

Koncepcja obiektywizacji testu zegara na potrzeby diagnostyki chorób neurodegeneracyjnych

Streszczenie. W artykule przedstawiona została propozycja wykorzystania inżynierskich technik akwizycji danych i metod głębokiego uczenia do obiektywnej analizy obrazów tworzonych w trakcie wywiadu przez pacjentów z zaburzeniami neurodegeneracyjnymi.

Abstract. The article presents a proposal of using engineering data acquisition techniques and deep learning methods for an objective analysis of images created during the history-taking in patients with neurodegenerative disorders. (**Concept of objective assessment of the clock-drawing test for medical diagnosis of neurodegenerative diseases.**)

Słowa kluczowe: rozpoznawanie obrazów, sieci konwolucyjne, diagnostyka medyczna.

Keywords: images recognition, convolution network, medical diagnostics.

Wprowadzenie

Choroby neurodegeneracyjne to grupa chorób o charakterze przewlekłym i postępującym których w większości przypadków podstawowym objawem jest zespół otępienny, tzw. demencja. Określa się ją jako upośledzenie możliwości myślenia wpływające na codzienne funkcjonowanie. Upośledzenie dotyka funkcji poznawczych, w tym pamięci, języka, umiejętności wzrokowo-przestrzennych, funkcji wykonawczych oraz osobowości. Otępienia neurodegeneracyjne klasyfikowane są na różne sposoby. Można podzielić je na odwracalne i nieodwracalne. Do grupy otępień odwracalnych zalicza się schorzenia przebiegające w niedoczynności tarczycy lub niedoborze witaminy B12. Do najczęściej występujących schorzeń z grupy otępień nieodwracalnych (przynajmniej na obecnym etapie rozwoju medycyny) należą choroba Alzheimera (AD) oraz otępienie czołowo-skroniowe (FTD). Rozróżnienie pomiędzy nimi, zwłaszcza w wieku przed-senioralnym, bywa trudne [1].

We współczesnej diagnostyce stanów otępienia dostrzec można z jednej strony specjalistyczne metody neuroobrazowania struktury i funkcji mózgu i z drugiej dążenie do stosowania prostych testów przesiewowych, możliwych do wykorzystania przez lekarza pierwszego kontaktu. Wyniki tych ostatnich mogą być jednak wypaczone subiektywną interpretacją, często wynikającą z braku odpowiedniego doświadczenia w wizualnej ocenie np. zachowania zdolności wzrokowo-przestrzennej pacjenta, będącej jednym z kryteriów diagnozy Lund-Manchester, używanej do różnicowania otępienia czołowo-skroniowego i choroby Alzheimera [2]. Jednocześnie stwierdza się, że żadna ze współczesnych technik neuroobrazowych nie jest w stanie dostatecznie dobrze uwidocznzyć mikro zmian w obrębie płatów mózgu, które wpływają na wyższe funkcje umysłowe człowieka [3]. Ponadto wykonanie tego typu badań jak również badań histologicznych i genetycznych wymaga kontaktu ze specjalistami różnych dziedzin oraz specjalistycznego sprzętu, a co za tym idzie – wymaga czasu. Ze względu na opóźnienie diagnozy choroba najczęściej przechodzi w kolejne stadium nim pacjent otrzyma prawidłowe rozpoznanie. Kliniczne różnicowanie chorób neurodegeneracyjnych z innymi różnicowymi chorobami mózgu dającymi objawy zbliżone do uszkodzeń płatów mózgu jest więc trudne, szczególnie na wczesnym etapie choroby. Podobne trudności występują również w przypadku próby odróżnienia chorych z otępieniem neurodegeneracyjnym od chorego z chorobą psychiczną, w przebiegu której występują podobne zmiany osobowości,

nastroju i zachowania [3]. Należy podkreślić, że przyczyny niektórych chorób z zakresu zaburzeń otępiennych nie zostały do końca poznane; być może spowodowane jest to faktem, iż pacjenci zgłaszają się do lekarza dopiero wtedy, gdy oni sami lub ich bliscy zauważą niepokojące objawy (jak w przypadku AD według [4]). Wcześniejsza diagnostyka mogłaby dostarczyć wielu cennych informacji i znacząco wzbogacić dotychczasową wiedzę, zarówno z zakresu etiologii danego zaburzenia, jak również jego leczenia. Otwarte pozostaje także pytanie, na ile objawy osób po 65 roku życia można uznać za patologiczne, a nie wynikające z naturalnego procesu starzenia się mózgu [5]. W starzejącym się społeczeństwie odpowiedź na to pytanie wydaje się być kluczowa.

