



INTERDYSCYPLINARNE CENTRUM MODELOWANIA
MATEMATYCZNEGO I KOMPUTEROWEGO

Autor: mgr inż. Norbert Kapiński

Autoreferat pracy doktorskiej:

**Proces gojenia ścięgna Achillesa oceniany przez fuzję
danych z wykorzystaniem głębokich sieci neuronowych**

Rozprawa doktorska przedłożona Radzie Naukowej Instytutu
Biocybernetyki i Inżynierii Biomedycznej

Kierujący pracą: dr hab. inż. Antoni Grzanka (Warszawski Uniwersytet Medyczny)
Promotor pomocniczy: dr Jakub Zieliński

Warszawa, luty 2021

Spis treści

Uzasadnienie wyboru tematu pracy	1
Cel i struktura pracy	2
Zbiór danych	3
Nowa metoda oceny ścięgna Achillesa	4
Charakterystyka i wyniki przeprowadzonych badań	5
Wnioski końcowe	8

Uzasadnienie wyboru tematu pracy

Wraz z występującym w ostatnich latach systematycznym wzrostem liczby obrazowań medycznych uwidacznia się potrzeba na komputerowe wspomaganie pracy radiologów oceniających badania obrazowe. W szczególności zastosowanie znajdują aplikacje usprawniające generowanie raportów, rozwiązania do personalizacji diagnostyki i narzędzia poprawiające jej jakość.

Niniejsza praca, w odpowiedzi na powyższe zagadnienia, przedstawia propozycję strukturyzacji i automatyzacji oceny gojenia ścięgna Achillesa widocznego w obrazowaniu Rezonansem Magnetycznym (w skr. RM). Badanie to, w kontekście przedmiotowego ścięgna, jest dokładną metodą wykorzystywaną do oceny zmian strukturalnych i morfologicznych w zakresie tkanek miękkich. Wedle obecnych standardów ocena tego badania jest subiektywna i niesparametryzowana, a zatem stanowi ciekawy temat badawczy związany z możliwościami komputerowego wspomaganie radiologów i usprawnienia ich pracy oraz w konsekwencji pomocy osobom ze schorzeniami ścięgna Achillesa.

Dodatkową motywację stanowi fakt dynamicznego rozwoju metod sztucznej inteligencji, a dokładniej ich podzbioru tj. głębokich sieci neuronowych, których rozwój znacząco przyspieszył od 2012 r. za sprawą innowacji w budowie sieci i szybkiej ewolucji możliwości sprzętowych. W szczególności konwolucyjne sieci neuronowe z uwagi na dużą skuteczność w problemach aplikowanych do przetwarzania obrazów znajdują zastosowanie w rosnącej liczbie narzędzi z certyfikacją medyczną dedykowanych dla radiologii.

Ograniczeniem wskazanych metod sztucznej inteligencji jest wymóg dużych, ustrukturyzowanych zbiorów danych, które mogą służyć do skutecznego uczenia się algorytmu. W toku prac nad wybraną problematyką, autor przedmiotowej rozprawy miał dostęp do unikatowego w skali światowej zbioru danych składającego się z 590 badań RM pacjentów po zerwaniu ścięgna Achillesa. Badania pochodziły z projektu START ("Wykorzystanie autologicznych mezenchymalnych komórek macierzystych w procesie regeneracji rekonstruowanego ścięgna Achillesa"), finansowanego przez Narodowe Centrum Badań i Rozwoju z krajowego programu STRATEGMED1.

Biorąc zatem pod uwagę występującą potrzebę w radiologii na wskazane inno-

wacje, dynamiczny rozwój możliwości metod sztucznej inteligencji w tym zakresie i dostęp do zasobów umożliwiających realizację pierwszych w skali światowej badań nad możliwością strukturyzacji i automatyzacji oceny ścięgna Achillesa widocznego w badaniu RM, autor zdecydował się na wybór przedstawionej w pracy tematyki.

