklasyfikowanych jako należące do tej klasy.

Wysoka wartość precision oznacza, że model rzadko popełnia błędy polegające na przypisywaniu przykładowi niewłaściwej klasy. Metryka ta jest szczególnie istotna w zadaniach, w których fałszywe alarmy są kosztowne, np. w diagnostyce medycznej czy systemach bezpieczeństwa.

## Czułość i swoistość

W ocenie modeli klasyfikacyjnych stosuje się także czułość (sensitivity, recall) oraz swoistość (specificity). Są to metryki mniej zależne od stopnia zrównoważenia danych niż sama dokładność.

Czułość określa, jak duża część przypadków danej klasy została poprawnie rozpoznana. Definiuje się ją wzorem:

$$SE = \frac{TP}{TP + FN}$$

gdzie TP (*True Positives*) to liczba próbek prawidłowo przypisanych do danej klasy (prawdziwie pozytywnych), a FN (*False Negatives*) – liczba przykładów należących do tej klasy, które model błędnie zaklasyfikował gdzie indziej (fałszywie negatywnych).

Z kolei swoistość pokazuje, jak dobrze model rozpoznaje przypadki nienależące do danej klasy. Opisuje ją wzór:

$$SP = \frac{TN}{TN + FP}$$

gdzie TN (*True Negatives*) oznacza liczbę prawidłowo sklasyfikowanych przypadków negatywnych (prawdziwie negatywnych), a FP (*False Positives*) – liczbę przykładów błędnie przypisanych do tej klasy (fałszywie pozytywnych).

W przypadku klasyfikacji wieloklasowej, takiej jak rozróżnianie czterech kategorii otępienia alzheimerskiego, wartości czułości i swoistości można obliczyć osobno dla każdej klasy, a następnie uśrednić (np. metodą macro averaging). Dzięki temu uzyskuje się pełniejszy obraz skuteczności modelu, również w odniesieniu do klas mniej licznych.

## Confusion Matrix

Macierz pomyłek (confusion matrix) jest narzędziem stosowanym do oceny jakości klasyfikatora. Przedstawia ona w formie tabeli porównanie przewidywanych przez model etykiet z etykietami rzeczywistymi. W przypadku klasyfikacji binarnej macierz ma rozmiar 2×2 i zawiera cztery wartości: True Positives (TP) – liczba poprawnie wykrytych przykładów pozytywnych, True Negatives (TN) – liczba poprawnie wykrytych przykładów negatywnych, False Positives (FP) – liczba przykładów błędnie uznanych za pozytywne oraz False Negatives (FN) – liczba przykładów błędnie sklasyfikowanych jako negatywne. W klasyfikacji wieloklasowej, użytej w mojej pracy na potrzeby przypadku choroby Alzheimera, macierz pomyłek przyjmuje postać kwadratowej tabeli, w której każda kolumna odpowiada klasom przewidzianym przez model, a każdy wiersz klasom rzeczywistym. Dzięki analizie macierzy pomyłek można łatwo zidentyfikować, które klasy są najczęściej mylone, a tym samym wskazać słabe strony klasyfikatora.

## F1-score

F1-score jest metryką służącą do oceny jakości klasyfikatorów, szczególnie przydatną w przypadku niezrównoważonych zbiorów danych. Wartość ta stanowi średnią harmoniczną dwóch innych miar – precyzji (precision) oraz czułości (recall). Dzięki temu F1-score uwzględnia zarówno odsetek poprawnie wykrytych przypadków w danej klasie, jak i dokładność wskazywania próbek należących do tej klasy. Wzór na tę metrykę ma postać:

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Wartość F1-score przyjmuje zakres od 0 do 1, gdzie 1 oznacza idealny wynik klasyfikacji, a wartości bliskie 0 wskazują na bardzo słabą skuteczność. W klasyfikacji wieloklasowej, takiej jak analiza różnych stadiów choroby Alzheimera, F1-score oblicza się osobno dla każdej klasy, a następnie uśrednia – najczęściej metodą macro lub micro averaging. Metryka ta pozwala zrównoważyć znaczenie zarówno precyzji, jak i czułości, dlatego jest szeroko stosowana w ocenie algorytmów uczenia maszynowego.

## Macro, micro i weighted average

W przypadku klasyfikacji wieloklasowej często stosuje się różne sposoby uśredniania metryk, takich jak Precision, Recall czy F1-score. Ma to na celu uzyskanie bardziej miarodajnych wyników, zwłaszcza przy nierównomiernym rozkładzie licznosci próbek w klasach.

- Macro average – oblicza średnią arytmetyczną metryk dla wszystkich klas, traktując je jednakowo, niezależnie od ich wielkości. Dzięki temu dobrze pokazuje skuteczność modelu w odniesieniu do każdej klasy, także tej najmniej licznej.
- Micro average – sumuje wartości prawdziwie i fałszywie sklasyfikowanych próbek ze wszystkich klas, a następnie wylicza metryki. Jest wrażliwe na klasy dominujące liczebnie i dobrze odzwierciedla ogólną skuteczność klasyfikacji.
- Weighted average – oblicza średnie ważone metryk, przydzielając wagę proporcjonalną do licznosci każdej klasy. Pozwala to lepiej oddać wpływ klas o większej liczbie próbek, a jednocześnie częściowo uwzględnia problem niezbalansowanych danych.

Zastosowanie tych trzech podejść umożliwia pełniejszą ocenę jakości klasyfikatora, zarówno w ujęciu globalnym, jak i pod względem równomierności działania dla każdej z klas.

## ROC

ROC (Receiver Operating Characteristic curve) to narzędzie używane do oceny jakości klasyfikatorów. Krzywa ROC przedstawia graficzną zależność między czułością (True Positive Rate, TPR) a 1 – swoistością (False Positive Rate, FPR) przy różnych progach decyzyjnych modelu. Dzięki temu pozwala ocenić, jak skutecznie klasyfikator odróżnia przykłady pozytywne od negatywnych w całym zakresie możliwych progów.

W klasyfikacji binarnej krzywa ROC biegnąca blisko lewego górnego rogu oznacza bardzo dobrą skuteczność modelu, natomiast linia bliska przekątnej ( $TPR \approx FPR$ ) wskazuje, że model działa losowo. Dla klasyfikacji wieloklasowej, takiej jak rozpoznawanie różnych stadiów choroby Alzheimera, krzywe ROC wyznacza się osobno dla każdej klasy, traktując ją w danym momencie jako klasę pozytywną (one-vs-rest).

Z krzywej ROC wylicza się także pole pod krzywą (AUC – Area Under the Curve), które przyjmuje wartości od 0 do 1. Wynik bliski 1 świadczy o bardzo dobrej zdolności modelu do odróżniania klas, a wynik 0.5 oznacza losowe działanie klasyfikatora.

## AUC

AUC (Area Under the Curve) to metryka obliczana na podstawie krzywej ROC, która w sposób liczbowy opisuje zdolność modelu do rozróżniania klas. Wartość AUC odpowiada polu powierzchni pod wykresem ROC i przyjmuje wartości w zakresie od 0 do 1. Im większe pole, tym lepsza skuteczność klasyfikatora.

Interpretacja jest intuicyjna:  $AUC = 0.5$  oznacza, że model nie przewyższa losowego zgadywania, AUC powyżej 0.8 wskazuje na dobrą jakość klasyfikacji, natomiast wartości bliskie 1.0 świadczą o bardzo wysokiej skuteczności.

W przypadku klasyfikacji wieloklasowej (np. rozróżnianie stadiów choroby Alzheimera) AUC wyznacza się osobno dla każdej klasy w podejściu one-vs-rest, a następnie można je uśrednić metodą macro lub micro averaging. Dzięki temu metryka ta daje całościowy obraz jakości modelu i pozwala porównywać różne architektury sieci neuronowych w jednolity sposób.

## Dodatkowe mechanizmy sterujące treningiem sieci

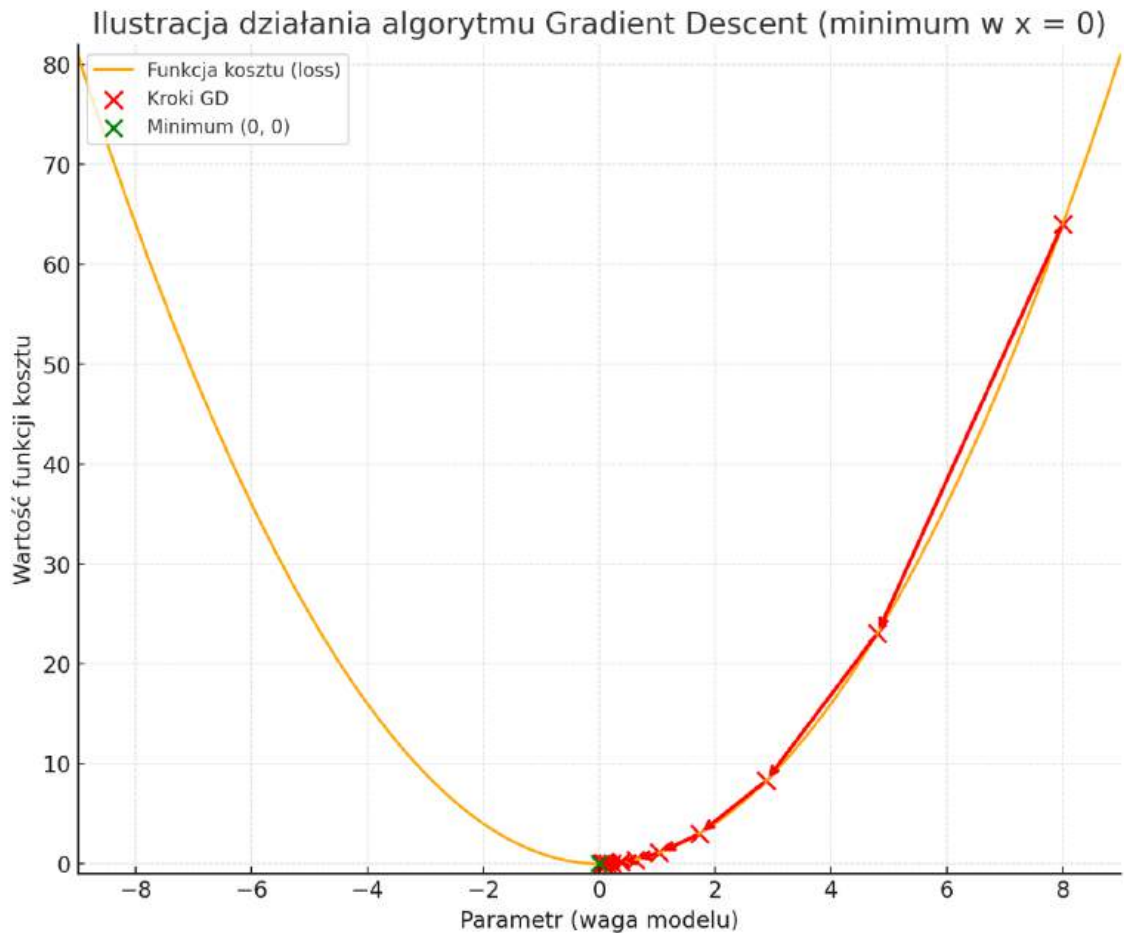
### EarlyStopping

Metoda wczesnego zatrzymania treningu. Działanie tego narzędzia polega na monitorowaniu wartości funkcji straty (`val_loss`) na zbiorze walidacyjnym. Jeżeli w kolejnych  $x$  epokach nie następuje poprawa jakości, proces uczenia zostaje zakończony, a model automatycznie przywraca wagi uzyskane w najlepszym momencie treningu. Rozwiązanie to pozwala uniknąć przeuczenia, polepsza generalizację i skraca czas obliczeń.

### ReduceLROnPlateau

Mechanizm ten dynamicznie dostosowuje tempo uczenia (learning rate). Algorytm obserwuje zmiany funkcji straty na zbiorze walidacyjnym i jeśli w ciągu  $x$  kolejnych epok nie nastąpi poprawa, tempo uczenia zostaje zmniejszone o połowę. Dzięki temu model ma możliwość bardziej precyzyjnej optymalizacji w końcowych etapach treningu, kiedy dalsza poprawa staje się trudniejsza do osiągnięcia. W celu poprawnego działania, wartość  $x$  musi być mniejsza niż wartość ustawiona w EarlyStopping.

## Gradient Descent



Rysunek 4. Działanie Gradient Descent. Źródło: opracowanie własne

Gradient Descent (spadek gradientowy) to jedna z podstawowych metod optymalizacji stosowanych w uczeniu maszynowym i głębokim uczeniu. Jego głównym celem jest minimalizacja funkcji kosztu (loss function), która mierzy różnicę pomiędzy przewidywaniami modelu a rzeczywistymi wartościami.

Algorytm opiera się na iteracyjnym dostosowywaniu wag modelu w kierunku przeciwnym do gradientu funkcji kosztu względem tych wag. Gradient wskazuje kierunek najszybszego wzrostu funkcji, a poruszanie się w stronę przeciwną pozwala stopniowo zbliżać się do jej minimum (Rysunek 4). Wielkość kroku wykonywanego w każdej iteracji określa learning rate – zbyt duży może powodować rozbieżność algorytmu, a zbyt mały spowalnia proces uczenia.

W praktyce stosuje się różne warianty spadku gradientowego, m.in.:

- Batch Gradient Descent – wykorzystuje cały zbiór danych w jednej iteracji, co daje stabilne, ale wolne uczenie.
- Stochastic Gradient Descent (SGD) – aktualizuje wagi na podstawie pojedynczej próbki, co przyspiesza obliczenia, ale wprowadza większą losowość.

- Mini-batch Gradient Descent – łączy zalety obu podejść, aktualizując wagi na podstawie niewielkich porcji danych (mini-batchy).

### **Optymalizator Adam**

Adam (ang. *Adaptive Moment Estimation*) łączy w sobie zalety dwóch innych podejść: momentum (wykorzystywanie średniej ważonej przeszłych gradientów, co pozwala wygładzić kierunek uczenia) oraz RMSProp (adaptacyjne skalowanie kroku uczenia w zależności od wartości gradientów). Dzięki temu Adam automatycznie dostosowuje tempo uczenia dla każdego parametru z osobna, co przyspiesza i stabilizuje proces treningu.

W odróżnieniu od klasycznego Gradient Descent, który aktualizuje wszystkie wagi w oparciu o ten sam stały learning rate i aktualny gradient, Adam bierze pod uwagę także historię gradientów i ich kwadratów, aby dobrać bardziej optymalny krok. W praktyce oznacza to, że Adam jest mniej wrażliwy na wybór hiperparametrów i często szybciej zbiega do minimum funkcji kosztu.

### **Learning Rate**

Jednym z najważniejszych hiperparametrów w procesie uczenia sieci neuronowych jest współczynnik uczenia (learning rate). Parametr ten określa, o jaką wartość będą aktualizowane wagi modelu w trakcie każdej iteracji algorytmu optymalizacji. Zbyt duży współczynnik uczenia może prowadzić do niestabilności treningu i „przeskakiwania” minimum funkcji straty, co uniemożliwia osiągnięcie dobrej jakości modelu. Z kolei zbyt mała wartość powoduje bardzo powolną zbieżność, a czasem utknięcie w minimum lokalnym. Współczynnik ten można ustawić ręcznie lub sterować nim dynamicznie za pomocą mechanizmu takiego jak ReduceLROnPlateau. Odpowiedni dobór learning rate ma kluczowe znaczenie dla skuteczności i szybkości uczenia sieci neuronowej.

### **Dodatkowa terminologia procesu uczenia**

#### **Batch size**

Batch size (rozmiar paczki) to liczba przykładów uczących, które są przetwarzane jednocześnie podczas jednej iteracji aktualizacji wag modelu. Ma on duże znaczenie dla procesu uczenia: małe batch size zwiększa losowość i może poprawić uogólnianie, ale spowalnia trening, natomiast duże batch size przyspiesza obliczenia na GPU, choć czasem prowadzi do gorszej generalizacji.

#### **Epoka**

Epoka (epoch) to jedno pełne przejście modelu przez cały zbiór treningowy. Oznacza to, że w trakcie jednej epoki każdy przykład z danych uczących został wykorzystany do aktualizacji wag co najmniej raz. Liczba epok decyduje o tym, ile razy model uczy się na wszystkich dostępnych danych – zbyt mała liczba może prowadzić do niedouczenia, a zbyt duża do przeuczenia.

## Przegląd literatury - przykładowe rozwiązania problemu

Helaly i współpracownicy<sup>25</sup> zaproponowali podejście oparte na konwolucyjnych sieciach neuronowych (CNN) do wieloklasowej i binarnej klasyfikacji choroby Alzheimera z wykorzystaniem obrazów MRI z bazy ADNI (Alzheimer's Disease Neuroimaging Initiative). W pracy opracowano dwie autorskie architektury sieci (2D-M2IC i 3D-M2IC) oraz zastosowano model transfer learning z dostrojeniem wag VGG19. Dane obejmowały 300 pacjentów (AD, EMCI, LMCI, NC), a po użyciu technik augmentacji danych (rotacje, odbicia, zmiany skali) zrównoważono zbiór i zwiększono liczbę przykładów do 48 000 obrazów (po 12 000 na klasę). Dane podzielono na zbiory treningowy, walidacyjny i testowy w proporcjach 80/10/10. Najwyższą skuteczność w klasyfikacji wieloklasowej uzyskał model VGG19 (fine-tuned) – 97% dokładności, natomiast w klasyfikacjach binarnych wyniki sięgały 98,47% (EMCI vs NC). Ocenę przeprowadzono z użyciem metryk takich jak Accuracy, Precision, Recall, F1-score, ROC, AUC. Zastosowaną funkcją straty była categorical cross-entropy, optymalizator: Adam, learning rate ustawiony na 0,0001. Finalny model składał się z 25,4 mln parametrów.

Modele trenowano przez 50 epok, z użyciem *batch size* równym 32, a czas treningu wynosił ok. 3 godziny dla 2D-M2IC, 10 godzin dla 3D-M2IC i 2 godziny dla VGG19. Eksperymenty przeprowadzono na komputerze z procesorem Intel Core i7 (10. generacji), pamięcią RAM 16 GB i kartą graficzną NVIDIA GTX 1050Ti (4 GB VRAM)

Architektura modelu 2D-M2IC, która bazowała na obrazach 2D, wygląda następująco: 3 warstwy konwolucyjne (3×3), 3 max-pooling, 2 warstwy dropout, flatten, 2 warstwy w pełni połączone, softmax na wyjściu. Uzyskała ona wynik dokładności 93,60%.

W badaniach Fultona i współpracowników<sup>26</sup>, zaproponowano podejście wykorzystujące zarówno dane demograficzne i kliniczne, jak i obrazy MRI do klasyfikacji choroby Alzheimera w oparciu o zbiór OASIS (Open Access Series of Imaging Studies). W pracy zestawiono dwa rodzaje modeli: gradient boosted machines (GBM) dla danych bezobrazowych oraz sieć ResNet-50 dla obrazów MRI.

GBM przewidywał obecność AD na podstawie wieku, płci, wykształcenia, statusu socjoekonomicznego i wyników testu MMSE, osiągając średnią dokładność 91,3%. Najważniejszym predyktorem okazał się **standaryzowany test przesiewowy** używany w diagnostyce zaburzeń poznawczych i otępień (MMSE).

ResNet-50 został wytrenowany na 16 661 obrazach MRI (51 wybranych przekrojów z 416 badań), z wykorzystaniem augmentacji (rotacje, przesunięcia, odbicia, zmiany skali). Dane podzielono w proporcji 80/20 na zbiór treningowy i walidacyjny. Funkcją straty była categorical cross-entropy. Model osiągnął 98,99% dokładności w klasyfikacji trzech klas nasilenia demencji

---

<sup>25</sup> Hossam A. Helaly, Mohamed Badawy i Ahmed Y. Haikal, „Deep Learning Approach for Early Detection of Alzheimer's Disease,” *Cognitive Computation* 14 (2022): 1711–1727, <https://doi.org/10.1007/s12559-021-09946-2>.

<sup>26</sup> Larry V. Fulton, David Dolezel, Jason Harrop, Yan Yan i Cathy P. Fulton, „Classification of Alzheimer's Disease with and without Imagery Using Gradient Boosted Machines and ResNet-50,” *Brain Sciences* 9, nr 9 (2019): 212, <https://doi.org/10.3390/brainsci9090212>.

(brak, bardzo łagodna, łagodna/umiarkowana) przy 133 epokach oraz 99,34% na zbiorze treningowym.

Trening ResNet-50 przeprowadzono na komputerze z procesorem Intel i7-6700, 64 GB RAM i kartą graficzną NVIDIA GTX 980M, a czas całkowity wyniósł ok. 10 godzin (152 epoki).

W artykule główną metryką raportowaną była accuracy – zarówno dla modeli bezobrazowych (GBM: średnio 91,3%) jak i obrazowych (ResNet-50: 98,99% na walidacji, 99,34% na treningu) oraz AUC (wyłącznie dla GBM) ~96%. Nie raportowali natomiast precyzji, czułości czy F1-score.

Naganandhini i Shanmugavadivu<sup>27</sup> opracowali zmodyfikowany klasyfikator drzewa decyzyjnego z dostrajaniem hiperparametrów (Decision Tree Classifier with Hyper Parameters Tuning – DTC-HPT) w celu diagnozy choroby Alzheimera na podstawie danych ze zbioru OASIS. Dane obejmowały informacje demograficzne i kliniczne, takie jak wiek, płeć, edukacja, status społeczno-ekonomiczny, MMSE, CDR (kliniczna skala nasilenia otępienia), eTIV (szacowana całkowita objętość wewnątrzczaszkowa) czy nWBV (znormalizowana objętość całego mózgu).

Model został zoptymalizowany m.in. pod względem głębokości drzewa, liczby węzłów liściowych i funkcji podziału (Gini, entropia). Klasyfikator osiągał średnią skuteczność 99,10%. Zbiór danych nie był odpowiednio zbalansowany, ponadto posiadał tylko 112 próbek. Oceny dokonano z użyciem czterech metryk: dokładność, precyzja, czułość i F1-score.

Hazarika i współpracownicy<sup>28</sup> przeprowadzili eksperymentalną analizę 20 popularnych modeli głębokiego uczenia (m.in. LeNet, AlexNet, VGG-16, VGG-19, Inception-V1–V3, ResNet-50/101/152, MobileNet, EfficientNet-B0/B7, Xception, NASNet, DenseNet-121) w klasyfikacji choroby Alzheimera z obrazów MRI pochodzących ze zbioru ADNI. Dane obejmowały ponad 210 pacjentów w trzech klasach: cognitively normal (CN), mild cognitive impairment (MCI) i AD. Po augmentacji (rotacje, odbicia lustrzane, transformacje skali) zbiór zwiększono do ponad 15 000 obrazów, rozdzielonych na zbiór treningowy, walidacyjny i testowy. Dodatkowo pacjentów pogrupowano wg wieku (60–69, 70–79 i 80+ lat), aby ocenić wpływ starzenia na skuteczność klasyfikacji.

Autorzy zbadali również wpływ wstępnej obróbki danych (m.in. usuwania czaszki) na jakość predykcji – wszystkie modele osiągały wyraźnie lepsze wyniki po jej wykonaniu. Trening prowadzono przez 50 epok, z batch size równym 32, z wykorzystaniem optymalizatora SGD i funkcji straty Sparse Categorical Cross-Entropy. Środowisko obliczeniowe obejmowało procesor Intel i7, 16 GB RAM i kartę graficzną NVIDIA 2 GB.

---

<sup>27</sup> S. Naganandhini i P. Shanmugavadivu, „Effective Diagnosis of Alzheimer’s Disease Using Modified Decision Tree Classifier,” *Procedia Computer Science* 165 (2019): 548–555, <https://doi.org/10.1016/j.procs.2020.01.049>.

<sup>28</sup> Rajarshi A. Hazarika, Debasish Kandar i Amit K. Maji, „An Experimental Analysis of Different Deep Learning Based Models for Alzheimer’s Disease Classification Using Brain Magnetic Resonance Images,” *Journal of King Saud University – Computer and Information Sciences* 34, nr 10 (2022): 8576–8598, <https://doi.org/10.1016/j.jksuci.2021.09.003>.

Najwyższą skuteczność osiągnął model DenseNet-121 – średnio 88,78% dokładności w klasyfikacji trzech klas, a po modyfikacji architektury (zastąpienie warstw konwolucyjnych operacjami depth-wise convolution) wynik wzrósł do 90,22%. Inne dobrze działające architektury to MobileNet-V2 (88,25%), InceptionResNet (86,33%) i VGG-19 (86,03%). W każdej ocenie raportowano metryki: Accuracy, Precision, Recall, F1-score oraz średnie czasy treningu na epokę (od 70 sekund dla LeNet do ponad 1000 sekund dla InceptionResNet).

## Zastosowane technologie

### Oprogramowanie

Algorytm stworzony w ramach pracy został zaimplementowany w języku Python. Wybór ten był uzasadniony szeroką dostępnością bibliotek dedykowanych analizie danych, uczeniu maszynowemu, budowie sieci neuronowych oraz przetwarzaniu obrazów, co znacząco usprawniło proces implementacji i eksperymentów. Jako środowisko programistyczne wykorzystano Jupyter Notebook działające w ramach platformy Anaconda Navigator.

### Biblioteki

W trakcie realizacji projektu korzystano z kilku popularnych bibliotek zewnętrznych, które można podzielić na dwie główne grupy: narzędzia do przetwarzania danych oraz rozwiązania związane z budową i trenowaniem modeli uczenia maszynowego.

Pierwszą kategorię stanowią pakiety ułatwiające przygotowanie i analizę danych. Zasadniczą strukturą używaną w pracy był DataFrame z biblioteki Pandas<sup>29</sup>, który umożliwia wygodne wczytywanie zbiorów, ich modyfikację oraz analizę w postaci tabelarycznej. W wielu operacjach wspierano się również biblioteką NumPy<sup>30</sup>, która dostarcza implementacji tablic wielowymiarowych oraz bogaty zestaw funkcji matematycznych i statystycznych, pozwalających na szybkie przetwarzanie dużych zbiorów numerycznych.

Drugą grupę narzędzi tworzą biblioteki związane z implementacją i trenowaniem sieci neuronowych. Kluczowym środowiskiem obliczeniowym był TensorFlow<sup>31</sup>, zawierający mechanizmy uczenia maszynowego oraz moduł Keras<sup>32</sup>, który dostarcza wysokopoziomowego interfejsu do budowania, kompilowania i trenowania sieci konwolucyjnych. Proces uczenia monitorowano z pomocą TensorBoard, narzędzia wizualizacyjnego stanowiącego rozszerzenie TensorFlow, pozwalającego śledzić przebieg treningu i zmiany wartości metryk w kolejnych epokach.

---

<sup>29</sup> Wes McKinney, „Data Structures for Statistical Computing in Python,” w: *Proceedings of the 9th Python in Science Conference* (2010), s. 51–56, <https://doi.org/10.25080/Majora-92bf1922-00a>.

<sup>30</sup> Charles R. Harris i in., „Array Programming with NumPy,” *Nature* 585, nr 7825 (2020): 357–362, <https://doi.org/10.1038/s41586-020-2649-2>.

<sup>31</sup> Abadi i in., *TensorFlow*.

<sup>32</sup> François Chollet i in., *Keras*, oprogramowanie dostępne na: <https://keras.io>.

Dodatkowo w projekcie wykorzystywano bibliotekę scikit-learn<sup>33</sup>. Służyła ona między innymi do dzielenia zbioru danych na części treningową, walidacyjną i testową, a także do implementacji funkcji pomocniczych wspierających proces ewaluacji modeli, takich jak generowanie macierzy pomyłek czy raportów klasyfikacji.

Ostatnim istotnym narzędziem była biblioteka Matplotlib<sup>34</sup>, wykorzystywana do wizualizacji wyników w postaci wykresów funkcji, histogramów czy graficznych prezentacji metryk, co ułatwiało analizę uzyskanych rezultatów.

## Środowisko sprzętowe

Do realizacji eksperymentów wykorzystano komputer wyposażony w procesor Intel Core i7-8700K, 16 GB pamięci RAM DDR4 3200 MHz oraz kartę graficzną MSI GeForce RTX 2060 Gaming Z z 6 GB pamięci VRAM, której moc obliczeniowa w operacjach FP32 wynosi około 6,5 TFLOPS. Dzięki użyciu GPU proces trenowania sieci neuronowych uległ znacznemu przyspieszeniu, co pozwoliło na sprawniejszą realizację obliczeń wymaganych przez modele konwolucyjne.