W tej sytuacji aktualnym wyzwaniem dla zespołów interdyscyplinarnych jest opracowanie możliwie najbardziej obiektywnych narzędzi umożliwiających postawienie jak najwcześniejszej prawidłowego rozpoznania, co pozwoliłoby na rozpoczęcie terapii oraz na postawienie diagnozy prognostycznej o dalszych zmianach stanu chorego, co ważne jest nie tylko dla niego samego (choćby z perspektywy psychoedukacji i włączenia do procesu leczenia psychoterapii), ale także dla jego bliskich. Stąd z jednej strony potrzeba poszukiwania innych, poza klinicznych wskaźników różnicujących [6] i z drugiej – potrzeba obiektywizacji istniejących metod z wykorzystaniem technik inżynierskich. Znaczenie tych technik dla przypadku zespołów otępiennych wydaje się wysoce istotne. Największe możliwości spowolnienia postępu zaburzeń występują w początkowej fazie, stąd wykrywanie przypadków o łagodnym lub umiarkowanym nasileniu MCI (ang. *Mild Cognitive Impairment*) wszelkimi dostępnymi metodami jest niezwykle cenne. Jednocześnie oprócz samej diagnozy istotne jest długofalowe monitorowanie przypadków z łagodnymi zaburzeniami poznawczymi wymuszone wysokim prawdopodobieństwem przejścia do wersji pełnoobjawowego otępienia. Szacuje się, że po 3-4 latach otępienie rozwija się w 50% przypadków [7], chociaż znane są też doniesienia mówiące o braku progresji zaburzeń MCI [8]. Środki inżynierskie mogłyby umożliwić nie tylko zobiektywizowanie istniejących metod, rozwój nowych czy też wpisanie się w nurt medycyny na odległość (telemedycyny), ale również zapewnić możliwość archiwizacji wyników na potrzeby oceny postępu choroby lub efektów terapii. Już sama systematyczna ocena historii choroby i obrazu klinicznego popartego testami może być istotnym narzędziem diagnostyki różnicowej [3].

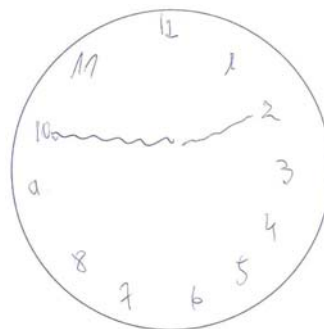
Problem badawczy

Wobec stwierdzonego pogorszenia funkcjonowania poznawczego w chorobach otępiennych [9], techniki inżynierskie mogą być zastosowane do obiektywnej analizy obrazów tworzonych przez pacjentów w trakcie wywiadu. Jednym z powszechnie stosowanych testów jest tzw. test zegara CDT (ang. *Clock-Drawing Test*). Jest to test przesiewowy służący do oceny procesów analizy i syntezy wzrokowo-przestrzennej, planowania, myślenia abstrakcyjnego (posługiwanie się pojęciem czasu jest zdolnością osobniczą) a związanych z funkcjonowaniem płata czołowego i kory skroniowo-ciemieniowej. Może wchodzić w skład różnych testów oceniających sprawność umysłową, ale jest też stosowany jako test samodzielny przede wszystkim z uwagi na fakt, że na jego wyniki jedynie w niewielkim stopniu wpływa pochodzenie etniczne, kultura i wykształcenie [10]. Ogólna idea testu polega na narysowaniu na papierze tarczy zegara, wpisaniu cyfr oznaczających godziny i ewentualnym naniesieniu wskazówek reprezentujących konkretne godziny.

Już sam fakt wykonania polecenia przez pacjenta jest dowodem na posiadanie przez niego odpowiednich umiejętności rozumienia ze słuchu. Test w sposób specyficzny wymusza jednocześnie występowanie kilku procesów interakcji umysłu z otoczeniem. Przede wszystkim pacjent powinien wykazać się umiejętnością pisania uporządkowanych liczb z jednoczesnym zwracaniem uwagi na ich poprawne wzajemne rozmieszczenie w przestrzeni i usytuowanie względem obrysu tarczy zegara. Do skoordynowania tych czynności potrzebne są funkcje planowania i organizacji z równoległym przetwarzaniem informacji wzrokowej. Jeżeli zlecane przez diagnostę zadanie obejmuje narysowanie wskazówek odpowiadających określonej godzinie, to potrzebne jest do tego również użycie pamięci pozwalające na przywołanie ich właściwego układu i odtworzenie go na tarczy. Za powyższe umiejętności odpowiadają różne obszary w obu półkulach mózgowych. Każda z czynności może być w sposób selektywny upośledzona co będzie skutkowało powstaniem różnych obrazów. I tak np. wskazanie czasu może mieć istotne znaczenie w wykryciu uszkodzeń mózgowych zlokalizowanych w jednej z półkul, czyli w ocenie tzw. pomijania stronnego [11]. Pacjent z pomijaniem lewostronnym poproszony o narysowanie dowolnej godziny będzie w sposób spontaniczny wybierał taką, przy której żadna ze wskazówek zegara nie będzie znajdowała się w lewej części tarczy – np. „15:30” lub „14:20”. Stąd wielu neurologów narzuca narysowanie takiej godziny, przy której każda ze wskazówek znajduje się w obszarze innej połowy tarczy, np. „11:10” lub „8:20”. Godziny takie dają dodatkowo możliwość oceny funkcjonowania płatów czołowych odpowiedzialnych za myślenie abstrakcyjne, gdyż np. w godzinie „11:10” wskazanie „10” musi zostać zdekodowane tak, aby wskazówka minutowa została ustawiona na cyfrze „2”. Próba rysowania zegara z powyższym wskazaniem przez osobę badaną, cierpiącą na uszkodzenie lewego płata czołowego, przedstawiona jest na rys. 1. Widoczne niepoprawne położenie i nieproporcjonalne wydłużenie wskazówki godzinowej, która występuje w lewej części tarczy, względem wskazówki minutowej świadczy o występującej dysfunkcji.