Cel i struktura pracy

W ramach prac w projekcie START powstała koncepcja automatyzacji procesu monitorowania gojenia się ścięgna przy pomocy metod przetwarzania obrazów i sztucznej inteligencji. Założono, że głębokie sieci neuronowe będą skuteczne do oceny procesów patofizjologicznych widocznych w obrazowaniu medycznym takich jak gojenie tkanki miękkiej, co stanowi hipotezę przedmiotowej pracy.

Za cel główny autor postanowił obrać opracowanie automatycznej metody oceny gojenia się ścięgna Achillesa, natomiast cele poboczne stanowiły:

1. Wybór efektywnego kosztowo i czasowo protokołu badania bazującego na technikach obrazowania medycznego, a dokładniej Rezonansu Magnetycznego.
2. Przetestowanie różnego rodzaju podejść związanych ze szkoleniem głębokich sieci neuronowych.
3. Porównanie wyników oceny nowej metody z wynikami klasyfikacji bazującej na danych z ultrasonografii.
4. Porównanie wyników oceny nowej metody z oceną funkcjonalną, rutynowo stosowaną do wspomagania rehabilitacji po urazie ścięgna.

W celu uporządkowanego i zrozumiałego przedstawienia podłoża oraz wyników badań, w pracy wprowadzono podział na Rozdziały. Po wstępie i opisie celu pracy, w Rozdziale 3 zostały opisane współczesne metody monitorowania gojenia się ścięgna Achillesa. Rozdział ten rozpoczyna się od omówienia podstaw anatomicznych, biomechanicznych oraz dynamiki procesu gojenia się przedmiotowego ścięgna. Następnie przedstawiony jest szczegółowy opis badań obrazowych i biomechanicznych wykorzystywanych do oceny stanu rehabilitującego się pacjenta.

W Rozdziale 4, omówione zostały szczegółowo konwolucyjne sieci neuronowe, które posłużyły jako rdzeń opracowanego rozwiązania. W szczególności, w opisie uwzględniono problemy z budową efektywnych modeli na bazie sieci neuronowych.

W Rozdziale 5 zaprezentowano nowatorską metodę automatycznej oceny procesu gojenia się ścięgna Achillesa. Co istotne, omówiono unikatowy zbiór danych oraz

wzorzec odniesienia, dzięki którym możliwa była realizacja przewidzianych badań i walidacja metody. Przedstawiono również eksperymenty wykorzystane do doboru komponentów i parametrów ostatecznego modelu oraz finalny wynik funkcjonowania opracowanego algorytmu.

W Rozdziale 6 opisano zestawienie wyników metody z wynikami oceny realizowanej przez inne podejścia. W szczególności porównano przedmiotową metodę z metodą uczącą się *explicite* modelować ocenę radiologa, opracowaną również pod kierownictwem autora tej pracy w ramach pobocznych działań. Następnie zaprezentowano zestawienie z wynikami metody działającej w oparciu o dane z ultrasonografii, również opracowanej pod kierownictwem autora tej pracy. Finalnie, porównano wyniki z oceną biomechaniczną realizowaną standardowo przez fizjoterapeutów. Wnioski z porównań natury praktycznej spisano w Rozdziale 7. Całość pracy zakończono podsumowaniem przedstawionym w Rozdziale 8.

Zbiór danych

W ramach projektu przebadano 59-ciu pacjentów po całkowitym zerwaniu ścięgna Achillesa i 27-ciu ochotników. Kryteria kwalifikacji i szczegóły dotyczące urazów opisano w pracy.

Podczas trwającej 12 miesięcy rehabilitacji pacjenci byli monitorowani z wykorzystaniem RM GE Signa HDxt 1.5T wyposażonego w cewkę Foot & Ankle dedykowaną do pomiarów w rejonie dolnej kończyny. Każde z badań RM było wykonane z użyciem 7-miu sekwencji i łącznie 10-ciu modalności.

