



POLSKA AKADEMIA NAUK
INSTYTUT BIOCYBERNETYKI I INŻYNIERII BIOMEDYCZNEJ

Praca doktorska

Proces gojenia ściegna Achillesa oceniany przez fuzję danych z
wykorzystaniem głębkich sieci neuronowych

Autor: mgr inż. Norbert Kapiński
Kierujący pracą: dr hab. inż. Antoni Grzanka
Promotor pomocniczy: dr Jakub Zieliński

Warszawa, wrzesień 2018

Streszczenie

The abstract will go here....

W tym miejscu można umieścić abstrakt pracy. W przeciwnym wypadku należy usunąć/zakomentować niniejszy fragment kodu.

Spis treści

1 Wstęp	1
2 Cel i przebieg pracy	2
3 Monitorowanie procesu gojenia ścięgna Achillesa	3
3.1 Ściegno Achillesa	3
3.1.1 Anatomia	4
3.1.2 Biomechanika	6
3.1.3 Urazy i czynniki im sprzyjające	7
3.1.4 Fazy gojenia, leczenie i rehabilitacja	10
3.2 Zastosowanie rezonansu magnetycznego	13
3.3 Zastosowanie ultrasonografii	18
3.4 Zastosowanie badań biomechanicznych	23
4 Konwolucyjne sieci neuronowe	27
4.1 Zarys historyczny	30
4.2 Szkolenie głębokich sieci neuronowych	33
4.2.1 Problem nadmiernego dopasowania	38
4.2.2 Problem redukcji wymiarowości	41
4.3 Przykłady współczesnych topologii	44
4.3.1 AlexNet	47

4.3.2	GoogLeNet	49
4.3.3	ResNet	52
4.3.4	Złożenia	55
4.4	Zastosowania w medycynie	56
5	Nowa metoda oceny procesu gojenia ścięgna Achillesa	61
5.1	Metodyka	62
5.1.1	zbiór danych	64
5.1.2	wzorzec odniesienia	67
5.2	Eksperymenty	70
5.2.1	Rozróżnienie ścięgna zdrowego i po zerwaniu	70
5.2.2	Obliczanie redukcja wymiarowości wektora cech z sieci neuronowej	72
5.2.3	Wybór sekwencji rezonansu magnetycznego	72
5.2.4	Optymalna konkatenacja cech	72
5.2.5	Fuzja z wykorzystaniem meta-regresji	72
5.3	Ocena procesu gojenia	72
6	Wyniki i walidacja	73
6.1	Ocena procesu gojenia z użyciem nowej metody	73
6.2	Porównanie z wynikami z rezonansu magnetycznego	73
6.3	Porównanie z wynikami ultrasonografii	73
6.4	Porównanie z wynikami badań biomechanicznych	73
7	Podsumowanie	74
Bibliografia		88
A AchillesDL: System komputerowego wspomagania oceny gojenia ścięgien i więzadeł		89

Spis rysunków

1.1	Podział przedstawiający różne rodzaje współczesnych głębokich sieci neuronowych.	1
3.1	Lokalizacja mięśnia trójdługiego łydki wraz ze ściegkiem Achillesa.	3
3.2	Schemat hierarchicznej budowy ściegna Achillesa.	5
3.3	Schemat generowania skurcza mięśni.	7
3.4	Schemat modelu Hilla.	7
3.5	Wykres przedstawiający proporcje występowania urazu ściegna Achillesa w różnych dyscyplinach sportowych na przykładzie amerykańskiego społeczeństwa.	9
3.6	Parametry ośrodków często mierzonych w badaniach USG.	20
4.1	Porównanie schematów przetwarzania danych z wykorzystaniem metod głębokiego uczenia się i innych algorytmów.	30
4.2	Topologia perceptronu.	31
4.3	Topologia perceptronu wielowarstwowego.	32
4.4	Topologia sieci LeNet.	33
4.5	Reprezentacja graficzna oceny krzyżowej.	39
4.6	Dwuwymiarowa przestrzeń cech wraz z zaznaczonymi poziomami wariancji i składowymi głównymi.	43

4.7	Błąd top-5 klasyfikacji obiektów w kolejnych latach uzyskiwany przez zwycięzców konkursu ILSVRC	46
4.8	Schemat architektury AlexNet.	47
4.9	Topologia architektury AlexNet z podziałem na dwa akceleratory GPU.	49
4.10	Topologia architektury AlexNet z podziałem na dwa akceleratory GPU.	50
4.11	Topologia architektury GoogleNet	50
4.12	Schemat funkcjonalny pojedynczego bloku w architekturze ResNet.	52
4.13	Topologia architektury ResNet-18.	54
4.14	Statystyki dotyczące publikacji medycznych zawierających słowa kluczowe związane z głębokim uczeniem się.	57
4.15	Porównanie automatycznej klasyfikacji retinopatii cukrzycowej i cukrzycowego obrzęku plamki z oceną panelu ekspertów.	58
4.16	Porównanie automatycznej klasyfikacji 3 chorób skóry z oceną ekspertów dermatologów.	58
5.1	Schemat automatycznej metody oceny procesu gojenia się ścięgna Achillesa.	63
5.2	Wizualizacja przykładowych danych RM w sekwencji PD dla zdrowego (po lewej) i gojącego się (po prawej) ścięgna Achillesa.	66
5.3	Wizualizacja przykładowych danych USG w kolejnych tygodniach po zszyciu ścięgna w przekrojach osiowych i strzałkowych.	68
5.4	Porównanie czasów treningu architektur na dwóch architekturach GPU: Pascal i Volta.	71

Spis tabel

4.1 Parametry architektury GoogleNet	51
5.1 Wyniki szkolenia sieci neuronowych z wykorzystaniem kroswalidacji z podziałem na 5 segmentów dla problemu klasyfikacji obrazów zdrowego i chorego ściegna.	70

Rozdział 1

Wstęp

Logistic regression — 1958

Hidden Markov Model — 1960

Stochastic gradient descent — 1960

Support Vector Machine — 1963

k-nearest neighbors — 1967

Artificial Neural Networks — 1975

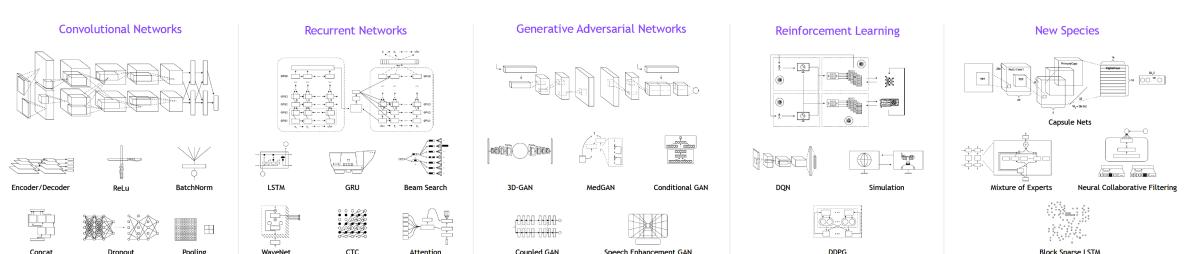
Expectation Maximization — 1977

Decision tree — 1986

Q-learning — 1989

Random forest — 1995

4% badań dotyczy oceny postępu w leczeniu [NVIDIA]



Rysunek 1.1: Podział przedstawiający różne rodzaje współczesnych głębokich sieci neuronowych.

Rozdział 2

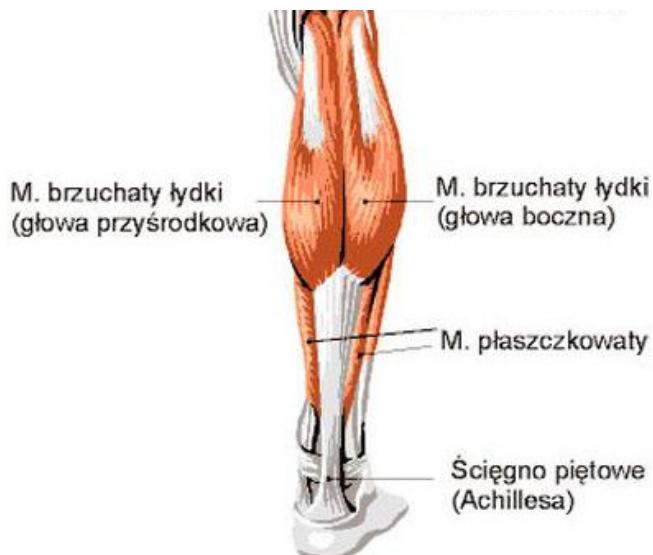
Cel i przebieg pracy

Rozdział 3

Monitorowanie procesu gojenia ścięgna Achillesa

3.1 Ścięgno Achillesa

Ścięgno Achillesa, nazywane również ściegnem piętowym, jest największym i zazwyczaj wytrzymującym największe obciążenia ściegkiem występującym w ciele ludzkim. Stanowi wspólne zakończenie mięśnia trójdługiego łydki, w którego skład wchodzą dwie głowy mięśnia brzuchatego i mięsień płaszczkowy. Całość struktury zlokalizowana jest w tylnym, powierzchniowym przedziale łydki co zobrazowano na Rysunku 3.1.



Rysunek 3.1: Lokalizacja mięśnia trójdługiego łydki wraz ze ściegkiem Achillesa.

Procesy patofizjologiczne zachodzące po urazie ściegna są ściśle związane z jego anatomią, biomechaniką, typem urazu i czynnikami jemu sprzyjającymi. Wszystkie te aspekty mają wpływ na możliwość monitorowania procesów gojenia się ściegna i jego wspomagania, dlatego zostaną szczegółowo omówione w kolejnych podsekcjach.

3.1.1 Anatomia

Z obu głów (brzuśców) mięśnia brzuchatego łydki wyrasta jedno szerokie, płaskie pasmo, które jest początkiem części brzuchatej ściegna Achillea. Następnie ściegno to łączy się z włóknami pochodząymi od mięśnia płaszczkowatego, które układają się stycznie do wcześniej powstałej struktury. Podążając w dół łydki, kształt ściegna Achillea ulega stopniowemu zwężeniu i zaokrągleniu, aż do punktu o minimalnej szerokości (około 4 cm nad przyczepem dolnym [42]). W rejonie samego przyczepu, znajdującego się na guzie piętowym, ściegno ponownie jest płaskie i szerokie.

Średnia długość ściegna Achillea to 15 cm (zakres to: 11 – 26 cm). Średnia szerokość w rejonie początku wynosi 6,8 cm (4,5 – 8,6 cm), w rejonie zwężenia to 1,8 cm (1,2 – 2,6 cm), a w miejscu przyczepu dolnego 3,4 cm (2,0 – 4,8 cm) [42, 67].

Ściegno jest zbudowane z tkanki łącznej, a dokładniej mówiąc z tkanki łącznej właściwej zbitej. Składa się w przeważającej części z *istoty międzymiędziorękowej* (nazywanej też *macierzą międzymiędziorękową*) zbudowanej z istoty podstawowej (ang. *ground substance*) oraz włókien. Dopełnienie struktury stanowią komórki takie jak fibroblasty, komórki tuczne, komórki plazmatyczne, histiocity i komórki napływowo.

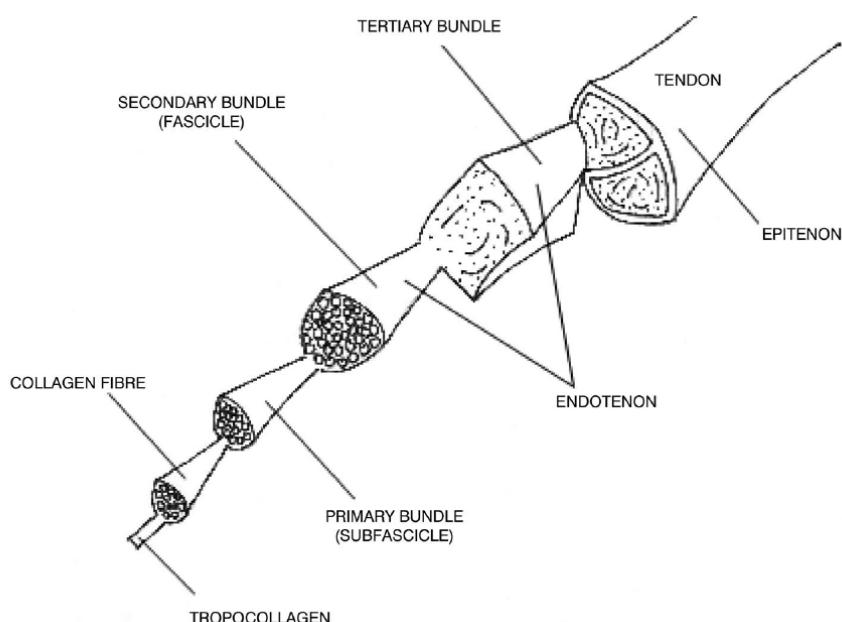
Istota podstawowa jest rodzajem żelu wiążącym duże ilości wody, włókien i komórek. Dokładniej mówiąc pełni ona rolę otoczki zapewniając możliwość transportu wody i substancji odżywczych do wnętrza struktury [112], gdzie znajdują się komórki poukładowane w wąskie pasy leżące pomiędzy włóknami [81]. Wśród włókien można wyróżnić trzy rodzaje: kolagenowe, siateczkowe i sprężyste.