## Metodologia

### Zbiór danych

W pracy wykorzystano zbiór danych udostępniony na platformie Kaggle<sup>35</sup>, przygotowany na podstawie oryginalnych danych z projektu OASIS<sup>36</sup>, który zawiera między innymi około 80 000 obrazów MRI mózgu. Obrazy zostały podzielone na cztery klasy w zależności od stopnia zaawansowania choroby Alzheimera. Projekt OASIS ma na celu dostarczenie wartościowego źródła do analizy oraz wykrywania wczesnych objawów choroby. Zbiór dostępny na Kaggle został wstępnie przetworzony (m.in. przycięcie, usunięcie czaszki i standaryzacja obrazów) oraz zaugmentowany.

Sztuczne zwiększanie liczby próbek (augmentacja danych) polegała na zastosowaniu różnorodnych transformacji obrazów MRI, m.in. takich jak obrót, odbicie lustrzane, zmiana skali czy przesunięcie, aby zwiększyć liczbę próbek i uzyskać bardziej zrównoważony rozkład klas. Poprzez zastosowanie takiego rozwiązania, model sieci neuronowej ma dostęp do większej liczby przykładów, co ogranicza ryzyko przeuczenia i poprawia jego zdolność do generalizacji. Dzięki temu zbiór jest gotowy do bezpośredniego wykorzystania w treningu sieci neuronowych.

---

<sup>33</sup> Fabian Pedregosa i in., „Scikit-Learn: Machine Learning in Python,” *Journal of Machine Learning Research* 12 (2011): 2825–2830, <http://jmlr.org/papers/v12/pedregosa11a.html>.

<sup>34</sup> John D. Hunter, „Matplotlib: A 2D Graphics Environment,” *Computing in Science & Engineering* 9, nr 3 (2007): 90–95, <https://doi.org/10.1109/MCSE.2007.55>.

<sup>35</sup> Kaggle: *Your Machine Learning and Data Science Community*, dostęp 11 stycznia 2025, <https://www.kaggle.com>.

<sup>36</sup> Daniel S. Marcus i in., „Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults,” *Journal of Cognitive Neuroscience* 19, nr 9 (2007): 1498–1507, <https://doi.org/10.1162/jocn.2007.19.9.1498>.

Do trenowania sieci neuronowej użyto obrazów 2D jako danych wejściowych. Klasyfikacja pacjentów w bazie OASIS została przeprowadzona na podstawie dostępnych metadanych oraz wartości Clinical Dementia Rating (CDR), co pozwala wyróżnić cztery klasy:

- moderate demented – umiarkowane otępienie,
- mild demented – łagodne otępienie,
- very mild demented – bardzo łagodne otępienie,
- non-demented – brak otępienia.

Podział ten umożliwia wykrywanie i badanie różnych etapów rozwoju choroby Alzheimera. Wstępnie, dane i rozkład klas wyglądał następująco:

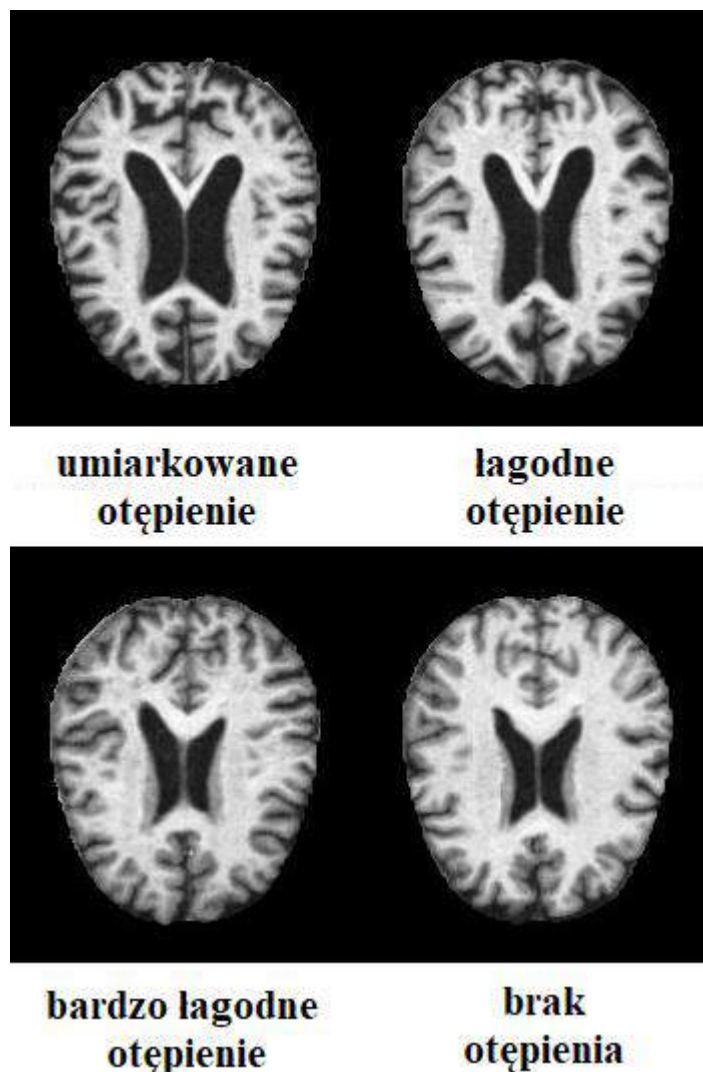
- non-demented: 3200, very mild demented: 2240, mild demented: 896, moderate demented: 64

Dzięki augmentacji, został powiększony o tyle obrazów:

- non-demented: 9600, very mild demented: 8960, mild demented: 8960, moderate demented: 6464

Całkowity rozmiar zbioru danych wynosi 382 MB, sumarycznie zawiera on 40 384 plików w formacie JPG o rozmiarach 176 x 208 pikseli (RGB).

Dane zostały podzielone w następujący sposób: 27 187 plików do zbioru treningowego, 6 797 do walidacyjnego i 6 400 do testowego. Co ważne, część testowa jest wyłącznie rzeczywistymi obrazami ze skanów mózgu po wstępnej obróbce, lecz bez augmentacji, dzięki czemu zostaną wykorzystane później w celu weryfikacji skuteczności modelu. Rozmiar obrazów został następnie przeskalowany do 224 x 224 pikseli.



Rysunek 5. Wizualizacja obrazów MRI. Źródło: opracowanie własne na podstawie obrazów użytych w pracy.

## Opis wykorzystanej architektury modelu

### Architektura i strojenie hiperparametrów

Do klasyfikacji obrazów MRI zastosowano model konwolucyjny definiowany w klasie `AlzheimersHyperModel`, co umożliwia strojenie i automatyczny dobór wybranych hiperparametrów architektury i procesu uczenia.

### Wejście i szkielet sieci

Model przyjmuje tensory poprzez warstwę wejściową. Rdzeń architektury stanowi sekwencyjna sieć z blokami konwolucyjnymi i warstwami w pełni połączonymi, zakończona klasyfikatorem softmax.

### Blok konwolucyjny (część ekstrakcji cech)

- Liczba warstw konwolucyjnych jest dobierana ze zbioru  $\{3, 4, 5, 6\}$ .

- Dla każdej warstwy konwolucyjnej dobierana jest liczba filtrów od 16 do 64 ze skokiem o 16, a rozmiar jądra jest ustalony na  $3 \times 3$  (aktywacja ReLU).
- MaxPooling2D ( $2 \times 2$ ) dodawany jest tylko po trzech ostatnich warstwach konwolucyjnych, co zapewnia stopniowe zwiększanie pola recepcyjnego bez zbyt wczesnej utraty rozdzielczości cech w pierwszych warstwach.
- Po części konwolucyjnej cechy są spłaszczane przez warstwę Flatten.

### **Część w pełni połączona (klasyfikator)**

- Liczba warstw w pełni połączonych jest dobierana z przedziału od 1 do 3.
- Rozmiar każdej takiej warstwy może wynosić 64, 96 lub 128 neuronów, z aktywacją ReLU.
- Po każdej warstwie stosowana jest regularyzacja Dropout z dobieranym współczynnikiem od 0.2 do 0.4 ze skokiem 0.1, ograniczając przeuczenie.
- Warstwa wyjściowa to warstwa w pełni połączona z funkcją aktywacji softmax, co realizuje wieloklasową klasyfikację.

### **Kompilacja i uczenie**

- Optymalizator: Adam z przestrzenią strojenia tempa uczenia (`learning_rate`) od  $1e-3$  do  $1e-4$ .
- Funkcja straty: `categorical_crossentropy`.
- Metryka: dokładność.
- `model.summary()` pozwala raportować finalnie wybraną przez tuner konfigurację.

### **Intencja doboru rozwiązań**

ReLU i rozmiar jądra  $3 \times 3$  to standard w mniej rozbudowanych konwolucyjnych sieciach neuronowych, który zapewnia stabilne uczenie i niski koszt obliczeniowy. Pooling tylko po końcowych konwolucjach pozwala na zachowanie szczegółów przestrzennych we wczesnym etapie oraz na redukcję wymiaru dopiero wtedy, gdy sieć ma bogatsze mapy cech.

Automatyczny dobór liczby warstw, filtrów, units, dropout i learning rate daje możliwość znalezienia kompromisu między pojemnością modelu a generalizacją, dopasowanego do konkretnego zbioru MRI. Natomiast dropout pomiędzy 0.2–0.4 to umiarkowana regularyzacja części w pełni połączonej, najbardziej podatnej na przeuczenie.

### **Trening modelu**

Automatyczne strojenie hiperparametrów zostało przeprowadzone z wykorzystaniem Bayesowskiej optymalizacji w Keras Tuner.

W pierwszym kroku definiuję klasę `CustomBayesianOptimization`, która dziedziczy po `kt.BayesianOptimization` i nadpisuje metodę `run_trial`. Dzięki temu każdy „trial” (pojedynczy eksperyment z konkretną kombinacją hiperparametrów) otrzymuje automatycznie callback `ModelCheckpoint` zapisujący najlepszą wersję wag (monitorowany parametr: `val_loss`) do unikalnego pliku, umożliwiając późniejszą analizę.

W początkowej fazie eksperymentów do strojenia hiperparametrów zastosowano metodę RandomSearch, która wybiera konfiguracje w sposób losowy. Po wstępnej analizie zdecydowano się jednak zastąpić ją podejściem CustomBayesianOptimization, opartym na optymalizacji bayesowskiej. Zastosowanie tej metody pozwoliło na bardziej efektywne przeszukiwanie przestrzeni hiperparametrów i przełożyło się na wzrost dokładności walidacyjnej modelu o ponad 2 punkty procentowe, w porównaniu z wynikami uzyskanymi przy użyciu Random Search.

W strojeniu hiperparametrów metodą Bayesowską nie sprawdzamy losowo możliwych kombinacji (co jest bardziej kosztowne). Zamiast tego buduje się model zastępczy (surrogate model), który aproksymuje zależność między hiperparametrami a wynikiem walidacyjnym (np. accuracy albo val\_loss). Innymi słowy, jest to uproszczony matematyczny model przewidujący, jak „dobry” będzie wynik dla danej konfiguracji.

Następnie algorytm używa tzw. funkcji akwizycji (acquisition function), która równoważy eksplorację (szukanie w nieznanach obszarach przestrzeni hiperparametrów) i eksploatację (szukanie tam, gdzie już mamy dobre wyniki).

Dzięki temu zamiast losowo testować kolejne konfiguracje (jak w Random Search), algorytm mądrze wybiera następne punkty, które z największym prawdopodobieństwem poprawią wynik. W efekcie proces strojenia hiperparametrów jest znacznie bardziej efektywny obliczeniowo niż losowe lub pełne przeszukiwanie siatki (GridSearch).

W kolejnym kroku konstruuje właściwy obiekt tunera, wskazując wcześniej zdefiniowany AlzheimersHyperModel, funkcję celu val\_accuracy, max\_trials=20 wyznaczający maksymalną liczbę różnych konfiguracji hiperparametrów oraz executions\_per\_trial=2 (każda konfiguracja uruchamiana jest dwukrotnie, a wynik uśredniany, co ogranicza wpływ losowości). Batch size został ustawiony na 32, a liczba epok wynosi 100.

Aby poprawić stabilność i efektywność treningu, wykorzystuję zestaw callbacków – odwołań do zastosowanych przeze mnie mechanizmów regulujących przebieg uczenia, omówionych w rozdziale 2.6.

Pierwszym z nich był EarlyStopping, z ustawieniem patience na 10 epok, co oznaczało, że jeśli po 10 epokach nie następuje poprawa jakości, proces uczenia zostaje zakończony.

Drugim zastosowanym mechanizmem była funkcja ReduceLROnPlateau, która dynamicznie dostosowuje tempo uczenia (learning rate). Minimalna wartość learning rate została ograniczona do  $1e-6$ .

Dodatkowo dołączam tensorboard\_callback (monitoring przebiegu treningu) oraz własny CleanUpCallback, który po zakończeniu uczenia uruchamia gc.collect() i K.clear\_session(). Ten ostatni krok jest istotny przy wielu kolejnych trialach: zwalnia pamięć GPU/CPU, bez czego istniałaby większa szansa na wystąpienie błędów „out of memory”.

Każdy trial przebiega według tych samych zasad trenowania, różniąc się jedynie architekturą i/lub hiperparametrami wybranymi przez tuner. Wraz z postępem strojenia, tuner aktualizuje swoje przekonania co do regionów przestrzeni hiperparametrów rokujących na najwyższą walidacyjną dokładność (val\_accuracy).

Łączny czas trenowania modelu wyniósł około 10 godzin.

### **Wykorzystane metryki**

Wykorzystano szereg metryk w modelu, takich jak: Accuracy, Precision (sensitivity), Recall, F1-score, Specificity, AUC. Najważniejszą z nich dla celu mojej pracy jest dokładność (accuracy).

### **Wyniki**

#### **Uzyskana architektura modelu**

Uzyskana architektura sieci po procesie uczenia składa się z pięciu warstw konwolucyjnych (Conv2D) z filtrami od 48 do 64 (pierwsze trzy warstwy mają filtr 48, następne trzy mają filtr 64). Następnie, zaczynając od 4 warstwy konwolucyjnej, przeplatane są warstwami max-poolingu w celu redukcji wymiarów. Po spłaszczeniu map cech (Flatten, 50 176 jednostek) dodano warstwę w pełni połączoną (Dense, 96 neuronów) oraz warstwę Dropout dla redukcji przeuczenia. Ostateczna warstwa wyjściowa (Dense, 4 neurony) realizuje klasyfikację wieloklasową (funkcja aktywacji softmax). Model zawiera łącznie 4 961 860 parametrów. Zdecydowaną większość stanowią parametry warstwy w pełni połączonej (ponad 4,8 milionów), co wskazuje na duże znaczenie etapu klasyfikacji w stosunku do ekstrakcji cech.

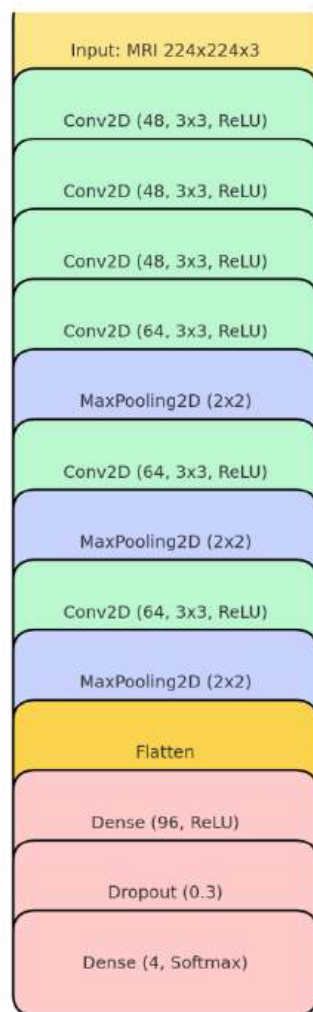
Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 224, 224, 48)	1344
conv2d_1 (Conv2D)	(None, 224, 224, 48)	20784
conv2d_2 (Conv2D)	(None, 224, 224, 48)	20784
conv2d_3 (Conv2D)	(None, 224, 224, 64)	27712
max_pooling2d (MaxPooling2D)	(None, 112, 112, 64)	0
conv2d_4 (Conv2D)	(None, 112, 112, 64)	36928
max_pooling2d_1 (MaxPooling2D)	(None, 56, 56, 64)	0
conv2d_5 (Conv2D)	(None, 56, 56, 64)	36928
max_pooling2d_2 (MaxPooling2D)	(None, 28, 28, 64)	0
flatten (Flatten)	(None, 50176)	0
dense (Dense)	(None, 96)	4816992
dropout (Dropout)	(None, 96)	0
dense_1 (Dense)	(None, 4)	388

=====  
Total params: 4,961,860  
Trainable params: 4,961,860  
Non-trainable params: 0

Rysunek 6. Struktura opracowanego modelu CNN przedstawiona w formie podsumowania warstw. Źródło: opracowanie własne.

Najlepsza architektura CNN (val\_accuracy = 0.902)



Rysunek 7. Graficzny schemat najlepszej architektury CNN opracowanego modelu. Źródło: opracowanie własne.

## Wizualizacja treningu (wykres metryk)

### Wizualizacja przebiegu uczenia z danych TensorBoard



Rysunek 8. Zmiana dokładności (accuracy) w trakcie uczenia modelu na zbiorze treningowym i walidacyjnym.  
Źródło: opracowanie własne.



Rysunek 9. Zmiana wartości funkcji straty (loss) w trakcie uczenia modelu na zbiorze treningowym i walidacyjnym.  
Źródło: opracowanie własne.

(Rysunek 8.) Przebieg zmian dokładności (accuracy) w kolejnych epokach. Na zbiorze treningowym obserwuje się systematyczny wzrost wartości metryki aż do poziomu bliskiego 1,0, co

wskazuje na bardzo dobre dopasowanie modelu do danych uczących. Krzywa dla zbioru walidacyjnego przebiega równoległe do krzywej treningowej i stabilizuje się na podobnym poziomie, co świadczy o prawidłowej generalizacji i braku wyraźnych oznak przeuczenia.

(Rysunek 9.) Przebieg zmian funkcji straty (loss) w kolejnych epokach. Na zbiorze treningowym wartość funkcji straty szybko maleje i zbliża się do zera, co potwierdza skuteczność procesu optymalizacji. Krzywa dla zbioru walidacyjnego pokazuje analogiczny trend spadkowy, przy czym wartości utrzymują się nieco powyżej krzywej treningowej, co jest spodziewane i wskazuje na poprawne uogólnienie modelu na dane niewidziane w trakcie uczenia. W okolicach ósmej epoki *val\_loss* zaczyna rosnąć, co sugeruje, że model w późniejszych epokach ulega przeuczeniu.

### Wyniki na zbiorze testowym

Tabela 1. Wyniki metryk na zbiorze testowym

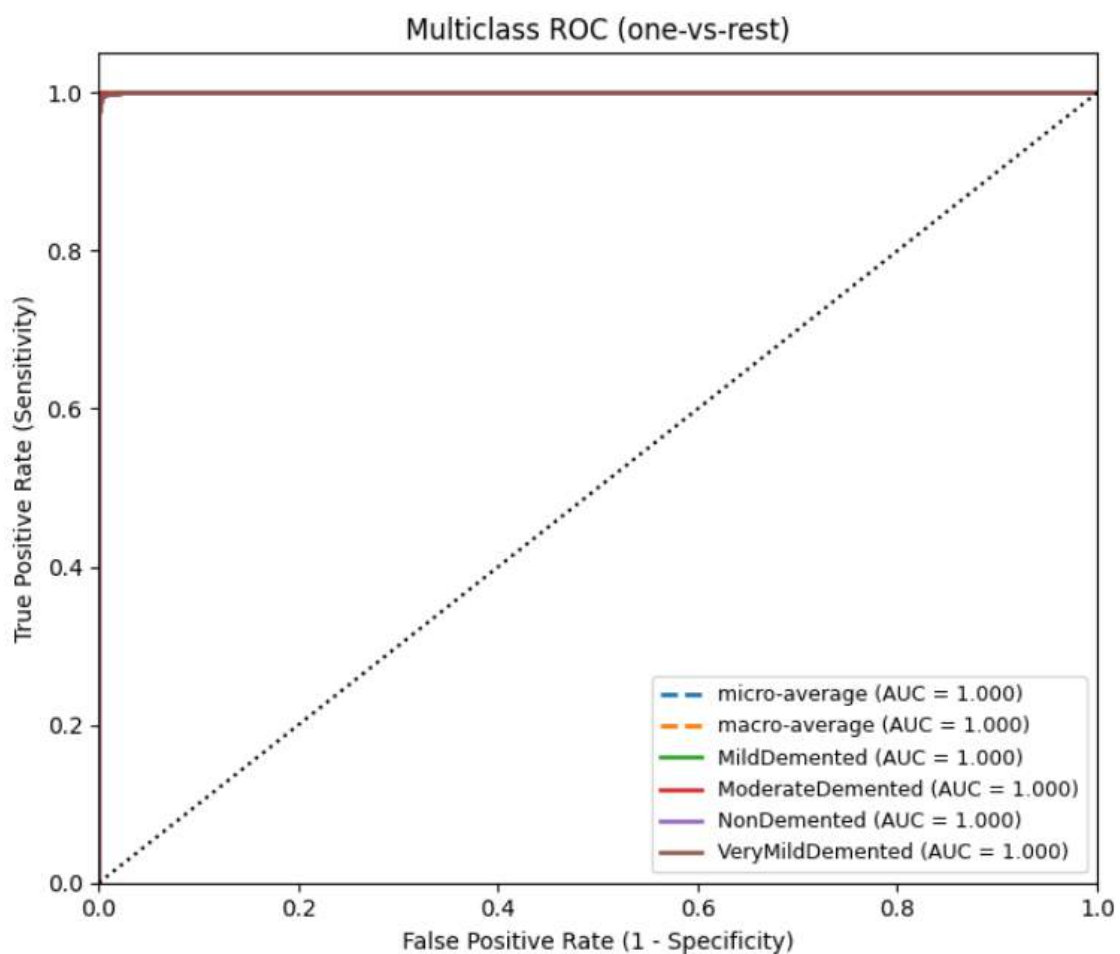
Klasa	Precyzja	Czułość	F1-score	Swoistość	AUC	Próbki	Dokładność
<b>Mild Demented</b>	100%	100%	100%	100%	100%	896	–
<b>Moderate Demented</b>	100%	100%	100%	100%	100%	64	–
<b>Non Demented</b>	99,56%	99,59%	99,58%	99,56%	99,99%	3200	–
<b>Very Mild Demented</b>	99,42%	99,38%	99,40%	99,69%	99,99%	2240	–
<b>Macro average</b>	99,75%	99,74%	99,74%	99,81%	100%	–	–
<b>Micro average</b>	99,58%	99,58%	99,58%	–	100%	–	–
<b>Weighted Average</b>	99,58%	99,58%	99,58%	–	–	–	–
–	–	–	–	–	–	–	99,58%

Źródło: opracowanie własne

Analiza metryk na zbiorze testowym potwierdza bardzo wysoką skuteczność modelu. Wynik dokładności (accuracy) wyniósł 99,58%. Klasy MildDemented oraz ModerateDemented zostały poprawnie sklasyfikowane – osiągnięto 100% precyzji, czułości, F1-score oraz swoistości. Dla klasy NonDemented wartości metryk były również bardzo wysokie (precyzja = 99,56%, czułość = 99,59%, F1 = 99,58%, swoistość = 99,56%). Klasa VeryMildDemented uzyskała bardzo zbliżone wyniki (precyzja= 99,42%, czułość = 99,38%, F1 = 99,40%, swoistość = 99,69%).

Średnie miary ogólne (Macro, Micro i Weighted average) kształtują się na poziomie ~99,6% co potwierdza równomiernie wysoką skuteczność klasyfikacji dla wszystkich klas.

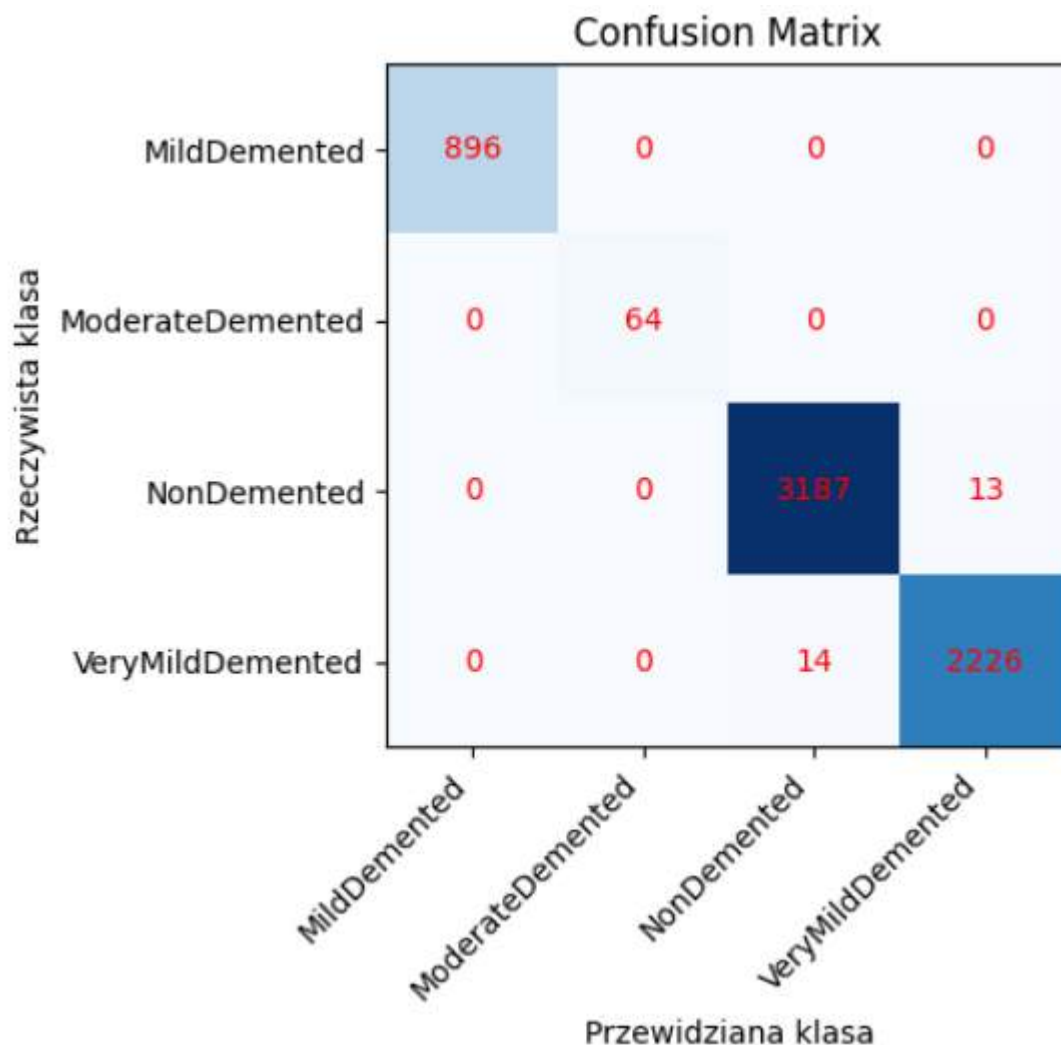
## Krzywa ROC



Rysunek 10. Krzywe ROC dla poszczególnych klas oraz uśrednione wyniki (micro i macro) wraz z wartościami AUC w opracowanym modelu. Źródło: opracowanie własne.

Rysunek 10 przedstawia wieloklasowe krzywe ROC (one-vs-rest) dla poszczególnych klas (NonDemented, VeryMildDemented, MildDemented, ModerateDemented), a także krzywe uśrednione (micro i macro). Wszystkie krzywe przebiegają bardzo blisko lewego górnego rogu wykresu, co świadczy o bardzo dokładnej zdolności modelu do rozróżniania klas. Uzyskane wartości AUC wynoszą 100% dla każdej klasy oraz dla uśrednień, co oznacza, że model osiąga maksymalną skuteczność w zakresie kompromisu pomiędzy czułością a swoistością.

## Macierz Pomyłek



Rysunek 11. Macierz pomyłek przedstawiająca wyniki klasyfikacji modelu dla czterech klas nasilenia otępienia. Źródło: opracowanie własne.

Na podstawie macierzy pomyłek można zauważyć, że model bardzo dobrze radzi sobie z rozpoznawaniem wszystkich klas. Klasa MildDemented została poprawnie sklasyfikowana we wszystkich 896 przypadkach, bez żadnych błędnych przypisań. Podobnie klasa ModerateDemented uzyskała 100% skuteczność (64 poprawne predykcje). W przypadku klasy NonDemented poprawnie rozpoznano 3187 obrazów, natomiast tylko 13 przypadków błędnie zakwalifikowano jako VeryMildDemented. Klasa VeryMildDemented została poprawnie przewidziana w 2226 przypadkach, przy zaledwie 14 błędach (nieprawidłowe przypisanie do NonDemented).