To swoiste próbkowanie różnych funkcji poznawczych jest podstawową przyczyną, dla której test zegara jest traktowany jako jeden z najlepszych testów przesiewowych [12]. Ma on istotne znaczenie w diagnostyce osób starszych, których pogarszająca się orientacja wzrokowo-przestrzenna przyczynia się do deficytu konstrukcyjnego ale nie musi pociągać problemów z pojmowaniem czasu, co

jest wskaźnikiem stanów otępienia. Test jest wykonywany w wersji zegara całkowicie tworzonych przez pacjenta w trakcie wywiadu lub w wersji, w której pacjent ma do dyspozycji kontur, który należy uzupełnić liczbami reprezentującymi godziny.



Rys. 1. Przykładowy obraz zegara narysowany przez osobę z patologią po otrzymaniu polecenia „narysuj godzinę 11:10”

Motywacją do stosowania tego drugiego podejścia jest uniknięcie przypadków rysowania tarczy, która jest zbyt mała aby pomieścić wszystkie godziny lub zbyt asymetryczna aby je w sposób poprawny rozmieścić [12]. Wynik testu rysowania zegara może być przedstawiony w sposób ilościowy za pomocą jednego ze stosowanych systemów oceny [13]. Na przykład w popularnym systemie Sunderlanda, którego czułość w wykryciu choroby Alzheimera szacuje się na 56÷79% a swoistość na 58-91%, uzyskuje się maksymalnie 10 punktów przy czym wartością progową, poniżej której przewiduje się wystąpienie patologii jest 6. Maksymalna liczba punktów przypisywana jest za zegar z prawidłowym wpisaniem cyfr i prawidłowym zaznaczeniem godziny, 9 punktów za drobne pomyłki w położeniu wskazówek, 8 za pomyłki wyraźnie widoczne, 7 za zdecydowanie źle położone wskazówki, 6 za brak wskazówek i poniżej 6 za rysunki błędne (grupowanie cyfr po jednej stronie, zaburzenie porządku cyfr, cyfry poza tarczą, brak powiązania wskazówek i cyfr itd.).

Podstawowy wyzwaniem jakie się pojawia w scharakteryzowanym powyżej teście zegara jest obiektywizacja uzyskiwanych wyników za pomocą podejścia technicznego. Jego opracowanie wymaga rozwiązania szeregu problemów inżynierskich i badawczych obejmujących kwestię akwizycji danych uwzględniającą ograniczone możliwości pacjentów w starszym wieku, opracowanie metod wykrywania lokalizacji i rozpoznawania pisanych odręcznie cyfr, opracowanie metod detekcji wskazówek i wreszcie wyznaczenie parametrów diagnostycznych – miar liczbowych zawierających ocenę jakości i poprawności położenia elementów zegara, być może z uwzględnieniem czasu ich tworzenia i używanej siły nacisku. W niniejszej pracy skupiono się na zadaniach, które stanowią podstawę dalszych badań, czyli na rozwiązaniu problemu akwizycji danych oraz na zadaniu rozpoznawania pisanych odręcznie cyfr.

Prace powiązane

Próba technicznej obiektywizacji testu zegara jest możliwa z wykorzystaniem urządzeń przeznaczonych do rejestracji pisma odręcznego, czyli skanerów lub tableatów graficznych. Pomysł wykorzystania tabletu w teście zegara nie jest nowy i pojawia się np. w [8] w zadaniu rozpoznawania łagodnych zaburzeń poznawczych MCI u osób starszych, jako próba telediagnostyki w geriatricz z tym, że jest on wykorzystany raczej jako rejestrator rysunku. Pacjent może, w domu lub gabinecie lekarza pierwszego

kontaktu, bez obecności specjalisty, zarejestrować obraz w sposób nienadzorowany a następnie przesłać go drogą elektroniczną do oceny eksperckiej. Badania opisane w [8] w zasadzie wykazały zgodność wersji tabletovej z klasyczną wersją papierową co do czułości i swoistości metody, ale wskazały również na szereg czynników utrudniających korzystanie z aplikacji, zwłaszcza w przypadku osób starszych nie obeznanych z techniką komputerową i nie przyzwyczajonych do pisania po dotykowym ekranie. Bardziej zaawansowane funkcjonalnie rozwiązania, jak np. opisywane w [14], automatyzują wyznaczanie wskaźnika ilościowego testu. Autorzy ww. pracy skorzystali z wersji testu bazującej na przygotowanym konturze zegara z zaznaczonymi na jego obwodzie punktami godzin, przy których pacjent nanosi rysikiem lub palcem na ekranie dotykowym tabletu lub telefonu komórkowego ocyfrowanie oraz dorysowuje wskazówki ilustrujące wymaganą godzinę. Wymuszone punktami na konturze zegara pozycjonowanie cyfr ułatwia ich wykrycie i rozpoznanie metodami przetwarzania obrazów pozwalając na ilościowe wyrażenie poprawności opisu.