Włókna kolagenowe i siateczkowe są zbudowane z fibrylnego białka – kolagenu, najczęściej występującego białka w organizmie człowieka, stanowiącego około 25% wszystkich białek. Makrocząsteczki kolagenu składają się ze zwiniętych łańcuchów polipeptydowych tworzących helisy. Dotychczas zidentyfikowano 20 typów kolagenu różniących się od siebie szczegółową budową helisy. Kolagen typu I stanowi około 90% wszystkich typów i jest podstawowym budulcem włókien kolagenowych tworzących zazwyczaj wiązki o grubości 50–100 μm . Włókna siateczkowe zbudowane są natomiast w

przeważającym stopniu z kolagenu typu III. Pozostałe typy kolagenu występują w znacząco mniejszym stopniu tworząc włókienka kolagenowe i siateczkowe o różnej grubości – od 10 do 300 nm (por. [107]).

Włókna sprężyste występują w postaci sieci i mają średnicę 0,2–10 μm . Zbudowane są z *elastyny*, rozciągliwego i sprężystego białka. Dzięki temu włókna te pod wpływem działania siły zewnętrznej mogą zwiększać swoją długość nawet o 50% (za. [107]).

W strukturze zdrowego ścięgna Achillesa widoczne są ściśle upakowane włókna sprężyste i kolagenowe. Włókna kolagenowe stanowią 70% masy suchego ścięgna i charakteryzuje się hierarchiczną strukturą zobrazowaną na Rys. 3.2.



Rysunek 3.2: Schemat hierarchicznej budowy ścięgna Achillesa.

Wyróżnić można: włókna, wiązki pierwszo, drugo i trzeciorządowe oraz ościegno [112]. W przeciwieństwie do innych ścięgien, ścięgno Achillesa nie posiada pochewki ścięgnistej, lecz jest otoczone ościegiem utworzonym z tkanki łącznej włóknistej. Struktura ta jest bogata w naczynia krwionośne i pełniąc funkcje transportowe jest bardzo ważnym elementem w procesie gojenia.

Oprócz otaczającej tkanki łącznej, ścięgno czerpie źródło unaczynienia z brzuśców mięśni brzuchatego i płaszczkowatego oraz z połączenia kostno-ścięgnistego. Najsłabsze unaczynienie występuje na poziomie ok. 4-5 cm powyżej górnego brzegu kości piętowej (por. [27]).

Unerwienie rejonu ściegna Achillesa zapewniają nerw piszczelowy, biegnący wzdłuż całego ściegna, a także nerw łydkowy, który krzyżuje się ze ściegkiem w odległości 8,7-12,4 cm proksymalnie od guza piętowego (por. [27]).

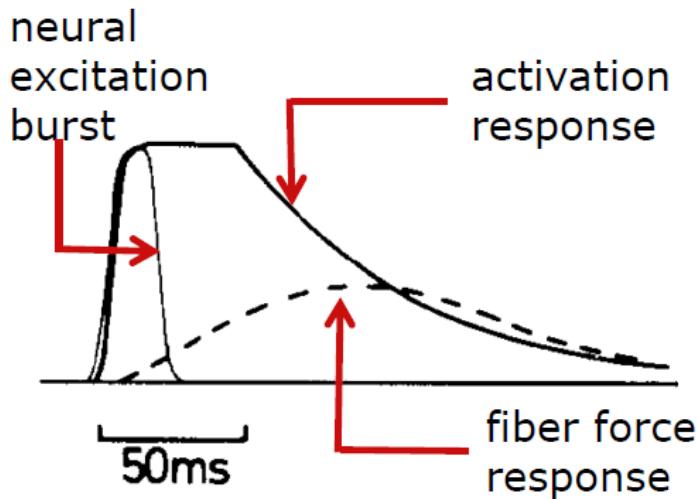
3.1.2 Biomechanika

Zadaniem ściegien jest transfer siły mięśniowej do układu szkieletowego. Pod względem mechanicznym ściegno piętowe jest najsilniejszym ściegiem całego ustroju. Dla przykładu podczas chodu maksymalne obciążenie ściegna Achillesa wynosi 500 N, przy biegu jest to 9000 N, natomiast podczas wyskoku może sięgać 12000 N, co stanowi równoważność 12–15 krotnej masy ciała. Podobne obciążenia wytrzymuje tylko ściegno właściwe rzepki [118].

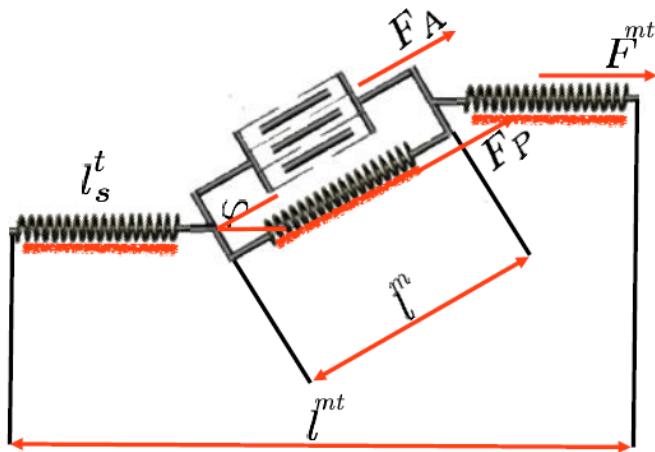
Głównym zadaniem ściegna Achillesa w trakcie prawidłowego chodu, biegu, czy wyskoku jest ruch *zginania podeszwowego stopy*, a zatem wyprost powodujący wspięcie na palce. Takie zadanie implikuje zwiększone ryzyko nadmiernego napięcia, a w konsekwencji przeciążenia, rozwoju stanu zapalnego, a nawet uszkodzenia ściegna [118].

Cały proces ruchu rozpoczyna się od centralnego układu nerwowego skąd wysyłane są impulsy nerwowe. Trafiają one do odpowiednich grup mięśniowych za pośrednictwem *nerwu ruchowego*, struktury przewodzącej impulsy wzbudzające proces skurcza mięśni. Na styku nerwu z mięśniem znajduje się *synapsa nerwowo-mięśniowa* (tzw. *płytką motoneuronalną*). Do niej właśnie na skutek impulsu wydzielany jest neuroprze-kaźnik, *acetylocholina*. W wyniku działania tej substancji następuje dalsze pobudzenie błony komórki mięśniowej i uwolnienie jonów wapnia Ca^{2+} . Cząsteczki te aktywują białka kurczliwe tj. *aktynę* i *miozynę*, a także pośrednio wywołują produkcję w mięśniu *adenozyno-5'-trifosforanu*, *ATP*, czyli substancji zapewniającej dostarczenie odpowiedniej energii chemicznej dla procesu.

Skurcz mięśni szkieletowych trwa 100–300 milisekund. Schemat czasów poszczególnych aktywacji przedstawiono na Rys. 3.3. Jak wcześniej zaznaczono, siła wygenerowana przez skurcz jest przekazywana za pomocą ściegien do układu szkieletowego. Dobrym przybliżeniem tego procesu jest *model Hill'a* [59], którego schemat przedstawiono na Rys. 3.4. Gdzie F_A oznacza siłę aktywnie wywoływaną skurczem mięśni, F_P to pasywna siła odpowiadająca za bezwładność tkanki miękkiej w mięśniu, a F^{mt} to wynikowa siła przekazywana przez ściegno, która zależy od długości ściegna l_t i kąta ϕ pomiędzy ściegiem a mięśniem.



Rysunek 3.3: Schemat generowania skurczu mięśni.



Rysunek 3.4: Schemat modelu Hilla.

To właśnie na skutek przekroczenia wartości granicznej siły F^{mt} dochodzi najczęściej do urazu ścięgna Achillesa. Dokładny opis tego problemu wraz z jego implikacjami został przedstawiony w kolejnej podsekcji.

3.1.3 Urazy i czynniki im sprzyjające

Uszkodzenie ścięgna Achillesa uznawane jest obecnie za chorobę cywilizacyjną (za [118]). Przykładowo, dla społeczeństwa amerykańskiego, urazy te występują u 18 na 100.000 osób rocznie [97]. Dodatkowo ryzyko ponownego zerwania ścięgna wynosi 20–40% [97].

Można wyróżnić dwa mechanizmy uszkodzenia ściegna Achillesa:

1. uraz bezpośredni – do którego zaliczyć można otwarte urazy spowodowane np. przecięciem ostrym przedmiotem takim jak szkło oraz urazy zamknięte spowodowane nagłym uderzeniem w napięte ściegno.
2. uraz pośredni – znacznie częściej spotykany niż uraz bezpośredni. Powstaje w wyniku nagłego, silnego skurcza mięśnia połączonego często z innymi siłami zewnętrznymi towarzyszącymi np. upadkowi.

O ile uraz bezpośredni jest najczęściej następstwem nieszczęśliwego wypadku, o tyle urazy pośrednie mają swoją przyczynę w rodzaju wykonywanych czynności oraz predyspozycji danej osoby.

Najczęściej do tego rodzaju urazów dochodzi podczas uprawiania sportu. Dotyczy to zwłaszcza sportu rekreacyjnego (około 70% przypadków [97, 118]). Aż 75% przypadków przytrafia się mężczyznom w przedziale wiekowym wahającym się od 30 do 50 roku życia [118]. Statystyki te mają związek z obniżeniem poziomu ukrwienia ściegna po 30 roku życia i trybem życia związany ze sporadycznym, ale nadmiernym obciążaniem ściegna przez osoby niewytrenowane [118].

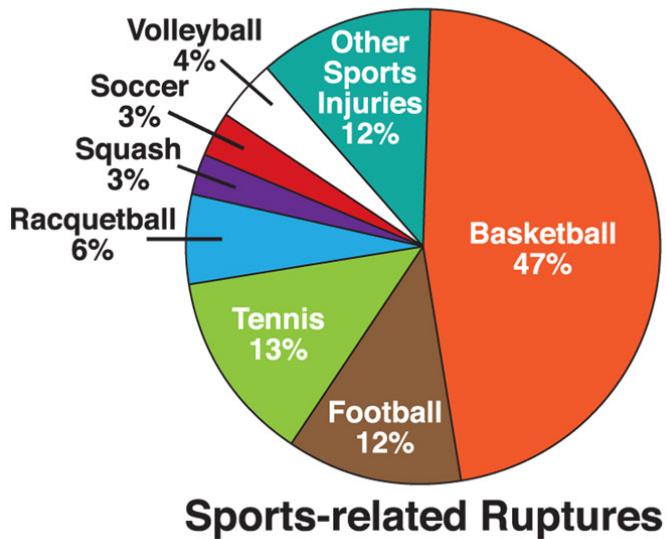
Dyscypliny sportowe, podczas których najczęściej dochodzi do urazu ściegna Achillesa to: piłka nożna, siatkówka, tenis, taniec, koszykówka, skoki, biegi, piłka ręczna, aerobic, pływanie, ping-pong, biegi, narciarstwo, balet. Proporcje są naturalnie uzależnione od popularności danego sportu w konkretnym kręgu ludzi. Dla przykładu na Rys. 3.5 zobrazowano udział poszczególnych dyscyplin w urazach ściegna Achillesa dla amerykańskiego społeczeństwa [97].

W Stanach Zjednoczonych sportem największego ryzyka jest koszykówka, natomiast w Europie jest to piłka nożna (około 60%). Według szacunków przedstawionych w [32] uraz ściegna Achillesa stanowi 5–10% wszystkich urazów atletycznych.

Powszechnie uważa się, że zdrowe ściegno piętowe jest bardzo ciężko zerwać z uwagi na jego dużą wytrzymałość. Dla przykładu wycinek ściegna o przekroju 1 cm^2 jest w stanie utrzymać masę 500–1000 kg [84]. Do urazów dochodzi zatem najczęściej, gdy ściegno jest zmienione patologicznie. Czynniki powodujące osłabienie ściegna Achillesa można podzielić na: zewnętrzne i wewnętrzne.

Do czynników wewnętrznych należy zaliczyć:

- mechaniczne uszkodzenie ściegna – uraz bezpośredni (np. częściowe przecięcie);



Rysunek 3.5: Wykres przedstawiający proporcje występowania urazu ścięgna Achillesa w różnych dyscyplinach sportowych na przykładzie amerykańskiego społeczeństwa.

- zmiany degeneracyjne – inwolucyjne (wiek powyżej 30 roku życia), na podłożu stanów zapalnych (przewlekłe stany zapalne), związane z mikrourazami;
- nadmierne przeciążenia – nieprawidłowa aktywność fizyczna;
- choroby metaboliczne i układowe – takie jak np. toczeń trzewny, dna moczanowa, nadczynność tarczycy, gruczolaki przysadki mózgowej, reumatoidalne zapalenie stawów, kolagenozy;
- odchylenia anatomiczne i biomechaniczne – anomalie takie jak: piszczel szpotawa, stopa koślawą, stopa płasko-koślawą, niestabilność stawu skokowego, koślawość lub szpotawość tyłstopia, osłabienie i przykurcz mięśni okołostawowych, ograniczenie ruchomości stawu skokowo-goleniowego i skokowo-piętowego, nierówność kończyn dolnych, nieprawidłowy stereotyp chodu, pogłębione krzywizny kręgosłupa, zespół ciasnoty przedziałów powięziowych w obrębie przedziału tylnego goleni;
- zaburzenia naczyniowe – schorzenia takie jak: hyperlipidemia, cukrzyca lub wywołane nadmiernym paleniem tytoniu;
- dysbalans mięśniowy;

- zaburzenia proprioceptywne;
- niepełne wyleczenie poprzednich obrażeń;
- przyjmowanie kortykosteroidów;

Do czynników zewnętrznych zaliczyć można:

- błędy w metodyce prowadzenia zajęć – np. nadmierny trening, trening jednostronny, trening wszechstronny, z przewagą intensywności nad wytrzymałością czynnościową aparatu więzadłowo-stawowo-mięśniowego, chęć nadrobienia zaległości treningowych;
- nagła zmiana dyscypliny sportowej lub aktywności rekreacyjnej;
- inne takie jak: nieodpowiednia nawierzchnia, niekorzystne warunki pogodowe, nieodpowiednie obuwie, zbyt krótkie przerwy między zawodami lub treningami.