Łącznie oznacza to bardzo wysoką skuteczność klasyfikacji stadiów otępień, z minimalną liczbą błędnych klasyfikacji, które występują wyłącznie pomiędzy klasami NonDemented i VeryMildDemented, co wskazuje na największe podobieństwo cech obrazowych właśnie pomiędzy tymi grupami.

## Porównanie wyników z literaturą

W części porównawczej wyników uwzględniono wyłącznie prace, które – podobnie jak niniejsze badanie – korzystały ze zbioru danych OASIS. Zdecydowano się pominąć publikacje oparte na bazie ADNI, ponieważ różnice w sposobie definiowania i grupowania klas diagnostycznych mogłyby utrudnić rzetelne zestawienie rezultatów.

Tabela 2. Porównanie metryk z literaturą.

<b>Model / Autorzy</b>	<b>Rozpoznawane klasy</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>	<b>F1-score</b>	<b>AUC</b>
<b>Mój model (CNN)</b>	Non Demented, VeryMild, Mild, Moderate	99,58%	99,58% (weighted)	99,58% (weighted)	99,58% (weighted)	100% (micro/macro)
<b>Fulton i inni, ResNet-50</b>	3 klasy: Non Demented, Very Mild, Mild/Moderate	98,99%	–	–	–	~96%
<b>Naganandhini &amp; Shanmugavadi, DTC-HPT</b>	Non Demented, Very Mild, Mild, Moderate	99,10%	98% (weighted)	99% (weighted)	99% (weighted)	–

Źródło: opracowanie własne.

Opracowany autorski model uzyskał najwyższą dokładność klasyfikacji (99,58%). Metryki takie jak Precision, Recall i F1-score osiągnęły zbliżone wartości na poziomie ~99,6% (liczone jako średnie ważone), a AUC zarówno w ujęciu mikro, jak i makro, wyniosło 100% (Tabela 2). Świadczy to o bardzo dokładnym rozróżnianiu poszczególnych grup pacjentów.

W pracy Naganandhini i Shanmugavadi<sup>37</sup>, w której wykorzystano klasyfikator drzewa decyzyjnego z dostrajaniem hiperparametrów (DTC-HPT), uzyskano dokładność na poziomie 99,10%. Metryki ważone (Precision = 98%, Recall = 99%, F1-score = 99%) były również bardzo wysokie. Jednak co do wykrywania poszczególnych klas, w klasie ModerateDemented przy pojedynczej próbce klasyfikator nie był w stanie poprawnie sklasyfikować przypadku (Precision i Recall = 0). (tabela x.x)

<sup>37</sup> Naganandhini i Shanmugavadi, „Effective Diagnosis of Alzheimer’s Disease Using Modified Decision Tree Classifier.”

Model ResNet-50 zastosowany przez Fultona<sup>38</sup> uzyskał dokładność 98,99%, lecz autorzy połączyli klasy, w efekcie uzyskując podział na trzy: NonDemented, Very Mild Demented, Mild/Moderate Demented. Raportowali również AUC powyżej 96%, lecz niestety nie uwzględnili szczegółowych metryk, takich jak Precision, Recall czy F1-score. Architektura ta składała się z 25,4 milionów parametrów, czyli aż pięciokrotnie więcej, w porównaniu do mojego modelu, składającego się z niecałych ~5 milionów.

Tabela 3. Porównanie metryk z modelem DTC-HPT.

Klasa	Próbki (mój model)	Precyzja (mój model)	Czułość (mój model)	F1-score (mój model)	Próbki (DTC)	Precyzja (DTC)	Czułość (DTC)	F1-score (DTC)
<b>Non Demented</b>	3200	99,56%	99,59%	99,58%	63	100%	100%	100%
<b>Very Mild AD</b>	2240	99,42%	99,38%	99,40%	40	100%	100%	100%
<b>Mild AD</b>	896	100%	100%	100%	8	89%	100%	94%
<b>Moderate AD</b>	64	100%	100%	100%	1	0%	0%	0%
<b>Macro average</b>	—	99,75%	99,74%	99,74%	—	72%	75%	74%
<b>Micro average</b>	—	99,58%	99,58%	99,58%	—	—	—	—

Źródło: opracowanie własne.

W przypadku modelu DTC-HPT metryki dla klas *Non-demented* oraz *Very Mild AD* osiągnęły wartości 100% (Precyzja, Czułość, F1-score), przy liczbie próbek odpowiednio: 63 i 40. Dla klasy *Mild AD* wyniki były nieco słabsze (Precyzja = 89%; F1-score = 94%), a w przypadku *Moderate AD* model całkowicie zawiódł (Precyzja, Czułość, F1-score = 0%) z powodu ekstremalnie małej liczności próbki (tylko 1 przypadek). W konsekwencji średnie zbalansowane (macro) osiągnęły znacznie niższe wartości (Precyzja = 72%; Czułość = 75%; F1-score = 74%), co pokazuje, że dostępna liczba przykładów w rozkładzie klas miała negatywny wpływ na wydajność modelu.

W przeciwieństwie do DTC-HPT, autorski model CNN uzyskał wysokie wyniki dla wszystkich klas, w tym także dla rzadkiej kategorii *ModerateDemented* (Precyzja, Czułość, F1-score = 100% przy 64 próbkach). Wartości metryk pozostały na poziomie ~99–100% niezależnie od liczności próbek, co przełożyło się również na znacznie wyższe średnie zbalansowane (macro) w porównaniu do modelu DTC.

<sup>38</sup> Fulton i in., „Classification of Alzheimer’s Disease with and without Imagery.”

## Wnioski

Opracowany model osiąga bardzo wysoki wynik dokładności (99,58%) dla klasyfikacji wieloklasowej. Wykorzystany szeroki wachlarz metryk, takich jak: dokładność, precyzja, czułość, F1-score, swoistość, AUC oraz krzywa ROC zapewnia kluczowe informacje do dokładnej oceny jakości modelu. Wszystkie miary wskazują na bardzo stabilne działanie, który skutecznie rozróżnia wszystkie cztery badane klasy, w tym także trudniejsze przypadki *Moderate AD*.

W zestawieniu z większymi sieciami głębokiego uczenia, takimi jak ResNet-50, a także z klasyfikatorem drzewa decyzyjnego z dostrajaniem hiperparametrów (DTC-HPT), opracowana architektura okazuje się nie tylko konkurencyjna jakościowo, ale również bardziej ekonomiczna obliczeniowo. W przeciwieństwie do ResNet-50 (dokładność = 98,99%), który posiada 25,4 milionów parametrów i wymaga większej mocy obliczeniowej do treningu, mój model zachowuje wyższą skuteczność przy pięciokrotnie mniejszej złożoności (~5 milionów parametrów).

Porównanie z DTC-HPT wskazuje, że choć dokładność tego modelu również była wysoka (99,10%), to jego skuteczność była silnie zależna od rozkładu klas. Przy pojedynczych próbkach w klasie *Moderate AD* DTC całkowicie zawiódł (Precision, Recall, F1-score = 0%), co skutkowało znacznym obniżeniem średnich zbalansowanych metryk (macro). Mój model, w przeciwieństwie do tego podejścia, utrzymał wysoką skuteczność niezależnie od liczności próbek w poszczególnych klasach, osiągając niemal 100% wyniki we wszystkich kategoriach.

## Podsumowanie

Podsumowując, uzyskane rezultaty wskazują, że opracowany przeze mnie model CNN stanowi wartościową i skuteczną alternatywę zarówno dla rozbudowanych architektur głębokiego uczenia, jak i klasycznych metod uczenia maszynowego. Kluczowym atutem mojej architektury jest wysoka dokładność w każdej z badanych klas, przy stosunkowo lekkim rozmiarze modelu (5 milionów parametrów), przez co obliczenia mogły zostać wykonane na sprzęcie konsumenckim. Dokładność, czyli najważniejsza metryka, którą wyznaczyłem we wstępie jako jeden z głównych celów mojej pracy, osiągnęła najwyższy wynik (99,58%) spośród porównywanych rozwiązań. Potwierdza to zasadność i użyteczność wykorzystania mojego modelu w zadaniach klasyfikacji choroby Alzheimera na podstawie obrazów MRI.

## Ograniczenia

Praca miała kilka istotnych ograniczeń, pomimo obiecujących rezultatów.

Po pierwsze, zastosowany został zbiór danych OASIS, podczas gdy większość badań porównawczych korzysta z bazy ADNI<sup>3940</sup>, co utrudnia bezpośrednie zestawienie wyników. Dodatkowo, niektóre porównywane prace bazowały na niewielkich i niezbalansowanych zbiorach

---

<sup>39</sup> Helaly, Badawy i Haikal, „Deep Learning Approach for Early Detection of Alzheimer’s Disease.”, s. 1719-1720.

<sup>40</sup> Gopi Battineni i in., „Artificial Intelligence Models in the Diagnosis of Late-Onset Dementias: A Systematic Review,” *Bioengineering* 9, nr 8 (2022): 370, <https://doi.org/10.3390/bioengineering9080370>.

danych, co mogło wpływać na jakość raportowanych wyników i utrudniać rzetelną ocenę w odniesieniu do uzyskanych rezultatów.

Kolejnym ograniczeniem był fakt, że wiele publikacji nie raportowało pełnego zestawu metryk, przez co porównania mogły obejmować jedynie wybrane wskaźniki.

Dodatkowo, w pracy wykorzystano wyłącznie obrazy MRI w formie 2D, pomijając inne dane udostępniane w ramach bazy OASIS, takie jak wiek, płeć, poziom wykształcenia, status społeczno-ekonomiczny, wyniki testów kognitywnych (MMSE) czy miary objętości mózgu (eTIV, nWBV). Ograniczyło to analizę jedynie do aspektu obrazowego, bez uwzględnienia czynników klinicznych i demograficznych, które mogłyby dodatkowo zwiększyć skuteczność predykcji.

### Kierunki dalszych badań

W przyszłych badaniach warto rozważyć rozszerzenie analiz o zbiór danych ADNI, który jest częściej wykorzystywany w literaturze, co umożliwiłoby bezpośrednie porównania z większą liczbą istniejących prac. Interesującym kierunkiem mogłoby być także połączenie danych z baz OASIS oraz ADNI, co pozwoliłoby na stworzenie bardziej zróżnicowanego i obszernego zbioru, a tym samym poprawę zdolności uogólniania modelu.

Istotnym aspektem jest również fakt, że zarówno OASIS, jak i ADNI udostępniają oprócz obrazów MRI także dane kliniczne, demograficzne, genetyczne, obrazy PET oraz biomarkery płynów ustrojowych. Wykorzystanie takich różnorodnych danych pozwoliłoby w przyszłości na tworzenie modeli multimodalnych, które lepiej odzwierciedlałyby rzeczywiste warunki diagnostyczne i mogłyby poprawić skuteczność predykcji.

### Bibliografia:

- Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Mark Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, ... Xiaoqiang Zheng. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Oprogramowanie dostępne na: <https://www.tensorflow.org>.
- Agarap, Abien Fred. „Deep Learning Using Rectified Linear Units (ReLU).” *arXiv preprint* arXiv:1803.08375 (2018). <http://arxiv.org/abs/1803.08375>.
- Barkhof, Frederik, Marieke Hazewinkel, Max Binnewijzend i Robin Smithuis. „Dementia – Role of MRI.” *Radiology Assistant*, 2012/2022. <https://radiologyassistant.nl/neuroradiology/dementia/role-of-mri>.
- Battineni, Gopi, Nandha Chintalapudi, Mohammed A. Hossain, Giuseppe Losco, Caterina Ruocco, Giovanni G. Sagaro, Enea Traini, Giulia Nittari i Francesco Amenta. „Artificial Intelligence Models in the Diagnosis of Late-Onset Dementias: A Systematic Review.” *Bioengineering* 9, nr 8 (2022): 370. <https://doi.org/10.3390/bioengineering9080370>.
- Bishop, Christopher M. *Pattern Recognition and Machine Learning*. Nowy Jork: Springer, 2006.
- Bridle, John. „Training Stochastic Model Recognition Algorithms as Networks Can Lead to Maximum Mutual Information Estimation of Parameters.” W: *Advances in Neural Information Processing Systems*, t. 2, red. D. Touretzky. Morgan-Kaufmann, 1989. <https://proceedings.neurips.cc/paper/1989/file/0336dcbab05b9d5ad24f4333c7658a0e-Paper.pdf>.

- Centrum e-Zdrowia. „11 sygnałów, które wskazują na Alzheimera.” *Pacjent.gov.pl – Serwis Rzeczypospolitej Polskiej*, 21 września 2022. <https://pacjent.gov.pl/aktualnosc/11-sygnalow-ktore-wskazuja-na-alzheimera>.
- Chollet, François i in. *Keras*. Oprogramowanie dostępne na: <https://keras.io>.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee i Kristina Toutanova. „BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.” W: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, s. 4171–4186. Association for Computational Linguistics, 2019. <https://arxiv.org/abs/1810.04805>.
- Fulton, Larry V., David Dolezel, Jason Harrop, Yan Yan i Cathy P. Fulton. „Classification of Alzheimer’s Disease with and without Imagery Using Gradient Boosted Machines and ResNet-50.” *Brain Sciences* 9, nr 9 (2019): 212. <https://doi.org/10.3390/brainsci9090212>.
- Goodfellow, Ian, Yoshua Bengio i Aaron Courville. *Deep Learning*. Cambridge, MA: MIT Press, 2016. <http://www.deeplearningbook.org>.
- Harris, Charles R., K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, ... Travis E. Oliphant. „Array Programming with NumPy.” *Nature* 585, nr 7825 (2020): 357–362. <https://doi.org/10.1038/s41586-020-2649-2>.
- Hazarika, Rajarshi A., Debasish Kandar i Amit K. Maji. „An Experimental Analysis of Different Deep Learning Based Models for Alzheimer’s Disease Classification Using Brain Magnetic Resonance Images.” *Journal of King Saud University – Computer and Information Sciences* 34, nr 10 (2022): 8576–8598. <https://doi.org/10.1016/j.jksuci.2021.09.003>.
- Helaly, Hossam A., Mohamed Badawy i Ahmed Y. Haikal. „Deep Learning Approach for Early Detection of Alzheimer’s Disease.” *Cognitive Computation* 14 (2022): 1711–1727. <https://doi.org/10.1007/s12559-021-09946-2>.
- Hinton, Geoffrey, Li Deng, Dong Yu, George Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath i Brian Kingsbury. „Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups.” *IEEE Signal Processing Magazine* 29, nr 6 (2012): 82–97. <https://doi.org/10.1109/MSP.2012.2205597>.
- Hunter, John D. „Matplotlib: A 2D Graphics Environment.” *Computing in Science & Engineering* 9, nr 3 (2007): 90–95. <https://doi.org/10.1109/MCSE.2007.55>.
- Kaggle. „Kaggle: Your Machine Learning and Data Science Community.” Dostęp 11 stycznia 2025. <https://www.kaggle.com>.
- Krizhevsky, Alex, Ilya Sutskever i Geoffrey E. Hinton. „ImageNet Classification with Deep Convolutional Neural Networks.” *Communications of the ACM* 60, nr 6 (2012): 84–90. <https://doi.org/10.1145/3065386>.
- Laraba, Sihem, Julien Tilmanne i Thierry Dutoit. *Deep Learning for Skeleton-Based Human Action Recognition*. Praca doktorska, Uniwersytet Mons, 2020. [https://www.researchgate.net/publication/344876266\\_Deep\\_Learning\\_for\\_Skeleton-Based\\_Human\\_Action\\_Recognition](https://www.researchgate.net/publication/344876266_Deep_Learning_for_Skeleton-Based_Human_Action_Recognition).
- LeCun, Yann, Léon Bottou, Yoshua Bengio i Patrick Haffner. „Object Recognition with Gradient-Based Learning.” W: *Shape, Contour and Grouping in Computer Vision*, red. David A. Forsyth, Joseph L. Mundy, Vincenzo Di Gesù i Roberto Cipolla, s. 319–345. Springer, 1999. [https://doi.org/10.1007/3-540-46805-6\\_19](https://doi.org/10.1007/3-540-46805-6_19).
- Li, Zhiyong, William T. Nash, Sean P. O’Brien, Yifei Qiu, Ritesh K. Gupta i Nick Birbilis. „cardiGAN: A Generative Adversarial Network Model for Design and Discovery of Multi-Principal Element Alloys.” *arXiv preprint arXiv:2202.00966* (2022). <https://doi.org/10.48550/arXiv.2202.00966>.
- Litjens, Geert, Thijs Kooi, Babak E. Bejnordi, Arnaud A. A. Setio, Francesco Ciompi, Mitko Ghahfoorian, J. A. W. M. van der Laak, Bram van Ginneken i Clara I. Sánchez. „A Survey on Deep Learning in Medical Image Analysis.” *Medical Image Analysis* 42 (2017): 60–88. <https://doi.org/10.1016/j.media.2017.07.005>.
- Marcisz, Aleksandra i Joanna Polańska. „Can T1-Weighted Magnetic Resonance Imaging Significantly Improve Mini-Mental State Examination-Based Distinguishing Between Mild Cognitive Impairment and Early-Stage Alzheimer’s Disease?” *Journal of Alzheimer’s Disease* 92, nr 3 (2023): 941–957. <https://doi.org/10.3233/JAD-220806>.
- Marcus, Daniel S., Tracy H. Wang, Jason Parker, John G. Csernansky, John C. Morris i Randy L. Buckner. „Open Access Series of Imaging Studies (OASIS): Cross-Sectional MRI Data in Young, Middle Aged, Nondemented,

- and Demented Older Adults.” *Journal of Cognitive Neuroscience* 19, nr 9 (2007): 1498–1507. <https://doi.org/10.1162/jocn.2007.19.9.1498>.
- McKinney, Wes. „Data Structures for Statistical Computing in Python.” W: *Proceedings of the 9th Python in Science Conference*, s. 51–56, 2010. <https://doi.org/10.25080/Majora-92bf1922-00a>.
  - Medscape. „Alzheimer Disease Imaging: Practice Essentials, Computed Tomography, Magnetic Resonance Imaging.” *Medscape*, 28 grudnia 2022. <https://emedicine.medscape.com/article/336281-overview>.
  - Naganandhini, S. i P. Shanmugavadivu. „Effective Diagnosis of Alzheimer’s Disease Using Modified Decision Tree Classifier.” *Procedia Computer Science* 165 (2019): 548–555. <https://doi.org/10.1016/j.procs.2020.01.049>.
  - Najwyższa Izba Kontroli. „NIK: o opiece nad chorymi na Alzheimera.” 16 maja 2017. <https://www.nik.gov.pl/najnowsze-informacje-o-wynikach-kontroli/nik-o-opiece-nad-chorymi-na-alzheimera.html>.
  - Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot i Édouard Duchesnay. „Scikit-Learn: Machine Learning in Python.” *Journal of Machine Learning Research* 12 (2011): 2825–2830. <http://jmlr.org/papers/v12/pedregosa11a.html>.
  - Romero, Jonathan, Robert Babbush, Jarrod R. McClean, Christian Hempel, Peter J. Love i Alán Aspuru-Guzik. „Strategies for Quantum Computing Molecular Energies Using the Unitary Coupled Cluster Ansatz.” *Quantum Science and Technology* 4, nr 1 (2018): 014008. <https://doi.org/10.1088/2058-9565/aad3e4>.
  - Rzecznik Praw Obywatelskich. *Sytuacja osób chorych na chorobę Alzheimera*. Wyd. 2. Warszawa: Biuro RPO, 2016. <https://www.rpo.gov.pl/pl/content/sytuacja-osob-chorych-na-chorobe-alzheimera>.
  - Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever i Ruslan Salakhutdinov. „Dropout: A Simple Way to Prevent Neural Networks from Overfitting.” *Journal of Machine Learning Research* 15 (2014): 1929–1958. <http://jmlr.org/papers/v15/srivastava14a.html>.
  - Tadeusiewicz, Ryszard i Mariusz Szaleniec. *Leksykon sieci neuronowych*. Wrocław: Projekt Nauka, Fundacja na rzecz Promocji Nauki Polskiej, 2015.
  - Vemuri, Prashanthi i Clifford R. Jack Jr. „Role of Structural MRI in Alzheimer’s Disease.” *Alzheimer’s Research & Therapy* 2, nr 4 (2010): 23. <https://doi.org/10.1186/alzrt47>.
  - Yamashita, Ryosuke, Masahiro Nishio, Richard K. G. Do i Kuni O. Togashi. „Convolutional Neural Networks: An Overview and Application in Radiology.” *Insights into Imaging* 9, nr 4 (2018): 611–629. <https://doi.org/10.1007/s13244-018-0639-9>.

Ewelina Bąk

## **Badanie pilotażowe semantycznej załamki N400 w zdaniach zakończonych homofonami w postaci ikon.**

### **Abstrakt**

Potencjał semantyczny N400 należy do potencjałów wywołanych (ang. Event – Related Potentials – ERPs), powstaje w odpowiedzi na bodziec językowy i charakteryzuje się wrażliwością na procesy znaczeniowe (semantyczne), występujące w trakcie analizy zdań. Najczęściej aktywowany jest w sytuacji, gdy bodziec rejestrowany przez mózg nie zgadza się z bodźcem przez niego oczekiwanym. W ramach badania obserwowano natężenie potencjału semantycznego w odpowiedzi na zdania zakończone homofonami - słowami tak samo brzmiącymi, lecz różniącymi się pisownią, pochodzeniem i znaczeniem - które przedstawiono w formie ikon. Grupę badawczą stanowili studenci posiadający podstawową znajomość języka angielskiego - który został wybrany ze względu na powszechność oraz ilość homofonów w słowniku. Każde zdanie prezentowane uczestnikom miało trzy warianty, w tym: próbę kontrolną, badawczą i zdanie bazowe. Aktywność elektryczną mózgu rejestrowano przy pomocy EEG. Praca zawiera wstęp teoretyczny, szczegółowy przebieg badania, analizę oraz wnioski.

**Słowa kluczowe:** EEG, potencjał wywołany (ERP), N400, homofony, ikony

### **Abstract**

The N400 semantic potential belongs to event-related potentials (ERPs). It arises in response to a linguistic stimulus and is characterized by sensitivity to semantic processing during sentence analysis. It most commonly occurs when the stimulus registered by the brain does not match the expected stimulus. This study examined the intensity of N400 semantic potential occurrence in sentences ending with homophones - words that sound the same but differ in spelling, origin, and meaning - represented in the form of icons. The research group consisted of students with basic English language skills. Each sentence presented to participants had three variations, including: control sample, research sample and basic sentence. All was displayed one word at a time. Brain activity was recorded using EEG. This paper includes a theoretical introduction, a detailed description of the study procedure, an analysis of the results, and the conclusions drawn from the research.

**Key words:** EEG, event-related potential (ERP), N400, homophones, icons

## Wstęp

Badanie zostało przeprowadzone w ramach pracy dyplomowej, realizowanej w formie projektu licencjackiego. Eksperyment z użyciem sprzętu do analizy sygnału bioelektrycznego (EEG actiCHamp), udostępnionego przez Laboratorium Kognitywistyczne Uniwersytetu Śląskiego przy ulicy Bankowej 14 w Katowicach, odbywał się pod nadzorem dr. Witolda Marzędy - adiunkta Uniwersytetu Śląskiego w Katowicach. Ze względu na krótki okres trwania projektu (tj. 1 semestr) oraz czasochłonność przeprowadzania badań z użyciem Elektroencefalografu (EEG), wykonano jedynie pilot eksperymentu. Główny cel projektu stanowiło zbadanie natężenia semantycznej załamki N400 w zdaniach zakończonych homofonami w postaci obrazkowej. Analiza zdań jest obiektem badań zarówno lingwistów jak i naukowców z dziedziny kognitywistyki, których interdyscyplinarne spojrzenie zgłębia rozumienie tego – specyficznie ludzkiego - procesu. Aktywność bioelektryczna mózgu, zainicjowana przez swoisty bodziec językowy, zauważalny już na poziomie podświadomym, może stanowić istotny element w rozumieniu mechanizmów analizy zdań, podprogowej percepcji bodźców i jej daleko idących skutków. Jednak, żeby zrozumieć wyjątkowość i specyfikę semantycznego potencjału N400, należy przybliżyć definicje podstawowych terminów używanych do opisywania aktywności bioelektrycznej mózgu.

Potencjały wywołane można zdefiniować jako swoiste bioelektryczne odpowiedzi mózgu na prezentowane bodźce, które mogą pochodzić z różnych modalności<sup>1</sup>. Nazwy poszczególnych potencjałów związane są z ich ładunkiem elektrycznym, gdzie N - Negative (pl. ujemny), P - Positive (pl. dodatni), oraz czasem występowania w milisekundach (ms).

Potencjały egzogenne; występujące do 100 ms po wprowadzeniu bodźca, związane z fizycznymi aspektami dochodzącej do mózgu informacji, nazywa się również potencjałami sensorycznymi, np. P50 - jest dodatnią aktywnością rejestrowaną do około 50 ms po wprowadzeniu bodźca, związaną z tzw. bramkowaniem sensorycznym – wczesnym filtrowaniem dochodzącej do mózgu informacji. Natomiast N100 osiągająca szczyt między 90 a 200 ms jest negatywnym odchyleniem aktywności elektrycznej mózgu występującym w trakcie procesu dopasowywania prezentowanego bodźca do bodźców wcześniej doświadczanych. Mózg „orientuje” się czy bodźce prezentowane są ze sobą powiązane. Potencjały te występują zawsze, niezależnie od specyfiki bodźca i są naturalną odpowiedzią mózgu na zmianę w otoczeniu, która wywołuje „zagęszczenie” uwagi i w konsekwencji skierowanie jej na dany sygnał. Potencjały egzogenne pełnią kluczową rolę w badaniach nad schizofrenią i innymi chorobami związanymi z deficytami funkcji poznawczych<sup>2</sup>.

Potencjały endogenne; nazywane również „poznawczymi”, klasyfikuje się do nich potencjały obserwowane powyżej 100 ms po wprowadzeniu bodźca, są związane z oceną i przetwarzaniem odebranej informacji. P200 lub P2 - osiąga szczyt między 150-250 ms po wprowadzeniu bodźca i jest związana z wczesnym skupianiem uwagi na bodźcu oraz szybkością jego identyfikacji. Potencjał ten jest często obserwowany w momencie wprowadzenia nowego typu bodźca do stałej powtarzalnej serii impulsów, np. w zadaniach typu oddball, które polegają na pokazywaniu dwóch rodzajów bodźców *standard* (często prezentowanego bodźca bazowego) oraz *target* (występującego rzadziej,

---

<sup>1</sup> Małgorzata Ganc, „Study system of late cortical responses”, *Nowa Audiofonologia* 3 (2020): 69-75.

<sup>2</sup> Marijn Lijffijt, Scott Lane D., Stacey Meier L. i in., „P50, N100, and P200 sensory gating: Relationships with behavioral inhibition, attention, and working memory”, *Psychophysiology* 5 (2009): 1059-1068.

zazwyczaj testowego bodźca). *Target* wywołuje porządane potencjały ERP<sup>3</sup>. Potencjał N200 lub N2 osiąga szczyt po około 200 ms po prezentacji, wiąże się z identyfikacją i rozróżnianiem bodźców, w momencie, gdy uwaga jest na nich świadomie skupiona. Jest rejestrowany w zadaniach typu go/no-go, które wymagają kontroli reakcji hamującej<sup>4</sup>.

P300 również występuje w zadaniach typu oddball, osiąga szczyt aktywności około 250-500 ms od momentu wprowadzenia<sup>5</sup> i jest związany z wewnętrzną aktualizacją kontekstu. Natomiast, negatywny odpowiednik tego potencjału - N300 - występuje przy bodźcach obrazowych, około 250-350 ms po ich wprowadzeniu<sup>6</sup>, w momencie dopasowywania danych wizualnych do przechowywanych danych semantycznych. Odzwierciedla proces wstępnego rozpoznania kategorii obrazu, niezależnie od jego pełnej integracji semantycznej.

N400 jest negatywną załamką, występującą około 400 ms (latencja od około 300 - 600 ms) po prezentacji bodźca<sup>7</sup>. Obserwuje się ją w odpowiedzi na semantyczną niezgodność słowa w zdaniu. Najczęściej jest rejestrowana na jego końcu, gdzie występuje również w zdaniach semantycznie poprawnych (w postaci niewielkiej amplitudy).

P600 może występować jako późniejsza reakcja na semantycznie niezgodne zdanie. W wielu badaniach nazywana potencjałem syntaktycznym, sygnalizującym nieprawidłowość składni w zdaniu, jednak może też stanowić odpowiedź na utrudniające przetwarzanie błędy gramatyczne<sup>8</sup>.