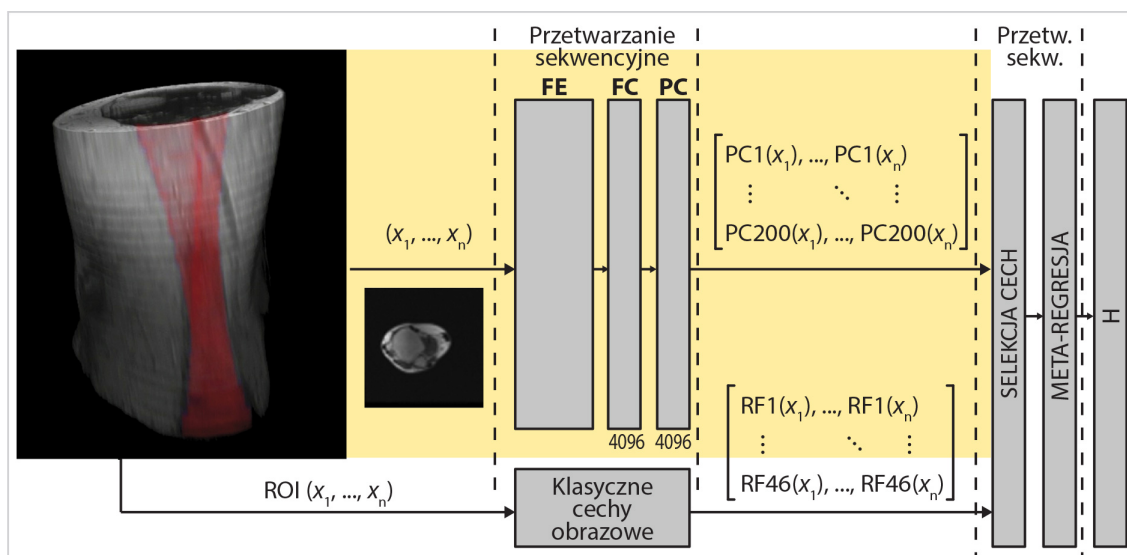
W grupie zdrowych ochotników przeprowadzono pojedyncze badanie, natomiast pacjentów skanowano 10-krotnie w odpowiednio zdefiniowanych odstępach czasowych. Pierwsze badanie odbyło się przed operacją, a następnych 9 odpowiednio w tygodniach: 1, 3, 6, 9, 12, 20, 26, 40 i 52 po operacji. Zbiory trójwymiarowe posłużyły do przygotowania dwuwymiarowych danych wejściowych dla wykorzystanych architektur sieci neuronowych. Finalna liczba tak utworzonych obrazów wyniosła 11.725 (oznaczonych jako obrazy zdrowego ścięgna) i 138.604 (oznaczonych jako obrazy chorego ścięgna). W zależności od wymogów eksperymentu liczby te były zmniejszane poprzez próbkowanie lub sztucznie powiększane z wykorzystaniem metod augmentacji danych.

Dodatkowo zgromadzono ustrukturyzowany opis radiologiczny dla 48-miu pa-

pacjentów w 10 krokach czasowych (480 ankiet z 6-cioma parametrami ocenionymi w skali 0–7). Czterech pacjentów (40 badań) zostało losowo wydzielonych na początku eksperymentów jako pacjenci testowi wykorzystani do celów wnioskowania i porównań wyników badań.

Nowa metoda oceny ścięgna Achillesa

W wyniku prac została opracowana komputerowa metoda oceny ścięgna Achillesa widocznego w badaniach RM, generująca numeryczny wynik w skali 0–7 dla 6-ciu zdefiniowanych w ramach projektu START radiologicznych parametrów opisujących strukturę i morfologię tkanek miękkich. Schemat generacji wartości pojedynczego parametru ilustruje rysunek poniżej.