Obecnie, ustalenie jednej optymalnej metody leczenia dla każdego rodzaju uszkodzenia ściegna Achillesa nie jest możliwe. Wybór uwarunkowany jest poziomem i stopniem zniszczenia ściegna, stanem zdrowia pacjenta oraz kwalifikacjami, doświadczeniem i możliwościami danego ośrodka. W kolejnej sekcji omówione zostaną współczesne metody leczenia tego schorzenia, natomiast w Rozdziale 5 zaprezentowana zostanie nowa metoda monitorowania procesu gojenia się ściegna Achillesa, która może być wykorzystana do optymalizacji rehabilitacji opisanego problemu.

3.1.4 Fazy gojenia, leczenie i rehabilitacja

Proces gojenia się ściegna Achillesa można podzielić na trzy zachodzące na siebie etapy. W pierwszej kolejności dochodzi do stanu zapalnego. Jest to tzw. *faza zapalna* (ang. *inflammatory phase*), która trwa w pierwszym tygodniu gojenia się. Następnie występuje *faza proliferacji* (ang. *proliferative phase*), trwająca od drugiego do szóstego tygodnia. Cały proces kończy się *fazą przebudowy* (ang. *remodeling phase*), która może trwać nawet do 18 miesiąca po urazie [41, 102, 112, 138].

Faza zapalna zaczyna się od razu po wypadku. W miejscu urazu tworzy się skrzep krwi oraz uwalniane są prozapalne substancje chemiczne, cytokiny, które w znaczącym stopniu odpowiadają za proces aktywacji i migracji komórek zapalnych z otoczenia

do miejsca zapalenia [77]. Komórki te, a dokładniej mówiąc leukocyty i makrofagi, rozpoczynają *fagocytzę*, czyli proces mający na celu usunięcie obumarłej tkanki oraz skrzepu [25, 77, 138]. Symultanicznie dochodzi do aktywacji i selekcji *tenocytów* (komórek ścięgna), które rozpoczynają odbudowę włókien ścięgna [138] oraz inicują drugą fazę całego procesu.

W fazie proliferacji poziom komórek zapalnych zaczyna spadać. Czynniki wzrostu, uwalniane m.in. przez makrofagi, nasilają migrację i rozmnażanie się tenocytów wokół miejsca urazu. Skutkiem tego jest wytworzenie nowych włókien kolagenu typu III. Nowo-powstałe struktury są początkowo losowo zorientowane [25, 41, 138]. Dlatego rozpoczyna się trzecia faza mająca na celu odtworzenie struktury zdrowego ścięgna.

W trakcie trwania ostatniej fazy, nazywanej również *remodelingiem* włókna kolagenu typu III zastępowane są kolagenem typu I. Pod wpływem naprężeń struktura włókien zmienia się i wyrównuje zgodnie z kierunkiem występujących sił [138]. Włókna pracujące niezgodnie z przewidzianym ruchem ulegają zerwaniu i fagocytacie, a te które wspomagają zadany ruch ulegają wzmacnieniu.

Pomimo całego wyżej opisanego procesu, ścięgno wygojone strukturą odbiega od ścięgna zdrowego. Charakteryzuje się nawet dwukrotnie zwiększoną powierzchnią przekroju osiowego oraz podatnością na ponowne zerwania z uwagi na zmniejszoną odporność na obciążenia.

Uraz ścięgna Achillesa można leczyć zachowawczo lub operacyjnie. Zachowawcze leczenie stosuje się najczęściej przy częściowym uszkodzeniu ścięgna lub u osób, które mają przeciwwskazania do leczenia chirurgicznego. Wówczas stosowane jest unieruchamianie kończyny np. poprzez włożenie w but ortopedyczny lub gips. Dodatkowo można również stosować preparaty z komórek macierzystych i czynników wzrostu [102]. Wybór leczenia zachowawczego wiąże się z długotrwającym, trwającym około 10 tygodni unieruchomieniem kończyny przed przystąpieniem do rehabilitacji.

Leczenie operacyjne stosowane jest najczęściej w przypadku uszkodzenia całkowitego. Tuż po urazie lub po przejściu najintensywniejszych procesów fazy zapalnej wykonuje się szycie ścięgna. Stosowane są metody operacji na otwartym ścięgnie lub przeszkode. Zabieg przezskórny polega na zbliżeniu kikutów ścięgna specjalnym szwem przeprowadzanym przez małe nacięcia. W teorii metoda ta umożliwia szybsze obciążanie nogi po operacji, skraca okres unieruchomienia i umożliwia szybszą rehabilitację. W praktyce jednak chirurgom rzadko udaje się uzyskać jakość zszycia porównywalną z metodą otwartą.

Rehabilitacja po urazie również podzielona jest ze względu na fazy gojenia. W pierwszej fazie terapia skupia się wokół ochrony miejsca urazu, odciążenia nogi, redukcji obrzęku i złagodzenia stanu zapalnego. Dokonuje się również zabiegów mających na celu wczesną mobilizację ściegna i przywrócenia ślizgu. Ochrona miejsca urazu może być wykonana przez zastosowanie dwu-częściowej łuski gipsowej lub buta ortopedycznego. Redukcję obrzęku wykonuje się poprzez np. masaż limfatyczny. Stan zapalny łagodzony jest chłodzeniem i okładami oraz lekami, natomiast ślizg przykładowo stymuluje się elektrostymulacją oraz mobilizacją mięśni łydki poprzez zginanie stopy.

W drugiej fazie zabiegi rehabilitacyjne są kontynuacją działań wykonywanych podczas pierwszej fazy. Szczególny nacisk przewidziany jest na przeciwdziałanie powstawania zrostów poprzez pracę nad ślizgiem ściegna oraz mobilizacją tkanek miękkich. W szczególności rehabilitanci poświęcają dużo uwagi na ćwiczenia dotyczące mobilizacji blizny i tkanek okalających. Nasilane są również zabiegi związane z elektrostymulacją celem utrzymania masy mięśniowej oraz wykonywane są ćwiczenia związane z stabilizacją stawu skokowego związane z ćwiczeniami izometrycznymi mięśni prostujących i pronujących stopę. W tej fazie stosuje się wciąż ochronę miejsca urazu oraz aplikuje się metody redukcji obrzęku.

Trzeci etap tj. przebudowa to najdłuższy z etapów gojenia się tkanek ściegnistych, podczas którego zachodzi najwięcej zmian w strukturze, odporności ściegna oraz jego funkcji. W tym okresie można wyszczególnić następujące cele rehabilitacji:

- przeciwdziałanie zrostom tkanek otaczających Achillesa;
- wzmacnienie mięśni łydki – brzuchatego i płaszczkowatego;
- przywrócenie pełnego zakresu ruchu i kluczowych funkcji ściegna;
- zwiększenie elastyczności tkanek wokół ściegna.

Oprócz podstawowych zabiegów ochronnych i fizjoterapii do działań włączane są terapie związane z ćwiczeniami czynnymi z oporem; stymulujące technikę prawidłowego chodu; ćwiczenia wzmacniające obręcz biodrową; ćwiczenia sensomotoryczne w różnych pozycjach ciała oraz ćwiczenia wzmacniające kończyny dolne na niestabilnym podłożu. Powoli wprowadzane są również aktywności sportowe takie jak rower stacjonarny czy stretching. Celem finalnym tego etapu jest powrót do pełnej sprawności.

Dużą nadzieję wiążę się z nowymi technikami bazującymi na zastrzykach z komórek macierzystych. Ciekawą propozycją wydaje się możliwość budowania biodegradowal-

nych rusztowań (ang. *scaffold*), implementowanych podczas operacji i wspomagających implementację komórek macierzystych i czynników wzrostu w odpowiednie miejsca [30].

Celem oceny tego rodzaju nowatorskich podejść jak również z uwagi na konieczność optymalizacji istniejących metod rehabilitacji i leczenia urazów ścięgna Achillesa potrzebne jest monitorowanie procesu gojenia. Najczęściej stosowane techniki zostaną omówione w kolejnych sekcjach.

3.2 Zastosowanie rezonansu magnetycznego

Obserwację gojącej się struktury ścięgna Achillesa umożliwiają techniki stosujące różnego rodzaju fale przenikające ciało ludzkie. Sposób w jaki propaguje się fala tworząc sygnał zależy od własności związanych z ośrodkiem propagacji oraz parametrów fali. Do najistotniejszych należą wyrażana w hercach (w skr. Hz) częstotliwość, czyli liczba pełnych cykli drgań w ciągu sekundy, amplituda tj. maksymalna wartość sygnału falowego oraz faza określająca w której części cyklu znajduje się sygnał. W technikach obrazowania medycznego wykorzystywana jest interpretacja zdarzeń wynikających ze zmian tych parametrów oraz reagującej materii.

Pierwszą opisaną w tej pracy metodą obrazowania stosowaną do monitorowania gojenia się ścięgna Achillesa jest obrazowanie z wykorzystaniem *Jądrowego Rezonansu Magnetycznego* nazywanego dla uproszczenia *Rezonansem Magnetycznym*, w skr. RM. Metoda ta bazuje na odpowiednim odczycie reakcji jąder atomów na stałe pole magnetyczne generowane przez magnesy (zwykle nadprzewodzące), zmienne pole magnetyczne skierowane prostopadle do osi stałego pola (generowane w dedykowanej do tego zadania cewce), zmienne pole lokalne wytwarzane przez sąsiednie jądra atomowe oraz związane z nimi cząstki elementarne.

Zrozumienie własności jądra atomowego i reakcji na magnetyzację było możliwe dzięki badaniom naukowym z początku XX wieku. Zapewniły one praktyczną możliwość wykorzystania procesów bazujących na zachowaniach obiektów o bardzo małych masach i wymiarach takich jak cząstki elementarne. Jako przykłady można wskazać prace Alberta Einsteina dotyczące ruchów Browna oraz efektu fotoelektrycznego; opis funkcji falowej zaproponowany przez Erwina Shrodingera; badania Nielsa Bohra, które doprowadziły do utworzenia utylitarnego modelu atomu zawierającego poziomy energetyczne; wyprowadzenie reguły nieoznaczoności przez Werner Heisenberga, czy opisanie przez Wolfganga Pauliego konfiguracji cząstek elementarnych w atomie.

W dużej mierze na bazie powyższych prac zdefiniowana została teoria mechaniki kwantowej umożliwiająca między innymi pogłębienie wiedzy na temat magnetycznej natury składników jądra atomowego tj. *protonów i neutronów*. Pierwszy ważny przełom nastąpił w 1933 r., kiedy to Frisch i Stern w [49] opisali moment magnetycznego protonu. Rok później, w [29] Breit i Rabi zaproponowali metodę rezonansową obserwacji własności magnetycznych jąder atomowych. W 1946 praca ta została udoskonalona i wykorzystana przez Felix Bloch'a do badania częstek w wodzie i parafinie [26]. W końcu dzięki pracom m.in. Lautenbur'a [72] i Mansfield'a [83] rozwinęły się metody lokalizacji przestrzennej powyższych zmian własności magnetycznych, co umożliwiło utylitarną prezentację wizualną wyników.

Opisując zagadnienie dokładniej, należy zwrócić uwagę na główną własność jądra atomowego umożliwiającą działanie RM, czyli *spin*. Można powiedzieć, że jądro posiada spin jeżeli nie ma jednocześnie parzystej liczby protonów i neutronów (za [91]). Dla przykładu jądro wodoru składające się z 1 protonu spełnia powyższy warunek.

Umieszczone w polu magnetycznym jądra atomowe posiadające spin wykonują ruch wirowy zwany *precesją*. Częstotliwość tej precesji jest proporcjonalna do natężenia pola magnetycznego i opisana jest *równaniem Larmora* 3.1:

$$\omega_0 = \gamma * \beta_0, \quad (3.1)$$

gdzie ω_0 to częstotliwość kołowa protonów, $f_l = \omega_0 / 2\pi$ to tzw. częstotliwość Larmora, γ to współczynnik żyromagnetyczny właściwy danemu protonowi, a β_0 to wartość natężenia danego pola magnetycznego. Dla przykładu częstotliwość Larmora dla protonu wodoru, gdzie $\gamma=42,57 \text{ MHz/T}$ dla pola $1,5\text{T}$ równa jest $63,8 \text{ MHz}$.

Jądra atomowe wirujące z częstotliwością Larmora na skutek odpowiedniej sekwencji impulsów magnetycznych ulegają swego rodzaju uporządkowaniu, tworząc rotujące w przybliżony sposób wektory magnetyzacji. Dodatkowo istotne jest, że wektory stowarzyszone z jądrami o wysokich poziomach energii odchylają się przeciwnie do zwrotu pola magnetycznego – antyrównolegle, a o niskich poziomach energii równolegle, czyli zgodnie ze zwrotem pola magnetycznego. W przeciętnym wycinku ciała ludzkiego jest znaczco więcej jąder o ustawieniu równoległym niż antyrównoległym. Przykładowo w objętości $1 \times 1 \times 5 \text{ mm}$ jest to liczba około 2×10^{15} częstek.

Sygnał RM pochodzi od różnicy magnetyzacji jąder ustawionych równolegle i antyrównolegle (różnicę tę oznaczamy jako M_0 – tj. całkowita magnetyzacja). Wyjściowo M_0

ma większą wartość zgodną ze zwrotem pola magnetycznego, czyli tzw. *magnetyzacji podłużnej*. Kierunek prostopadły do magnetyzacji podłużnej nazywany jest *magnetyzacją poprzeczną*.