Proponowana metoda

Prezentowane rozwiązania dają podstawy do badań nad opracowaniem nowego podejścia do diagnostyki za pomocą testu rysowania zegara. Potencjalnie dodatkowe możliwości otwiera nie tylko ilościowa ocena logicznej poprawności ale również ocena dynamiki tworzenia obrazu zegara rozumianej w kategoriach czasu jego tworzenia i siły nacisku pióra. Podstawowe znaczenie ma zastosowanie odpowiednio zaawansowanego technologicznie rejestratora, który z jednej strony uwzględniałby ograniczenia ludzi starszych a z drugiej zapewniał pozyskanie możliwie wszechstronnej informacji o sposobie tworzenia obrazu. Możliwości takich nie mają tradycyjne skanery obrazów, ale zapewniają je współczesne tablety graficzne przeznaczone do zastosowań profesjonalnych. Jednym z tabletów graficznych nadającym się do tego typu zadań jest tablet Intuos Pro Paper Edition firmy WACOM, który umożliwia nie tylko standardową obsługę poprzez piórko dotykowe tworzące obraz cyfrowy, ale również jest wyposażony w specjalny długopis pozwalający pisać na kartce papieru i odzwierciedlać zapis na ekranie komputera w formie cyfrowej. Obszar aktywny tabletu ma wymiary 311 x 216 mm. Za pomocą specjalnego klipsu możliwe jest przypięcie do jego powierzchni kartki papieru w formacie A4. Używany do pisania długopis to bezprzewodowe pióro atramentowe, kompatybilne z tabletem. W związku z powyższym tablet graficzny umożliwia wykonanie testu zegara w warunkach niemal identycznych do warunków występujących podczas wykonywania testu ręcznie na kartce papieru – rys. 2.



Rys. 2. Ilustracja tworzenia obrazu zegara z wykorzystaniem tabletu graficznego z opcją pisania na kartce papieru

Dodatkowo tablet można podłączyć do komputera bezprzewodowo z wykorzystaniem interfejsu Bluetooth, dzięki czemu pacjent nie jest rozpraszany przewodami [15]. Tablety firmy Wacom korzystają ze standardowego interfejsu WinTab, służącego do pozyskiwania danych z tabletów graficznych. Dzięki temu możliwa jest bezpośrednia komunikacja z tabletem i akwizycja danych z poziomu np. programu Matlab.

W trakcie pisania na powierzchni tabletu do komputera przekazywane i zapisywane są informacje dotyczące pióra, służące do późniejszej analizy. Tablet umożliwia rejestrację danych takich jak nacisk pióra na powierzchnię i jego odchylenie od pionu z uwzględnieniem współrzędnych lokalizacji i momentów próbkowania odbywającego się z częstotliwością 200Hz. Rozdzielczość pomiaru nacisku jest 13-bitowa, co oznacza 8192 poziomy nacisku [15]. Informacja o nacisku może być wykorzystana do opracowania liczbowych miar diagnostycznych różnicujących przypadki występujące w chorobach neurodegeneracyjnych.

Na rys. 3 przedstawiono wyniki dwukrotnego pisania przez tę samą osobę na tablecie graficznym cyfry 9, kolejno z małym i dużym naciskiem pióra. Przedstawione obrazy nie przenoszą informacji o dynamice ich tworzenia, jeśli jednak uwzględni się przedstawione na rys. 4 dane zwracane przez układ pomiarowy tabletu, to pojawiają się dodatkowe możliwości interpretacyjne. Przede wszystkim widoczny jest czas (ok. 0.8s) tworzenia charakterystycznego dla cyfry 9 owalu, który jest reprezentowany przez przebiegi czasowe współrzędnych położenia rysika, zbliżone kształtem do jednego okresu funkcji sinus i kosinus. Widoczna jest też ok. 2-krotnie większa wartość nacisku pióra przy tworzeniu obrazu B oraz zupełnie inna jego zmiana w czasie niż w przypadku A.

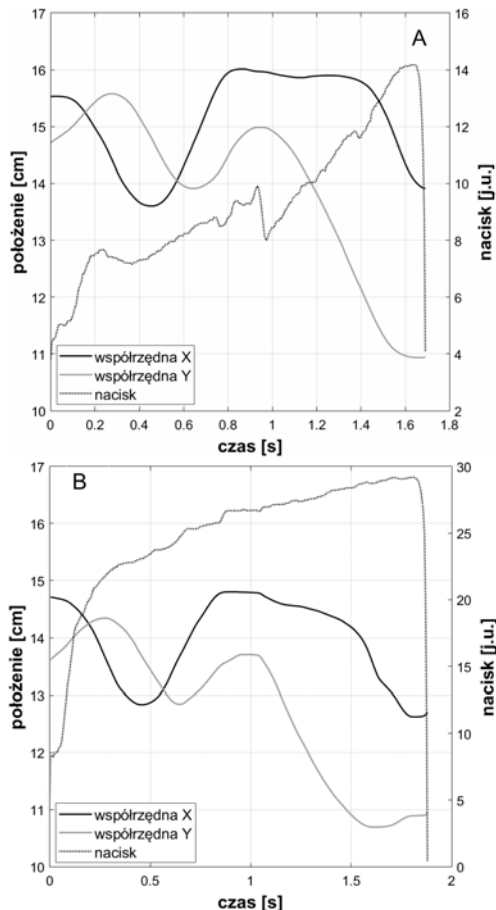


Rys. 3. Wyniki pisania na tablecie graficznym cyfry 9 przez tę samą osobę: A – mały nacisk pióra, B – duży nacisk pióra