Późne pozytywne potencjały (ang. Late Positive Components - LPC), jak i FN400 (negatywna aktywność obserwowana w płatach czołowych), osiągają szczyt między 400 a 800 ms po wprowadzeniu bodźca - zazwyczaj tzw. sondy pamięci - i są związane z procesem przypominania informacji oraz rozpoznawania wcześniej widzianych elementów. Amplituda LPC rośnie wraz z siłą śladu pamięciowego i dokładnością decyzji<sup>9</sup>.

Oprócz powyższej klasyfikacji, kluczowym do zrozumienia badania terminem jest homofonia. Posługując się definicją Słownika Języka Polskiego PWN: Homofonia to identyczność fonetyczna wyrazów różnych pod względem pisowni i znaczenia.<sup>10</sup> . Wobec tego, homofonami będą wyrazy brzmiące tak samo, jednak posiadające inne znaczenie i/bądź pisownię. Za przykład można podać parę wyrazów „lud” i „lód”, przy wymowie których nie istnieje zauważalna różnica, natomiast w pisowni oraz pod względem semantyki wyrazy te nie są zbieżne.

---

<sup>3</sup> Takashi Furutsuka, „Effects of Rapid Attention Switching on the NI-P2 Amplitude of the Visual Event-Related Potentials”, *Research and Clinical Center for Child Development: Annual Report 11* (1989): 55-64.

<sup>4</sup> Jonathan Folstein, Cyma Van Petten, „Influence of cognitive control and mismatch on the N2 component of the ERP: A review”, *Psychophysiology* 1 (2008): 152-170.

<sup>5</sup> John Polich, Jose R Craido, „Neuropsychology and neuropharmacology of P3a and P3b”, *International Journal of Psychophysiology* 2 (2006): 172-185.

<sup>6</sup> Kara D. Federmeier, Kutas Maria, „Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP)”, *Annual Review of Psychology* 1 (2011): 621-647.

<sup>7</sup> Kutas Maria, Hillyard Steven, „Reading Senseless Sentences: Brain Potentials Reflects Semantics Incongruity”, *Science* 207 (1980): 203-205.

<sup>8</sup> Osterhout Lee, McLaughout Judith, Bersick Michael, „Event-related brain potentials and human language”, *Trends in Cognitive Sciences* 6 (1997): 203-209.

<sup>9</sup> Finnigan Simon, „ERP 'old/new' effects: memory strength and decisional factor(s)”, *Neuropsychologia* 13 (2002): 2288-2304.

<sup>10</sup> Słownik Języka Polskiego PWN, dostęp 23.06.2025, [sjp.pwn.pl/slowniki/homofony](http://sjp.pwn.pl/slowniki/homofony)

Badanie przeprowadzone w ramach tejże pracy jest zmodyfikowaną replikacją badania Marty Kutas i Stevena A. Hillyard z 1980 r. W tym, już klasycznym badaniu, po raz pierwszy zaobserwowano i opisano komponent N400. Przeprowadzono trzy wersje tego eksperymentu, w każdym z nich wyświetlone zostało 160 zdań, każde składało się z 7 słów i było prezentowane po jednym słowie na raz. Wyświetlane przez 100 ms, z częstotliwością jedno słowo na sekundę. Każde zdanie poprzedzał slajd, na którym widniał ciąg pięciu liter „X” (XXXXX). W dwóch pierwszych wersjach eksperymentu 25% zdań kończyło się semantycznie niezgodnym wyrazem, którego siła była modyfikowana.

Przykład: „He took a sip from the waterfall” (tł. „Wziął łyk z wodospadu.”) w eksperymencie nr 1 reprezentuje umiarkowaną niezgodność, „He took a sip from the transmitter.” (tł. „Wziął łyk z nadajnika”) w eksperymencie nr 2 - tutaj niezgodność jest silniejsza. Natomiast w trzeciej wersji eksperymentu 25% słów znajdujących się na końcu zdania nie wykazywały niezgodności semantycznej, lecz zostały powiększone.

Łącznie przebadano 32 osoby w wieku między 18 a 35 rokiem życia. W trzech wariantach eksperymentu brało udział, kolejno 11, 12 i 9 uczestników. Nie przedstawiono uprzednio przedmiotu badań, by odpowiedź neuronalna była w jak największym stopniu spontaniczna, a załamek mogła zaistnieć. Aktywność elektryczną mózgu rejestrowano przy pomocy Elektroencefalografu (EEG). Użyto montażu, składającego się ze stosunkowo niewielkiej liczby elektrod o lokalizacjach: frontalnej/czołowej (Fz), centralnej/środkowej (Cz), ciemieniowej (Pz). Elektrody stanowiące referencją znajdowały się na wyrostkach sutkowych, band-pass ustawiono na 0 - 40 Hz. Artefakty wynikające z ruchu gałek ocznych i powiek zostały odrzucone z ostatecznego wyniku, między innymi dzięki elektrodom pod prawym okiem. W wynikach zaobserwowano charakterystyczne negatywne załamanie, o latencji od 250 – 400 ms, którego amplituda zwiększała się wraz z siłą niezgodności słowa występującego na końcu zdania.

Na przestrzeni lat przeprowadzono już wiele badań nad komponentem N400. Między innymi modyfikowano je pod względem kanałów doprowadzających bodźce, łączono wyświetlanie obrazów z odtwarzaniem haseł i obserwowano jak niezgodność bodźców z różnych modalności wpływa na badany potencjał<sup>11</sup>. Sprawdzano również natężenie N400 u dzieci<sup>12</sup> oraz w kontekście poziomu znajomości języka<sup>13</sup>. W poniższym eksperymencie, przetestowano natomiast jakie natężenie negatywnego potencjału semantycznego wywoła zdanie w języku angielskim zawierające na końcu ikonę, która:

- a) jest niespodziewaną zmianą formy przekazu informacji (po ciągu słów występuje obrazek),
- b) jej semantyczna treść nie jest oczywista (obrazki symbolizują słowa, będące homofonami).

---

<sup>11</sup> Holcomb Phillip, Neville Helen, „Auditory and Visual Semantic Priming in Lexical Decision: A Comparison Using Event-related Brain Potentials”, *Language and Cognitive Processes* 4 (1990): 281-312.

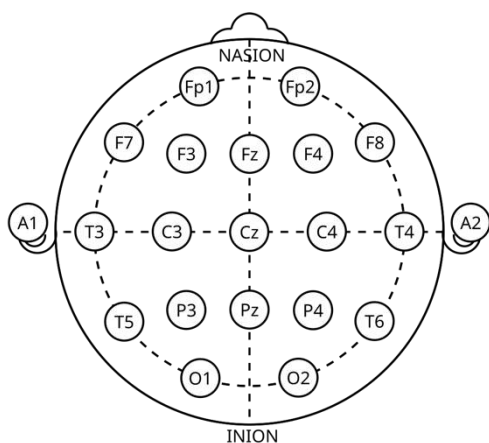
<sup>12</sup> Benau Erik, Morris Joanna, „Semantic Processing in Children and Adults: Incongruity and the N400”, *Journal of Psycholinguistic Research* 3 (2011) 225-239.

<sup>13</sup> Newman Aaron, Tremblay Antoine, Nichols Emily, i in., „The Influence of Language Proficiency on Lexical Semantic Processing in Native and Late Learners of English.” *Journal of Cognitive Neuroscience* 5 (2012): 1205-1223

## Metody

Eksperyment realizowany w trakcie projektu w dużym stopniu bazuje na badaniu M.Kutas i S. A. Hillyard. z 1980 r. Zakłada występowanie negatywnej załamki (N400) na końcu zdania, gdzie występuje umiarkowanie (ang. moderate) niespodziewane słowo, a przede wszystkim jest wykonane przy użyciu niewielkiej ilości elektrod. Typowy System 10-20 (składający się zazwyczaj z 19 elektrod), ograniczony został do 9 kluczowych lokalizacji: Fz, Cz, Pz (linia środkowa), F7 i F8 (czołowo-skroniowe), T5 i T6 (skroniowo-ciemieniowe), O1 i O2 (potyliczne). Dodatkowo 10 elektroda, umieszczona za prawym uchem (right mastoid - oznaczenie TP10 lub A2), będącą referencją oraz 11 - uziemiająca (ground Fpz lub GND). Artefakty powstałe w trakcie badania zostały usunięte „ręcznie” w programie EEGLAB. Natomiast, podobnie jak w eksperymencie z 1980 r. najbardziej miarodajne okazały się być elektrody umieszczone w linii środkowej (Fz, Cz, Pz). Wobec tego, również jak to miało miejsce w podsumowaniu M.Kutas i S. A. Hillyard, w wynikach umieszczone zostały tylko dane zebrane z tych właśnie trzech kluczowych lokalizacji.

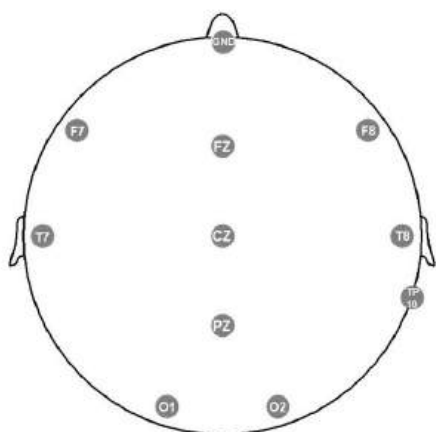
Typowy montaż EEG - System 10 - 20:



Źródło:

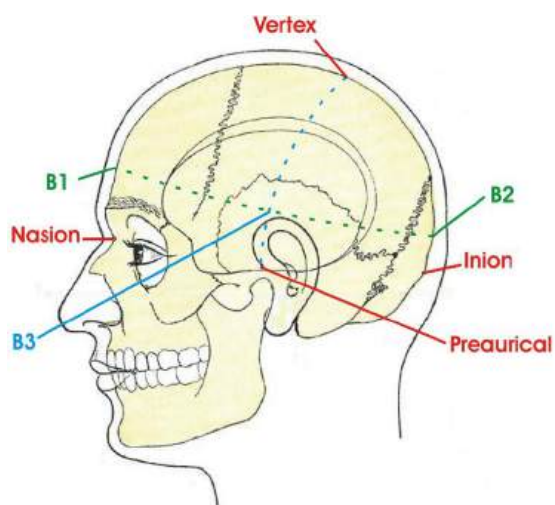
[https://en.wikipedia.org/wiki/10%E2%80%9320\\_system\\_%28EEG%29#/media/File:21\\_electrodes\\_of\\_International\\_10-20\\_system\\_for\\_EEG.svg](https://en.wikipedia.org/wiki/10%E2%80%9320_system_%28EEG%29#/media/File:21_electrodes_of_International_10-20_system_for_EEG.svg)

Montaż zastosowany w badaniu:



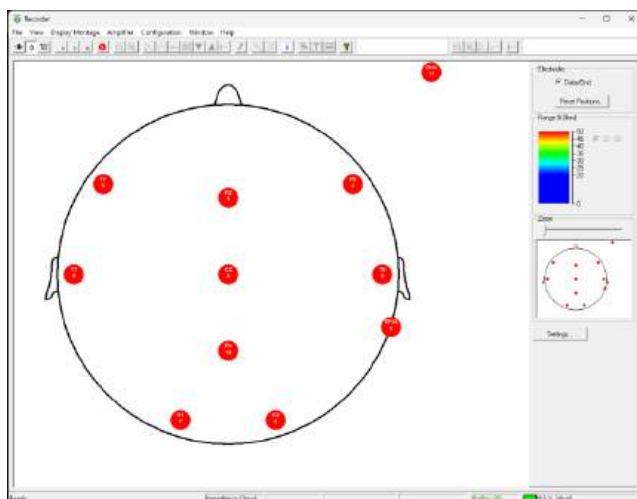
Źródło: Schemat montażu (zrzut ekranu) z BrainVisionRecorder, zmodyfikowane

Lokalizacja punktu CZ (szczyt głowy), oznaczona na obrazku pomocniczym jako *Vertex* została wyznaczona dzięki miarom zebranych w płaszczyźnie czołowej; od punktu B3 (obraz pomocniczy) nad lewym uchem do analogicznie wyznaczonego punktu nad uchem prawym, w płaszczyźnie strzałkowej; punkty *Nasion* i *Inion*. Punkt *Vertex* położony jest na przecięciu obu długości, dokładnie w ich połowie.



Źródło: <https://brain.fuw.edu.pl/edu/index.php/Plik:Czaszka.png> dostęp 23.06.25

Opór elektryczny był monitorowany w programie BrainVisionRecorder. Żel aplikowano na elektrody do momentu uzyskania pożądanej przewodności, tzn. możliwie najniższego oporu elektrycznego (do 10 kOm). Dodatkowe ustawienia aktywnych elektrod: zakres ustawiono od 50 kOm (Bad level) do 10 kOm (Good level) ma on odniesienie do skali kolorów, która sygnalizowała stopień przewodności elektrod.



Źródło: zrzut ekranu z programu BrainVisionRecorder

Tabela 1. Protokół badania

	UCZESTNIK	WYBÓR ODPOWIEDZI
1.	WIEK	(PODAJ)
2.	WADA WZROKU	TAK / NIE
3.	CZY NOSISZ OKULARY KOREKCYJNE PRZEPISANE NA TĄ WADE	TAK / NIE
4.	RĘCZNOŚĆ	PRAWORĘCZNY / LEWORĘCZNY
5.	POZIOM ZMĘCZENIA	SKALA OD 1 - 5, GDZIE: 1 - NIE CZUJĘ SIĘ ZMĘCZONY 5 - CZUJĘ SIĘ BARDZO ZMĘCZONY
6.	POZIOM ZNAJOMOŚCI JĘZYKA ANGIELSKIEGO	SUBIEKTYWNY WYBÓR, WEDŁUG SKALI CEFR: 1. POCZĄTKUJĄCY A1/A2 2. ŚREDNIOZAAWANSOWANY B1/B2 3. ZAAWANSOWANY C1/C2

7.	LOKALIZACJA CZ	(MIERZONE PRZEZ PROWADZĄCEGO)
8.	KLAUSTROFOBIA	TAK / NIE
9.	LĘK PRZED CIEMNOŚCIĄ	TAK / NIE

W przypadku pytania nr 6 - przy ustalaniu poziomu znajomości języka - posłużono się klasyfikacją CEFR (ang. Common European Framework of Reference for Languages), według której, osoby prezentują umiejętności językowe zgodne z opisem<sup>14</sup> :

- A1/A2 - poziom początkujący; posługiwanie się językiem obcym w stopniu podstawowym, używanie potocznych zwrotów i układanie prostych zdań,
- B1/B2 - poziom średniozaawansowany; posługiwanie się językiem w stopniu zaawansowanym, bezproblemowa komunikacja i zrozumienie innych użytkowników języka obcego,
- C1/C2 - poziom zaawansowany; biegła znajomość języka obcego.

Poziom znajomości języka nie był weryfikowany, klasyfikacja pełniła funkcję wyłącznie pomocniczą. Pytanie 8 i 9 zadano ze względu na warunki, w których przeprowadzono badanie. Ciemne, wyciszone i względnie niewielkie pomieszczenie mogło być przeciwwskazaniem dla osób cierpiących na klaustrofobię bądź lęk przed ciemnością.

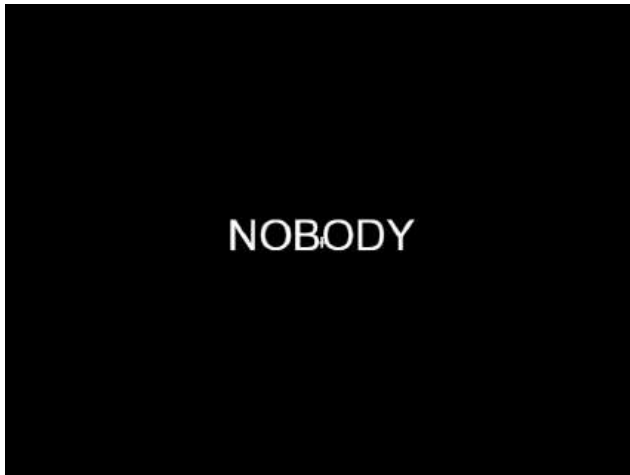
Tabela 2. Protokół; zestawienie danych uczestników

UCZESTNIK NR	1	2	3	4	5	6	7	8	9	10	11	12	13
WIEK	21	20	19	21	19	19	20	21	21	21	20	22	20
WADA WZROKU	TAK	TAK	TAK	TAK	NIE	NIE	TAK	NIE	NIE	NIE	TAK	NIE	NIE
OKULARY KOREKCYJNE	TAK	TAK	TAK	TAK	NIE	NIE	TAK	NIE	NIE	NIE	TAK	NIE	NIE

<sup>14</sup> BritishCouncilPolska, dostęp 23.06.2025, [www.britishcouncil.pl](http://www.britishcouncil.pl)

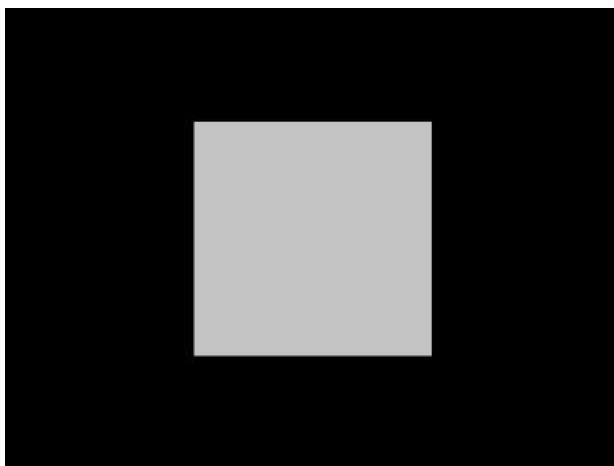
<b>RĘCZNOŚĆ</b> P - praworęczny L - leworęczny	P	P	P	P	P	P	P	P	L	P	P	P	P
<b>POZIOM ZMĘCZENIA</b>	2	3	2	2	4	2	2	1	3	2	2	1	3
<b>POZIOM ZNAJOMOŚCI JĘZYKA</b>	C1	B2	C1	B2	B1	B2	C2	C1	B2	C1	B2	C2	C2
<b>LOKALIZACJA CZ (W CM)</b>	16 17,5	16 17,5	16 18	20 18,5	15 17	15 17,5	16 17	16,5 18,5	15,5 19	16 19	15 18	16 18,5	16,5 18,5
<b>KLAUSTROFOBIA</b>	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE
<b>LĘK PRZED CIEMNOŚCIĄ</b>	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE	NIE

W trakcie eksperymentu uczestnik przez 25 min przebywa sam w wyciszonym i zaciemnionym pomieszczeniu, gdzie czyta, wyświetlane słowo po słowie, zdania. Wpierw zostaje wyświetlona krótka instrukcja, którą wyłącza dowolnym przyciskiem na klawiaturze. Następnie, w przypadkowej kolejności, wyświetlane zostaje 81 zdań (27 zdań wyjściowych powtarza się trzykrotnie), po jednym słowie na raz, każde słowo widnieje na ekranie przez około 900 ms, z częstotliwością jedno na sekundę. Poniżej znajduje się przykładowa klatka z eksperymentu ze słowem prezentowanym i krzyżem fiksacyjnym w tle.



Źródło: zrzut ekranu eksperymentu w programie PsychoPy

Między zdaniem znajduje się dłuższa pauza trwająca ok. 4 sekundy. Dodatkowo po każdym zdaniu wyświetla się szary kwadrat, sygnalizujący koniec sekwencji wyświetlania słów dla każdego zdania.



Źródło: zrzut ekranu eksperymentu w programie PsychoPy

Tabela 3. Lista zdań podstawowych







Indeks	Treść
1	TO BE OR NOT TO BE
2	BE CAREFUL WHAT YOU WISH FOR
3	DO NOT WORRY MY DEAR
4	HE WAS HIT WITH A BAT
5	HE IS NOWHERE TO BE SEEN
6	I'M NOT HUNGRY I JUST ATE
7	WHAT A BEAUTIFUL TALE
8	TELL ME WHY
9	NOBODY KNOWS

Z 9 zdań podstawowych utworzono 27 zdań wyjściowych. Bodźcem w każdym zdaniu jest słowo końcowe, podzielono je na trzy typy: S2, S3 i S4. Nazwy grup bodźców odpowiadają nazwom wysyłanych przez program markerów. Obrazy typu S3 – homofony, wybrano przy pomocy internetowego słownika homofonów anglojęzycznych<sup>15</sup>.

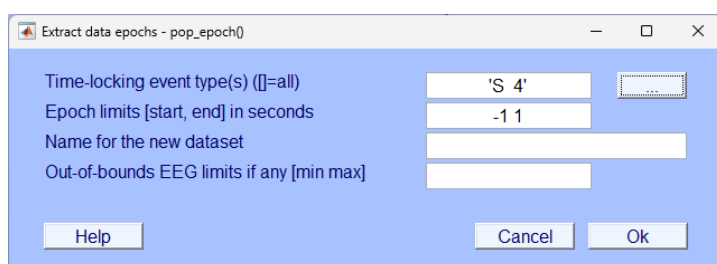
Tabela 4. Lista bodźców z podziałem na typy

S2 - ZWYKLE SŁOWA NA KOŃCU ZDANIA	S3 - HOMOFONY OBRAZKOWE	S4 – PRZYPADKOWE OBRAZKI (gr. KONTROLNA)
BE		
FOR		
DEAR		

<sup>15</sup> Homophone Dictionary, dostęp 23.06.2025, [www.homophone.com](http://www.homophone.com)

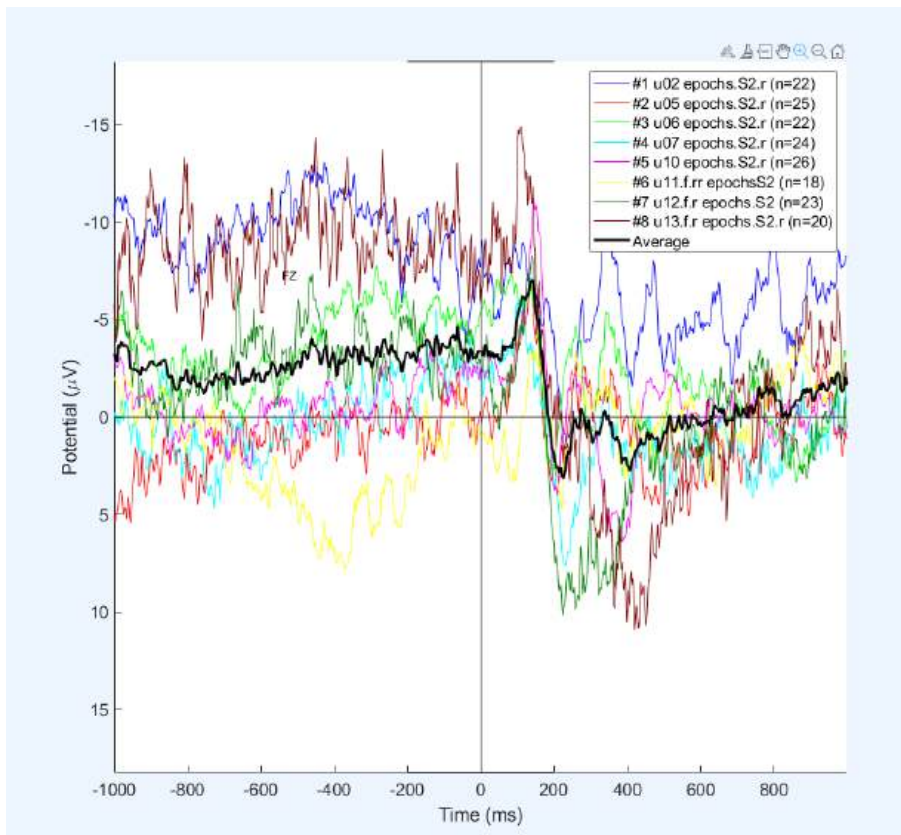
BAT		
SEEN		
ATE		
TALE		
WHY		
KNOWS		

Surowy sygnał został zaimportowany w formacie .vhdr do programu EEGLAB i przefiltrowany przy pomocy Basic FIR i parametrze dolnej granicy częstotliwości (pass - band) - 0,1 Hz. Artefakty wywołane przez mruganie, zaciskanie szczęki, ruchy głowy i inne ruchy, np. związane ze zmianą pozycji ciała, wycięto. Z przefiltrowanych wyników wyekstrahowano trzy epoki: 'S 2', 'S 3' i 'S 4', każda znajdowała się w oknie czasowym od 1 sekundy przed (-1) do 1 sekundy po (1) określonym markerze.



Źródło: zrzut ekranu z programu EEGLAB

Otrzymane krzywe aktywności, z uwzględnieniem wartości średniej zestawiono na jednym wykresie.



### Legenda:

#1 - numer datasetu,

u02 - numer uczestnika,

### epochs.S2.r

- wyodrębniona epoka, po redukcji wadliwych wartości r, zawierająca markery 'S 2', czyli zwykle słowa na końcu zdania,

### (n=22)

- ilość uwzględnionych zdarzeń (zarejestrowanych momentów wystąpienia danego bodźca).

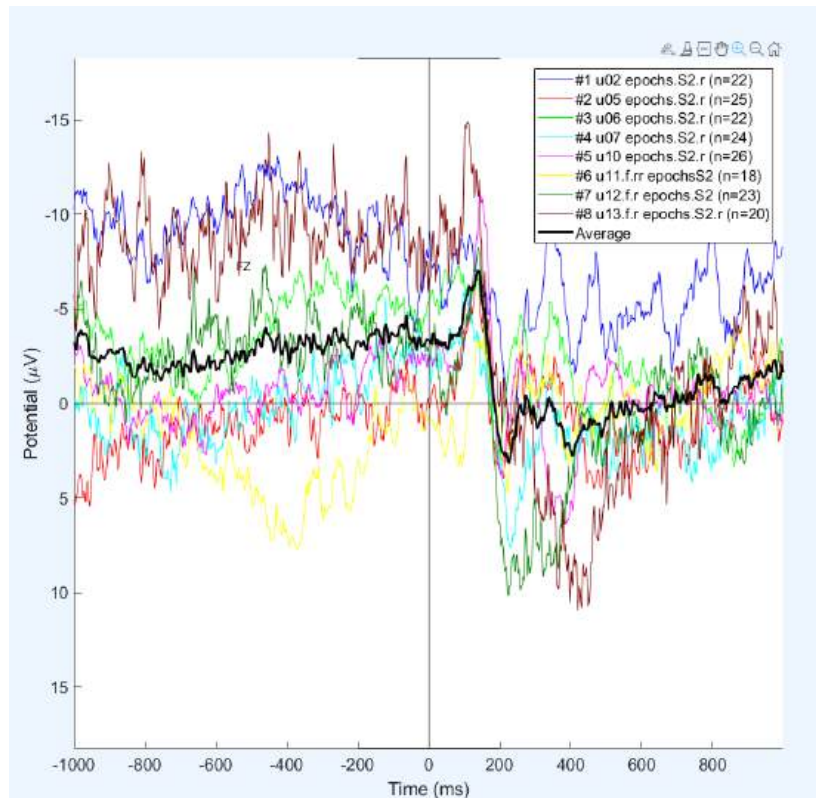
Przykładowe zestawienie aktywności zebranej z elektrody FZ w granicach epoki S2 dla 8 uczestników. Źródło: Zrzut ekranu programu EEGLAB.

Ze względu na znaczącą ilość artefaktów ruchowych, zwarcie kluczowych elektrod bądź niemiernodajny sygnał wywołany przesunięciem elektrod wraz z czepkiem, z grupy 13 uczestników odrzucono wyniki pięciorga z nich, tj. uczestników numer: 1, 3, 4, 8 i 9.

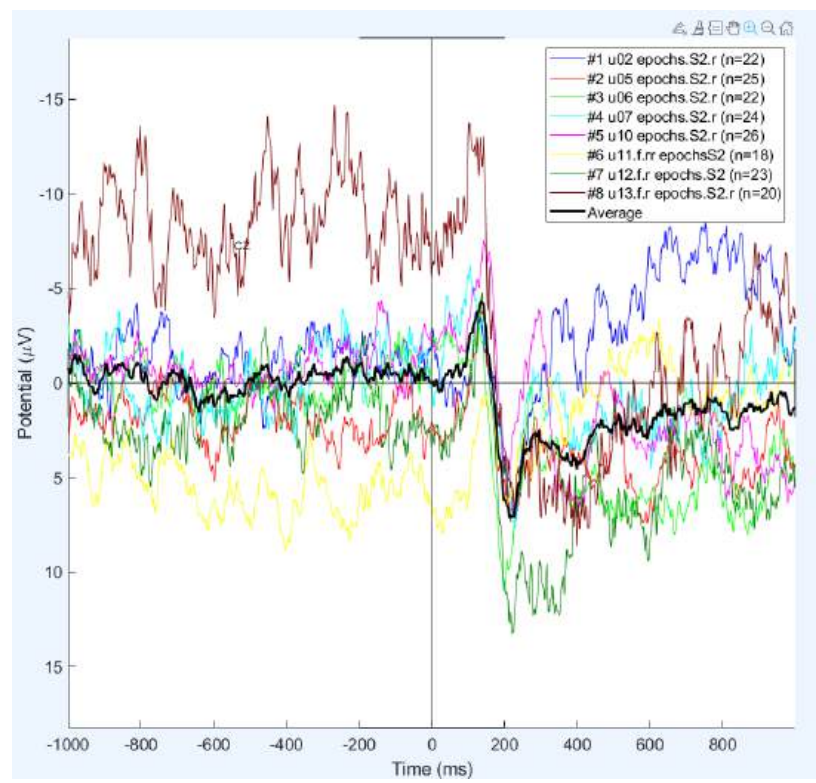
## Wyniki

Zestawienie wszystkich aktywności z elektrod FZ, CZ, PZ, dla 8 uczestników.