Rysunek 1: Schemat automatycznej metody oceny pojedynczego parametru procesu gojenia się ścięgna Achillesa.

Dane wejściowe stanowi trójwymiarowe badanie RM obszaru podudzia. Badanie dzielone jest na n obrazów będących przekrojami poprzecznymi względem osi długiej ścięgna. Z każdego przekroju dwójako ekstrahowane są zestawy cech. Po pierwsze, wykorzystywany jest ekstraktor cech będący złożeniem warstw konwolucyjnych i jednej warstwy gęstej wytrenowanej sieci neuronowej na wyjściu (po redukcji z wykorzystaniem metody analizy czynników głównych) generującej 200 wartości. Po drugie, dla obszaru ścięgna widocznego na każdym z przekrojów, wyliczane są cechy statystyczne i teksturalne, których łączna liczba wynosi 46. Z pośród tak utworzonego zestawu 246 wartości, dla każdego z 6-ciu radiologicznych parametrów, wybierane

są z wykorzystaniem metody LASSO optymalne podzbiory o liczebności mniejszej niż 20. Podzbiory są kolejno procesowane z wykorzystaniem regresji opartej o algorytm wektorów nośnych, której wyniki dla poszczególnych przekrojów łączone są z wykorzystaniem średniej trymowanej odrzucającej rezultaty skrajne i generującej pojedynczą wartość parametru dla całego badania 3D RM.

Charakterystyka i wyniki przeprowadzonych badań

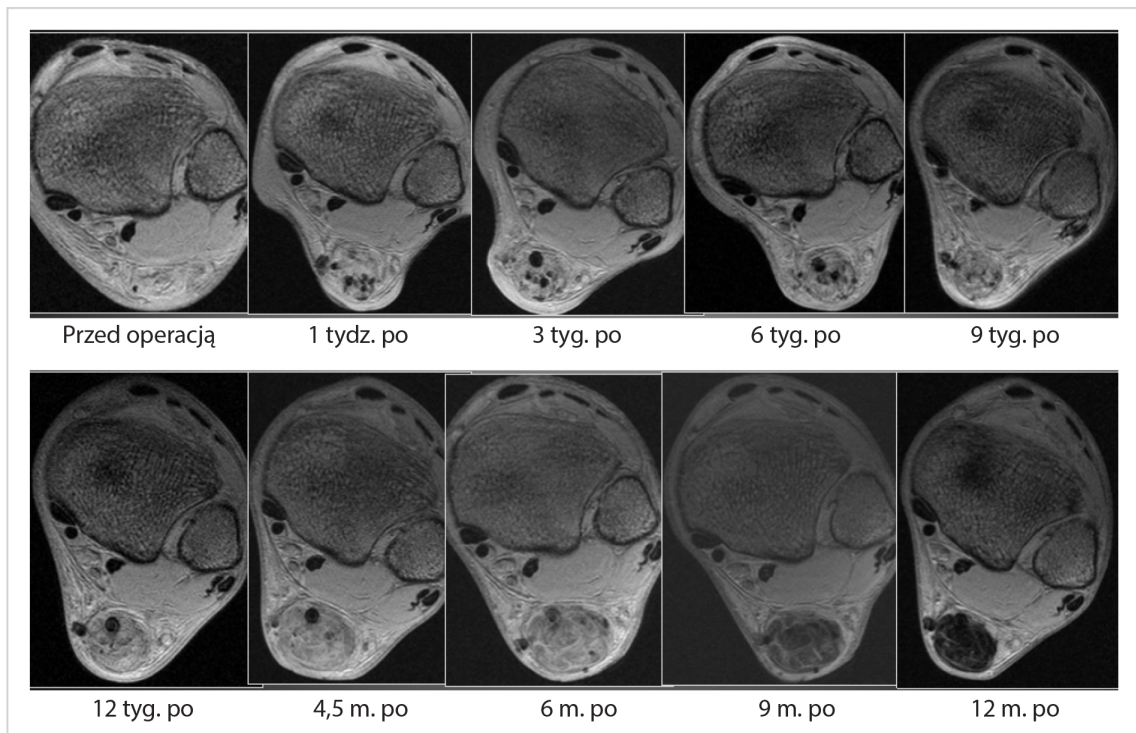
[Najpierw charakterystyka/ potem wyniki i ich omówienie]