Podstawą ilościowego wyrażenia poprawności tworzenia zegara w czasie testu realizowanego przez urządzenie techniczne jest automatyczne rozpoznawanie pisanych przez pacjenta cyfr. Nie jest to jednak zadanie trywialne z uwagi na silny osobniczy i kulturowy wpływ na wykształconą umiejętność pisania. Klasyczne podejście do rozpoznawania pisma polega na zastosowaniu swojej inżynierii cech, która bazuje na eksperckim wskazaniu wektorów atrybutów ilościowych bezpośrednio z jego cyfrowej postaci a następnie na zastosowaniu jednej z metod uczenia maszynowego. Uzyskiwane tym sposobem dokładności rozpoznawania wynoszą ok. 90% i zależą od zastosowanego aparatu normowania obrazów, ekstrakcji cech i metody klasyfikacji [16].

Jakościowo odmienne postępowanie jest możliwe dzięki współczesnym rozwiązaniom sieci neuronowych dostosowanych w swej strukturze i metodzie uczenia do pracy nie z wektorami cech, jak to ma miejsce w klasycznym przypadku, ale bezpośrednio z całymi obrazami [18]. Cechy charakterystyczne obrazów, tworzone są w sposób automatyczny poprzez wskazanie sieci przykładowych obrazów trenujących wraz z ich etykietami przynależności do klas. Za generację cech wewnątrz struktury sieci odpowiadają rozmieszczone w warstwach neurony. Neurony w pierwszej warstwie są połączone z dedykowanymi dla nich lokalnymi fragmentami obrazu,

zwanyymi polami recepcyjnymi. Ponieważ neurony, w procesie agregowania informacji z pól recepcyjnych, wykonują operację iloczynu skalarnego nauczonych wag z wartościami pikseli, to zestawienie wszystkich ich odpowiedzi tworzy wewnątrz sieci kolejny obraz, tzw. mapę cech, której powstawanie może być ekwiwalentnie przyrównane do operacji splotu, czyli realizacji iloczynu skalarnego przesuwającej się po obrazie wejściowym maski filtru.



Rys. 4. Surowe dane – szeregi czasowe – zwrócone przez tablet graficzny w trakcie tworzenia obrazów przedstawionych na rys. 5: A – lekki nacisk pióra, B – mocny nacisk pióra

Stąd często do określenia tego rodzaju sieci neuronowych używa się pojęcia sieci konwulucyjnej. Mapy cech powstałe przy rozpoznawaniu cyfr 2 i 8 przez wytrenowaną w wyniku realizacji niniejszej pracy sieć przedstawia rys. 5. Tworzone wewnątrz sieci obrazy są poddawane działaniu kolejnych warstw konwulucyjnych, których liczba i rozmiary są dobierane przez użytkownika do rozwiązania postawionego zadania rozpoznawania. Warstwy wypracowujące decyzję są klasycznymi warstwami perceptronów wielowarstwowych MLP, które są zasilane wszystkimi wartościami końcowych map cech i z tego powodu są one nazywane warstwami pełnego połączenia. Efekt końcowy działania nauczonej sieci obserwuje się na wyjściu realizującym regresję logistyczną (warstwa SoftMax), które podaje wynik rozpoznawania w kategoriach logiki rozmytej wyznaczając prawdopodobieństwo zakwalifikowania obiektu do określonej kategorii – rys. 5. W strukturze wewnętrznej sieci wykorzystuje się nowe rozwiązania funkcji aktywacji, które zapobiegają zjawisku zaniku gradientu funkcji celu, obserwowanemu przy stosowaniu funkcji sigmoidalnych, i znacząco przyspieszają obliczenia [19]. Celem redukcji wymiaru danych wewnątrz sieci stosuje się różne rodzaje

filtrów statystycznych (tzw. pooling ukierunkowany na wyznaczenie wartości maksymalnej, minimalnej lub średniej z maski o zadanym rozmiarze). Zjawisku przeuczenia sieci zapobiega się techniką tzw. dropout, która polega na zerowaniu losowo wybranych danych wejściowych.