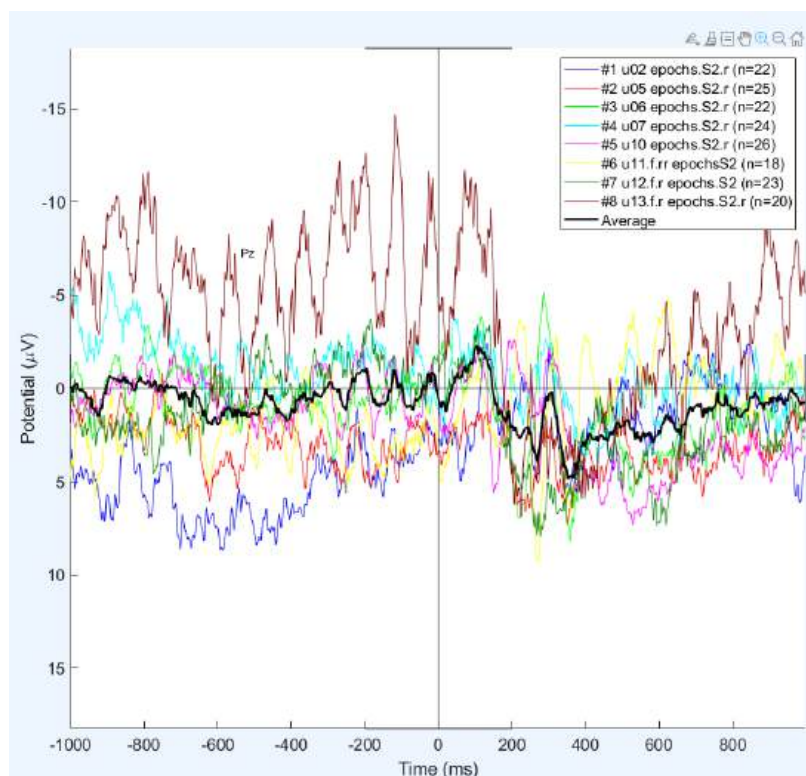
Dla bodźców typu S2 (zwykle słowa na końcu zdania):



FZ

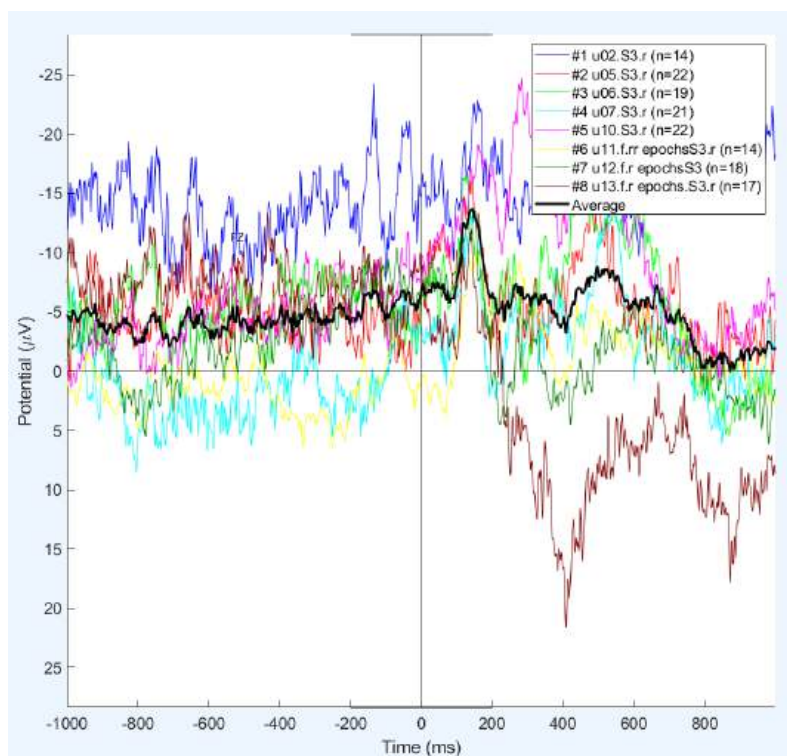


CZ

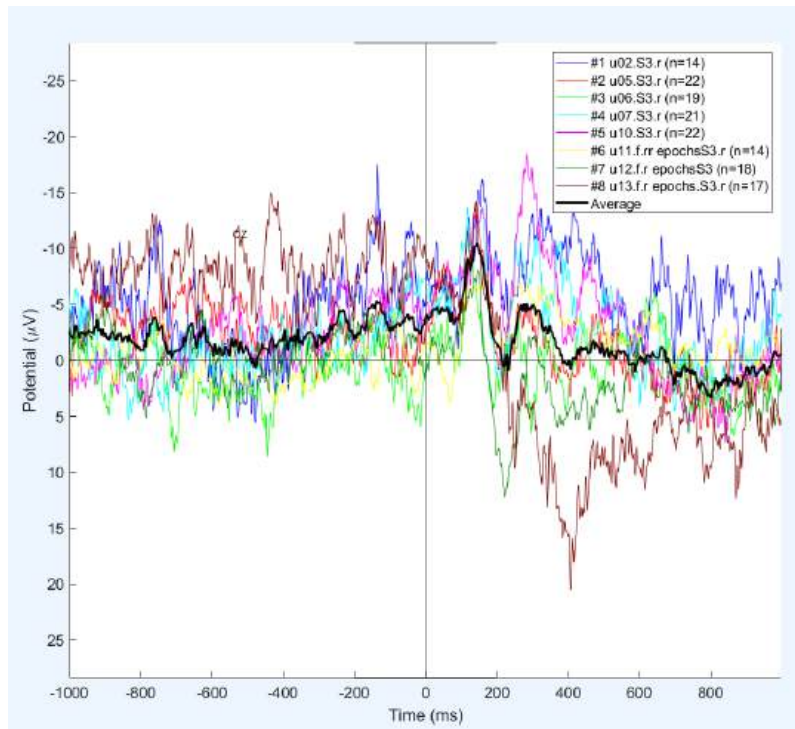


PZ

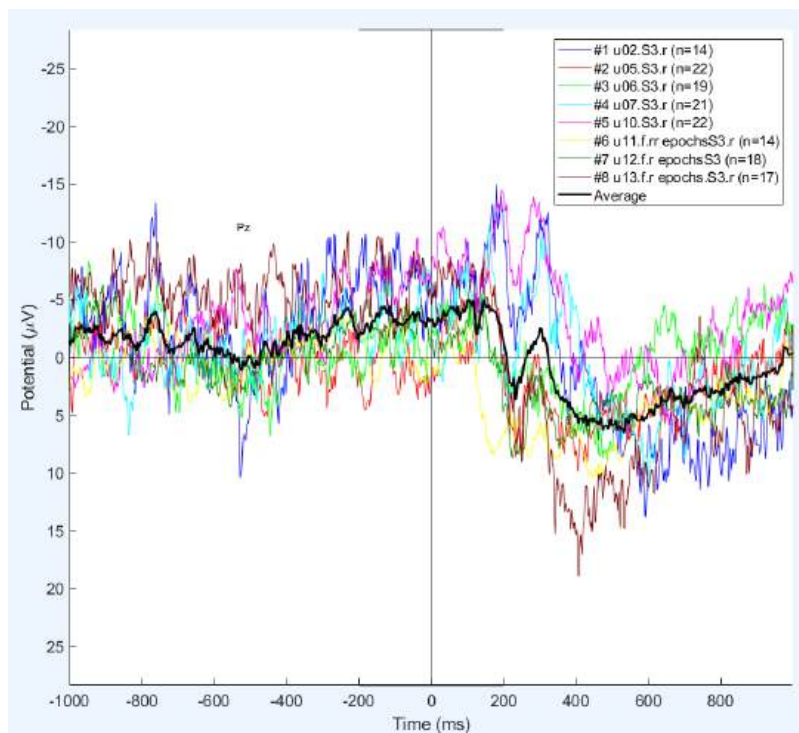
Dla bodźców typu S3 (próba badawcza - homofony obrazkowe):



FZ

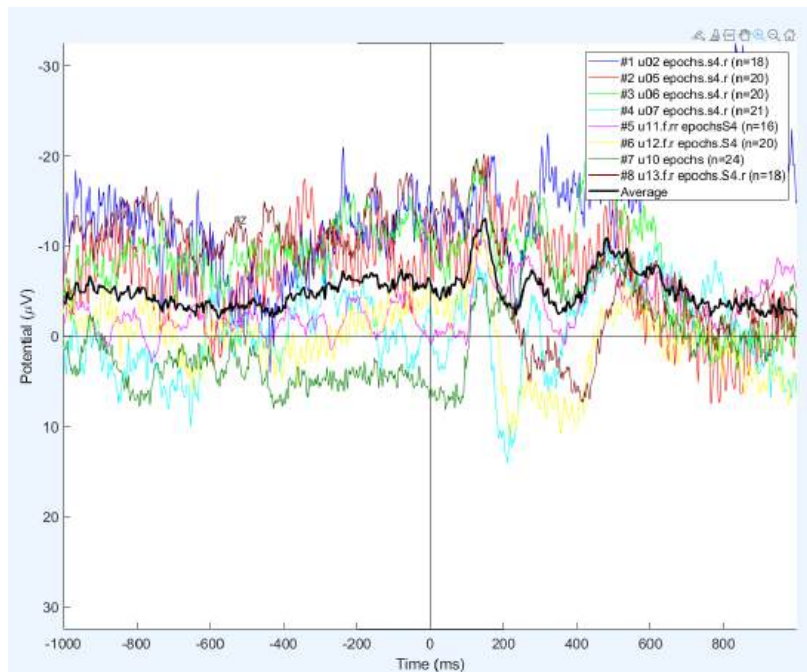


CZ

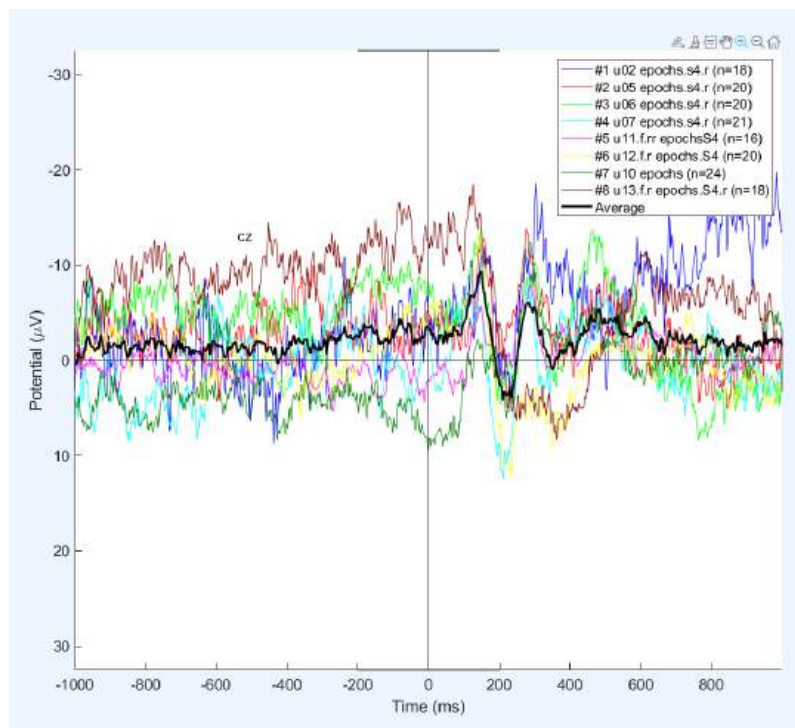


PZ

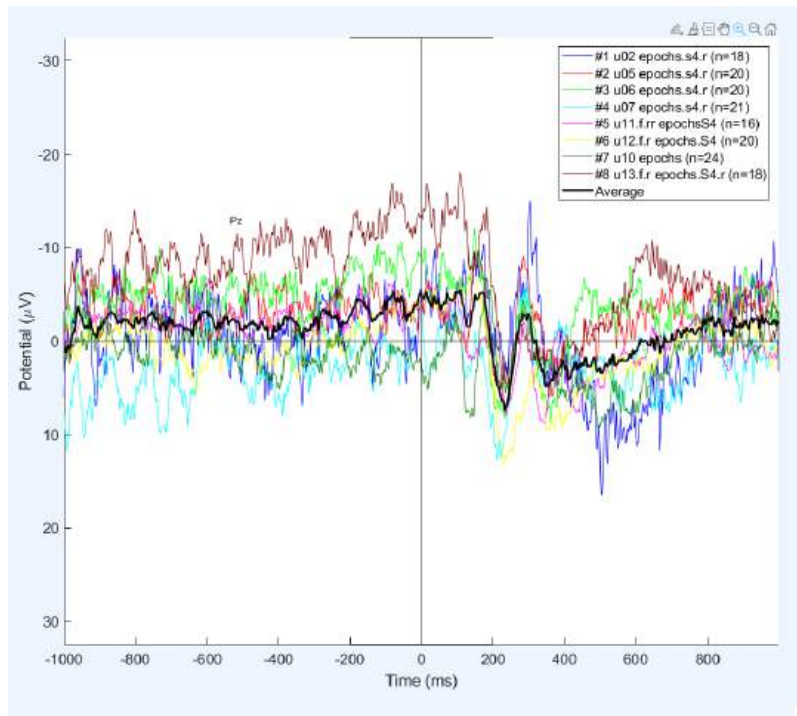
Dla bodźców typu S4 (próba kontrolna - przypadkowe obrazki):



FZ



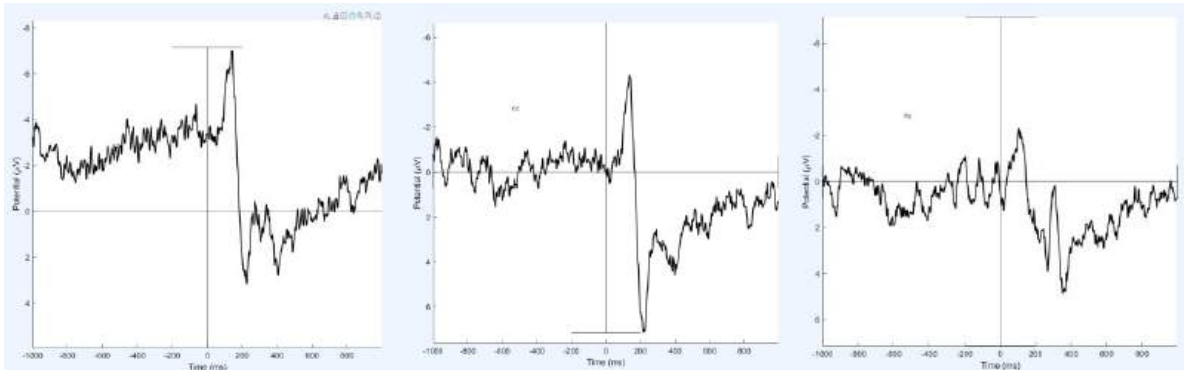
CZ



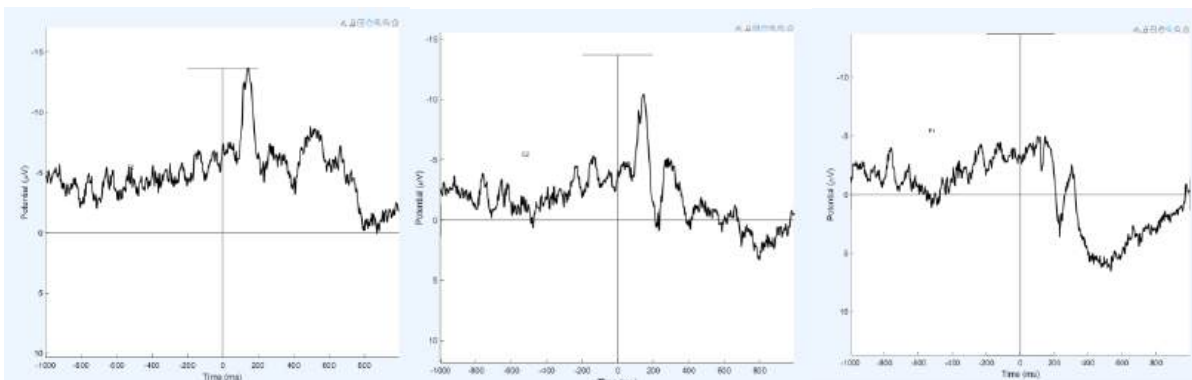
PZ

Średnie uzyskane z wyników; elektrody FZ, CZ, PZ, dla bodźców typu:

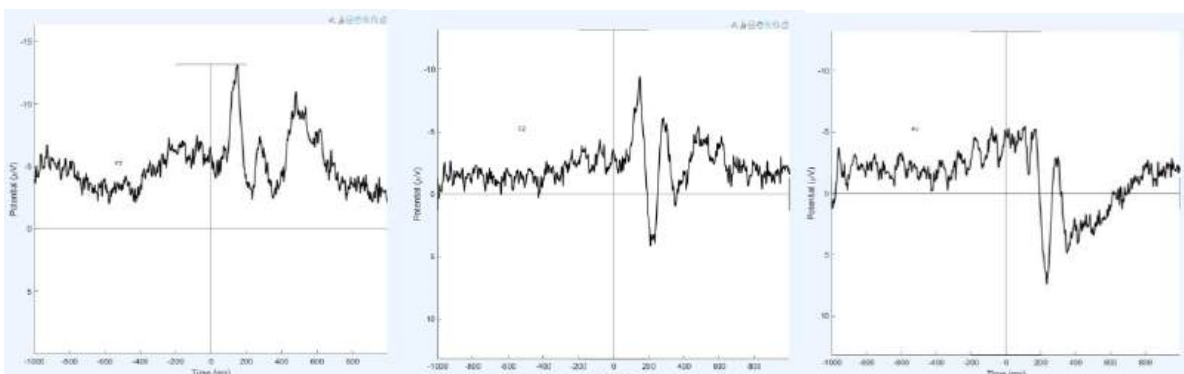
S2 (zwykle słowa na końcu zdania):



S3 (próba badawcza – homofony obrazkowe):



S4 (próba kontrolna – przypadkowe obrazki):

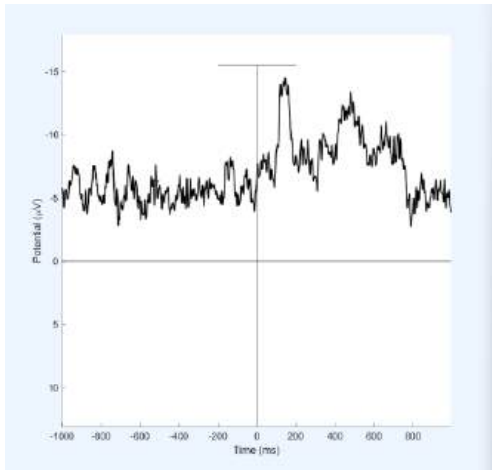


Zestawienie pod kątem poziomu znajomości języka angielskiego i jego wpływ na natężenie potencjału N400. Grupa B2 - wyniki uczestników, którzy deklarowali znajomość języka do poziomu

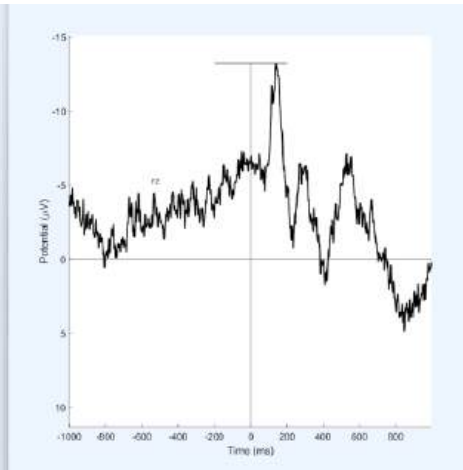
B2 włącznie oraz Grupa *CI+* zawierająca krzywe uczestników deklarujących zaawansowaną lub biegłą znajomość języka angielskiego.

Dla bodźców typu S3 (homofony obrazkowe), kolejno elektrody FZ, CZ, PZ:

*B2*

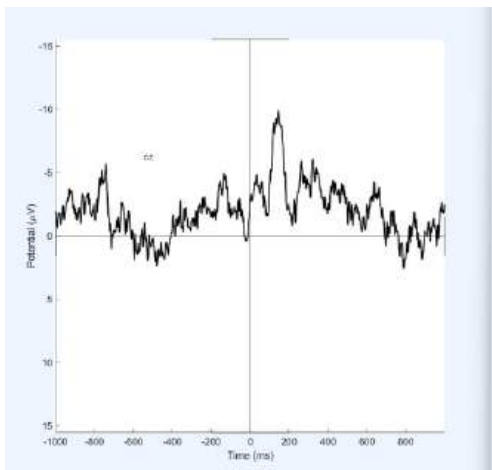


*CI+*

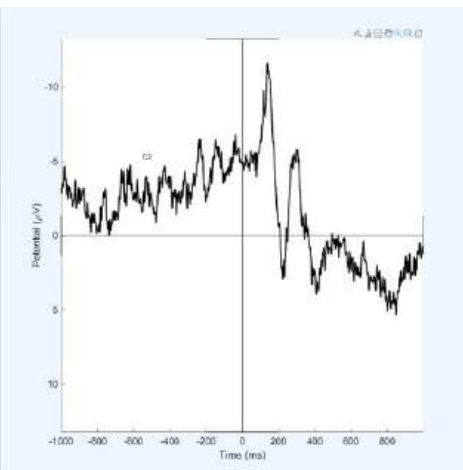


FZ

*B2*



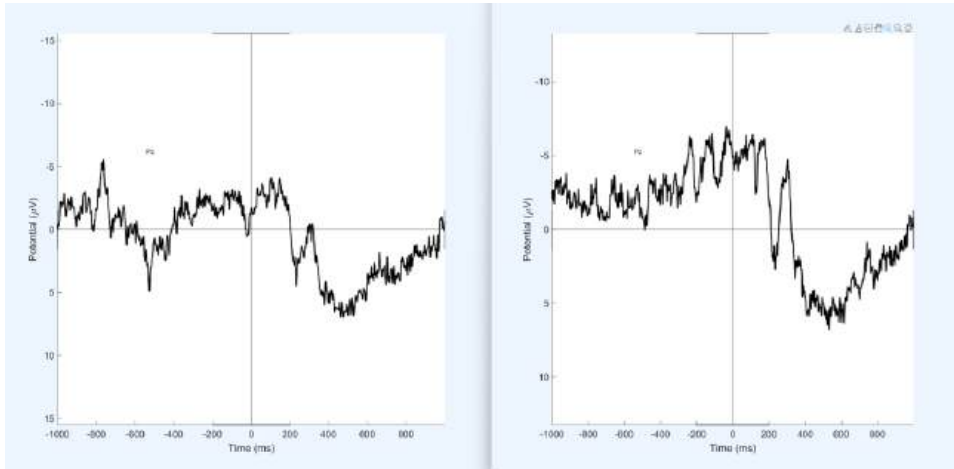
*CI+*



CZ

*B2*

*CI+*

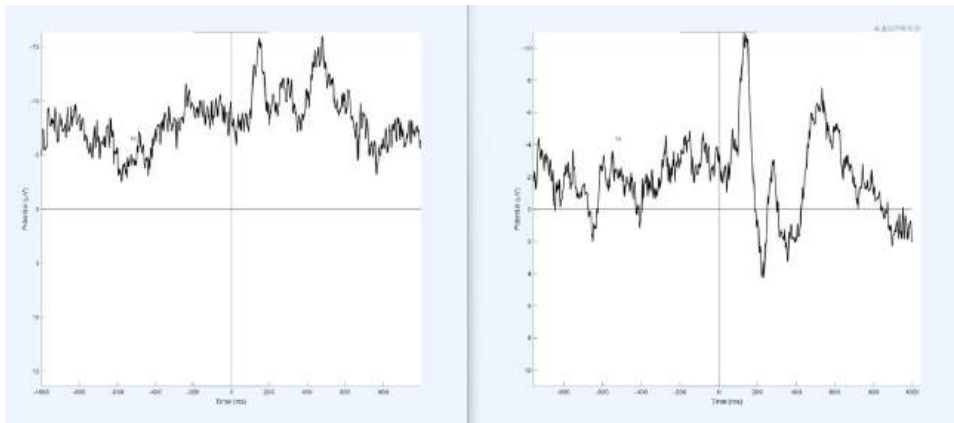


PZ

Dla bodźców typu S4 (przypadkowe obrazki), kolejno elektrody FZ, CZ, PZ:

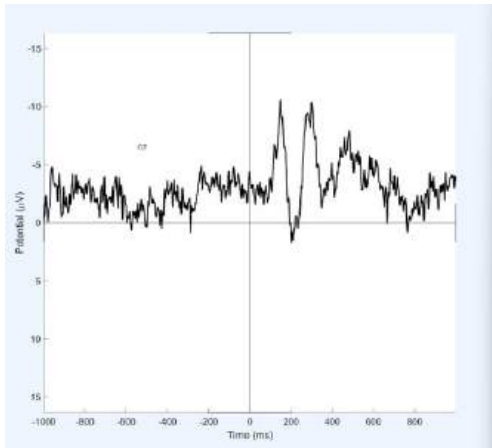
*B2*

*CI+*

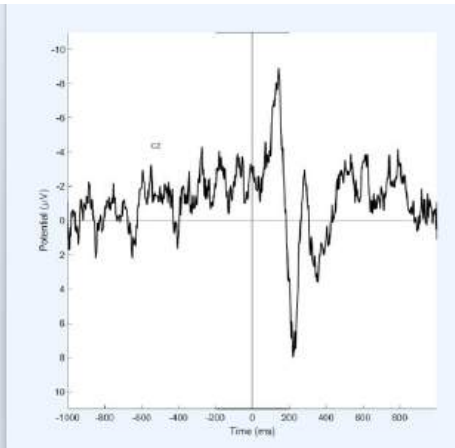


FZ

*B2*

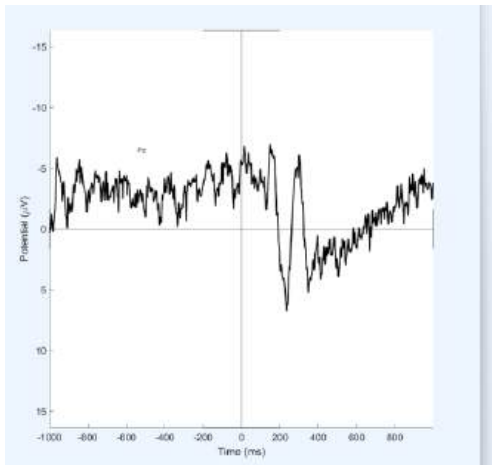


*CI+*

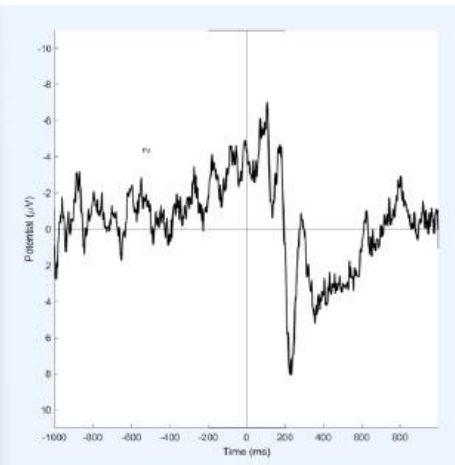


*CZ*

*B2*



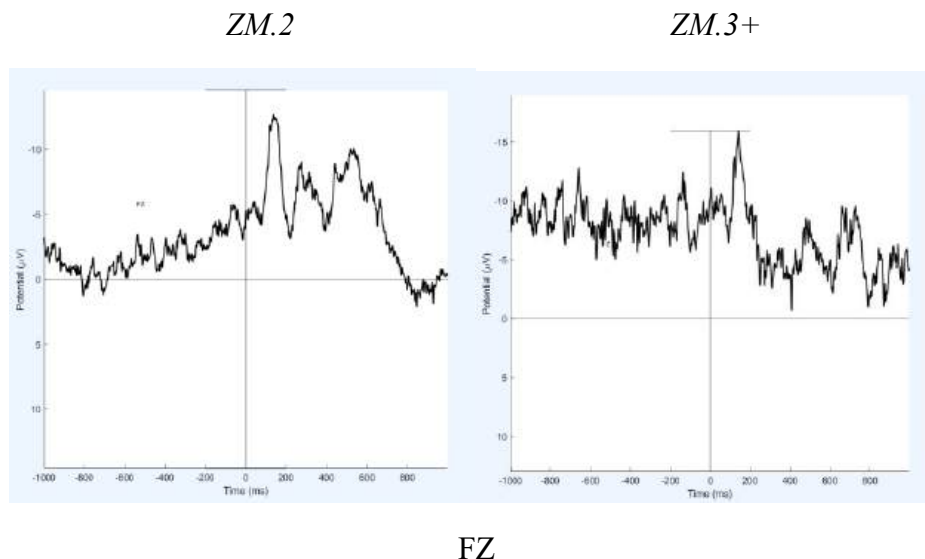
*CI+*



*PZ*

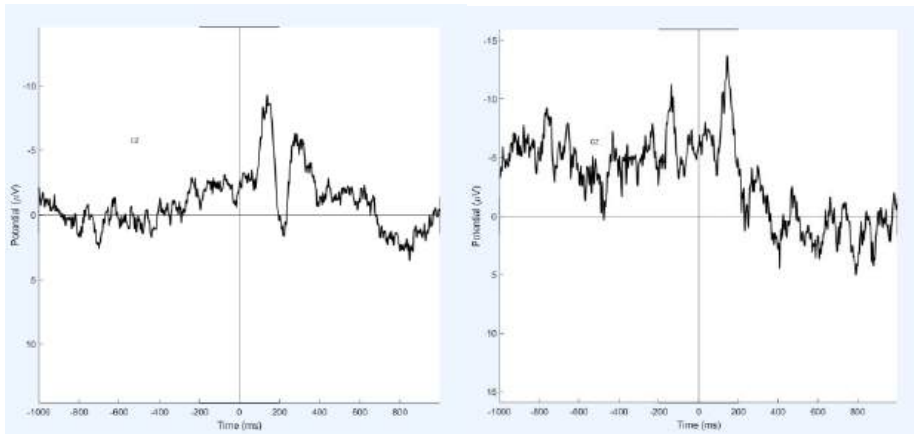
W trakcie analizy wyników zauważono różnice w morfologii krzywych. Amplitudy skrajnych aktywności gr. *C1+* mają większą wartość oraz w przebiegu fali widoczniej odznaczają się potencjały. Może to sygnalizować większe zrozumienie słów (w tym lepsze zarejestrowanie niezgodności w zdaniu, tj. Wyraźniejszej załamki N400) grupy deklarującej zaawansowaną znajomość języka angielskiego. Aktywność bioelektryczna grupy *B2* charakteryzowała się delikatnie bardziej ujemnymi wartościami i większym spłaszczeniem ich fali (mniejszymi amplitudami). Ogólne zmęczenie organizmu można zaobserwować w aktywności bioelektrycznej mózgu. Wobec tego otrzymane wyniki podzielono ponownie na dwie grupy, tj. *ZM.2*: Uczestnicy deklarujący poziom zmęczenia do 2 włącznie (gdzie 1 świadczy o braku zmęczenia) oraz *ZM.3+*: Uczestnicy określający swój poziom zmęczenia na 3 i wyżej (gdzie 5 świadczy o bardzo silnym zmęczeniu). Zestawienie średnich wyników z kluczowych elektrod, kolejno: FZ, CZ, PZ.