Zmianę M_0 wykonuje się przy pomocy odpowiednio sterowanej cewki nadającej impulsy o częstotliwości fal radiowych (tzw. *impulsy RF* – od ang. *Radio Frequency*). Impuls RF jest krótkotrwały i ma taką samą częstotliwość, co częstotliwość precesji. Dzięki temu zachodzi zjawisko *rezonansu magnetycznego*, czyli wymiany energii między jądrami atomowymi, a falą radiową. Cząstka pochłania wówczas energię *fotonu*, tj. nośnika oddziaływania elektromagnetycznego, i uzyskuje wyższy poziom energetyczny, odchylając wektor magnetyzacji i zaburzając tym samym uporządkowanie. Następuje zmniejszenie M_0 w kierunku podłużnym oraz wzrost magnetyzacji poprzecznej. Po wyłączeniu impulsu RF magnetyzacja wraca do wartości początkowej. Czas potrzebny na zwiększenie magnetyzacji podłużnej do stanu z przed działania impulsu nazywany jest czasem *relaksacji podłużnej*, *relaksacji spin – sieć* lub *czasem T1*. Natomiast czas potrzebny na obniżenie magnetyzacji poprzecznej nazywany jest czasem *relaksacji poprzecznej*, *relaksacji spin – spin* lub *czasem T2*. Relaksacja spin–spin ma przede wszystkim znaczenie w kontekście oddziaływań z sąsiednimi jądrami i występującą na skutek tego dyfuzją fazy obniżającą wartość magnetyzacji poprzecznej. Jednak z uwagi na globalny powrót do namagnesowania wzdłuż linii pola czas relaksacji T2 nigdy nie jest dłuższy niż czas relaksacji T1.

Lokalizacja zmian magnetyzacji w poszczególnych *wokselach* (tj. wycinkach przestrzeni 3D o określonych wymiarach) odbywa się przy pomocy celowanego różnicowania fazy i częstotliwości wektorów magnetyzacji. Operacja ta wykonywana jest przy pomocy sygnału z dodatkowych cewek gradientowych. Warstwa, w której dokonywany jest pomiar kodowana jest przy pomocy nałożenia na pole β_0 pola gradientowego G_{zz} , co umożliwia wzbudzenie jąder atomowych w tych wokselach, dla których współrzędna z wyraża się wzorem:

$$\Delta z = \frac{\Delta\omega}{\gamma G_{zz}}, \quad (3.2)$$

gdzie $\Delta\omega$ to szerokość widmowa pobudzającego jądra impulsu RF. Współrzędna y kodowana jest przy pomocy zróżnicowania fazy, rotujących w różnych obszarach wektorów magnetyzacji, na skutek działania pola gradientowego G_{yz} — tzw. *gradientu kodowania fazy*. Zróżnicowanie fazy wyraża się wzorem:

$$\phi = \gamma G_{yz} y T_y, \quad (3.3)$$

gdzie T_y oznacza czas włączenia pola gradientowego G_{yz} . Ostatecznie, współrzędna x kodowana jest w wyniku zróżnicowanie częstotliwości precesji wektorów magnetyzacji poprzez włączanie pola gradientowego G_{xz} (zob. [5]).

Nieprzetworzona informacja z wokseli o współrzędnych x, y, z nazywana jest *przestrzenią K* i koduje ona składniki częstotliwości sygnału RM. Do transformacji z przestrzeni częstotliwości na przestrzeń obrazu używana jest operacja matematyczna zwana *odwrotną transformacją Fouriera*. Przykład takiego działania można znaleźć w [14].

Główną informacją jaką otrzymujemy za pomocą obrazowania metodą jądrowego rezonansu magnetycznego jest rozkład gęstości jąder atomowych. Jądra wodoru mają największy spośród jąder atomowych współczynnik żyromagnetyczny oraz występują w znaczącej liczbie, w prawie każdej z tkanek ludzkich (por. [91]). Dlatego ich rozkład jest najczęściej mierzony. Jedynie w szczególnych przypadkach stosuje się obrazowanie z wykorzystaniem częstotliwości rezonansu odpowiednich dla innych pierwiastków takich jak fosfor [120], sód [119] czy węgiel [44].

Dodatkowo, obserwacja zmiany sygnału RM w czasie, często dostarcza istotnych danych na temat właściwości fizyko-chemicznych tkanek. Z uwagi na ten fakt, dla uwydatnienia interesujących informacji, stosuje się szereg ustaleń urządzeń do rezonansu magnetycznego pogrupowanych w *sekwencje RM* oraz ich podgrupy nazywane w tej pracy *modalnościami*. Do badań opisanych w dalszej części tego manuskryptu zostało użytych 7 sekwencji, w tym jedna z 4 modalnościami. Zostaną one pokrótko scharakteryzowane poniżej:

- *T1 zależne* – obraz rekonstruowany jest na podstawie sygnału RM zmierzonym po odchyleniu impulsem RF, w momencie gdy wartość magnetyzacji podłużnej wynosi 63% wartości początkowej. W takim przypadku sygnał rezonansu magnetycznego MR_s można obliczyć z zależności:

$$MR_s \sim \gamma_{pd} * [1 - e^{-TR/T_1}], \quad (3.4)$$

gdzie γ_{pd} to gęstość protonowa tkanki, a TR to *czas repetycji* określany przez użytkownika.

- *T2 zależne* – obraz rekonstruowany jest na podstawie sygnału RM zmierzonym po odchyleniu impulsem RF, w momencie gdy wartość magnetyzacji poprzecznej wynosi 37% wartości początkowej. W takim przypadku sygnał rezonansu magne-

tycznego MR_s można obliczyć z zależności:

$$MR_s \sim \gamma_{pd} * e^{-TE/T2}, \quad (3.5)$$

gdzie TE to *czas echo* zdefiniowany przez użytkownika.

- *PD* (od ang. *Proton Density*) – obraz rekonstruowany na podstawie sygnału RM przy bardzo krótkim TE i bardzo długim TR . Wówczas sygnał jest wprost proporcjonalny do γ_{pd} :

$$MR_s \sim \gamma_{pd}. \quad (3.6)$$

- *T2 mapping* – w tej sekwencji wartość sygnału RM zależna od $T2$ (zob. 3.5) mierzona jest dla $8 TE$. Pozwala to na wyliczenie czasu T2 (zob. [99]).
- *T2 * GRE* (od ang. *Gradient Echo*) – jest to przykład *sekwencji gradientowych*. W tym przypadku pole gradientowe jest użyte w celu przyspieszenia dyfuzji fazy, a następnie przywrócenia jej z użyciem takiego samego pola o przeciwnym zwrocie. Całość ma na celu przyspieszenie procesu zbiegania wartości sygnału RM zależnego od $T2$ (zob. 3.5) do 37% wartości początkowej na podstawie której rekonstruowany jest obraz.
- *T2 * GRE TE_MIN* (od ang. *Minimal Time Echo*) – obraz rekonstruowany jest na podstawie sygnału RM $T2$ zależnego (zob. 3.5) przy minimalnym czasie TE i tylko dla fragmentu przestrzeni K.
- *3D FSPGR* (od ang. *Three-dimensional Fast Spoiled Gradient Echo*) – jest to szybka sekwencja gradientowa z usuwaną (ang. *spoiled*) pozostałością magnetyzacji poprzecznej. W wyniku sekwencji impulsów RF otrzymywane są 4 modalności:
 - In Phase Ideal – sygnał mierzony w czasie, gdy protony należące do tłuszcza i wody są w zgodnej fazie;
 - Out Phase Ideal – sygnał mierzony w czasie, gdy protony należące do tłuszcza i wody są w antyfazie;
 - Fat Ideal – sygnał mierzony przy maksymalnej wartości sygnału od protonów tłuszcza i minimalnej od protonów wody;
 - Water Ideal – sygnał mierzony przy maksymalnej wartości sygnału od protonów wody i minimalnej od protonów tłuszcza.

Powyższe modalności są możliwe do realizacji, gdyż w sekwencji 3D FSPGR sygnał RM odczytywany jest przy dwóch różnych czasach TE. Przy czym czasy TE są tak dobrane, aby sygnały od wody i tłuszczu wzmacniały się (In Phase Ideal) lub osłabiały (Out Phase Ideal). Jest to możliwe gdyż przy zadanym polu magnetycznym częstotliwość Larmora wodoru w wodzie jest minimalnie inna niż w tłuszczach (fakt ten nazywany jest również przesunięciem chemicznym i jest wykorzystywany w spektroskopii NMR [76]). Modalności Water Ideal i Fat Ideal są odpowiednią kombinacją liniową dwóch wcześniejszych – taką aby zminimalizować wpływ tłuszczu (Water Ideal) lub wody (Fat Ideal) na sygnał RM.

Rezonans magnetyczny może być wykorzystany do monitorowania gojenia się ścięgien i więzadeł, gdyż tkanki miękkie człowieka zawierają dużo atomów wodoru. W kontekście ścięgna Achillesa RM w szczególności pozwala radiologom na obserwację: ciągłości ścięgna w płaszczyźnie strzałkowej (ang. *sagittal plane*); uszkodzeń śródścięgnistych objawiających się przerwaniem w naturalnym warkoczu ułożonym z tkanek; pogrubienia ścięgna ocenianego w przekrojach osiowych (ang. *axial planes*); jednorodności, wrzecionowatości i innych nieregularności kształtu; ostrości ścięgna i jego rozgraniczenia od tkanek otaczających wraz ze zmianami występującymi na brzegach; obrzęku w okalających ścięgno tkankach; jednorodności charakteryzującej się podobieństwem przekrój sąsiednich oraz liczby ewentualnych zrostów.

Ograniczenia związane ze stosowaniem RM wynikają głównie z oddziaływaniami silnego pola magnetycznego na przedmioty z nim reagujące, zawierające ferromagnetyki w składzie np. odłamki metalowe w ciele badanego, wszczepiony stymulator serca, endoprotezy stawu biodrowego, stawu kolanowego itp. Wrażliwość pacjenta na hałas lub skłonności klaustrofobiczne również są wskazaniami do wykluczenia z badania. Ponadto urządzenie do rezonansu magnetycznego jest bardzo drogie, nieprzenośne i wymagające specjalnego otoczenia pracy. W celu wyeliminowania tych ograniczeń można posłużyć się inną metodą obrazowania medycznego opisaną w kolejnej sekcji.

3.3 Zastosowanie ultrasonografii

Kolejną z metod obrazowania medycznego jest *Ultrasonografia*, w skr. *USG* (ang. *Ultrasonography, US*). Bazuje ona na efektach związanych z propagacją w tkankach *ultradźwięków*, tj. fal akustycznych o częstotliwościach powyżej 20 kHz. W praktyce, z uwagi na parametry związane z rozdzielczością obrazu, są to fale o częstotliwości rzędu

kilku MHz.

Propagacja fal akustycznych w przyrodzie była tematem rozważań myślicieli takich jak Pitagoras, Arystoteles czy Galileusz, którzy ugruntowali pole badań pod kolejne osiągnięcia matematyczno-inżynierijne. W tej kwestii, do jednego z przełomów doszło w 1822 roku, kiedy to szwajcarski inżynier Daniel Colladen oraz matematyk Charles-Francois Sturm wyznaczyli przybliżoną prędkość rozchodzenia się fali akustycznej w wodzie. Badanie wykonano na Jeziorze Genewskim symultanicznie mierząc czas jaki potrzebny był dźwiękowi podwodnego wystrzału i sygnałowi dzwonka rozchodzącego się w powietrzu do przebycia drogi pomiędzy dwoma łódkami oddalonymi o 10 mil. Wyliczona wartość wyniosła wówczas 1435 m/s nie różniąc się znaczaco od dzisiaj przyjmowanej estymacji równej 1480 m/s. Analogicznie, stosując nowoczesne techniki, wyliczane są obecnie prędkości rozchodzenia się fali w innych ośrodkach takich jak powietrze, tkanki miękkie, kości itp.

58 lat po eksperymencie na Jeziorze Genewskim, w 1880 roku bracia Curie opisali w [62] *efekt piezoelektryczny*, czyli zjawisko polegające na pojawienniu się ładunku elektrycznego pod wpływem naprężeń mechanicznych w krysztale o anizotropowej budowie, takiej jak ma np. kwarc. W przypadku odwrotnym, przyłożenie napięcia do odpowiedniego kryształu generuje drgania (tzw. *odwrotny efekt piezoelektryczny*).

Efekty te są wykorzystywane w *głowicy aparatu usg*, przyrządu do generowania i odbierania ultradźwięków. Przykładowo, polaryzowanie kryształu piezoelektrycznego krótkim impulsem elektrycznym $\sigma_1(t)$ pobudza go do drgań na własnej częstotliwości rezonansowej. Zakładając, że ów kryształ ma kształt walca o grubości $d=0,64$ mm, to będzie stanowił rezonator półfalowy, w którym wystąpi drganie rezonansowe o długości fali $\lambda = 2d$ czyli $\lambda = 1,28$ mm. Jeżeli wykonany jest z tytanianu baru, dla którego prędkość propagacji drgań wynosi $c=4460$ m/s, to częstotliwość drgań własnych tego kryształu wyniesie:

$$f = \frac{c}{\lambda} = \frac{4460\text{m/s}}{1,28\text{mm}} = 3,5\text{MHz}. \quad (3.7)$$

Odwrotnie, powracająca fala tzw. *echo* wygeneruje impuls elektryczny $\sigma_2(t)$ w skutek drgań wywołanych w krysztale. Echo jest naturalnie falą różniącą się od sygnału nadawanego, a zmiany te są w przeważającym stopniu efektem zjawisk takich jak: *odbicia, załamania, dyfrakcja, rozpraszanie i pochłanianie* (zob. [82]).