Eksperymenty podzielono na dwa etapy. W pierwszym zrealizowano zagadnienia związane z doбором parametrów komponentów nowej metody i oceną jej skuteczności działania. W drugim porównano skuteczność nowej metody z innymi koncepcjami oceny gojenia ścięgna Achillesa.

W ramach pierwszego etapu wyłoniono, z pośród 10-ciu modalności RM analizowanych w pracy, jedną sekwencję będącą najlepszym kandydatem na dane wejściowe. W tym celu zrealizowano zarówno badania związane z analizą wizualną jak i testami ilościowymi. Wybrana sekwencja to T2* GRE TE_MIN, w której obraz ścięgna najlepiej różnicuje etapy gojenia się. Taki proces widoczny jest na rysunku poniżej.

Na przestrzeni całego procesu gojenia się ścięgna, zmiany uwidaczniają się w postaci silnie rozróżnialnych odcieni szarości w obszarze ścięgna, co nie występuje w innych sekwencjach. Jednocześnie zbadano, iż dodawanie kolejnych sekwencji nie miało istotnie statycznego wpływu na polepszenie jakości predykcji wartości parametrów radiologicznych, a skutkowało wydłużeniem czasu badania pacjenta.

W ramach badań nad doбором parametrów nowej metody, po pierwsze, zrealizowano wyliczenie parametrów ekstraktora cech poprzez zadanie treningu z celem binarnego podziału obrazów pacjentów na zdrowe i chore ścięgna. Przetestowano 3 architektury sieci neuronowych i na podstawie analizy ich dokładności klasyfikacji oraz czasu treningu wybrano sieć AlexNet. Takie podejście umożliwiło zakodowanie w ekstraktorze charakterystycznych cech dla tkanki zdrowej i patologicznej oraz finalne pogrupowanie ich w wektor o wymiarze 4096. Analiza tego wektora dla 10-ciu wybranych losowo pacjentów, uwidoczniała możliwość redukcji do 200 czynników głównych z zachowaniem blisko 99% poziomu wariancji. W kolejnym kroku podjęto próbą fuzji tych cech z 46-cioma cechami wybranymi na podstawie analizy literatury, ekstrahowanymi z obszaru ścięgna. W tym celu wykorzystano własność metody



Rysunek 2: Proces gojenia się ścięgna Achillesa widoczny w obrazach zrekonstruowanych na podstawie danych z sekwencji T2* GRE TE_MIN.

LASSO związaną z zerowaniem się współczynników mających marginalny wpływ na ostateczny wynik zadania. Zrealizowano badanie, dla którego wyniki metody LASSO (tj. błąd średniokwadratowy powiększony o funkcję kary) dla każdego pacjenta zestawiono w krzywe łączące kolejne punkty obrazowania. Kryterium selekcji współczynników była najlepsza korelacja wzorca odniesienia z krzywymi, którą uzyskano dla $\alpha = 0,1$. W rezultacie, dla 6-ciu parametrów wzorca odniesienia uzyskano różniące podzbiory n predyktorów, gdzie $n < 20$. Wykorzystując podzbiory i stosując metodę generalizacji stosów (stacking) wytrenowano oraz porównano szereg meta-regresorów: maszynę wektorów nośnych, wielowarstwowy perceptron, drzewa losowe jak i regresję liniową oraz nieliniową. W wyniku badania wybrano algorytm bazujący na wektorach nośnych, który dla każdego przekroju badania 3D oblicza ocenę z zakresu 0–7 dla zadanego parametru radiologicznego. Dla finalnej oceny badania 3D wybrano prostą metodę uśrednienia z odrzuceniem wartości skrajnych ocen poszczególnych przekrojów.