Dla bodźców typu S3 (homofony obrazkowe):



ZM.2

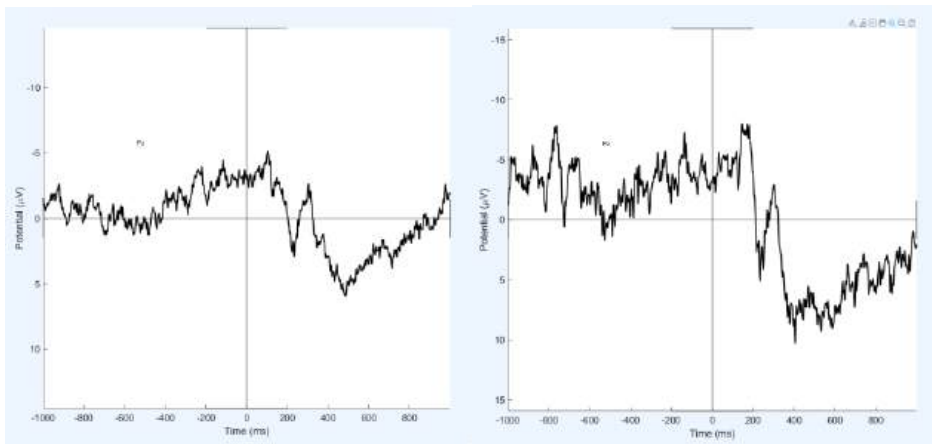
ZM.3+



CZ

ZM.2

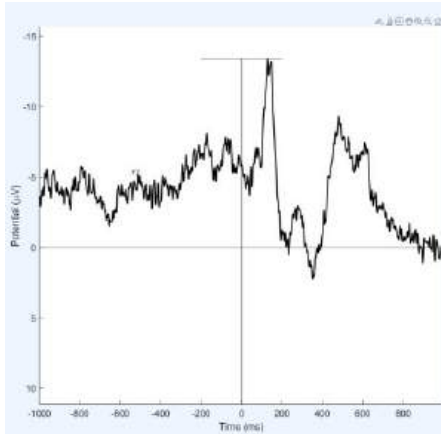
ZM.3+



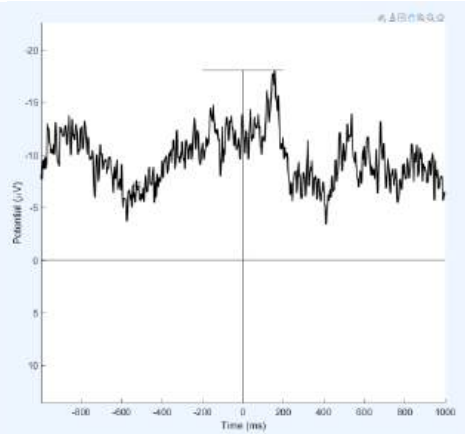
PZ

Dla bodźców typu S4 (przypadkowe obrazki):

ZM.2

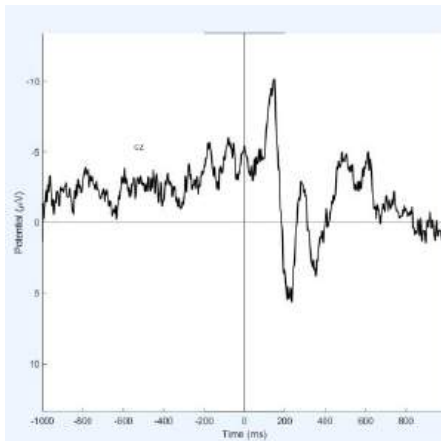


ZM.3+

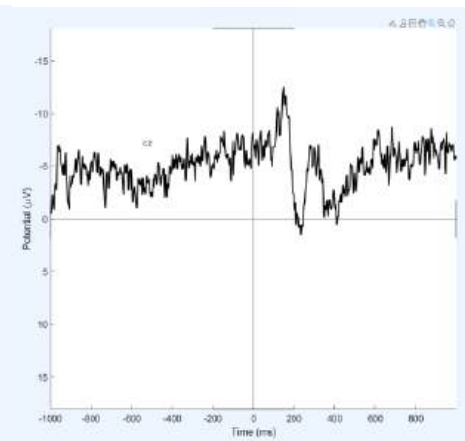


FZ

ZM.2



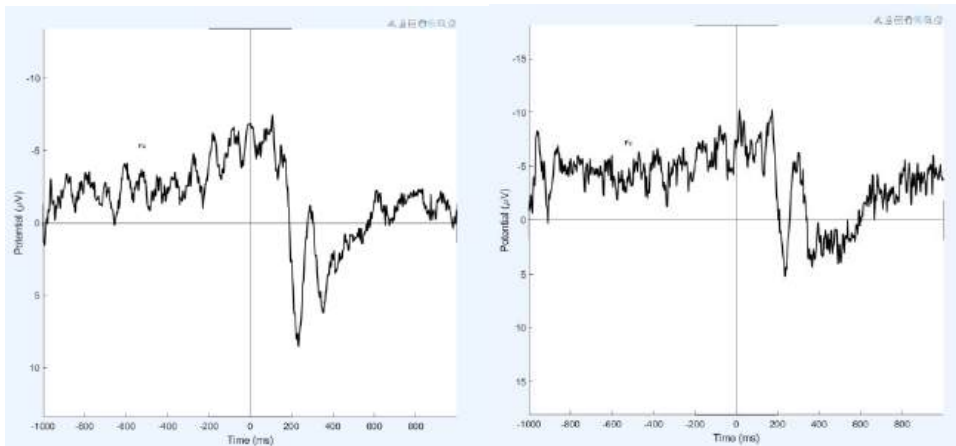
ZM.3+



CZ

ZM.2

ZM.3+



PZ

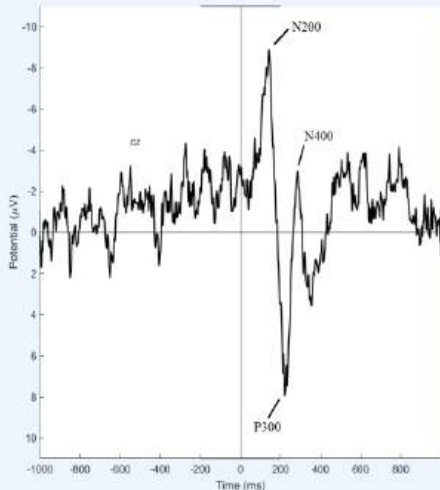
W przypadku osób deklarujących większe zmęczenie można zauważyć niewielkie różnice w przebiegu aktywności bioelektrycznej mózgu. Amplituda przyjmuje nieco mniejsze wartości a sama aktywność jest stosunkowo bardziej ujemna. Potencjał N400 delikatnie bardziej zauważalny jest w grupie ZM. 2, co może świadczyć o większym zaangażowaniu uczestników deklarujących mniejsze zmęczenie.

## Wnioski

W wynikach zaobserwowano obecność załamki N400. W przypadku próby badawczej, tj. Zdań zakończonych homofonami w postaci obrazkowej, potencjał N400 charakteryzował się niewielką aktywnością, zazwyczaj nie przekraczającą -5 mikrowoltów ( $\mu\text{V}$ ). Dla próby kontrolnej, tj. Zdań zakończonych obrazem o dużej niezgodności semantycznej, aktywność również nie przekraczała -5 mikrowoltów, jednak zaobserwowano znaczą różnicę w morfologii między obydwojma próbami. W przypadku próby kontrolnej S4, szczególnie w grupie *CI+*, obserwujemy wyraźnie zaznaczone potencjały N200, P300 i szukane N400.

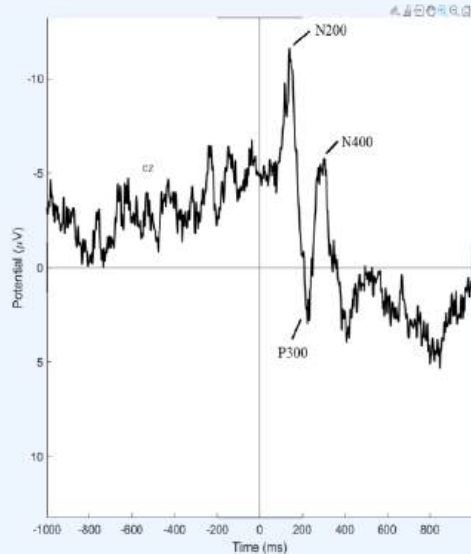
Bodziec typu S4 (przypadkowe obrazki),  
Elektroda CZ, gr. CI+

W przypadku próby badawczej S3 (homofony obrazkowe), średnia aktywności z tej samej elektrody CZ charakteryzuje się mniejszą amplitudą między potencjałem N200 a P300. Można zauważyć potencjał N400 oraz (w przeciwieństwie do próby S4) późniejsze pozytywne aktywności.



Bodziec typu S3 (homofony obrazkowe),  
Elektroda CZ, gr. CI+

W konfrontacji z wynikami badań M.Kutas i A.Hillyarda z 1980 r. można zauważyć, że zdania o nieznacznej niezgodności (ang. moderate) - tutaj homofony obrazkowe - generują mniejszą aktywność potencjału N400. Natomiast te posiadające silną niezgodność - bodźce próby kontrolnej - odznaczają się większym negatywnym załamaniem. Dodatkowo, testując zmiany aktywności bioelektrycznej przy bodźcu typu S2 - będącym zwykłym słowem - nie zauważono załamki N400.



Poniżej znajduje się ostateczne zestawienie wyników osób deklarujących lepszą znajomość języka angielskiego (CI+) z elektrody CZ, dla poszczególnych typów bodźców:

S2

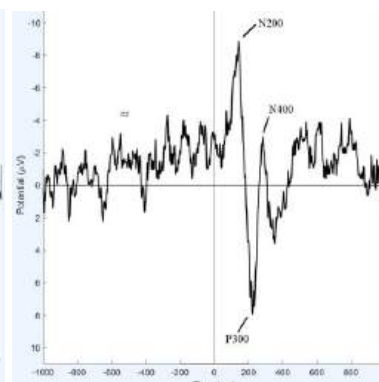
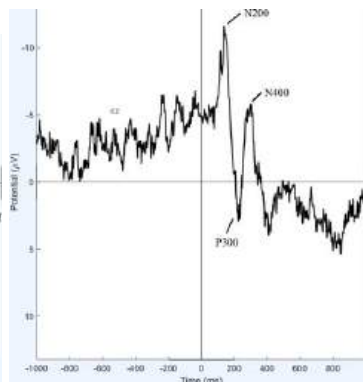
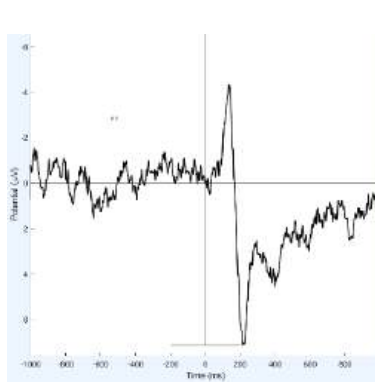
S3

S4

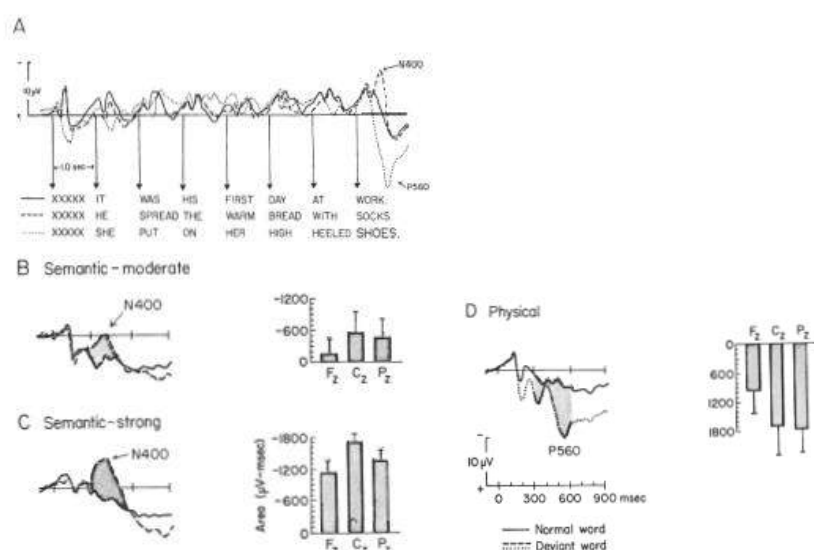
Zwykłe słowo

Homofon obrazkowy

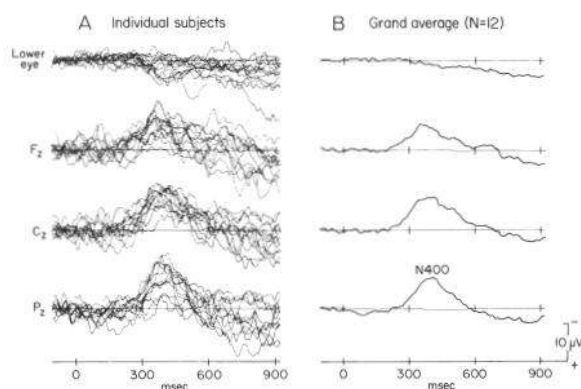
Przypadkowy obrazek



## Wyniki badania M.Kutas z 1980 r.



Źródło; Kutas M., Reading Senseless Sentences: Brain Potentials Reflects Semantics Incongruity, 1980, nr 207, s. 203, Fig. 1.



Źródło; Kutas M., Reading Senseless Sentences: Brain Potentials Reflects Semantics Incongruity, 1980, nr 207, s. 204, Fig. 2.

W odniesieniu do badań M.Kutas, bodźcem S2 jest normalne słowo (ang. Normal word), S3 – słowo umiarkowanie semantycznie niezgodne (ang. Deviant word semantic - moderate), S4 – słowo silnie semantycznie niezgodne (ang. Deviant word semantic - strong).

Pilot eksperymentu wykonanego w ramach projektu licencjackiego stanowił kolejny element badań nad złożonością przetwarzania bodźców językowych - będący jednym z głównych zainteresowań nauk kognitywnych, a także dostarczył interesujących informacji na temat przetwarzania semantycznego zdań i obrazów. Ze względu na wielkość próby badawczej, pilot ten jest jedynie wstępem do dalszych badań nad procesem przetwarzania homofonów i jej wpływu na potencjał semantyczny N400.

## Bibliografia

- Benau, Erik M., Joanna Morris, J. W. Couperus, „Semantic Processing in Children and Adults: Incongruity and the N400”, *Journal of Psycholinguistic Research* 3 (2011): 225-239.
- Federmeier, Kara D. „Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres”, *Neuropsychologia* 7 (2002): 730-747.
- Finnigan, Simon, „ERP 'old/new' effects: memory strength and decisional factor(s)”, *Neuropsychologia* 13 (2002): 2288-2304.
- Folstein, Jonathan R., Cyma Van Petten, „Influence of cognitive control and mismatch on the N2 component of the ERP: A review”, *Psychophysiology* 1 (2008): 152-170.
- Furutsuka, Takashi, „Effects of Rapid Attention Switching on the NI-P2 Amplitude of the Visual Event-Related Potentials”, *Research and Clinical Center for Child Development: Annual Report* 11 (1989): 55-64.
- Ganc, Małgorzata, „Study system of late cortical responses”, *Nowa Audiofonologia* 3 (2020): 69-75.
- Holcomb, Phillip J., Helen J. Neville, „Auditory and Visual Semantic Priming in Lexical Decision: A Comparison Using Event-related Brain Potentials”, *Language and Cognitive Processes* 4 (1990): 281-312.
- Homophone Dictionary at Homophone. Dostęp 23.06.2025. [www.homophone.com](http://www.homophone.com)
- Kutas, Maria, Kara D. Federmeier, „Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP)”, *Annual Review of Psychology* 1 (2011): 621-647.
- Kutas, Maria, Steven A. Hillyard, „Event-related brain potentials to grammatical errors and semantic anomalies”, *Memory & Cognition* 5 (1983): 539-550.
- Kutas, Maria, Steven A. Hillyard, „Reading Senseless Sentences: Brain Potentials Reflects Semantics Incongruity”, *Science* 207 (1980): 203-205
- Lijffijt, Marijn, Scott D. Lane, Stacey L. Meier i in., „P50, N100, and P200 sensory gating: Relationships with behavioral inhibition, attention, and working memory”, *Psychophysiology* 5 (2009): 1059-1068.
- Newman, Aaron.J., Antoine Tremblay, Emily S. Nichols, i in., „The Influence of Language Proficiency on Lexical Semantic Processing in Native and Late Learners of English”, *Journal of Cognitive Neuroscience* 5 (2012): 1205-1223.
- Osterhout, Lee, Judith McLaughout, Michael Bersick, „Event-related brain potentials and human language”, *Trends in Cognitive Sciences* 6 (1997): 203-209.
- Polich, John, Jose R. Criado, „Neuropsychology and neuropharmacology of P3a and P3b”, *International Journal of Psychophysiology* 2 (2006): 172-185.
- Severance, Charles R, „Python for Everybody. Exploring Data in Python 3. Wydanie trzecie” Andrzej Wójtowicz, 2023.
- Silva-Pereyra, Juan, Maritza Rivera Gaxiola, Patricia K. Kuhl, „An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing”, *Cognitive Brain Research* 2-3 (2005): 247-258.
- Słownik języka polskiego PWN, Dostęp 23.06.2025, [www.sjp.pwn.pl](http://www.sjp.pwn.pl).
- Van Petten, Cyma, Marta Kutas, „Interactions between sentence context and word frequency in event-related brainpotentials”, *Memory & Cognition* 4 (1990): 380-393.

Paweł Ficek

## **Czy posiadanie dziecięcej twarzy sprawia, że wydajemy się zabawniejsi? Raport z projektu.**

**Abstrakt:** Wśród teorii tłumaczących funkcję poczucia humoru u gatunku ludzkiego nietrudno znaleźć takie, które łączą stan rozbawienia z redukcją napięcia mięśniowego i sygnalizowaniem bezpieczeństwa. Niektóre badania sugerują, że zbliżone efekty powinna dawać interakcja z osobą, której twarz zawiera zestaw cech wpisujący się w tak zwany *Kindchenschema*, szczególnie u młodych kobiet. Biorąc pod uwagę te teorie postanowiłem zbadać, czy istnieje jakakolwiek korelacja pomiędzy postrzeganiem danej twarzy jako zabawną, a posiadaniem przez niej cech *Kindchenschema*. W tym celu przeprowadziłem badanie z wykorzystaniem obrazów twarzy wygenerowanych przez AI, a następnie zmodyfikowanych w programie graficznym oraz elektrod EDA i EMG w celu wykrycia stanu rozbawienia u badanych poprzez monitorowanie ich przewodnictwa skórniego i ruchów mięśni twarzy.

**Słowa kluczowe:** rozbawienie, Kindchenschema, baby face, EDA, EMG

**Abstract:** Among various theories explaining the role of the sense of humor in humans, it is easy to find those that link the state of amusement with a reduction in muscle tension and the signaling of safety. Some research suggests that interacting with a person whose face exhibits a set of features that fit into the so-called *Kindchenschema* should yield similar effects, especially among young women. Considering these theories, I decided to examine whether there is any correlation between perceiving a face as funny and it possessing *Kindchenschema* features. To this end, I conducted a study using AI-generated facial images, which were later modified in graphic software, along with EDA and EMG electrodes to detect amusement in participants by monitoring their skin conductivity and facial muscle movements.

**Keywords:** amusement, Kindchenschema, baby face, EDA, EMG

## Wstęp

Do podjęcia tematyki poczucia humoru z perspektywy kognitywistycznej zainspirowała mnie popularnonaukowa książka Łukasza Jacha pod tytułem „Od ucha do ucha. Homo sapiens się śmieje”<sup>1</sup>. Nieopisaną pomocą była dla mnie również praca magisterska Krzysztofa Holvika<sup>2</sup>, która zainspirowała mnie do przeprowadzenia badania psychofizjologicznego i wykorzystania elektrod EDA i EMG do monitorowania stanu emocjonalnego badanych. Przedstawiona w dalszej części pracy metodologia wykorzystania tych elektrod oraz analizy i interpretacji zarejestrowanych przez nie wyników jest w gruncie rzeczy zapożyczona z eksperymentu przeprowadzonego w ramach wspomnianej pracy magisterskiej.

Jeśli udałooby się zaobserwować istotną korelację pomiędzy występowaniem na twarzy cech charakterystycznych dla *babyface* a oceną tej twarzy jako zabawną mogłoby to potwierdzać słuszność teorii o rozluźniającej funkcji poczucia humoru u człowieka. Ponadto, grupa badawcza bardziej zróżnicowana pod względem zarówno płci, jak i wieku dawałaby możliwość analizy w oparciu o właśnie te zmienne, gdyż efekt *babyface* powinien być bardziej zauważalny u młodych kobiet<sup>3</sup>.

Praca składa się zasadniczo z dwóch części. W pierwszej, teoretycznej, zaprezentowane zostaną teorie dotyczące funkcjonalnego aspektu poczucia humoru skupiając się w szczególności na teorii rozluźnienia (przez naturę tematu ta część pracy będzie miała charakter bardziej filozoficzny), a następnie scharakteryzuję efekt *babyface*. Część teoretyczną kończy opis fizjologicznych markerów emocji i możliwości wykrywania ich za pomocą sygnałów EDA i EMG. W drugiej części został szczegółowo opisany przebieg przeprowadzonego badania, włącznie z fazą przygotowawczą i analizą danych, a w ostatnim rozdziale znajduje się próba wyciągnięcia wniosków z badania.

## 1. Podstawy teoretyczne

### 1.1. Przegląd teorii na temat funkcji poczucia humoru

Nie powinno ulegać wątpliwości, że poczucie humoru jest inherentną cechą gatunku ludzkiego i nie przysługuje ono jedynie wybranym kulturom czy okresom w historii. Badania antropologiczne dowiodły, iż we wszystkich znanych nam społecznościach ludzkich na Ziemi występuje poczucie humoru, choć jest ono oczywiście mocno zróżnicowane przez wpływ lokalnych kultur<sup>4</sup>. Prawa doboru naturalnego wskazywałyby więc na oczywisty wniosek, że posiadanie przez człowieka takiej właśnie cechy musi pełnić jakąś funkcję natury biologicznej, psychologicznej lub społecznej. Próby wyjaśnienia funkcji pełnionej przez poczucie humoru podjęli się badacze z dziedzin takich jak filozofia, psychologia, antropologia, neurobiologia i innych, a jednak wciąż brak konsensusu w tej sprawie. Obszar ten nie jest szczególnie intensywnie badany, choć takie dociekania mogłyby rzucić

---

<sup>1</sup> Łukasz Jach, *Od ucha do ucha. Homo sapiens się śmieje*, Zrozum (Wydawnictwo Poznańskie, 2021).

<sup>2</sup> Krzysztof Holvik, „Reakcje emocjonalne użytkowników języka polskiego na ironię i emotikony: analiza psychofizjologiczna z wykorzystaniem EMG i EDA” (Uniwersytet Śląski, 2024).

<sup>3</sup> Reiner Sprengelmeyer et al., „The Cutest Little Baby Face: A Hormonal Link to Sensitivity to Cuteness in Infant Faces”, *Psychological Science* 20, nr 2 (2009): 149–54.

<sup>4</sup> Matthew Gervais i David Sloan Wilson, „The Evolution and Functions of Laughter and Humor: A Synthetic Approach”, *The Quarterly Review of Biology* 80, nr 4 (2005): 395–430.

nowe światło na działanie ludzkiego umysłu, gdyż, jak opisane jest w dalszej części pracy, jest to zjawisko bardzo skomplikowane i wielopłaszczyznowe.

Funkcjonalna strona poczucia humoru u człowieka stanowiła wyzwanie intelektualne dla filozofów już w starożytnej Grecji, choć użycie słowa humor (a konkretnie anielskiego *humour*) w znanym nam współcześnie kontekście pojawia się dopiero w XVIII wieku. We wpisie *Philosophy of Humor*<sup>5</sup> autor przedstawia perspektywę, według której śmiech i poczucie humoru, mimo pozytywnych skojarzeń w życiu codziennym, przez filozofów postrzegane były raczej negatywnie. Sokrates w dziełach Platona przypisywał śmiechowi zdolność do kształtowania ludzkiego charakteru w zależności od przedmiotu żartu, skupiając się na negatywnym i demoralizującym wpływie nieprzyzwoitych dowcipów. W średniowieczu śmiech kojarzył się z kolei z brakiem kontroli i nie przystawał wysoko postawionym urzędnikom kościelnym. W epoce nowożytnej T. Hobbes sformułował podejście, które później nazwano teorią wyższości (*superiority theory*). Zgodnie ze swoim poglądem na człowieka jako jednostkę z natury indywidualistyczną i nastawioną na rywalizację, śmiech według Hobbes'a miałby być oznaką wygranej i dlatego sprawiać nam przyjemność, bycie wyśmiewanym z kolei oznaką przegranej i sprawiać dyskomfort. Teoria ta jest jednak niewystarczająca do wytłumaczenia funkcji poczucia humoru, bo poczucie wyższości nie jest jedynym powodem, dla którego ludzie się śmieją. Czasem śmiejemy się przez własne słabości, czego eksperymentalnie dowiódł chociażby L. Deckers<sup>6</sup>.