Zjawiska te zależą od częstotliwości fali, która propaguje się w ośrodku o pewnej

3.3 Zastosowanie ultraso

Rozdział 3: Monitorowanie procesu gojenia ścięgna Achillesa

impedancji akustycznej ośrodka Z wyrażanej jako:

$$Z = \rho c = \sqrt{\epsilon \rho}, \quad (3.8)$$

gdzie c , to prędkość rozchodzenia się fali, ρ to gęstość ośrodka, a ϵ to *moduł odkształcalności objętościowej*, tj. parametr opisujący jak zmieni się objętość ośrodka pod danym ciśnieniem. Parametry wybranych ośrodków zestawiono w Tabeli 3.6.

Ośrodek	Moduł Sprężystości [kg/m/s ² × 10 ⁻⁸]	Gęstość [kg/m ³]	Prędkość [m/s]	Impedancia akustyczna- [kg/m ² /s × 10 ⁻⁸]
Powietrze	0,0000134	1,2	330	0,0004
Woda (20°C)	2,19	1 000	1 480	1,48
Rtęć	28,5	13 600	1 450	20,0
Tkanki miękkie	2,51	1 060	1 540	1,63
Tkanka tłuszczowa	2,0	952	1 450	1,38
Wątroba	2,54	1 060	1 550	1,64
Mięśnie wzdłużnie	2,74	1 080	1 592	1,70
Mięśnie poprzecz.	2,80	1 080	1 610	1,74
Mózg	2,40	994	1 550	1,55
Śledziona	2,59	1 045	1 578	1,64
Krew	2,47	1 057	1 057	1,62
Kość	32,0	1 912	4 080	7,8
Pluca	0,169	400	650	0,26
Soczewka oka	3,03	1 142	1 620	1,85
Ciało szkliste	2,31	1 000	1 520	1,52
Ciało wodnistre	2,25	1 000	1 500	1,50
Kwarc	88,0	2 650	5 750	15,2
Tytanian baru	107,0	5 400	4 460	24,0
Olej mineralny	2,1	969	1 477	1,43

Rysunek 3.6: Parametry ośrodków często mierzonych w badaniach USG.

Dla przykładu, do zjawiska załamania lub odbicia dochodzi kiedy fala pada na granice dwóch ośrodków o różnych impedancjach akustycznych Z_1 i Z_2 . Dla fali prostopadłej zależność ta opisana jest następująco:

$$R = \frac{I_r}{I_0} = \left(\frac{Z_1 - Z_2}{Z_1 + Z_2} \right)^2, \quad (3.9)$$

gdzie I_r to natężenie fali padającej, a I_0 odbitej. Natomiast R , czyli współczynnik odbicia, jest parametrem, który rośnie wraz ze wzrostem kąta odchylenia od kierunku prostopadłego, aż do całkowitego odbicia.

Z kolei do rozpraszań bądź pochłaniania fali dochodzi kiedy to fala pokonuje daną drogę w ośrodku o pewnej Z , co zapisywane jest następująco:

$$I = I_0 e^{-\gamma x}, \quad (3.10)$$

gdzie γ to współczynnik osłabienia zależny od Z , a x to droga przebyta przez falę. Efekt ten można korygować poprzez dobór odpowiedniego I_0 .

Analiza amplitudy i częstotliwości sygnału nadanego i echa umożliwia rekonstrukcję obrazu USG. W przypadku najczęściej stosowanych w praktyce rekonstrukcji przekrojów dwuwymiarowych (tzw. *tryb B*) współczesny tor budowania prezentacji wizualnej (tzw. *beamforming*) wygląda następująco:

1. Głowica ultradźwiękowa emisuje impulsy w postaci wąskiej wiązki w ścisłe określonym kierunku. Wiązka jest efektem interferencji sygnału z N przetworników zawierających kryształy piezoelektryczne.
2. Echa z danego kierunku pozwalały na obliczenie pojedynczego promienia akustycznego, który jest iloczynem charakterystyk nadawanego i odbieranego sygnału.
3. Wszystkie promienie, których we współczesnych aparatach może być do kilkuset (zob. [7]), służą do formowania obrazu, który tworzony jest we *współrzędnych biegunowych* (r, θ) w przypadku głowic mechanicznych sektorowych, wieloelementowych convex czy fazowych lub we współrzędnych prostokątnych (x, y) w przypadku głowic mechanicznych lub wieloelementowych liniowych¹.

Tryb B umożliwia również wizualizację obrazów dynamicznych. Przykładowo jeżeli na obraz składa się 400 promieni i każdy odsłuchiwanym jest do głębokości 15 cm, to czas gromadzenia danych dla ośrodka o średnim $c=1500$ m/s wynosi $2\frac{2\times 15}{1500\text{m/s}} \times 400 = 0,08$ s, czyli 12 obrazów na sekundę. Częstotliwość tę można zwiększać, zmniejszając liczbę promieni lub głębokość obserwacji.

Innym często stosowanym trybem rekonstrukcji obrazu jest *tryb D* bazujący na *efekcie Dopplera*, do którego dochodzi w przypadku przechodzenia fali przez ośrodek przesuwający się względem głowicy. Zmienia się wówczas częstotliwość fali, co wyrażone jest następującym wzorem:

$$f_r = 2f_o \frac{v}{c} \cos(\theta), \quad (3.11)$$

gdzie f_r to zmiana częstotliwości fali nadawanej f_0 , zależna od kąta θ pomiędzy falą i ośrodkiem poruszającym się z prędkościami rozchodzenia się fali w obu ośrodkach tj. v i c . Dla przykładu, im większa prędkość komórek przesuwających się w monitorowanym

¹szczegółowy opis głowic USG i ich charakterystyk można znaleźć w [15]

3.3 Zastosowanie ultrasoRozdział 3: Monitorowanie procesu gojenia ścięgna Achillesa

ciele pacjenta, tym większa jest f_r . Dlatego tryb D z powodzeniem jest wykorzystywany np. do monitorowania przepływu krwi w dużych naczyniach takich jak tętnice.

Rozwinięciem trybu D jest tryb *Power D* (od ang. *Power Doppler*). Gdzie zamiast przesunięcia częstotliwości interpretowana jest moc sygnału Dopplerowskiego. Za pomocą odpowiedniego parametru nazywanego *gain* można wzmacniać sygnał Power D uwidaczniając np. informacje pochodzące od przepływów krwi z niewielką prędkością w charakteryzujących się małą średnicą naczyniach krwionośnych. Jest to również możliwe, gdyż tryb Power D ma nawet trzykrotnie większą czułość niż standardowy tryb D [21]. Wynika to z faktu, że Power D nie uwzględnia informacji o prędkości czy kierunku przepływów, a jest sumą wszystkich możliwych przesunięć częstotliwości w zadanym fragmencie obrazowanego obiektu.

W kontekście ścięgna Achillesa tryb Power D może służyć do oceny unaczynienia ścięgna w kolejnych etapach gojenia, które jak wiadomo z sekcji 3.1.4 zmienia się w czasie. Tryb B natomiast może być użyteczny do obrazowania struktury tkanek miękkich. W praktyce wykorzystywana jest zwłaszcza możliwość zobrazowania ukierunkowania struktur włókien ścięgnistych na podstawie czego radiolog może wnioskować o fazie gojenia. Składowa czasowa jest interesująca z perspektywy m.in. fizjoterapeuty oceniającego ślizg w ścięgnie przy wykonywaniu odpowiednich ruchów np. zginania podeszwowo-grzbietowego stopy.

W porównaniu do rezonansu magnetycznego, w kwestii ograniczeń, należy zwrócić uwagę na fakt, że fale akustyczne używane w USG nie propagują się dobrze przez kości i gazy. Dlatego RM jest częściej rekomendowany do oceny struktur umiejscowionych w otoczeniu lub składających się w większości z takich ośrodków np. płuca.

Z użyciem RM możliwe jest uzyskanie obrazów o lepszej jakości detali. Sam czas badania jest natomiast dłuższy i w większości przypadków niemożliwe jest obrazowanie w czasie rzeczywistym, co z kolei jest naturalne dla techniki USG.

W wymiarze finansowym istotny jest fakt, że aparat do USG kosztuje nawet 10 razy mniej niż aparatura do RM. Jak wspomniano jednak, z uwagi na jakość detali, uzykiwane obrazy są trudniejsze do interpretacji, co przekłada się na koszty wyszkolenia kadry.

W tym ostatnim kontekście przydatne mogą okazać się nowe rozwiązania w warstwie sprzętowej i oprogramowania. Do pierwszej grupy należy zaliczyć zastąpienie przetworników z piezoelektrykami, przetwornikami budowanymi w technologii MEMS

np. cMUT, czy pMUT (zob. [10]) oraz układy pozwalające przetwarzać surowy sygnał ultradźwiękowy (zob. [16]). Dzięki przetwornikom MEMS można wytworzyć cały układ generujący drgania w krzemie w jednym procesie technologicznym razem z dedykowanym układem do zadanej aplikacji (ang. *Application-Specific Integrated Circuit* w skr. *ASIC*). Takie podejście znaczaco redukuje koszty oraz implikuje możliwość miniaturyzacji urządzeń.

Do drugiej grupy należą algorytmy sztucznej inteligencji pozwalające wydobyć i zinterpretować interesującą informację z niskiej jakości obrazów. Przykładowo w [36] algorytmy sztucznej inteligencji zostały użyte do określenia orientacji włókien mięśniowych, a w [13] do segmentacji komór serca w czasie rzeczywistym.

Obiecujący jest też rozwój metod charakterystyki tkanki na podstawie ultrasonografii (ang. Ultrasonography Tissue Characterization, w skr. *UTC*). W 2003 r., w [130], Bakker et. al. przedstawili pracę pokazującą w jakim stopniu obraz USG jest mieszaną echa związanego z budową tkanki, a w jakim z *interferencji*, czyli nakładania się fal. W zależności od stabilności echa można zatem wnioskować czy obserwowana struktura składa się głównie z dużych, stałych struktur wywołujących stabilne echo, czy np. płynów lub małych włókienek powodujących zmienne interferencje. Jako referencji dla studiowanego echa używa się badań histologicznych (zob. [129]). Na tej podstawie sklasyfikowano 4 różne rodzaje echa i jak zasygnalizowano w pracach [131] i [58] informacja o proporcjach występowania tych rodzajów może być użyta do oceny struktury ścięgna Achillesa. Metoda ta nie została jednak jeszcze w pełni zwalidowana i możliwości wnioskowania na jej podstawie są wciąż niejasne (zob. [58]).

USG i inne techniki obrazowania medyczne nie są jedynymi metodami oceny gojenia się ścięgna Achillesa. W kolejnej sekcji zostały opisane techniki oceny funkcji ścięgna, które samodzielnie jak i w połączeniu z analizą obrazową stanowią wartościową информацию diagnostyczną.

3.4 Zastosowanie badań biomechanicznych

W poprzednich sekcjach zostały opisane dwie najczęściej wykorzystywane obrazowe metody monitorowania procesu gojenia się ścięgna Achillesa. Komplementarnie, podczas rehabilitacji może być również wykonana *ocena funkcjonalna*, technika weryfikująca w jakim stopniu dany element (tkanka, narząd, organizm) może realizować swoje zadania. W przypadku monitorowania gojenia się ścięgna Achillesa najczęściej

3.4 Zastosowanie badań biomechaniki Monitorowanie procesu gojenia ścięgna Achillesa

w tym celu stosuje się *ocenę biomechaniki*, czyli badania pozwalające wnioskować na temat właściwości mechanicznych elementów składowych organizmów żywych.

Najbardziej zaawansowane i dokładne metody oceny biomechaniki stosowane współcześnie możliwe są do wykonania przy użyciu urządzeń pomiarowych takich jak:

- *Komputerowa analiza ruchu* (ang. *Motion Capture*) – narzędzie wykorzystujące systemy czujników do zapisu informacji o zmianach położenia obiektu rejestrowanego np. pacjenta. Do wiodących rozwiązań należy zaliczyć systemy firmy Vicon [17], czy BTS [3].
- *Płyty dynamometryczne* (ang. *Force Plates*) – narzędzie wykorzystywane do pomiaru sił reakcji podłożu w trzech prostopadłych płaszczyznach. Dzięki temu można określić sumaryczny udział mięśni w generowaniu sił odpowiadających za balans ciała, ruch w danym kierunku oraz przecistawianie się sile grawitacji. Do wiodących rozwiązań należą płyty firmy Kistler [8].
- *Elektromiografia*, w skr. EMG (ang. *Electromyography*) – narzędzie do pomiaru pobudzeń poszczególnych grup mięśniowych podczas ruchu. Wykorzystywane jest m.in. do określenia rozkładu sił zmierzonych przez płyty dynamometryczne na poszczególne mięśnie.

Synchronizacja danych z powyższych urządzeń umożliwia konstrukcję modeli układu mięśniowo-szkieletowego i symulacje funkcji poszczególnych grup mięśniowych oraz ścięgien przy zadanych problemach.

Danymi stosowanymi do uszczegółowienia takich modeli są np.: wymiary poszczególnych segmentów ciała (goleń, udo, tors) tzw. pomiary antropometryczne; maksymalne siły izometryczne mierzone z użyciem systemów takich jak Bidex; geometria kości mierzonych np. z pomocą tomografii komputerowej (w skr. TK) ew. RM; lokalizacja przyczepów mięśniowych określanych przy pomocy RM lub USG; środki masy poszczególnych segmentów określanych np. przy użyciu badania *DXA* (od ang. *Double X Ray Absorption*); skład włókien mięśniowych widocznych w USG.