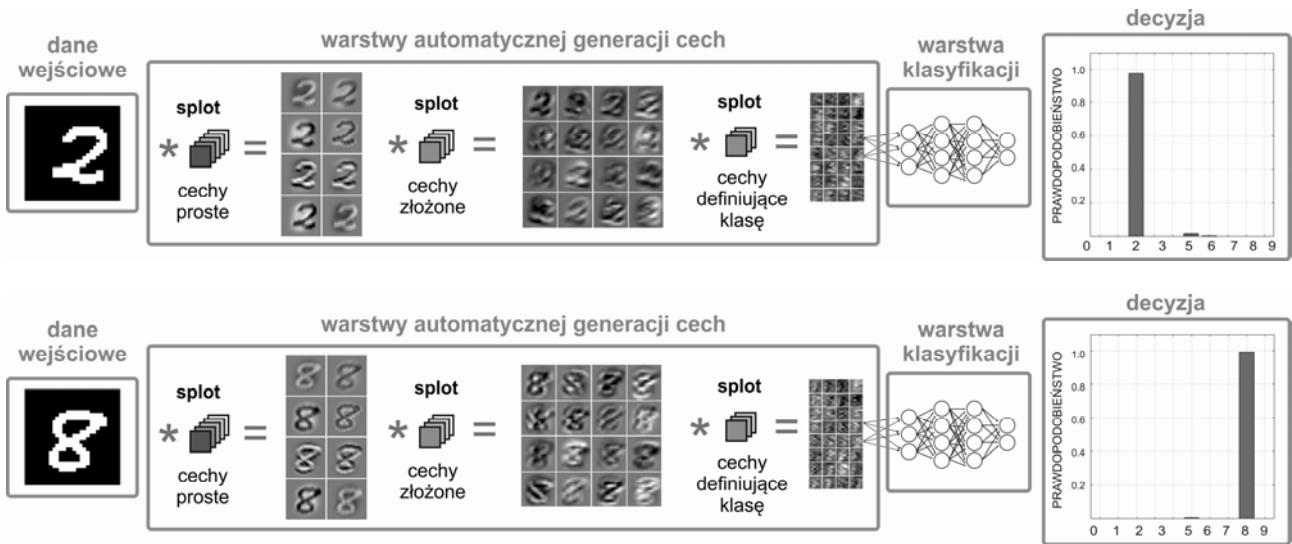
Wyniki

Efektom podjętych prac jest adaptacja do dalszych badań, przewidywanych do wykonywania w środowisku Matlab, scharakteryzowanego powyżej tabletu graficznego oraz uzyskanie wstępnej ilościowej oceny wykorzystania konwulucyjnej sieci neuronowej do rozpoznawania pisanych na nim cyfr. Podstawowym elementem występującym na rysunku zegara są cyfry. W związku z tym w pierwszej kolejności opracowano metodę akwizycji pisanych odręcznie cyfr na powierzchni tabletu i odtworzenia na ekranie komputera zarejestrowanych obrazów. Po uruchomieniu opracowanego oprogramowania osoba pisząca jest proszona o zapisanie kolejno cyfry od 0 do 9. Dane pochodzące z każdej cyfry są zapisywane w oddzielnych plikach. Zawierają one następujące informacje: numer próbki, czas rejestracji w [ms], położenie pióra wzdłuż osi X oraz Y tabletu, oraz nacisk w jednostkach umownych. W Tabeli 1 zaprezentowano fragment zebranych przykładowych danych. Tak otrzymane dane mogą posłużyć do zobrazowania pisma pacjenta na powierzchni tabletu oraz analizę samego procesu rysowania.

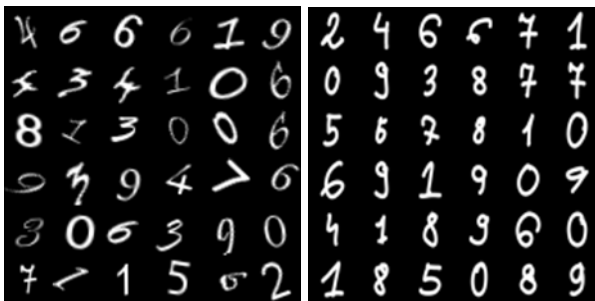
Tabela 1. Przykładowy zestaw danych rejestrowanych przez wykorzystany tablet graficzny

numer próbki	znacznik czasu [ms]	współrzędna x [cm]	współrzędna y [cm]	nacisk [j.u.]
1	0	5,696	13,038	3876
2	5	5,698	13,038	4088
3	10	5,701	13,038	4712
4	14	5,703	13,038	4836
5	15	5,707	13,038	4836
6	20	5,710	13,038	4784
7	29	5,713	13,040	4776
8	25	5,716	13,041	4772
9	30	5,719	13,041	4696
10	44	5,720	13,041	4292

Otrzymane dane następnie są przetwarzane. W pierwszej kolejności zostają usunięte tzw. puste punkty, czyli momenty, w których próbka nie została pobrana. Następnie na podstawie wartości nacisku usuwane są wszystkie punkty zebrane w momencie gdy pióro znajdowało się nad powierzchnią tabletu. Obraz cyfry jest tworzony jako wykres zależności wartości współrzędnych X i wartości współrzędnych Y za pomocą punktów o odpowiedniej grubości. Ostatecznie utworzony zostaje obraz przedstawiający białą cyfrę na czarnym tle w skali szarości. Podejście „neuronowe” do rozwiązania zadania rozpoznawania wymaga dostępu do odpowiednio licznego zbioru danych reprezentujących próbki uczące. W pracy wykorzystano do tego celu fragment standardowej bazy danych MNIST zawierającej pisane odręcznie cyfry [17], dostępny w docelowym środowisku badawczym Matlab. Obrazy występujące w tej bazie są obrazami normowanymi tak, aby środek ciężkości ich oryginalnej postaci zamkniętej w kwadracie 20x20 pikseli znajdował się w środku obrazu o wymiarach 28x28 pikseli. Do testów użyty został zbiór 100 obrazów (po 10 dla każdej cyfry od 0 do 9) zarejestrowanych wykorzystanym tabletem. Obrazy te, celem zbadania zdolności uogólniania sieci, poza zmianą rozmiaru do wartości 28x28 pikseli, nie były poddane procesowi normalizacji. Zestawienie przykładowych obrazów zbioru uczącego i testującego przedstawia rys. 6.