Niedostatki teorii wyższości dostrzegano już w XVIII wieku. Jedną z teorii, która miałaby w bardziej holistyczny sposób wytłumaczyć funkcję śmiechu i poczucia humoru była (i nadal jest) teoria niedorzeczności (*incongruity theory*), którą postulował chociażby Immanuel Kant. Głosi ona, że zabawnym wydaje nam się to, co jest niezgodne z naszymi przewidywaniami i oczekiwaniami opartymi na naszych schematach myślowych. Przyjemność płynąca z żartowania nie wynika z poczucia wyższości nad przedmiotem żartu, ale ze świadomości sprawności naszego umysłu do brania udziału w tej „grze myśli”. Słabość tego modelu leży w fakcie, iż sama tylko niezgodność treści ze schematami myślowymi odbiorcy może doprowadzić do zgoła innych reakcji niż rozbawienie. Może to być chociażby groteskowe obrzydzenie<sup>7</sup>, choć tego typu emocje również mogą sprawiać przyjemność bez elementu humorystycznego. Teoria ta została poddana bardzo licznym badaniom eksperymentalnym i nadal jest jednym z najpopularniejszych tłumaczeń funkcji poczucia humoru.

Inną często poruszaną w tym dyskursie kwestią jest rola poczucia humoru w zacieśnianiu więzów społecznych i doborze partnera lub partnerki. W takim ujęciu poczucie humoru i gry słowne to jedynie bardziej zaawansowany rodzaj fizycznej zabawy, takiej, w jaką angażują się ludzkie dzieci i młode innych gatunków. Cieleśne oznaki rozbawienia, takie jak uśmiech, głośny śmiech czy łaskotanie innych wykształciły się na podstawie udawanej agresji okazywanej podczas zabawy. Takie

---

<sup>5</sup> John Morreall, „Philosophy of Humor”, w *The Stanford Encyclopedia of Philosophy*, red. Edward N. Zalta i Uri Nodelman (Metaphysics Research Lab, 2024), <https://plato.stanford.edu/archives/fall2024/entries/humor/>.

<sup>6</sup> Lambert Deckers, „On the Validity of a Weight-Judging Paradigm for the Study of Humor”, *Humor: International Journal of Humor Research* 6, nr 1 (1993): 43–56.

<sup>7</sup> Mike W. Martin, „Humour and Aesthetic Enjoyment of Incongruities”, *The British Journal of Aesthetics* 23, nr 1 (1983): 74–85.

zachowania występują u różnych hominidów, na przykład goryli<sup>8</sup>. W takiej psychologiczno-ewolucyjnej perspektywie humor gra bardzo ważną rolę w relacjach społecznych, czego nie uwzględniają pozostałe teorie. Według badań świadomość bycia postrzeganym drastycznie zmienia nasz sposób reagowania na treści humorystyczne. Człowiek śmieje się wtedy częściej<sup>9</sup> i nieświadomie dostosowuje brzmienie swojego śmiechu kodując w nim informacje między innymi o swoim statusie społecznym<sup>10</sup>. Biegłe posługiwanie się humorem w interakcjach z innymi może sygnalizować również wysoką sprawność umysłową<sup>11</sup>. Ponadto rodzaj humoru, jakim się posługujemy umożliwia szyfrowanie informacji w opowiadanych dowcipach i tym samym komunikowanie przynależności do danej grupy<sup>12</sup>. Ponadto, poczucie humoru ma również znaczenie dla doboru płciowego. Mężczyźni bieglej się nim posługujący są bardziej atrakcyjni w oczach kobiet, co może się wiązać ze wspomnianym sygnalizowaniem sprawności umysłowej<sup>13</sup>.

## 2.2. Rozluźniająca funkcja poczucia humoru

W poprzednich akapitach celowo pominięta została jedna z dwóch głównych, powstałych w XVIII wieku teorii stojących w kontrze do teorii wyższości. Chodzi tu o teorię rozluźnienia (*release theory*) i to jej poświęcony będzie niniejszy podrozdział. Oprócz podstaw filozoficznych wspomniane będą różnorakie badania zajmujące się tą kwestią.

W swojej początkowej, XVIII wiecznej fazie teoria rozluźnienia bazowała na panującym w ówczesnej nauce przekonaniu, że nerwy łączące mózgowie z resztą ciała działały na zasadzie transportu różnorodnych cieczy i gazów zwanymi „duchami zwierzęcymi” (*animal spirits*). Według Lorda Shaftesbury’ego rozbawienie miało służyć ulotnieniu owych duchów i, co za tym idzie, zmniejszeniu ciśnienia w neuronach. Stąd właśnie zmiana znaczenia słowa *humour* w języku angielskim z nawiązującego do medycyny starożytnej Grecji na znane nam obecnie. Wraz z rozwojem nauk neuronalnych szczegóły działania tej teorii zostały oczywiście odstawiona do lamusa. Pierwotna myśl Lorda Shaftesbury’ego na temat rozluźniającej funkcji poczucia humoru przetrwała jednak w dyskursie akademickim przez następne wieki.

Najbardziej chyba znanym myślicielem, który pochylił się nad zagadnieniem uwalniania zakumulowanej energii psychicznej poprzez użycie humoru był Zygmunt Freud. W swoim dziele „*Der Witz und seine Beziehung zum Unbewußten*” („*Dowcip i jego stosunek do nieświadomości*”)<sup>14</sup> postuluje on, że użycie humoru może uwalniać energię tworzącą napięcie na trzy sposoby. Po pierwsze, człowiek pozbywa się w ten sposób energii, która inaczej stłumiłaby emocje wyrażane

<sup>8</sup> Jan van Hooff, „A Comparative Approach to the Phylogeny of Laughter and Smiling”, *Nonverbal Communication*, 1972, 209–41.

<sup>9</sup> Robert Provine, „Laughter: A Scientific Investigation”, *American Scientist*, nr 84 (2000): 38–45.

<sup>10</sup> Christopher Oveis et al., „Laughter Conveys Status”, *Journal of Experimental Social Psychology* 65 (2016): 109–15.

<sup>11</sup> Daniel P. Howrigan i Kevin B. MacDonald, „Humor as a Mental Fitness Indicator”, *Evolutionary Psychology* 6, nr 4 (2008): 653–66.

<sup>12</sup> Thomas J. Flamson i Gregory A. Bryant, „Signals of Humor: Encryption and Laughter in Social Interaction”, w *Developments in Linguistic Humour Theory* (John Benjamins Publishing Company, 2013).

<sup>13</sup> Gil Greengross i Geoffrey Miller, „Humor Ability Reveals Intelligence, Predicts Mating Success, and Is Higher in Males”, *Intelligence* 39, nr 4 (2011): 188–92.

<sup>14</sup> Sigmund Freud, *Dowcip i jego stosunek do nieświadomości*. Tłumaczenie Robert Reszke, tłum. Robert Reszke (Sen: KR, 1993).

właśnie poprzez humor. W manierze typowej dla Freuda podaje on przykład żartów na tematy seksualne jako ujście dla energii libido poprzez obejście zinternalizowanej blokady. Po drugie, za pomocą rozbawienia pozbyć można się energii zebranej na potrzeby myślenia, która okazuje się niepotrzebna. Po trzecie, tym sposobem uchodzi również energia emocji, które później okazują się nieadekwatne. Przykładowo współczucie komuś w sytuacji, która okazuje się kończyć bez żadnej szkody dla niego może być powodem do rozbawienia. Taki model, posługujący się enigmatycznym pojęciem energii, nie przystaje do współczesnej wiedzy psychologicznej i neurologicznej. Istnieją jednak dowody na istnienie rozluźniającej funkcji humoru w pewnych sytuacjach.

Badania socjologiczne wykazały, że użycie treści humorystycznych w interakcjach społecznych może pomóc w łagodzeniu konfliktów<sup>15</sup>. Skuteczność takiej taktyki zależy od rodzaju humoru i statusu poszczególnych członków grupy, co nawiązuje do koncepcji przedstawionych w poprzednim podrozdziale. Ponadto, śmianie się w sytuacjach społecznych podnosi próg bólu<sup>16</sup>, a wysoki poziom poczucia humoru jest skorelowany z niższym poziomem ogólnego stresu i lęku<sup>17</sup>. Humor może być narzędziem do radzenia sobie z przytłaczającą rzeczywistością, funkcjonuje wtedy jako zawór bezpieczeństwa i chroni przed negatywnymi konsekwencjami rozumowania. Innymi słowy, nie pozwala na poważne analizowanie rzeczy, które nie powinny być w ten sposób pojmowane<sup>18</sup>. Wysoki poziom poczucia humoru u jednostki podnosi samoocenę, a także pozwala reagować w bardziej pozytywny sposób nie tylko na pozytywne, ale i negatywne wydarzenia w życiu<sup>19</sup>.

Rozluźniającą funkcję poczucia humoru można jednak rozumieć bardziej dosłownie niż prezentują to przytoczone dotychczas przykłady. Badania wykazują bowiem, że podczas śmiechu obniża się pobudliwość mięśniowa<sup>20</sup>, zmniejsza się więc zdolność jednostki do podejmowania jakichkolwiek działań. Efekt ten jest mocniejszy w przypadku śmiechu szczerego, ale występuje również przy śmiechu nieszczerym. Późniejsze badanie wykazało, iż efekt ten wynika rzeczywiście z samego stanu rozbawienia, a nie jedynie z fizjologicznych procesów związanych ze śmianiem się<sup>21</sup>.

### 2.3. Efekt *babyface*

Efekt *babyface* (efekt twarzy niemowlęcia) po raz pierwszy został opisany i scharakteryzowany przez etologa Konrada Lorenza pod niemiecką nazwą *Kindchenschema* (schemat dziecięcości), co w literaturze anglojęzycznej często tłumaczone jest bezpośrednio jako *baby-schema*. Opisał on zestaw cech fizjonomicznych, które powodują, że dana twarz jest postrzegana jako bardziej urocza niż twarz pozbawiona takich cech i wywołuje u ludzi odruchy opiekuńcze, a także przyciąga uwagę. Wspomniane cechy upodobniają daną twarz do ludzkiego niemowlęcia, więc są to przykładowo

---

<sup>15</sup> Neal R. Norrick i Alice Spitz, „Humor as a Resource for Mitigating Conflict in Interaction”, *Journal of Pragmatics* 40, nr 10 (2008): 1661–86.

<sup>16</sup> Robin I. M. Dunbar et al., „Social Laughter Is Correlated with an Elevated Pain Threshold”, *Proceedings of the Royal Society B: Biological Sciences* 279, nr 1731 (2012): 1161–67.

<sup>17</sup> Millicent H. Abel, „Humor, Stress, and Coping Strategies”, *HUMOR* 15, nr 4 (2002): 365–81.

<sup>18</sup> Wallace Chafe, „Humor as a disabling mechanism”, *American Behavioral Scientist* 30, nr 3 (1987): 16–25.

<sup>19</sup> Rod A. Martin et al., „Humor, Coping with Stress, Self-Concept, and Psychological Well-Being”, *Humor: International Journal of Humor Research* 6, nr 1 (1993): 89–104.

<sup>20</sup> Sebastiaan Overeem et al., „Weak with Laughter”, *The Lancet* 354, nr 9181 (1999): 838.

<sup>21</sup> Sebastiaan Overeem et al., „Is Motor Inhibition during Laughter Due to Emotional or Respiratory Influences?”, *Psychophysiology* 41, nr 2 (2004): 254–58.

stosunkowo duże oczy, niewielki nos, pełne usta i wąski podbródek<sup>22</sup>. Koncepcję tę udało się potwierdzić metodą eksperymentalną. Mimo, iż badacze nie doszli do konsensusu w sprawie jednolitej listy cech fizjonomicznych składających się na wywoływanie *Kindchenschema*, wyniki licznych badań potwierdzają teorię Lorenza.

Istoty niebędące ludzkimi dziećmi, ale posiadające wspomniane cechy wyzwalają w ludziach odruchy podobne do takich, jakie mieliby oni w przypadku kontaktu z prawdziwym niemowlęciem<sup>23</sup>. Fenomen ten jest obserwowalny nie tylko w przypadku, gdy to dorośli są obserwatorami, ale również już u trzyletnich dzieci<sup>24</sup>. *Kindchenschema* może więc charakteryzować twarze zwierząt, bohaterów kreskówek, zabawek, jak również, co ważne w kontekście niniejszej pracy, dorosłych ludzi<sup>25</sup>. Efekt ten, z uwagi na jednoznacznie pozytywny afekt *babyface*, często wykorzystywany jest w branży marketingowej. Z tego też powodu duża liczba prac badawczych na jego temat wykonywana jest właśnie z perspektywy i na potrzeby przemysłu reklamowego, przykładowo<sup>26</sup>.

Biorąc pod uwagę przytoczone przeze mnie właściwości efektu *babyface*, jak i założenia teorii o rozluźniającym działaniu humoru dostrzegam między nimi wspólny mianownik w postaci ogólnego uspokojenia organizmu i zogniskowaniu na sobie uwagi odbiorcy. Na tej bazie sformułowałem hipotezę będącą podstawą pod projekt badawczy, która głosi, że dzięki wspomnianej funkcji rozluźniającej twarze ze znamionami *babyface* będą w kontekście komicznym wydawały się badanym zabawniejsze niż inne twarze w takim samym kontekście.

## 2.4. Wykorzystanie EDA i EMG w badaniu emocji

Do roli narzędzia badającego poziom rozbawienia badanych wybrane zostało urządzenie badające sygnały EDA i EMG. Wskaźniki te pozwalają na monitorowanie takich zmiennych fizjologicznych jak przewodnictwo skórne i aktywność mięśniowa. Słusznym jest jednak poddawanie w wątpliwość stopnia korelacji pomiędzy wspomnianymi markerami a rzeczywistym występowaniem u badanego danych emocji. Psychologia emocji, jak również neuronalne i fizjologiczne korelaty emocjonalności nadal kryją przed badaczami wiele tajemnic i bez wątpienia nie da się ich sprowadzić do dwóch zmiennych mierzonych na powierzchni ciała. Mając to na uwadze, poziom pobudzenia i mimika twarzy są w jakiś sposób powiązane z emocjonalnością człowieka i są wykorzystywane do badania stanu emocjonalnego osób w sposób ilościowy. Nie występuje w ich przypadku konieczność polegania na relacji badanego, która siłą rzeczy może dotyczyć tylko dostatecznie silnych i uświadomionych emocji. Jest to więc, mimo wszystko, przydatne narzędzie skorzystano z dostępności do odpowiedniego sprzętu na uniwersytecie.

---

<sup>22</sup> Konrad Lorenz, „Die angeborenen Formen moeglicher Erfahrung”, *Zeitschrift für Tierpsychologie* 5, nr 2 (1943): 235–409.

<sup>23</sup> Stephen Jay Gould, „Mickey Mouse meets Konrad Lorenz”, *Nat Hist* 88 (1979): 30–36.

<sup>24</sup> Marta Borgi et al., „Baby Schema in Human and Animal Faces Induces Cuteness Perception and Gaze Allocation in Children”, *Frontiers in Psychology* 5, nr 411 (2014).

<sup>25</sup> Leslie A. Zebrowitz et al., „Neural Activation to Babyfaced Men Matches Activation to Babies”, *Social Neuroscience* 4, nr 1 (2009): 1–10.

<sup>26</sup> Alexis Yim et al., „Do Salespeople’s Online Profile Pictures Predict the Number of Online Reviews? Effect of a Babyface”, *European Journal of Marketing* 57, nr 7 (2023): 1886–911.

Sygnal EDA jest miarą aktywności elektrodermalnej (sam skrót EDA pochodzi od angielskiego *electrodermal activity*), czyli aktywności elektrycznej skóry. W ramach pomiaru EDA bierze się pod uwagę zarówno aspekt toniczny, jak i fazowy aktywności elektrodermalnej. Aktywność toniczna jest stosunkowo stała i nie obserwuje się na jej wykresie reakcji badanego na bodziec, natomiast aktywność fazowa jest krótkotrwała i da się na jej podstawie obserwować reakcje w czasie trwania bodźca<sup>27</sup>. Z tego względu w procesie analizy opisanego później badania brano pod uwagę była jedynie aktywność fazowa. EDA jest bardzo popularnym sposobem pomiaru poziomu pobudzenia emocjonalnego i procesów uwagowych<sup>28</sup>. Pocenie się (w szczególności) dłoni, które jest efektem pobudzenia emocjonalnego lub stanu skupienia uwagi, zwiększa przewodnictwo na powierzchni skóry i jest zauważalne na sygnale EDA. Aktywność elektrodermalna może informować jednak jedynie o poziomie pobudzenia emocjonalnego, a nie o jego pozytywnym lub negatywnym nacechowaniu. Dlatego w projekcie rejestrowany również sygnał EMG na mięśniach mimicznych twarzy.

Skrótem EMG określa się elektromiografię, czyli badanie elektrycznej aktywności mięśni szkieletowych<sup>29</sup>. Istnieją różne techniki wykonywania pomiaru EMG, z których najdokładniejszy jest pomiar z użyciem elektrod igłowych wkłuwanych domięśniowo. W niniejszym projekcie użyty został pomiar powierzchniowy, przy którym wykorzystuje się nieinwazyjną metodę przyklejania elektrod na powierzchnię skóry bezpośrednio nad badanym mięśniem<sup>30</sup>.

Badanie sygnału EMG na mięśniach mimicznych jest jedną z najpopularniejszych metod badań psychofizjologicznych. Nie ulega wątpliwości, że w gatunku ludzkim wyraz twarzy, czyli specyficzne układy napięcia i rozluźnienia mięśni mimicznych, jest bardzo silnie skorelowany z emocjonalnością. Tematyka neurologicznego podłoża ekspresji emocji za pomocą mimiki oraz dedukcji o stanach mentalnych na jej podstawie jest oczywiście niesamowicie obszerna. W badaniach elektromiograficznych mimiki przyjęło się badać sygnał z dwóch mięśni twarzy, z czego jeden z nich jest miernikiem afektu pozytywnego, drugi z kolei afektu negatywnego. Są to, odpowiednio, mięsień jarzmowy większy (*zygomaticus major*) oraz mięsień marszczący brwi (*corrugator supercilii*)<sup>31</sup>. Taka aproksymacja jest w zupełności wystarczająca na potrzeby niniejszego badania i nad tymi właśnie mięśniami umieszczone zostały elektrody.

Szczegółowy opis wykorzystania elektrod do pomiaru EDA i EMG na potrzeby przeprowadzonego badania znajduje się w podrozdziale 3.4 (s. 16).

### 3. Badanie

---

<sup>27</sup> Wolfram Boucsein, *Electrodermal activity* (Springer Science & Business Media, 2012).

<sup>28</sup> Michael E. Dawson et al., „The electrodermal system”, w *Handbook of psychophysiology*, t. 2 (2007).

<sup>29</sup> Tytus Sosnowski, „Psychofizjologia”, w *Psychologia: Podręcznik akademicki*, t. 1 (Gdańskie Wydawnictwo Psychologiczne, 2000).

<sup>30</sup> Carlo J. De Luca et al., „Decomposition of Surface EMG Signals”, *Journal of Neurophysiology* 96, nr 3 (2006): 1646–57.

<sup>31</sup> John T. Cacioppo et al., „Electromyographic Activity over Facial Muscle Regions Can Differentiate the Valence and Intensity of Affective Reactions”, *Journal of Personality and Social Psychology* 50, nr 2 (1986): 260; Gerger Gernot et al., „Empathy, Einfühlung, and Aesthetic Experience: The Effect of Emotion Contagion on Appreciation of Representational and Abstract Art Using fEMG and SCR”, *Cognitive Processing* 19 (2018): 147–65.

### 3.1. Hipoteza

Hipoteza, którą zweryfikować ma niniejsze badanie opiera się na opisanych wcześniej teoriach o rozluźniającej funkcji poczucia humoru oraz koncepcji *Kindchenschema* Konrada Lorenza. Eksperyment ma na celu sprawdzenie, czy ze względu na aspekt rozluźnienia (*release*) występujący w obydwu tych teoriach da się zauważyć istotną korelację pomiędzy posiadaniem przez daną, prezentowaną w komediowym kontekście, twarz cech twarzy dziecka a uznaniem jej za bardziej zabawną, niż inne prezentowane twarze.

**Hipoteza 0:** Nie ma istotnej różnicy pomiędzy twarzami posiadającymi cechy *Kindchenschema*, a twarzami nieposiadającymi ich, jeśli chodzi o ocenę ich zabawności w komediowym kontekście.

**Hipoteza 1:** Twarze charakteryzujące się *Kindchenschema* prezentowane w komediowym kontekście są zabawniejsze niż twarze pozbawione takich cech.

### 3.2. Grupa badawcza

Przed przeprowadzeniem eksperymentu na większej grupie badawczej wykonane zostało badanie pilotażowe na grupie 8 osób. Miało ono na celu lepsze zapoznanie się ze sprzętem wykorzystanym do pomiaru EDA i EMG oraz zidentyfikowanie błędów w przygotowanym programie do wyświetlania bodźców. W jego efekcie pula przygotowanych materiałów została zmniejszona o jedną trzecią, gdyż uczestnicy zwracali uwagę na znużenie oglądaniem twarzy pod koniec eksperymentu, co było zauważalne na zmniejszających się amplitudach sygnału EDA. Z tego względu osoby te nie wchodziły w zbiór grupy badawczej.

Właściwa grupa badawcza liczyła 30 osób, lecz ze względu na duże zakłócenia w sygnale (a w niektórych przypadkach ze względu na niepoprawne przygotowanie sprzętu) 8 z nich nie była brana pod uwagę przy analizie. Z 22 pozostałych osób wszystkie mieściły się w przedziale wiekowym od 18 do 25 lat, 16 (72,7%) stanowiły kobiety, 5 (22,7%) stanowili mężczyźni, a jedna osoba (4,6%) zadeklarowała się jako osoba niebinarna. Zebrane zostały również dane na temat spożycia kofeiny przed badaniem, gdyż może ona wpłynąć chociażby na sygnał EDA i habituację oglądania bodźców<sup>32</sup>, jednak ze względu na małą dostępność osób chętnych do wzięcia udziału w badaniu czynnik ten nie został ostatecznie wzięty pod uwagę.

### 3.3. Przygotowane materiały

Badanie przeprowadzone zostało w laboratorium wyposażonym między innymi w komputer z dwoma monitorami oraz urządzenie marki PLUX Biosignals służące do rejestracji różnorodnych sygnałów fizjologicznych przy użyciu elektrod. Badani proszeni byli o zajęcie miejsca przy jednym z monitorów, gdzie po podłączeniu elektrod wyświetlana była prezentacja z przygotowanymi obrazami przedstawiającymi twarze. Dokładniejszy opis procesu montowania elektrod zamieszczony został w następnym podrozdziale (s. 16), ten natomiast skupia się na przygotowanej prezentacji. Na

---

<sup>32</sup> Robyn A. Davidson i Barry D. Smith, „Caffeine and Novelty: Effects on Electrodermal Activity and Performance”, *Physiology & Behavior* 49, nr 6 (1991): 1169–75.

drugim monitorze równocześnie wyświetlał się program OpenSignals, który w czasie rzeczywistym tworzył wykresy rejestrowanych sygnałów.

Bodźce wizualne, a więc obrazy przedstawiające twarze, przygotowane zostały przy wykorzystaniu strony internetowej <https://thispersonnotexist.org/> wykorzystującej model SI StyleGan3<sup>33</sup>. Pozwoliło to na jednolitość stylu, oświetlenia, wymiarów i rozdzielczości materiałów przy bardzo niewielkim wysiłku. Badani zostali uświadomieni, że żadne z oglądanych przez nich zdjęć nie przedstawia prawdziwej osoby.

Łącznie przygotowano zostały 24 obrazy. Osiem z nich po wygenerowaniu nie zostało w żaden sposób zmodyfikowanych i przedstawiało twarze nieodbiegające w znaczący sposób od normy (fig. 1). Bodźce z tej grupy stanowiły punkt odniesienia względem pozostałych, zmodyfikowanych obrazów. Pozostałe materiały poddane zostały obróbce graficznej przy użyciu programu GIMP. W celu nadania im charakteru komediowego twarze na nich przedstawione zostały przerobione na karykatury, a więc niektóre ich cechy zostały wyolbrzymione. W przypadku ośmiu karykatur były to cechy niezwiązane z *Kindchenschema* (fig. 2), natomiast u pozostałych ośmiu to właśnie te cechy zostały wyolbrzymione (fig. 3). W każdej z trzech grup połowa przygotowanych obrazów twarzy to twarze kobiece, a połowa przedstawia twarze męskie. Wszystkie twarze wyglądają jak osoby dorosłe w różnym wieku.

Prezentacja przygotowana została w programie lab.js. Służy on do konstruowania eksperymentów psychologicznych ze szczególnym uwzględnieniem nauk behawioralnych i kognitywnych<sup>34</sup>. Program ten oparty jest na językach HTML i JavaScript, ale posiada wizualny interfejs pozwalający na intuicyjne zaprojektowanie ogólnego kształtu prezentacji oraz jej najprostszych funkcji takich jak wyświetlanie tekstu czy stworzenie pętli wyświetlania obrazów. Niektóre z funkcji, które są możliwe do zaprogramowania w językach HTML i JavaScript nie mają swojej reprezentacji na wizualnym interfejsie, istnieje jednak opcja samodzielnego dodania ich do badania poprzez wpisanie linijek kodu w odpowiednich oknach w programie.

Wyświetlana badanym prezentacja składała się z ekranu instrukcji, powtarzającej się 24 razy pętli wyświetlającej obrazy i pytanie o subiektywny poziom rozbawienia wywołany daną twarzą oraz ekranu z podziękowaniem za wzięcie udziału w badaniu.

W instrukcji, którą badany widział na początku zawarte były informacje na temat ogólnej struktury badania (a więc że wyświetlana będzie sekwencja obrazów przedstawiających ludzkie twarze) oraz na temat systemu oceny zabawności danej twarzy. Badany był również informowany, że twarze zostały wygenerowane przez model SI i nie przedstawiają rzeczywistych osób. W razie braku wątpliwości i ewentualnych pytań badany samodzielnie przechodził do dalszej części eksperymentu za pomocą klawisza spacji.

Każda iteracja pętli rozpoczynała się od ekranu, na którym widniał napis „Naciśnij dowolny klawisz na klawiaturze”. Miało to na celu zapobiegnięcie sytuacji, w której na monitorze wyświetla

---

<sup>33</sup> Tero Karras et al., „Alias-Free Generative Adversarial Networks”, w *Advances in Neural Information Processing Systems*, t. 34, red. M. Ranzato et al. (Curran Associates, Inc., 2021), [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/076ccd93ad68be51f23707988e934906-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/076ccd93ad68be51f23707988e934906-Paper.pdf).

<sup>34</sup> Felix Henninger et al., „lab.js: A free, open, online study builder”, *Behavior Research Methods* 54, nr 2 (2022): 556–73.

się już kolejny bodziec w momencie, gdy badany dalej jest pod emocjonalnym wpływem poprzedniego. Następnie na środku monitora wyświetlał się przez 500 ms krzyż fiksacji, po którym następowała prezentacja bodźca. Kolejność wyświetlanych obrazów została ustalona w sposób losowy przed zaprogramowaniem eksperymentu i była taka sama dla każdego z badanych. Czas wyświetlania się bodźca został ustawiony na 3 s, po których następowała trwająca 800 ms przerwa, podczas której nic nie wyświetlało się na ekranie. Okres przerwy służył wygaszeniu fazowego sygnału EDA przed pojawieniem się następnego ekranu. Na nim badany proszony był o subiektywną ocenę zabawności zobaczonej przed chwilą twarzy (fig. 4). Ocena przebiegała na zasadzie ustawienia suwaka w wybranym miejscu za pomocą strzałek na klawiaturze na arbitralnej skali od „neutralne” przez „całkiem zabawne” po „wyjątkowo zabawne”. W celu umożliwienia analizy odpowiedzi rejestrowane były na ilościowej skali od 1 do 100 punktów. Po wybraniu odpowiedzi badany zatwierdzał ją przyciskiem spacji, co uruchamiało następną iterację pętli (lub, oczywiście, uruchamiało ekran z podziękowaniem w przypadku 24. iteracji).

Kategoria subiektywnej i świadomej oceny została wprowadzona do eksperymentu ze względu na ograniczenia wykorzystanych metod badania psychofizjologicznego. Jak opisałem wcześniej, sygnały EDA i EMG mogą informować o poziomie pobudzenia emocjonalnego i o pozytywnym lub negatywnym nacechowaniu reakcji emocjonalnej, nie są jednak w stanie jednoznacznie wskazać na rodzaj odczuwanej przez badanego emocji.