Z uwagi na dużą liczbę możliwych do zmierzenia parametrów, ich integracja odbywa się w modelach komputerowych zaimplementowanych w różnego rodzaju oprogramowaniu do symulacji biomechanicznych. Do najczęściej używanych modeli należą opisane w [63] Gait2392 i Gait 2354 oraz obecnie najbardziej złożony – *AnyBody Full Body Model* [22].

Historia komputerowo wspomaganego, kompleksowego modelowania biomechaniki ruchu sięga wczesnych lat 90-tych ubiegłego wieku, kiedy to Delp i Loan przedstawili oprogramowanie SIMM [39]. Obecnie SIMM jak również inne oprogramowania komercyjne takie jak: Visual 3D (Cmotion Inc.) [18], Anybody (Anybody Technology) [2], czy Adams (MSC Software Corp.) [1], dostarczają narzędzi do wartościowych symulacji np. chodu [117], biegu [54] jak również konsekwencji różnych zabiegów chirurgicznych [51] i chorób [111]. Istnieją również narzędzia otwarte, do których należą szeroko wykorzystywany OpenSim [38] rozwijany na Uniwersytecie w Stanford, czy też Human Motion [124] wywodzący się z instytutu badawczego RIKEN w Japonii.

Powyżej przedstawione kompleksowe badania w praktyce realizowane są rzadko z uwagi na wysokie koszty. Dla przykładu w Polsce, ośrodki wyposażone w sprzęt pomiarowy takiej klasy to np. Instytut "Pomnik – Centrum Zdrowia Dziecka", Warszawski Uniwersytet Medyczny, Akademia Wychowania Fizycznego imienia Józefa Piłsudskiego, czy komercyjna placówka Fizjofit w Gliwicach. Żeby obniżyć koszty badania stosuje się wybiórcze podejście i selekcję parametrów pomiarowych uznanych przez ekspertów dziedzinowych za wystarczające do analizy zadanego problemu. Dla przykładu badania stosowane do oceny biomechaniki ściegna Achillesa w placówce Carolina Medical Center (gdzie realizowane były badania wykorzystywane w tej pracy) składają się z następujących pomiarów (zob. [102]):

1. Pomiar *ATRS* (od ang. *Achilles Tendon Total Rupture Score*) – oceniany jest w skali od 0 do 100 [93] poziom ograniczenia, z którymi pacjenci borykają się w następstwie urazu.
2. Pomiar stabilograficzny na platformie dynamometrycznej – mierzony są wychylenia środka ciężkości. Pacjent ma za zadanie utrzymanie równowagi na niestabilnym podłożu. Badanie realizowane jest boso z oczami otwartymi. Wykonane są dwie próby po 30 sekund kolejno na prawej i lewej kończynie dolnej na dynamicznej platformie dynamometrycznej Biodek Balance System. Wyniki zostają porównane między kończynami.
3. Pomiar stabilograficzny na platformie statycznej – mierzona jest droga wychylenia środka masy pacjenta w trakcie stania jednonóż na platformie dynamometrycznej, statycznej.
4. Pomiar sił reakcji na ścieżce podometrycznej – mierzony jest rozkład sił nacisków podeszwowej strony stóp na podłożę (jedynie w kierunku prostopadłym do pod-

3.4 Zastosowanie badań Biomechanicznych Monitorowanie procesu gojenia ścięgna Achillesa

łoża). Pomiar wykonywany jest podczas stania swobodnego, wspień na palce oraz przysiadu boso bez odrywania pięt. Dokonywana jest również analiza chodu (3 przejścia) i biegu (5 przebiegnięć). Boso oraz w obuwiu sportowym. Na podstawie sił reakcji wyliczane są parametry: rotacja podudzia [deg], długość kroków [cm], udział fazy podparcia [%], udział fazy przenoszenia [%], maksymalna siła na pięcie [N] oraz maksymalna siła na palcach [N].

5. Pomiar skoczności i mocy (tzw. siły dynamicznej) kończyn dolnych – mierzona jest moc maksymalna P_{max} i średnia P_m , maksymalna wysokość uniesienia h_{max} i obniżenia k środka masy ciała przed odbiciem. Wykonywane są wyskoki pionowe z miejsca na platformie dynamometrycznej. Realizowane są dwie próby obuńóż oraz na prawej i lewej kończynie dolnej w obuwiu sportowym. W celu pełnego zaangażowania kończyn dolnych pacjent podczas badania trzyma ręce na biodrach.
6. Pomiary momentów sił mięśni stawu skokowego – mierzone są maksymalne wartości momentu siły mięśni zginaczy podeszwowych i grzbietowych stawu skokowego [Nm] oraz deficit pomiędzy operowaną i zdrową kończyną dolną [%]. Momenty sił mierzone są w dwóch pozycjach tj. z wyprostowanym oraz zgiętym do 50 stopni stawem kolanowym [95]. Pomiar realizowany jest w warunkach izometrii i izokinetyki w trzech prędkościach kątowych 60°/s (5 powtórzeń), 120°/s (8 powtórzeń) oraz 180°/s (10 powtórzeń) przy wykorzystaniu urządzenia Humac Norm (USA). Przed badaniem osoba odbywa 5 minutową rozgrzewkę na steperze.

Wymienione wyżej badania zostały określone przez ortopedów i fizjoterapeutów jako wystarczające do oceny przywracania funkcji ścięgna Achillesa po rekonstrukcji.

Wraz z oceną strukturalną realizowaną poprzez badania obrazowe i wiedzą ekspercką informacja tak zgromadzona może służyć do subiektywnego monitorowania procesu gojenia się ścięgna. Do skutecznej obiektywizacji tego procesu potrzebne są jednak dodatkowe metody bazujące na agregacji ilościowych współczynników i automatycznym wnioskowaniu na ich podstawie. Do tej grupy należą algorytmy sztucznej inteligencji opisane w kolejnym rozdziale.

Rozdział 4

Konwolucyjne sieci neuronowe

Konwolucyjne sieci neuronowe (ang. *Convolutional Neural Networks*, CNN) są biologicznie inspirowanymi sztucznyimi sieciami neuronowymi. Należą do zbioru *głębokich sieci neuronowych* (ang. *Deep Neural Networks*, DNN), które z kolei są podzbiorem *systemów uczących się* (ang. *Machine Learning Systems*) tj. algorytmów nie wymagających explicite programowania do ustalenia swoich parametrów.

Pierwsze matematyczne formalizmy dotyczące CNN zostały zaproponowane już w latach 40-tych XX wieku, natomiast fundamentalne inspiracje biologiczne dały badania nad korą wzrokową kotów Hubel'a i Wiesel'a z lat 50-tych i 60-tych (zob. [136]). M.in. dzięki pracom tych neurofizjologów ustalono, że kora wzrokowa zawiera złożone układy komórek, które odpowiadają za przetwarzanie informacji z wybranych regionów pola widzenia, sumarycznie pokrywając je w całości. Komórki kory wzrokowej działają zatem jak lokalne filtry przestrzeni wejściowej zaprojektowane, tak aby wydobyć istotne informacje z naturalnych obrazów. Dla przykładu reagują na orientację linii, kształty i kolory.

W odróżnieniu od naturalnych struktur biologicznych, sieci konwolucyjne operują przeważnie na reprezentacji cyfrowej obrazu. Zapisywana jest ona w postaci tablicy liczb tzw. *macierzy* o współrzędnych (x,y) , gdzie x to kolumna macierzy, a y wiersz. W przypadku obrazów trójwymiarowych dochodzi jeszcze składowa z , a zamiast macierzy użyty jest wówczas *tensor*, czyli uogólniona reprezentacja.

W obrazach cyfrowych dla każdego punktu (x,y) , tzw. *piksela*, kodowana jest informacja o wartości funkcji obrazowej $I(x,y)$. Na tej podstawie dzielimy obrazy na:

- binarne – gdzie kodowane są jedynie dwie możliwe wartości I (0 lub 1);

- monochromatyczne – gdzie kodowana jest informacja o natężeniu jednej barwy (najczęściej są to odcienie szarości lub brązu tzw. *sepia*);
- kolorowe – gdzie kodowane są wartości natężenia składowych koloru.

Do zakodowania informacji o wartości funkcji I w obrazie binarnym potrzebny jest jeden bit na punkt. Odcienie obrazu monochromatycznego lub pojedyncze składowe punktu obrazu kolorowego kodowane są przy pomocy 8–16 bitów, co uzależnione jest od zakresu wartości występujących w danych lub precyzji wymaganej do obliczeń. Przykładowo, obrazy medyczne przetwarzane w tej pracy należą do grupy obrazów monochromatycznych i są zapisywane w reprezentacji 16 bitowej z uwagi na zakres wartości występujących w sygnale RM.

W sieciach konwolucyjnych procesy zachodzące w korze wzrokowej są modelowane przy użyciu odpowiednio dobranych parametrów i funkcji służących do ekstrakcji istotnych informacji obrazowych i ich przetwarzania. Parametry wykorzystywane do wstępnej ekstrakcji informacji grupowane są w *filtre*, które realizują funkcję *splotu* maski K z kolejnymi fragmentami funkcji I . K jest najczęściej macierzą kwadratową o wymiarze N , gdzie zazwyczaj $N = 1, 3, 5, 7, 9, 11$ z uwagi na rosnącą wraz z N złożonością obliczeniową. Równanie splotu dyskretnego można zapisać następująco:

$$I'(x, y) = \sum_{j=-n}^n \sum_{i=-k}^k I(x - j, y - i) K(j, i), \quad (4.1)$$

gdzie I' jest nową funkcją obrazową powstałą po filtracji, a i i j to kolejne współrzędne maski filtru w odniesieniu do jego punktu centralnego. Dla wartości brzegowych podczas realizacji splotu, współrzędne poza sygnałem przyjmują z reguły wartość 0. Rzadziej stosuje się inne metody takie jak: odbicie wartości I obrazu poza jego granicami; powtórzenie wartości I obrazu bez odbicia; modyfikację wymiaru maski filtru na brzegu obrazu, tak by nie wychodziła poza jego granicę.

W kontekście sieci konwolucyjnych, pojedynczy splot filtru z obrazem wejściowym nazywany jest *cechą* (zob. [105]). W ostatnich latach powstało wiele algorytmów do wizualizacji cech umożliwiających wydajną analizę wyników działania sieci konwolucyjnych np.: Saliency Maps [113], GradCam [109] czy CNNVIS [80]. Zasada działania tych algorytmów bazuje na selekcji cech, które mają zasadniczy wpływ na wynik końcowy. Grupy cech obliczane na tym samym obszarze obrazu wejściowego nazywane są *mapami cech* (ang. *feature maps*). Zazwyczaj mapy cech zawierają hierarchicznie upo-

rządkowaną strukturę składającą się z cech od najniższego do najwyższego poziomu komplikacji (np. od charakterystycznych skupisk pikseli, przez obiekty składowe do finalnego kształtu). Do cech najniższego rzędu zaliczyć można:

- *krawędzie* – ciągi punktów o gwałtownych zmianach $I(x,y)$;
- *roggi* – przecięcia dwóch krawędzi;
- *grzbiety* lub *doliny* – odpowiednio lokalne maksimum lub minimum funkcji I ;
- *skupiska* (ang. *blobs*) – obszary jednorodnych wartości funkcji I , różniących się znacząco od najbliższego otoczenia;
- *tekstury* – charakterystyczne, przestrzenne ułożenie wartości funkcji I w powtarzające się wzory.

Wszystkie cechy wyższego rzędu są kombinacją powyższych składowych podstawowych tworząc bardziej skomplikowane struktury odpowiadające charakterystycznym obiektom znajdującym się w obrazach wejściowych.

Na podstawie cech, na końcu toru przetwarzania danych z wykorzystaniem sieci neuronowych, realizowane jest wnioskowanie, gdzie w zależności od problemu można wyszczególnić następujące zadania:

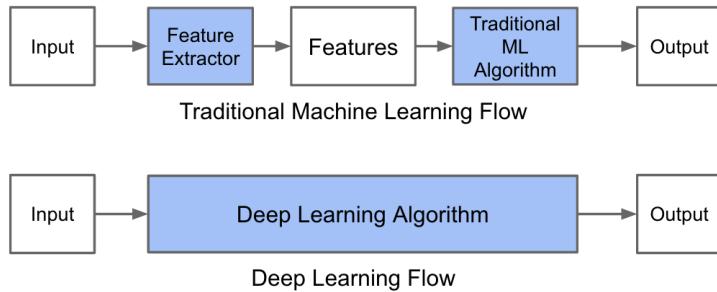
- *segmentację* – podział obrazu na spójne fragmenty, najczęściej wiążące się z wyodrębnieniem *obiektu* z *tła* na podstawie ustalonego progu lub miary np.:

$$\begin{aligned} I(x, y) \geq T_p &\Rightarrow (x, y) \in "obiekt" \\ I(x, y) < T_p &\Rightarrow (x, y) \in "tlo", \end{aligned}$$

przy czym metody ustalania wartości progowej T_p lub wielu wartości T_p^1, \dots, T_p^n są obszarem szerokich badań (zob. [19]);

- *klasyfikację* – przyporządkowanie obiektu do odpowiedniej klasy (np. tkanka zdrowa lub patologiczna);
- *detekcję* – binarne rozróżnienie traktujące o tym czy obiekt znajduje się w obrazie czy nie;
- *śledzenie* – detekcja lub też klasyfikacja obiektów w kolejnych krokach czasowych.