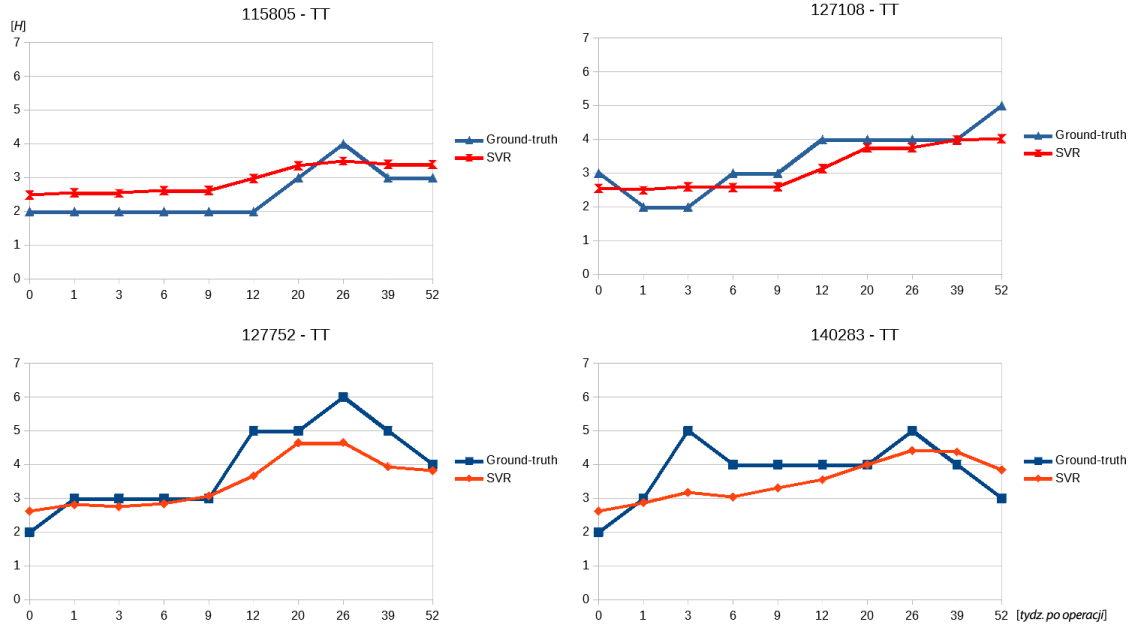
Opracowaną metodę (SVR) zwalidowano z wykorzystaniem zbioru testowego tj. 4-ech pacjentów (40 badań). Do oceny wykorzystano miary średniego błędu absolut-

Tabela 1: Wyniki oceny procesu gojenia z wykorzystaniem proponowanej metody.

Model		SCT	TT	STE	TE	TU	TisE
SVR	MAE	$1,05 \pm 0,06$	$0,56 \pm 0,03$	$0,75 \pm 0,04$	$0,91 \pm 0,05$	$0,91 \pm 0,04$	$0,94 \pm 0,05$
	MAX-AE	2,62	1,82	1,92	2,54	2,01	2,38
	Corr	0,85	0,85	0,31	0,72	0,65	0,80

nego (MAE) z informacją na temat błędu średniej, maksymalnego błędu absolutnego (MAX-AE) i średniej korelacji obliczonej z wykorzystaniem transformacji Z-Fishera (Corr). Ocenę wykonano dla wszystkich sześciu parametrów radiologicznych o nazwach: SCT, TT, STE, TE, TU, TisE. Szczegóły znajdują się w tabeli poniżej.

Dla każdego z pacjentów testowych przeanalizowano również krzywe gojenia, których przykład znajduje się na rysunku poniżej.



Rysunek 3: Ocena parametru TT.