Rys. 5. Ilustracja wybranych elementów wykorzystanej w pracy struktury sieci oraz map cech generowanych przy rozpoznawaniu zarejestrowanych tabletem graficznym obrazów kolejno cyfry 2 i cyfry 8



Rys. 6. Przykładowe obrazy zbioru uczącego (po lewej) i testującego (po prawej)

Dobrana w sposób eksperymentalny struktura sieci neuronowej zawierała 3 warstwy konwulcyjne oddzielone warstwami redukującymi oraz warstwami aktywacji typu ReLU. Do oceny porównawczej uzyskanych wyników użyta została prosta metoda rozpoznawania typu najbliższe sąsiedztwo dla cech będących uszeregowanymi do postaci wektora wartościami wszystkich pikseli obrazu. Uzyskane macierze błędów rozpoznawania dla metody najbliższego sąsiedztwa oraz dla sieci konwulcyjnej przedstawiono na rys. 7.

Macierz błędów

	0	1	2	3	4	5	6	7	8	9	
0	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
1	0 0.0%	9 90.0%	0 0.0%	0 0.0%	1 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	10.0%
2	0 0.0%	1 1.0%	7 70.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 20.0%	0 0.0%	0 0.0%	70.0%
3	0 0.0%	0 0.0%	0 0.0%	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%
4	0 0.0%	0 0.0%	0 0.0%	0 0.0%	9 90.0%	0 0.0%	0 0.0%	1 10.0%	0 0.0%	0 0.0%	90.0%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7 70.0%	2 20.0%	0 0.0%	1 10.0%	0 0.0%	30.0%
6	1 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7 70.0%	2 20.0%	0 0.0%	0 0.0%	70.0%
7	0 0.0%	2 20.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	8 80.0%	0 0.0%	0 0.0%	20.0%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 10.0%	0 0.0%	9 90.0%	0 0.0%	10.0%
9	0 0.0%	0 0.0%	0 0.0%	1 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	9 90.0%	10.0%
	90.9%	75.0%	100%	90.9%	90.0%	100%	70.0%	88.9%	64.3%	100%	85.0%
	9.1%	25.0%	0.0%	9.1%	10.0%	0.0%	30.0%	11.1%	35.7%	0.0%	15.0%
	0	1	2	3	4	5	6	7	8	9	

klasa zadana

Rys. 7. Macierze błędów dla dwóch metod rozpoznawania cyfr zarejestrowanych tabletem wykorzystanym sąsiedztwa (u góry) i sieci konwulcyjnej (na dole)

Macierz błędów

	0	1	2	3	4	5	6	7	8	9	
0	10 100.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 2.0%	0 0.0%	3 3.0%	2 2.0%	58.8%
1	0 0.0%	4 40.0%	1 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4 40.0%	2 20.0%	1 10.0%	33.3%
2	0 0.0%	1 10.0%	9 90.0%	1 10.0%	1 10.0%	0 0.0%	0 0.0%	1 10.0%	1 10.0%	0 0.0%	64.3%
3	0 0.0%	1 10.0%	0 0.0%	8 80.0%	1 10.0%	4 40.0%	0 0.0%	1 10.0%	1 10.0%	3 30.0%	42.1%
4	0 0.0%	3 30.0%	0 0.0%	0 0.0%	3 30.0%	0 0.0%	0 0.0%	2 20.0%	0 0.0%	0 0.0%	37.5%
5	0 0.0%	0 0.0%	0 0.0%	1 10.0%	1 10.0%	3 30.0%	2 20.0%	0 0.0%	0 0.0%	1 10.0%	62.5%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 10.0%	2 20.0%	6 60.0%	0 0.0%	1 10.0%	0 0.0%	60.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 20.0%	0 0.0%	0 0.0%	2 20.0%	0 0.0%	2 20.0%	33.3%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 10.0%	0 0.0%	0 0.0%	2 20.0%	0 0.0%	66.7%
9	0 0.0%	1 10.0%	0 0.0%	0 0.0%	1 10.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 10.0%	33.3%
	100%	40.0%	90.0%	80.0%	30.0%	30.0%	60.0%	20.0%	20.0%	10.0%	48.0%
	0.0%	60.0%	10.0%	20.0%	70.0%	70.0%	40.0%	80.0%	80.0%	90.0%	52.0%
	0	1	2	3	4	5	6	7	8	9	

klasa zadana

Podsumowanie

Przeprowadzone prace wskazują na potencjalne korzyści płynące z połączenia zaawansowanego technicznie tabletu graficznego z oprogramowaniem wykorzystującym sieci konwulcyjne do analizy obrazów tworzonych przez pacjentów z zaburzeniami neurologicznymi w trakcie badania diagnostycznego. Oprócz realizacji zadania archiwizacji, możliwości profesjonalnego tabletu otwierają zupełnie nową perspektywę analizy obrazów, w których wartość kluczową mogą mieć informacje o dynamice ich tworzenia. Zbadane możliwości wykorzystania sieci konwulcyjnej do rozpoznawania pisanych odręcznie cyfr na potrzeby testu zegara również dowodzą ich użyteczności – skuteczność rozpoznawania zarejestrowanych tabletem obrazów, które w dodatku nie były wstępnie przetwarzane celem dostosowania do obrazów z referencyjnej bazy danych, w sposób znaczący przewyższyła skuteczność rozpoznawania cyfr za pomocą jednej z klasycznych metod. Uzyskane wyniki stanowią podstawę do dalszych prac nad wyznaczeniem ilościowych parametrów diagnostycznych na potrzeby obiektywizacji testu zegara oraz ewentualnie

innych testów, w których ocenie podlegają manualne możliwości pacjenta.