Jeszcze do niedawna tor przetwarzania danych w większości badań wykorzystujących uczenie się maszyn, operacje wyliczenia cech i wnioskowanie końcowego zawierał w dwóch osobnych krokach¹. Porównanie tego schematu z obecnie funkcjonującym podejściem realizowanym w algorytmach głębokiego uczenia się przedstawiono na Rys. 4.1. W nowym podejściu zarówno ekstrakcja cech jak i ostateczne wnioskowanie na ich pod-



Rysunek 4.1: Porównanie schematów przetwarzania danych z wykorzystaniem metod głębokiego uczenia się i innych algorytmów.

stawie realizowane jest w jednym kroku, co nazywane jest paradygmatem *end-to-end learning*. Na przestrzeni lat wprowadzano kolejne składowe, które finalnie utworzyły obecnie funkcjonujące podejście. W kontekście poznania tych komponentów, w kolejnej sekcji omówiono dokładniej zarys historyczny przedstawiający ewolucję nowego paradygmatu.

4.1 Zarys historyczny

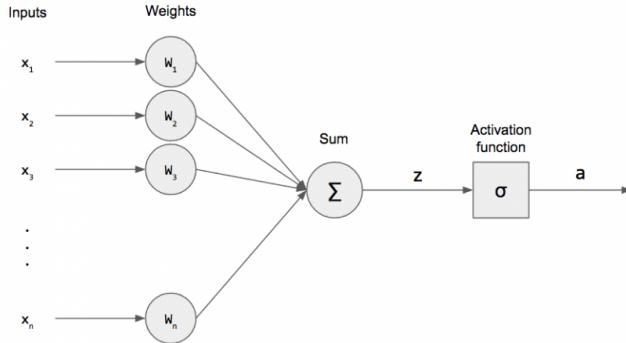
Pierwszy formalny model neuronu został zaproponowany przez Warrena McCulloch i Waltera Pitts w roku 1943 (zob. [89]). Była to bramka logiczna, której wyjście stawało się aktywne w momencie, gdy liczba aktywnych wejść przekroczyła pewien zdefiniowany próg. Taka zależność sygnału wyjściowego y od sygnałów wejściowych $x_1 \dots x_n$ została potem nazwana *funkcją aktywacji neuronu*, którą zapisujemy jako:

$$y = f(x_1, x_2, \dots, x_n) \quad (4.2)$$

W modelu neuronu McCulloch-Pitts można było modyfikować parametr progu, nie istniała natomiast możliwość uczenia się takiej architektury. Problem ten rozwiązano

¹Są wyjątki od tej reguły jak np. [110], inkrementalny algorytm SVM.

w 1957 proponując sztuczną sieć neuronową zawierającą wiele neuronów z ważonymi połączeniami między sobą (zob. [101]). Sieć nazwano perceptronem, co było implikacją zamiłowania jego twórcy Franka Rosenblatta do aplikacji związanych z percepcją, zwłaszcza mowy czy pisma. Schemat sieci pokazano na Rys. 4.2.



Rysunek 4.2: Topologia perceptronu.

Zastosowanie dodatkowej jednowymiarowej tablicy współczynników, czyli *wektora wag* $w = (w_1, \dots, w_n)$ dało możliwość uczenia się poprzez adaptacyjną zmianę wartości poszczególnych jego elementów. W modelu Rosenblatta zastosowano ponadto progową funkcję aktywacji z progiem T_a :

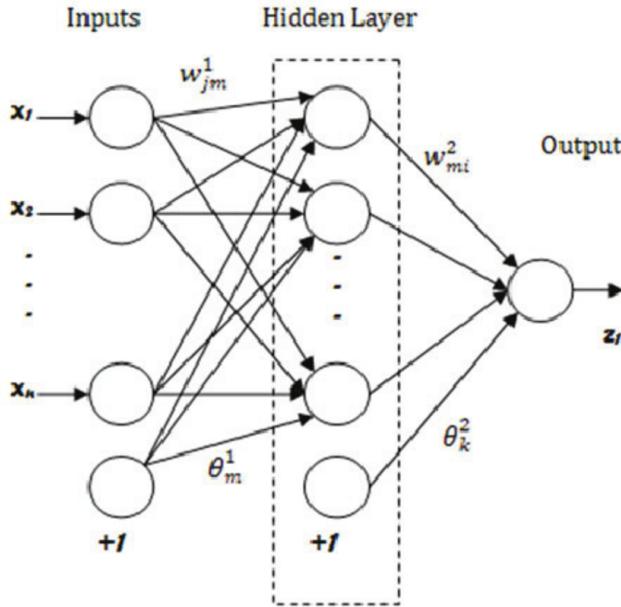
$$y(z) = \begin{cases} 0, & n < T_a, \\ 1, & n \geq T_a, \end{cases} \quad (4.3)$$

gdzie z to suma ważona wyjść poszczególnych neuronów:

$$z = \sum_{i=1}^n w_i x_i \quad (4.4)$$

Dla przykładu, w dwuwymiarowej przestrzeni sygnałów wejściowych x_1, x_2 działanie perceptronu można opisać jako wyliczenie funkcji liniowej rozdzielającej obserwacje o_1 od o_2 . W trzech wymiarach będzie to płaszczyzna, a w n -wymiarach, hiperplaszczyzna. Z uwagi na swój liniowy charakter perceptron proponowany przez Rosenblatta miał wiele ograniczeń, które trafnie sformułowali w 1969 roku Marvin Minsky i Seymour Papert w książce *Perceptrons* (zob. [88]). Autorzy opublikowali listę problemów, których nie można było rozwiązać z użyciem perceptronu m.in. do najszerzej dyskutowanych należał przykład związany z brakiem możliwości modelowania funkcji *XOR*, aktywującej wyjście przy aktywnym jednym z dwóch wejść.

Po latach intensywnych prac, część z opisanych przez Minsky-Papert problemów udało się rozwiązać za sprawą pracy Davida Rumelharta, Geoffa Hintona i Ronalda Williams, którzy opublikowali w 1986 roku pracę [103], traktującą o perceptronach wielowarstwowych. Schemat takiej sieci zaprezentowano na Rys. 4.3.



Rysunek 4.3: Topologia perceptronu wielowarstwowego.

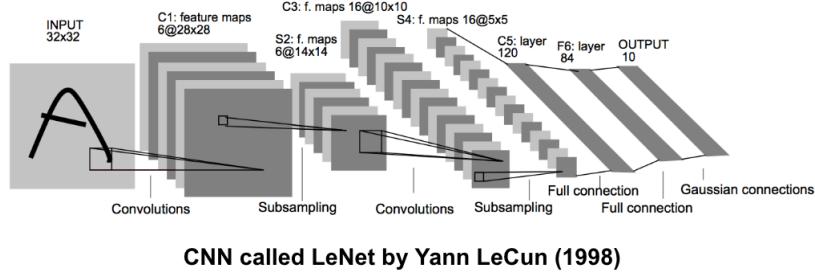
Spośród najważniejszych innowacji wprowadzonych w perceptronie wielowarstwowym wyszczególnić można zastosowanie w praktyce nowego algorytmu uczenia się sieci, który został opisany w kolejnej sekcji jak również nowej, sigmoidalnej funkcji aktywacji:

$$y(x) = \frac{1}{1 + e^{-x}} \quad (4.5)$$

Z wykorzystaniem sieci wielowarstwowych możliwe stało się modelowanie funkcji XOR jak i innych problemów nieliniowych o praktycznym wymiarze.

Kolejny przełom nastąpił w 1989 roku kiedy to Yann LeCunn, uczeń Geoffa Hintona, zaprezentował swoje wyniki dotyczące klasyfikacji odręcznego pisma z użyciem sieci wielowarstwowych (zob. [75]). Finalnie, badania te doprowadziły do przedstawienia w 1998 roku pierwszej sieci konwolucyjnej nazwanej LeNet (zob. [74]). Architekturę tej sieci przedstawiono na Rys. 4.4.

Sieć składała się z 7 warstw i zawierała około 60,000 parametrów. Oryginalnie, sygnał wejściowy sieci stanowił obrazek o wymiarach 32×32 . W architekturze LeNet



Rysunek 4.4: Topologia sieci LeNet.

zaobserwować można było dwie podstawowe składowe współczesnych sieci konwolucyjnych wykorzystywanych do klasyfikacji obrazów tj.:

- *ekstraktor cech* – część zawierająca m.in. filtry, służące do automatycznej ekstrakcji wektora cech w .
- *klasyfikator* – część wykorzystywana do zadania wnioskowania końcowego na podstawie w .

Przy użyciu takiej architektury możliwe stało się zastosowanie paradygmatu end-to-end learning rozumianego w tym kontekście jako znalezienie możliwie dobrej transformacji, która surowe obrazy przekształca bezpośrednio w ostateczną klasyfikację. Dokładny opis szkolenia się sieci konwolucyjnych i problemów z tym związanych został przedstawiony w kolejnej sekcji.

4.2 Szkolenie głębokich sieci neuronowych

Większość algorytmów szkolenia głębokich sieci neuronowych obejmuje zadanie *optymalizacji*, rozumiane jako minimalizację, bądź maksymalizację *funkcji celu* $f(x)$ przez zmianę x . W literaturze można też znaleźć inne nazwy funkcji celu takie jak *kryterium*, *funkcja kosztów*, *funkcja strat* lub *funkcja błędów* (por. [52]).

Podczas zadania optymalizacji bardzo często wykorzystuje się *pochodną funkcji* oznaczaną jako $f'(x)$ lub $\frac{\delta y}{\delta x}$, gdyż niesie ona informacje o nachyleniu funkcji w punkcie x . W praktyce funkcje celu są wielowymiarowe dlatego wykorzystywane są *pochodne cząstkowe* informujące o nachyleniu w poszczególnych wymiarach. Wektor zawierający pochodne cząstkowe funkcji nazywany jest *gradientem* i oznaczany jest jako $\nabla f(x)$.

W praktyce wykorzystywane są również: macierz pochodnych cząstkowych tzw. *macierz Jacobiego* oznaczana jako J oraz macierz *drugich pochodnych* cząstkowych (tj. pochodnych pochodnych) tzw. *macierz Hesego* oznaczana jako H .

Z uwagi na szybkie znajdowanie lokalnych minimów funkcji celu, to właśnie metody optymalizacji bazujące na wartości gradientu są najczęściej używane w szkoleniu głębokich sieci neuronowych. Inne metody, niegradientowe (zob. [126]), przeważnie nie są w tym kontekście wystarczająco efektywne.

Znalezienie lokalnego minimum zwykle nie jest równoważne z otrzymaniem najlepszego możliwego rozwiązania, jednak w praktyce uznawane jest za satysfakcyjujące. Dla potwierdzenia można przytoczyć następujące fakty wynikające z wielu badań (por. [34]):

1. Dla sieci neuronowych o dużych rozmiarach większość lokalnych minimów charakteryzuje się podobnymi wartościami, przekładającymi się na porównywalny efekt wnioskowania końcowego.
2. Prawdopodobieństwo znalezienia lokalnego minimum, którego implikacją będą niezadowalające rezultaty wnioskowania przy użyciu sieci maleje wraz ze wzrostem rozmiaru sieci.
3. Próba znalezienia globalnego minimum bardzo często prowadzi do problemu nadmiernego dopasowania, omówionego dokładniej w sekcji 4.2.1.

Podsumowując, metody gradientowe są wydajne obliczeniowo i prowadzą do znalezienia wielu satysfakcyjujących rozwiązań, które mogą być wykorzystane do rozwiązania praktycznych problemów.

Zasada działania metod gradientowych bazuje na obieraniu w kolejnych krokach iteracji następujących wartości funkcji f , przesuwającą się w kierunku spadku gradientu:

$$x' = x - \epsilon \bigtriangledown f(x), \quad (4.6)$$

gdzie ϵ to *szynkość uczenia się*, parametr określający wielkość kroku.

Funkcja celu w przypadku praktycznych zadań optymalizacyjnych, wykorzystujących głębokie uczenie się jest *funkcją złożoną*, a zatem efekt jej działania jest równoważny operacjom wykonywanym przez kilka lub więcej funkcji po kolei. W takich

sytuacjach do obliczeń spadku gradientu wykorzystywane są funkcje, których pochodne są znane.

Aplikuje się je do tzw. *reguły łańcuchowej*. Przypuśćmy, że $v = g(p)$ i $u = f(g(v)) = f(q)$, gdzie p i q to wektory. Wówczas regułę łańcuchową można zapisać jako:

$$\frac{\delta u}{\delta p_i} = \sum_j \frac{\delta u}{\delta q_j} \frac{\delta q_j}{\delta p_i}, \quad (4.7)$$

co w zapisie wektorowym równoważne jest z równaniem:

$$\nabla_x z = \left(\frac{\delta q}{\delta p} \right)^T \nabla_q u, \quad (4.8)$$

gdzie $\frac{\delta q}{\delta p}$ to macierz Jacobiego.

Regułę łańcuchową zadaną w 4.8 prosto uogólnia się do zmiennych tensorowych (zob. [52], str. 205) i stosuje się w różnych meta-algorytmach służących do szkolenia sieci. Przykładem jest *algorytm propagacji wstecznej*, który oblicza regułę łańcuchową w wydajnej kolejności stosując działania w grafie takim jak topologia perceptronu wielowarstwowego (zob. [52]).

Proces szkolenia się sieci ma na celu najlepsze możliwe przybliżenie docelowej klasyfikacji bazując na danych przykładach, czyli *zbiorze uczącym* U . Algorytmy optymalizacyjne używane do szkolenia głębokich sieci neuronowych zazwyczaj działają pośrednio, optymalizując pewną miarę wydajności P , która jest zdefiniowana na *zbiorze testowym*, zawierającym przykłady inne niż w U . Często bierze się również pod uwagę jeszcze dodatkowy podzbiór, rozłączny z U i T – *zbior walidacyjny*, który ma pomóc w wyborze najlepszych algorytmów i wartości parametrów wpływających na jakość szkolenia sieci.