Ekran końcowy wyświetlał tekst zawierający podziękowanie za wzięcie udziału w badaniu i prośbę o poinformowanie prowadzącego o zakończeniu eksperymentu. Po naciśnięciu przycisku k na klawiaturze na ekranie pojawiała się opcja pobrania arkusza z danymi na temat danej sesji. Czynność tę wykonywał przeprowadzający badanie a uczestnicy nie byli o niej informowani. Poniżej zamieszczam wykres przedstawiający przebieg prezentacji na przykładzie pojedynczej iteracji pętli wraz z czasem wyświetlania ekranów (fig.5).

Tak zaprojektowany w programie lab.js eksperyment został wyeksportowany do plików o rozszerzeniach .js oraz .html, co pozwoliło uruchomić go na komputerze laboratoryjnym przy użyciu przeglądarki bez potrzeby łączenia się z siecią. Po zakończeniu danej sesji program sporządzał arkusz kalkulacyjny o rozszerzeniu .csv, który zawierał między innymi dane na temat czasu wyświetlania każdego z ekranów oraz wybranych przez uczestników ocen dla poszczególnych bodźców. Czas trwania właściwego badania wahał się w zależności od uczestnika, zazwyczaj wynosił jednak około czterech minut. Dzięki temu udało się uniknąć efektu habituacji bodźców, o którym wspomniałem wcześniej.

### **3.4. Użycie elektrod**

Do rejestracji sygnałów EDA i EMG wykorzystane zostało urządzenie marki PLUX Biosignals, do którego podłączone zostały przewody z przyłączonymi elektrodami. Urządzenie to łączyło się z komputerem bezprzewodowo poprzez standard Bluetooth. Do rejestracji i zapisu odbieranych sygnałów użyte zostało dedykowane oprogramowanie OpenSignals (r)evolution. Maksymalna liczba kanałów obsługiwanych przez urządzenie wynosi osiem (nie licząc kanału dla elektrody referencyjnej), z czego w opisywanym projekcie użytych zostały cztery. Widoczny na fig. 6 przewód w dolnej części urządzenia służył do jego ładowania i nie był podpięty w trakcie trwania sesji badania.

- Kanał 1: LUX
- Kanał 2: EDA
- Kanał 3: EMG nad mięśniem jarzmowym większym
- Kanał 4: EMG nad mięśniem marszczącym brwi

Do kanału pierwszego wpięty był przewód zakończony czujnikiem światła. Zastosowanie zarejestrowanego przez niego poziomu natężenia światła w czasie ma charakter pragmatyczny, służył on bowiem do synchronizacji czasowej pozostałych wykresów na etapie analizy oraz jako marker do oznaczania na wykresie czasu pojawienia się bodźców z danej grupy. Podczas wyświetlania ekranu z bodźcem w lewym górnym rogu pojawiał się kwadrat w kolorze czarnym, szarym lub jasnoszarym, w zależności od rodzaju bodźca. Czujnik zamontowany był do monitora w tym właśnie miejscu.

Do kanału drugiego wpięty został przewód dedykowany rejestracji sygnału EDA. Posiada on dwa zakończenia, do których można przymocować elektrody. W tym celu wykorzystane zostały elektrody Ag/AgCl (chlorosrebrowe) przymocowane do rzepów marki Biosignals. Elektrody te umieszczone zostały na opuszkach palców wskazującego i środkowego niedominującej ręki. Tego typu elektrody nie wymagają nakładania żelu i mogą być stosowane wielokrotnie. Nie ma konsensusu w sprawie wpływu mycia rąk przez badanego przed rejestracją sygnału EDA w opisany wyżej sposób, w niniejszym badaniu uczestnicy nie byli więc o to proszeni. Jednostką pomiarową EDA jest mikrosiemens ( $\mu\text{S}$ ).

Kanały trzeci i czwarty obsługiwały sygnały EMG. Przewody dedykowane rejestracji sygnału EMG również posiadają po dwa zakończenia, dlatego do rejestracji używane były cztery elektrody. Były to samoprzylepne, jednorazowego użytku elektrody piankowe Ag/AgCl o średnicy 24 mm marki Kendall H124SG. Pokryte są one przewodzącym hydrożelem, przez co nie jest konieczne stosowanie dodatkowego żelu. W badaniu w charakterze elektrody referencyjnej również wykorzystane zostały takie właśnie elektrody. Kanał trzeci obsługiwał sygnał EMG rejestrowany nad mięśniem jarzmowym większym, czwarty natomiast nad mięśniem marszczącym brwi. W te miejsca przyklejane były po dwie elektrody możliwie jak najbliżej siebie, a więc w odległości mniej więcej 20 mm. Elektroda referencyjna przyklejana była w górnej części czoła jak najbliżej linii włosów. Umieszczenie elektrod ilustruje fig. 8. Zarejestrowany sygnał wyrażany jest w miliwoltach (mV).

Po przyklejeniu wszystkich elektrod w odpowiednich miejscach robiono pięciominutową przerwę, podczas której uczestnicy proszeni byli o zachowanie względnego bezruchu. Umożliwiło to lepsze związanie żelu adhezyjnego i zmniejszenie ryzyka odklejenia się elektrod w trakcie trwania właściwego badania. Sygnały EDA i EMG próbkowane były podczas rejestracji z częstotliwością 2000 Hz, a do pomiaru aktywności elektrodermalnej wykorzystany był prąd o częstotliwości 16 Hz.

Przed wyświetleniem prezentacji i przystąpieniem do właściwego badania przeprowadzany był test sprzętu. Po połączeniu urządzenia PLUX Biosignals z komputerem oraz uruchomieniu programu OpenSignals badani proszeni byli kolejno o wzięcie głębokiego oddechu, uśmiechnięcie się oraz zmarszczenie brwi. Dopiero w przypadku braku jakichkolwiek nieprawidłowości widocznych na wykresach kończono test, rozpoczynano właściwą rejestrację sygnałów i uruchamiano wyświetlanie prezentacji na monitorze uczestnika badania. Pomimo to w puli wyników końcowych znalazły się przypadki skrajne, które trzeba było wykluczyć z analizy.

Po zakończeniu testu badani proszeni byli o zajęcie możliwie jak najwygodniejszej pozycji ciała, gdyż ruch w trakcie badania mógłby pozostawić wyraźny ślad na sygnałach EDA i EMG. Dłoń niedominującą z przymocowanymi za pomocą rzepów elektrodami do pomiaru EDA osoby badane mogły trzymać w bezruchu na blacie biurka lub własnym kolanie. Z uwagi na bardzo krótki czas trwania prezentacji uczestnicy na ogół nie mieli trudności z zastosowaniem się do tych wymagań.

W trakcie trwania badania sygnały z wszystkich czterech kanałów wyświetlały się w czasie rzeczywistym na wykresach w programie OpenSignals (r)evolution. Były one widoczne jedynie na monitorze osoby prowadzącej badanie. Po otrzymaniu od badanego informacji o zakończeniu prezentacji nagrywanie zostawało zatrzymane, a dane psychofizjologiczne zapisywane były w formacie EDF. Tak jak wspomniano wcześniej, pobierane również były dane na temat prezentacji wraz z ocenami wprowadzonymi przez badanych w arkuszach CSV. Demontaż elektrod wymagał jedynie odpięcia rzepów i odklejenia elektrod z twarzy, w dodatku żel adhezyjny był na tyle silny, że zbyt gwałtowne odklejenie elektrod mogłoby wywołać u uczestnika duży dyskomfort. Z tych powodów badani byli proszeni o zrobienie tego we własnym zakresie.

### 3.5. Przygotowanie danych do analizy

Dane zebrane w programie OpenSignals (r)evolution zapisane w formacie EDF poddane zostały obróbce w programie AcqKnowledge. Przed przystąpieniem do obróbki pliki konwertowane były do plików .acq, gdyż na plikach o takim rozszerzeniu pracuje się w programie AcqKnowledge.

Sygnał LUX został przeskalowany za pomocą funkcji *Rescale* w taki sposób, aby możliwe było automatyczne oznaczenie czasu pojawienia się bodźców na ekranie. Skalowanie to w głównej mierze opierało się na zamianie wartości z ujemnych na dodatnie i odwrotnie, aby znaczniki pojawienia się obrazu twarzy jawiły się jako wzrost natężenia światła a nie jego spadek. Oznaczanie bodźców odbyło się przy wykorzystaniu funkcji *Digital Input to Stim Events*, która tworzyła nowy marker bodźca w każdym miejscu, w którym rosnący wykres natężenia światła przekraczał wskazaną wartość. Następnie każdemu z owych markerów przypisywana była wartość od 1, 2 lub 3 odpowiadająca przynależności bodźca do grupy 1, 2 lub 3 (fig. 1-3).

Sygnał EDA został poddany decymacji, wygładzeniu, przefiltrowaniu, a finalnie wydzielono sygnał fazowy od tonicznego i to na nim przeprowadzono właściwą analizę. Decymacja odbyła się przy użyciu funkcji *Resample Waveform*, która pozwoliła na zmniejszenie liczby próbek na sekundę z oryginalnie rejestrowanych 2000 na 15,625. Wygładzenie zostało przeprowadzone na podstawie mediany pięciu kolejnych próbek. Następnie zastosowano filtrowanie dolnoprzepustowe o skończonej odpowiedzi impulsowej (FIR) przy częstotliwości odcięcia 1,5 Hz. Filtrowanie górnoprzepustowe przeprowadzone zostało podczas wyodrębniania sygnału fazowego od tonicznego i ustawione zostało na 0,5 Hz. Powstały w ten sposób wykres zmian fazowego EDA w czasie przechowywany był na nowym, osobnym od tonicznego EDA kanale w programie.

Sygnały z obydwu kanałów EMG zostały poddane identycznemu procesowi obróbki. Najpierw zastosowany został filtr środkowoprzepustowy o przedziale częstotliwości od 20 do 500 Hz. Następnie zmieniono wartości ujemne na dodatnie, co określa się mianem prostowania. Ostatnim krokiem było wygładzenie wykresu, przy czym tym razem wykorzystano wartość średnią.

### 3.5. Analiza

Analiza danych fizjologicznych odbyła się w programie AcqKnowledge z wykorzystaniem funkcji *Stim Response Analysis*. Dzięki wcześniejszemu wykorzystaniu sygnału LUX do automatycznego stworzenia markerów czasu pojawienia się bodźców i odpowiedniemu ich pogrupowaniu analiza wymagała jedynie określenia długości trwania okna, dla którego sprawdzane będą wybrane parametry i opóźnienia analizy po rozpoczęciu wyświetlania bodźca. Opóźnienie to ustawione zostało na 0.2 s, aby uwzględnić czas reakcji badanych i czas potrzebny na przetworzenie prezentowanych materiałów. Długość okna analizy wynosiła 3 s, gdyż tyle wynosił czas wyświetlania bodźców na ekranie.

Dla każdego przedziału mierzone były:

- Średnia wartość fazowego EDA z utworzonego kanału 6,
- Maksymalna wartość EMG nad mięśniem jarzmowym większym z kanału 3,
- Maksymalna wartość EMG nad mięśniem marszczącym brwi z kanału 4.

Wyniki analizy zapisywane były w arkuszu kalkulacyjnym o rozszerzeniu *.ods*. Pliki AcqKnowledge, na których wykonywana była obróbka i analiza nie były zapisywane, gdyż nadpisałoby to oryginalne pliki EDF, przez co niemożliwa byłaby powtórna analiza w przypadku zauważenia błędu. Plik zawierający wyniki został przeformatowany w taki sposób, aby możliwe było bezpośrednio zaimportowanie go do programu do analizy statystycznej. Dodana została również kolumna zawierające subiektywne oceny zabawności twarzy podane przez uczestników w trakcie badania zapisane w plikach z wynikami z programu *lab.js*.

Każdy z wyników został zaetykietowany na skali nominalnej, co przydzieliło go do jednego z trzech zbiorów: Norm, Mod oraz Baby. Zbiór Norm zawierał wyniki zarejestrowane przy bodźcach z grupy pierwszej, zbiór Mod z grupy drugiej, zbiór Baby natomiast z grupy trzeciej.

## 4. Wyniki

Po obróbce sygnałów i poddaniu wybranych przedziałów analizie dane z arkusza kalkulacyjnego zaimportowane zostały do programu Jasp. Służy on do przeprowadzania analiz statystycznych na zbiorach danych. Kolejne kolumny zawierały zbiór wyników wszystkich uczestników grupy badawczej. Zamieszczona poniżej lista przedstawia etykiety nadane poszczególnym kolumnom.

EDA – średnia wartość fazowego sygnału EDA dla określonego okna czasowego wyrażona w mikrosiemensach.

EMGp – maksymalna wartość sygnału EMG mierzonego nad mięśniem jarzmowym większym (p od słowa policzek, gdzie umieszczana była elektroda) dla określonego okna czasowego wyrażona w miliwoltach.

EMGb – maksymalna wartość sygnału EMG mierzonego nad mięśniem marszczącym brwi (b od słowa brew, nad którą umieszczana była elektroda) dla określonego okna czasowego wyrażona w miliwoltach.

Ocena – średnia wartość ocen zabawności przyznanych przez uczestników badania bodźcom z danej grupy.

Tak nazwane zbiory danych poddane zostały analizie statystycznej. Narzędziem wybranym do jej realizacji był test t dla niezależnych wyników. Pozwala on na sprawdzenie, czy pomiędzy średnimi wartościami dwóch zbiorów istnieje istotna statystycznie różnica („SPSS tutorials: Independent samples t test.” 2025). Porównywane będą zawsze dwa zbiory danych, a odpowiadające sobie bodźce (np. pierwszy bodziec z grupy pierwszej i pierwszy bodziec z grupy drugiej) nie są ze sobą skorelowane, dlatego właśnie test t dla niezależnych wyników wydał się najlepszą opcją.

#### 4.1. EDA i EMG

Jako pierwsze przeanalizowane zostało, czy zmodyfikowane obrazy ze zbiorów Mod i Baby rzeczywiście wydały się badanym zabawniejsze, niż zwykle twarze z grupy Norm. Jako hipotezę alternatywną przyjęto więc, że dla kolumn EDA i EMGp wartości zbiorów Mod i Baby są istotnie większe od tych ze zbioru Norm i odwrotnie dla kolumny EMGb. Hipotezę alternatywną udało się potwierdzić dla kolumny EDA ( $p = 0.02$ ), jednak nie dla obydwu kolumn z wynikami EMG ( $p = 0.116$  dla EMGp oraz  $p = 0.592$  dla EMGb). Był to pierwszy przejaw problemu, który ujawnił się we wszystkich kolejnych testach. Różnice między wynikami EMG okazały się nieistotnie statystycznie w każdym przypadku. Dociekania na temat możliwych powodów takiego stanu rzeczy zawarte są w rozdziale piątym (s. ).

W następnej kolejności przeprowadzony został test t dla niezależnych wyników między zbiorami Mod oraz Baby, a więc wyników na podstawie bodźców z grup drugiej i trzeciej. Wykazanie istotnej statystycznie wyższości wyników dla zbioru Mod stanowiłoby potwierdzenie hipotezy 1 postawionej w podrozdziale 3.1 (s. 10). Hipoteza alternatywna dla testu t brzmiała, że wyniki z kolumny Baby są istotnie większe niż te z kolumny Mod dla EDA i EMGp i odwrotnie dla EMGb. Poniżej zamieszczone są tabele i wykresy przedstawiające wyniki testu (fig. 9 - 11).

Jak można z nich wyczytać, nie udało się jednoznacznie potwierdzić hipotezy alternatywnej dla żadnej z kolumn. Dla kolumny EDA  $p = 0.061$ , dla EMGp  $p = 0.367$ , a dla EMGb  $p = 0.449$ .

## 4.2. Oceny podane przez badanych

Na podstawie arkuszy CSV pobranych z programu lab.js sporządzona została tabela (tab. 1), na której w poszczególnych kolumnach zapisana jest średnia ocen, które dany uczestnik przypisał bodźcom z danej grupy.

L. P.	Norm	Mod	Baby
1	15	69	78,625
2	6,75	41,25	33,375
3	28,5	51,375	78
4	8	37,5	47,625
5	18,75	59,125	72,875
6	18,25	68	79,375
7	5,5	50	51,25
8	16	27,875	39,25
9	6,5	47	47,875
10	34	55,75	40,125
11	4,5	24,625	47,375
12	1	43,25	36
13	1,25	42,25	50,25
14	1	7,125	1
15	32,375	49,25	40,875
16	28,125	53,125	74,375
17	22,75	46,75	27,625
18	12,75	71,5	84,25
19	18,375	47,5	56,125
20	21,25	30,75	30,5
21	6,875	24,875	48,375
22	4,625	59,625	38,625
Średnia	<b>14,19</b>	<b>45,80</b>	<b>50,17</b>

Tabela 1: Uśrednione oceny zabawności obrazów podane przez badanych.

Na tabeli widać wyraźnie różnicę w ocenie obrazów niezmodyfikowanych i zmodyfikowanych, karykatury ewidentnie wydały się badanym bardziej zabawne ( $p = <0.001$ ). Różnica między zbiorem Mod i Baby jest już jednak dużo bardziej subtelna, niż w przypadku różnicy fazowych sygnałów EDA ( $p = 0.218$ ).

## 5. Dyskusja

Pierwszym tematem, który poruszony zostanie w dyskusji jest niekonkluzywność otrzymanych wyników pomiaru EMG. Przyczyn takiego stanu rzeczy może być wiele i nie sposób jednoznacznie wskazać na konkretny powód. Jednym z czynników mogło być zastosowanie samoprzylepnych elektrod. Po przyklejeniu na twarz mogą one hamować reakcje mimiczne na bodźce, ale jest to niestety nieodłączny problem badań psychofizjologicznych wykorzystujących tę metodę. Problematyczny mógł być również rozmiar elektrod, gdyż przez swoją średnicę nie mogły one być przyklejone tak blisko siebie, jak jest to robione w innych badaniach. Wykluczam błąd na etapie rejestracji sygnału, gdyż w przypadku pozostałych kanałów przebiegała ona w większości przypadków bezproblemowo. Kolejnym etapem, na którym mogły zostać popełnione błędy była obróbka sygnałów w programie AcqKnowledge. Wadliwa mogła być zarówno przyjęta metoda, jak i przedziały czasowe wykorzystane do analizy. Badania EMG na powierzchni skóry twarzy są sprawdzoną metodą badań psychofizjologicznych, a za nieudaną próbę ich zastosowania w opisywanym eksperymencie winę ponosi wyłącznie autor pracy.

Wyniki rejestracji EDA wydają się z kolei interesujące. Zgodnie z przewidywaniami, fazowy sygnał EDA badanych był istotnie wyższy dla obrazów zmodyfikowanych. Co jednak ważniejsze, zarejestrowane wyniki dla obrazów przedstawiających twarze z cechami *Kindchenschema* również były wyższe od tych pozbawionych takich cech, choć nieco poniżej granicy istotności statystycznej. Wynik analizy statystycznej jest bardzo zbliżony do owej granicy. Pomimo więc, że badanie jest oczywiście nierozstrzygające, to interesującym byłoby zreplicowanie go z poprawnym wykorzystaniem pomiaru EMG, większą i bardziej różnorodną grupą badawczą, po przemodelowaniu procesu obróbki i analizy w programie AcqKnowledge oraz z wzięciem pod uwagę płci i wieku badanych. Odkrycie istotnej statystycznie różnicy w fazowym sygnale EDA na korzyść zbioru Baby byłoby dowodem na silniejszą reakcję emocjonalną badanych na twarze dorosłych noszące znamiona *Kindchenschema*, poprawnie zarejestrowane i przeanalizowane sygnały EMG zaś dostarczyłyby informacji na temat pozytywnego lub negatywnego afektu reakcji.

Wyniki subiektywnych ocen również są niejednoznaczne. Niewielka przewaga na rzecz zbioru Baby okazała się nieistotna statystycznie w teście  $t$ , można by się jednak spierać o jakość przygotowanych bodźców i przyjętą metodologię. Po zakończeniu badania niektórzy z uczestników raportowali, że obrazy twarzy wywoływały u nich niepokój lub poczucie bycia obserwowanym. Szczególnie intensywnie było to widoczne u osoby numer 13, co wyraźnie odbiło się na przyznanej przez nią ocenach (tab. 1).

Podsumowując, badanie przeprowadzone w ramach projektu okazało się nie potwierdzać przyjętej hipotezy, jednak przy lepszym jego przeprowadzeniu wyniki mogłyby okazać się inne i ewentualnie przekroczyć granicę błędu statystycznego. Gdyby po replikacji hipoteza alternatywna nadal pozostałaby niepotwierdzona oznaczałoby to najpewniej, iż nie ma żadnego związku pomiędzy

posiadaniem przez daną twarz cech charakterystycznych dla *Kindchenschema* a odbieraniem jej przez innych jako zabawniejszą.

## 6. Podziękowania

Podczas planowania projektu, przeprowadzania badania i pisanie pracy otrzymałem ogromną pomoc od dra Witolda Marzędy, Krzysztofa Holvika i Mateusza Hadamczyka, za co jestem im bardzo wdzięczny.

## 7. Bibliografia

- Abel, Millicent H. „Humor, Stress, and Coping Strategies”. *HUMOR* 15, nr 4 (2002): 365–81.
- Borgi, Marta, Irene Cogliati-Dezza, Victoria Brelsford, Kerstin Meints, i Francesca Cirulli. „Baby Schema in Human and Animal Faces Induces Cuteness Perception and Gaze Allocation in Children”. *Frontiers in Psychology* 5, nr 411 (2014).
- Boucsein, Wolfram. *Electrodermal activity*. Springer Science & Business Media, 2012.
- Cacioppo, John T., Richard E. Petty, Mary E. Losch, i Hai Sook Kim. „Electromyographic Activity over Facial Muscle Regions Can Differentiate the Valence and Intensity of Affective Reactions”. *Journal of Personality and Social Psychology* 50, nr 2 (1986): 260.
- Chafe, Wallace. „Humor as a disabling mechanism”. *American Behavioral Scientist* 30, nr 3 (1987): 16–25.
- Chiappini, Emilio, Giorgia Silani, Johan N. Lundström, i Sebastian Korb. „Facial Electromyography in Food Research in a Behavioral and MR Setting”. W *Basic Protocols on Emotions, Senses, and Foods*. Springer, 2023.
- Davidson, Robyn A., i Barry D. Smith. „Caffeine and Novelty: Effects on Electrodermal Activity and Performance”. *Physiology & Behavior* 49, nr 6 (1991): 1169–75.
- Dawson, Michael E., Anne M. Schell, i Diane L. Fillion. „The electrodermal system”. W *Handbook of psychophysiology*, t. 2. 2007.
- De Luca, Carlo J., Alexander Adam, Robert Wotiz, L. Donald Gilmore, i S. Hamid Nawab. „Decomposition of Surface EMG Signals”. *Journal of Neurophysiology* 96, nr 3 (2006): 1646–57.
- Deckers, Lambert. „On the Validity of a Weight-Judging Paradigm for the Study of Humor”. *Humor: International Journal of Humor Research* 6, nr 1 (1993): 43–56.
- Dunbar, Robin I. M., Rebecca Baron, Anna Frangou, et al. „Social Laughter Is Correlated with an Elevated Pain Threshold”. *Proceedings of the Royal Society B: Biological Sciences* 279, nr 1731 (2012): 1161–67.
- Flamson, Thomas J., i Gregory A. Bryant. „Signals of Humor: Encryption and Laughter in Social Interaction”. W *Developments in Linguistic Humour Theory*. John Benjamins Publishing Company, 2013.
- Freud, Sigmund. *Dowcip i jego stosunek do nieświadomości. Tłumaczenie Robert Reszke*. Przetłumaczone przez Robert Reszke. Sen: KR, 1993.
- Gernot, Gerger, Matthew Pelowski, i Helmut Leder. „Empathy, Einfühlung, and Aesthetic Experience: The Effect of Emotion Contagion on Appreciation of Representational and Abstract Art Using fEMG and SCR”. *Cognitive Processing* 19 (2018): 147–65.
- Gervais, Matthew, i David Sloan Wilson. „The Evolution and Functions of Laughter and Humor: A Synthetic Approach”. *The Quarterly Review of Biology* 80, nr 4 (2005): 395–430.
- Gould, Stephen Jay. „Mickey Mouse meets Konrad Lorenz”. *Nat Hist* 88 (1979): 30–36.
- Greengross, Gil, i Geoffrey Miller. „Humor Ability Reveals Intelligence, Predicts Mating Success, and Is Higher in Males”. *Intelligence* 39, nr 4 (2011): 188–92.
- Henninger, Felix, Yury Shevchenko, Ulf K. Mertens, Pascal J. Kieslich, i Benjamin E. Hilbig. „lab.js: A free, open, online study builder”. *Behavior Research Methods* 54, nr 2 (2022): 556–73.
- Holvik, Krzysztof. „Reakcje emocjonalne użytkowników języka polskiego na ironię i emotikony: analiza psychofizjologiczna z wykorzystaniem EMG i EDA”. Uniwersytet Śląski, 2024.

- Howrigan, Daniel P., i Kevin B. MacDonald. „Humor as a Mental Fitness Indicator”. *Evolutionary Psychology* 6, nr 4 (2008): 653–66.
- Jach, Łukasz. *Od ucha do ucha. Homo sapiens się śmieje*. Zrozum. Wydawnictwo Poznańskie, 2021.
- Karras, Tero, Miika Aittala, Samuli Laine, et al. „Alias-Free Generative Adversarial Networks”. W *Advances in Neural Information Processing Systems*, t. 34, zredagowane przez M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang, i J. Wortman Vaughan. Curran Associates, Inc., 2021. [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/076ccd93ad68be51f23707988e934906-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/076ccd93ad68be51f23707988e934906-Paper.pdf).
- Kent State University Libraries. „SPSS Tutorials: Independent Samples t-Test”. 2025. <http://libguides.library.kent.edu/SPSS/IndependentTTest>.
- Lorenz, Konrad. „Die angeborenen Formen moeglicher Erfahrung”. *Zeitschrift für Tierpsychologie* 5, nr 2 (1943): 235–409.
- Martin, Mike W. „Humour and Aesthetic Enjoyment of Incongruities”. *The British Journal of Aesthetics* 23, nr 1 (1983): 74–85.
- Martin, Rod A., Nicholas A. Kuiper, L. Joan Olinger, i Kathryn A. Dance. „Humor, Coping with Stress, Self-Concept, and Psychological Well-Being”. *Humor: International Journal of Humor Research* 6, nr 1 (1993): 89–104.
- Morreall, John. „Philosophy of Humor”. W *The Stanford Encyclopedia of Philosophy*, zredagowane przez Edward N. Zalta i Uri Nodelman. Metaphysics Research Lab, 2024. <https://plato.stanford.edu/archives/fall2024/entries/humor/>.
- Norrick, Neal R., i Alice Spitz. „Humor as a Resource for Mitigating Conflict in Interaction”. *Journal of Pragmatics* 40, nr 10 (2008): 1661–86.
- Oveis, Christopher, Aleksandr Spectre, Pamela K. Smith, Mary Y. Liu, i Dacher Keltner. „Laughter Conveys Status”. *Journal of Experimental Social Psychology* 65 (2016): 109–15.
- Overeem, Sebastiaan, Gert J. Lammers, i J. Gert van Dijk. „Weak with Laughter”. *The Lancet* 354, nr 9181 (1999): 838.
- Overeem, Sebastiaan, Taal Walter Taal, E. Öcal Gezici, Gert Jan Lammers, i J. Gert van Dijk. „Is Motor Inhibition during Laughter Due to Emotional or Respiratory Influences?” *Psychophysiology* 41, nr 2 (2004): 254–58.
- Provine, Robert. „Laughter: A Scientific Investigation”. *American Scientist*, nr 84 (2000): 38–45.
- Sosnowski, Tytus. „Psychofizjologia”. W *Psychologia: Podręcznik akademicki*, t. 1. Gdańskie Wydawnictwo Psychologiczne, 2000.
- Sprengelmeyer, Reiner, David I. Perrett, Elena C. Fagan, et al. „The Cutest Little Baby Face: A Hormonal Link to Sensitivity to Cuteness in Infant Faces”. *Psychological Science* 20, nr 2 (2009): 149–54.
- Van Hooff, Jan. „A Comparative Approach to the Phylogeny of Laughter and Smiling”. *Nonverbal Communication*, 1972, 209–41.
- Yim, Alexis, Bradley Price, Raj Agnihotri, i Annie Peng Cui. „Do Salespeople’s Online Profile Pictures Predict the Number of Online Reviews? Effect of a Babyface”. *European Journal of Marketing* 57, nr 7 (2023): 1886–911.
- Zebrowitz, Leslie A., Victor X. Luevano, Philip M. Bronstad, i Itzhak Aharon. „Neural Activation to Babyfaced Men Matches Activation to Babies”. *Social Neuroscience* 4, nr 1 (2009): 1–10.