Nową metodę porównano z dwiema innymi propozycjami oceny gojenia się ścięgna opracowanymi w grupie pod kierownictwem autora tej rozprawy, który odpowiadał za sformułowanie założeń eksperymentów i ich analizę, oraz z metodą bazującą na testach biomechanicznych realizowaną u partnera klinicznego. Pierwsze dwa podejścia wytrenowano w koncepcji end-to-end tak, aby modelowały *explicite* ocenę radiologa wykorzystując (1) poetykietowane obrazy RM, (2) poetykietowane obrazy ultrasonografii (USG). Wyniki przedstawione są w tabeli poniżej.

opracowana metoda została porównana z ATRS i deficyty sił mięśniowych bę-

Tabela 2: Porównanie wyników wnioskowania z wykorzystaniem zbioru testowego dla metody proponowanej tj. SVR oraz metody wyszkolonej w paradygmacie end-to-end. Pogrubieniem oznaczono najlepsze rezultaty, a kolorem czerwonym istotne statystycznie różnice w liczonych średnich ($p < 0,05$).

Model		SCT	TT	STE	TE	TU	TisE
Inception-v3 _e	MAE	1,12±0,08	0,80±0,04	1,40±0,07	0,89±0,05	1,08±0,04	0,69±0,07
	MAX-AE	2,14	1,01	2,13	1,18	1,44	0,78
	Corr	0,82	0,77	0,05	0,59	0,52	0,77
SVR	MAE	1,05±0,06	0,56±0,03	0,75±0,04	0,91±0,05	0,91±0,04	0,94±0,05
	MAX-AE	2,62	1,82	1,92	2,54	2,01	2,38
	Corr	0,85	0,85	0,31	0,72	0,65	0,80

dące rezultatem badań biomechanicznych. Badania te były realizowane w ramach protokołu monitorowania rehabilitacji pacjentów opracowanego w klinice Carolina Medical Center.

Można zaobserwować, że wskazania radiologa słabo korelują z wynikami badań biomechanicznych oraz ATRS, co jest zgodne z obecną praktyką kliniczną. Fakt ten jest znany ortopedom i ekspertom dziedzinowym, którzy w swojej praktyce spotykają pacjentów z patologiami ścięgna Achillesa z symptomami bólu, deficytami funkcjonalnymi jak również z brakiem tych objawów (tzw. pacjentów asymptomatycznych). Dlatego zalecane jest wykonywanie badań radiologicznych, które dostarczają komplementarnych informacji.

Interesującym rezultatem jest również wynik istotnej, niewielkiej korelacji oceny automatycznej z ATRS. Można wnioskować, że eliminacja losowości wynikającej z czynnika ludzkiego, spowodowała usystematyzowanie oceny i uwspólnienie pewnych informacji, które obecne są również w ustrukturyzowanym teście ATRS. Wynik ten wyznacza ciekawy kierunek dalszych prac, mających na celu stworzenie komplementarnego testu maksymalizującego informację dla lekarza prowadzącego, przy jednoczesnym minimalizowaniu objętości protokołu monitorowania pacjentów. Powyższe wnioski w opinii autora tej pracy stanowią bardzo duże pole do optymalizacji obecnej praktyki klinicznej w zakresie diagnostyki ścięgna Achillesa i są dalej dyskutowane w kolejnym rozdziale dotyczącym możliwości zastosowań praktycznych przedmiotowego rozwiązania.

Wnioski końcowe

Tabela 3: Porównanie wyników oceny automatycznej, bazującej na danych USG i RM, dla pacjentów ze zbioru testowego. Pogrubieniem oznaczono najlepsze wyniki. Kolorem czerwonym oznaczono poprawę w stosunku do kolejnego wyniku istotną statystycznie z $p < 0,05$.