Praca współfinansowana ze środków MON na realizację badań podstawowych w ramach grantu badawczego No. GBMON/B9/13-996/2018.

Autorzy: mgr inż. Kamila Jadczyk, Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, ul. Gen. Sylwestra Kaliskiego 2, 00-908 Warszawa, E-mail: kamila.jadczyk@wat.edu.pl;

dr hab. inż. Jacek Jakubowski Wojskowa Akademia Techniczna, Instytut Systemów Elektronicznych, ul. Gen. Sylwestra Kaliskiego 2, 00-908 Warszawa, E-mail: jacek.jakubowski@wat.edu.pl;

dr hab. n. med. Kazimierz Tomczykiewicz, Warszawski Uniwersytet Medyczny, Katedra i Klinika Neurologii, ul. Banacha 1a, 02-097 Warszawa, E-mail: kazimierz.tomczykiewicz@awf-bp.edu.pl;

LITERATURA

- [1] Smid J., Adoni T., Behind the faces: Alzheimer's disease and emotional blindness, *Arq Neuropsychiatr*, vol. 73(5), (2015), 381-382.
- [2] Pąchalska M., Otepienie czołowo-skroniowe. Ujęcie interdyscyplinarne, *Krakowska Akademia im. Andrzeja Frycza Modrzewskiego*, Kraków, (2012).
- [3] Olszewski H., Otepienie czołowo-skroniowe. Ujęcie neuropsychologiczne, *Oficyna Wydawnicza Impuls*, Kraków, (2008).
- [4] Barcikowska M., Otepienie, w: Gajewski P. *Interna Szczeklika 2012. Medycyna Praktyczna*, Kraków (2012), 2058–2063.
- [5] Youdim, M. B., & Bakhle, Y. S., Monoamine oxidase: isoforms and inhibitors in Parkinson's disease and depressive illness, *British journal of pharmacology*, 147 (S1), (2006).
- [6] Kruk A., Różne oblicza otepienia czołowo-skroniowego, *Neuropsychologia.org*, (2016), <http://neuropsychologia.org/rozne-oblicza-otepienia-czolowo-skroniowego>.
- [7] Pąchalska M., et. al. Kliniczny Test Funkcji Wykonawczych – zrewidowany w diagnostyce różnicowej depresji, ze-społu lekkich zaburzeń poznawczych oraz otepienia typu Alzheimera, *Psychogeriatrya Polska*, 1 (2), (2004), 119-144.
- [8] Bednorz A., et. al., Zastosowanie tabletovej wersji Testu Rysowania Zegara do rozpoznawania łagodnych zaburzeń poznawczych (MCI) u osób starszych, jako próba telediagnostyki w geriatryi, *Geriatrya*, vol. 11, (2017), 5-14.
- [9] Onofri E., et al., Dysgraphia in Relation to Cognitive Performance in Patients with Alzheimer's Disease, *Journal of Intellectual Disability - Diagnosis and Treatment*, vol. 1, (2013), 113-124.
- [10] Krzywiński S., Test rysowania zegara, *Postępy Psychiatrii i Neurologii*, No 4, suplement 1(2), (1995), 21-30.
- [11] Domańska Ł., Ocena pomijania stronnego metodami angażującymi procesy wzrokowe, *Przegląd psychologiczny*, tom 42, nr 1-2, (1999), 49-56.
- [12] Freedman M. et. al., Clock Drawing – A Neuropsychological Analysis, *Oxford University Press*, (1994).
- [13] Strauss E., et. al., A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary, Third Edition, *Oxford University Press*, (2006).
- [14] Nirjon S., et. al., MOBI-COG: A Mobile Application for Instant Screening of Dementia Using the Mini-Cog Test, *5th Conference on Wireless Health*, Bethesda, USA, (2014).
- [15] <https://intuospro.wacom.pl/intuos-pro-paper.htm>
- [16] Shamim et.al., Handwritten Digit Recognition using Machine Learning Algorithms, *Global Journal of Computer Science and Technology: D Neural & Artificial Intelligence*, vol. 18 (1), (2018), 17-23.
- [17] <http://yann.lecun.com/exdb/mnist/>
- [18] Bengio Y., Courville A., and Goodfellow I., Deep Learning – systemy uczące się, *Wydawnictwo Naukowe PWN SA*, Warszawa, (2018).
- [19] Glorot X., Bordes A., Bengio Y., Deep Sparse Rectifier Neural Networks, *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, FL, USA, (2011), 315-323..