Model		USG – p. strzałkowy					
		SCT	TT	STE	TE	TU	TisE
Inception-v3 _{eus}	MAE	0,81 ±0,19	0,63±0,03	0,56 ±0,09	0,85±0,10	0,54±0,02	0,87±0,14
	MAX-AE	1,59	1,79	1,7	1,25	1,38	1,69
	Corr	0,80	0,77	0,31	0,52	0,69	0,62
ResNet-50 _{eus}	MAE	0,88±0,16	0,65±0,07	0,66±0,04	0,83±0,12	0,75±0,06	0,93±0,11
	MAX-AE	1,49	1,26	1,74	1,48	1,78	1,71
	Corr	0,60	0,55	0,25	0,55	0,34	0,56
		USG – p. poprzeczny					
Inception-v3 _{euo}	MAE	0,84±0,27	0,75±0,7	0,58±0,05	0,83±0,05	0,53 ±0,08	0,83 ±0,15
	MAX-AE	2,8	1,46	1,51	1,27	1,63	1,65
	Corr	0,69	0,68	0,45	0,51	0,66	0,68
ResNet-50 _{euo}	MAE	0,92±0,18	0,76±0,16	0,68±0,04	0,81 ±0,08	0,65±0,10	0,94±0,05
	MAX-AE	2,01	1,56	1,61	1,69	1,43	1,58
	Corr	0,55	0,57	0,35	0,44	0,39	0,61
		Rezonans magnetyczny					
SVR	MAE	1,05 ± 0,06	0,56 ±0,03	0,75 ± 0,04	0,91 ± 0,05	0,91 ± 0,04	0,94 ± 0,05
	MAX-AE	2,62	1,82	1,92	2,54	2,01	2,38
	Corr	0,85	0,85	0,31	0,72	0,65	0,80

Tabela 4: Korelacja badań biomechanicznych i testu ATRS z badaniami radiologicznymi ocenionymi automatycznie i przez radiologa. Oznaczone wsp. korelacji są istotne z $p < 0,05$ dla $N=30$.

	ocena automatyczna	ocena radiologa
WW	0,1197	0,0706
WZ	0,1576	0,1931
ZW	0,3429	0,3092
ZZ	0,3100	0,2996
ATRS	0,3854	0,2088

W ramach pracy osiągnięto wszystkie założone cele. Zrealizowane badania zostały pozytywnie ocenione przez panel ekspertów Narodowego Centrum Badań i Rozwoju z rekomendacją do ich komercjalizacji. Przedmiotowe rozwiązanie, bazujące na głębokich sieciach neuronowych, może być podstawą do budowy narzędzi komputerowo wspomaganej radiologii, w szczególności w zakresie bieżących potrzeb tj. przyspieszenia generowania raportów, polepszenia personalizacji diagnostyki i poprawy jej jakości.

Wstępne szacunki przedstawione w pracy, bazujące na liczbie wykonywanych rocznie opisów badań RM ścięгна Achillesa w USA i Europie Zachodniej wskazują, że z wykorzystaniem wspomaganie generowania raportów w zakresie przedmiotowe-

go badania, sumarycznie, możliwa jest oszczędność do 83 tys. godzin pracy radiologów rocznie. Opracowane metody mogą również wspomóc wczesną detekcję zmian strukturalnych ścięgna i zredukować finalną liczbę zerwań, których np. w grupie sportowców (najliczniejszej biorąc pod uwagę problemy ze ścięgnem Achillesa) jest obecnie ok. 1 mln. rocznie. Finalnie, należy również podkreślić, iż certyfikowane narzędzie zaprojektowane na bazie przedstawionych algorytmów, żywotnie wpisałoby się w ideę poprawy jakości diagnostyki, umożliwiając gromadzenie ustrukturyzowanych danych, porównań między pacjentami i holistycznego wnioskowania będącego katalizatorem innowacji. Autor ma nadzieję na realizację kolejnych etapów w ramach działań związanych z założoną w Styczniu 2020 r. spółką spin-off Uniwersytetu Warszawskiego.