W procesie szkolenia zmniejszana jest wartość funkcji kosztów f w oparciu o U , a celem jest poprawa P . Algorytmy optymalizacyjne wykorzystujące cały zbiór U do liczenia wartości gradientu nazywane są *pakietowymi* lub *deterministycznymi*, gdyż przetwarzają jednocześnie wszystkie przykłady szkoleniowe. Te, które używają jednego przykładu na raz są nazywane *stochastycznymi*. Natomiast w praktyce przy szkoleniu głębokich sieci stosowane są algorytmy *minipakietowe*, wykorzystujące więcej niż jeden przykład, ale mniej niż cały zbiór. Z reguły są to liczby z przedziału 8–256. Takie podejście zapewnia kompromis między szybkością obliczeń i dokładnością estymacji wartości gradientu.

Przykładem algorytmu minipakietowego jest *stochastyczny spadek gradientu* (ang.

stochastic gradient descent, SGD). Bazuje on na założeniu, że estymację gradientu można otrzymać wyliczając średnią gradientu z minipakietu m przykładów. Kolejne kroki algorytmu można zapisać następująco:

1. Wybierz wartość parametru szybkości uczenia się ϵ_k .
2. Wybierz minipakiet złożony z m przykładów ze zbioru szkoleniowego.
3. Oblicz estymację gradientu $g = \frac{1}{m} \nabla \sum_{i=1}^m L(x_i, y_i, f)$, gdzie L to funkcja strat dla jednego przykładu o wejściowej wartości próbki x_i i oczekiwany wyjściu y_i .
4. Zastosuj aktualizację wartości funkcji celu równą $\epsilon_k g$.
5. Jeżeli kryterium stopu nie zostało spełnione wróć do kroku 2.

Kryterium stopu jest najczęściej określone liczbą iteracji lub satysfakcjonującą wartością funkcji f . Kwestia optymalnego wyboru parametru ϵ_k zależy od problemu i najczęściej stosowane są metody empiryczne, przy czym zazwyczaj ϵ_k maleje wraz ze zbliżaniem się do satysfakcjonującego rozwiązania.

Parametry algorytmów szkoleniowych nazywane są *hiperparametrami*, gdyż nie są wyznaczane bezpośrednio w procedurze uczenia się. Strategia nadawania hiperparametrom wartości początkowych jest silnie dyskutowana w literaturze (zob. [66]) i jej kompleksowy opis wykracza poza zakres tej pracy. Warto jednak wspomnieć o algorytmach z adaptacyjną szybkością uczenia się, gdyż jest to jeden z najtrudniejszych do ustawienia hiperparametrów, a jednocześnie bardzo istotny. Są to m.in.:

- Adaptive Gradient Algorithm (AdaGrad) [45] – wykorzystywany do indywidualnej adaptacji szybkości uczenia się wszystkich parametrów modelu, skalując ją odwrotnie proporcjonalnie do pierwiastka kwadratowego sumy wszystkich historycznych kwadratów gradientów.
- Root Mean Square Propagation (RMSProp) [108] – modyfikacja AdaGrad, w której zamiast akumulacji gradientu wykorzystuje się wykładniczo ważoną ruchomą średnią z gradientu.
- Adaptive moments algorithm (Adam) [65] – W porównaniu z RMSProp, Adam poza momentem pierwszego rzędu (tj. średnią) wykorzystuje również moment drugiego rzędu (tj. *wariancję*). Dokładniej rzecz biorąc, w algorytmie liczona jest

wykładnicza ruchoma średnia gradientu i kwadrat z gradientu oraz parametry β_1 i β_2 , które kontrolują zakres liczenia średnich.

W momencie pisania tej pracy algorytm Adam jest najczęściej rekomendowanym jako domyślna metoda szkolenia głębokich sieci neuronowych, dlatego poniżej zamieszczono jego dokładny opis:

1. Wybierz wartość początkową $\epsilon, \beta_1, \beta_2$ oraz ustaw początkowe wartości zmiennych momentu 1 i 2 stopnia $s = 0$ i $r = 0$, wartość kroku czasowego $t = 0$ i stałą σ używaną do stabilizacji numerycznej.
2. Wybierz minipakiet złożony z m przykładów ze zbioru szkoleniowego.
3. Oblicz estymację gradientu $g = \frac{1}{m} \nabla \sum_{i=1}^m L(x_i, y_i, f)$ i zwięksź t o 1
4. aktualizuj estymację pierwszego momentu. $s \mapsto \beta_1 s + (1-\beta_1)g$
5. aktualizuj estymację drugiego momentu. $r \mapsto \beta_2 r + (1-\beta_2)g \odot g$
6. skoryguj obciążenie momentu pierwszego rzędu $s = \frac{s}{1-\beta_1^t}$
7. skoryguj obciążenie momentu drugiego rzędu $r = \frac{r}{1-\beta_2^t}$
8. Zastosuj aktualizację wartości funkcji celu równą $-\epsilon \frac{s}{\sqrt{r+\sigma}}$.
9. Jeżeli kryterium stopu nie zostało spełnione wróć do kroku 2.

Ocena procesu szkolenia się sieci polega na obliczeniu odpowiednich miar i współczynników odzwierciedlających przybliżenie zbioru T albo W przez znalezione rozwiązańe. W powszechnym problemie klasyfikacji binarnej, występującym również w tej pracy, w kontekście oceny efektywności algorytmów można wyszczególnić następujące parametry:

- *Fałszywie pozytywna klasyfikacja* (*FP* od ang. *False Positive*) – liczba obserwacji zaklasyfikowanych jako pozytywne, a należących do klasy obserwacji negatywnych.
- *Fałszywie negatywna klasyfikacja* (*FN* od ang. *False Negative*) – liczba obserwacji zaklasyfikowanych jako fałszywie negatywne, a należących do klasy obserwacji pozytywnych.

- *Prawdziwie pozytywna klasyfikacja* (*TP* od ang. *True Positive*) – liczba wyników poprawnie zaklasyfikowanych jako pozytywne.
- *Prawdziwie negatywna klasyfikacja* (*TN* od ang. *True Negative*) – liczba wyników poprawnie zaklasyfikowanych jako negatywne,

oraz miary:

- *Dokładność klasyfikacji* (ang. *Accuracy*) – $ACC = \frac{TP+TN}{TP+TN+FP+FN}$.
- *Czułość klasyfikacji* (ang. *Sensitivity*) – $TPR = \frac{TP}{TP+FN}$.
- *Swoistość klasyfikacji* (ang. *Specificity*) – $TNR = \frac{TN}{TN+FP}$.
- *Precyzja klasyfikacji* (ang. *Precision*) – $PPV = \frac{TP}{TP+FP}$.
- *Nieprecyzyjność klasyfikacji* (ang. *Negative precision*) – $NPV = \frac{TN}{TN+FN}$.

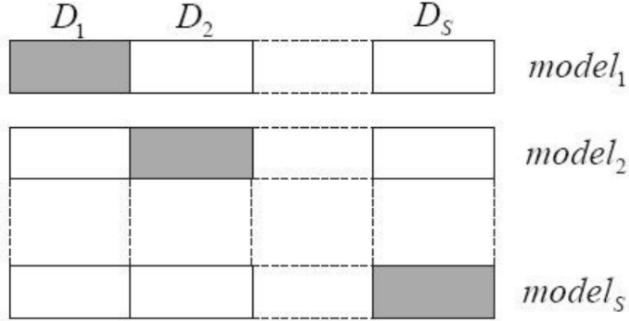
W problemach empirycznych dąży się do maksymalizowania TP, TN, ACC, TPR, TNR i PPV oraz minimalizowania FP, FN i NPV. Prawidłowe podejście polega również na wybraniu możliwie efektywnej architektury sieci. Problemy z tym związane zostaną omówione w kolejnych podsekcjach.

4.2.1 Problem nadmiernego dopasowania

Dążenie do najlepszego możliwego przybliżenia zbioru U wprowadza niepożądane zjawisko zwane *nadmiernym dopasowaniem*. Wtedy to dokładność klasyfikacji zbioru U jest wysoka lub nawet bezbłędna, natomiast znacznie niższa jest dokładność klasyfikacji zbioru testowego T i walidacyjnego W . W praktyce oznacza to, że model staje się mało użyteczny, gdyż wnioskowanie na nowych danych charakteryzuje się niską dokładnością.

Zatem ogólnym dążeniem w procesie uczenia się sieci jest osiągnięcie *maksymalnej generalizacji* klasyfikacji. Sieć o wysokim współczynniku generalizacji lepiej klasyfikuje ogólny zbiór wektorów wejściowych niż sieć, która ma niski współczynnik generalizacji i jest nadmiernie dopasowana do zbioru U . Właściwym działaniem jest więc ustalenie maksymalnie ogólnych, dostatecznych warunków poprawnej klasyfikacji, dzięki którym wzrasta prawdopodobieństwo, że przykład z poza zbioru U będzie poprawnie klasyfikowany.

W tym celu można zastosować metodę oceny krzyżowej (ang. *cross-validation*). Metoda ta polega na podziale zbioru uczącego na s segmentów D , z których każdy w innej iteracji służy jako zbiór testujący i walidacyjny, a pozostałe segmenty pełnią rolę zbioru uczącego. Podział zobrazowany jest na Rysunku 5.1. Stosując metodę oceny krzyżowej



Rysunek 4.5: Reprezentacja graficzna oceny krzyżowej.

dla różnych modeli sieci można stwierdzić, który z nich spełnia najlepiej kompromis między dobrą klasyfikacją zbioru U i wysoką generalizacją.

Kombinacja predykcji wielu różnych modeli jest bardzo wydajną metodą do poprawiania generalizacji i zmniejszenia błędu klasyfikacji na zbiorach innych niż U (zob. [23, 28]). Jednak współczesne sieci neuronowe, których przykłady zostały opisane w dalszych sekcjach mogą zawierać miliony parametrów i ich optymalizacja jest wymagająca obliczeniowo, dlatego w praktyce ogranicza się liczbę segmentów obierając $s \in \langle 5, 10 \rangle$.

Innym podejściem zaproponowanym w 2012 w [61] jest technika *dropout*, której główna idea bazuje na zerowaniu wyjścia neuronów sieci z prawdopodobieństwem 0,5 przy każdej iteracji treningu sieci. Neurony, które są w ten sposób tymczasowo dezaktywowane nie mają wpływu w danej iteracji na predykcję sieci i nie są uwzględniane przy wstecznej propagacji gradientu. Podejście to można porównać do treningu w każdej iteracji różnych modeli sieci. Dla przykładu w [70] wykazano, że metoda dropout wymaga jedynie 2 razy więcej iteracji do przybliżenia zbioru U , przy tym uzyskując znacznie lepszą generalizację.

Kluczowym składnikiem potrzebnym do treningu sieci i maksymalizacji generalizacji jest odpowiedni rozmiar zbioru danych. Jest to problem szeroko dyskutowany, gdyż zwłaszcza w danych medycznych istnieje szereg ograniczeń związanych z dostępem i akwizycją odpowiedniego materiału badawczego (np. ograniczenia prawne, związane z prywatnością lub z etyką). W przypadku, gdy zgromadzenie odpowiedniego zbioru

danych jest niemożliwe, pewnym rozwiązaniem problemu jest zastosowanie metod jego sztucznego powiększania (ang. *data augmentation*).

Dla obrazów stosuje się metody *afiniczych przekształceń* zgodne z definicją algebraiczną:

$$\mathbf{o} \mapsto \mathbf{A}\mathbf{o} + \mathbf{b}, \quad (4.9)$$

gdzie A jest macierzą przekształcenia liniowego, a b wektorem przesunięcia. Jako przykłady takich przekształceń dla dwuwymiarowych obrazów można wymienić:

- *rotację* – obrót obrazu o kąt θ , gdzie:

$$A = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (4.10)$$

- *odbicie lustrzane* – odwrócenie kolejności pikseli w każdym wierszu, gdzie:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad (4.11)$$

- *skalowanie* – zmiana rozmiaru obrazu o S , gdzie:

$$A = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \quad (4.12)$$

- *translację* – przesunięcie punktów obrazu o wektor b .

Analogiczne równania istnieją dla obrazów trójwymiarowych (zob. [60]). W określonych przypadkach używane są również nieafiniczne przekształcenia. Np. w [100] wykorzystano nieliniowe deformacje do powiększenia zbioru 30 obrazów mikroskopowych przedstawiających macierz komórkową i uzyskano znaczco lepsze wyniki detekcji komórek niż istniejące w 2015 najlepsze rozwiązania.

W przypadku danych medycznych należy szczególnie zwrócić uwagę, aby powiększony zbiór zawierał dane przypominające w rzeczywistości występujące przypadki np. nieduże obroty występujące u pacjentów skanowanych rezonansem magnetycznym lub niewielkie skalowania rozmiaru kości widocznych w Tomografii Komputerowej. Szeroką dyskusję prowadzi się również na temat wykorzystania sztucznie generowanych zbiorów danych o czym więcej można przeczytać w [79, 115].

4.2.2 Problem redukcji wymiarowości

Duży rozmiar wektora cech prowadzi do problemu nazwanego *przekleństwem wymiarowości* (ang. *curse of dimensionality*). Określenie zostało po raz pierwszy sformułowane przez Richarda Bellmana w latach 50-tych XX wieku. Naukowiec ten podczas swojej pracy obserwował algorytmy doskonale działające w 3 wymiarach, a prezentujące znaczaco gorsze wyniki w hiperprzestrzeni (zob. [24]).

Problem przekleństwa wymiarowości ma dwie główne przyczyny: (1) nie wszystkie cechy są jednakowo znaczące w kontekście rozróżnienia danych; (2) w miarę wzrostu rozmiaru przestrzeni cech, liczba obserwacji w zbiorze uczącym potrzebnych do wiarygodnego oszacowania funkcji wyjściowej rośnie wykładniczo.

Problem (1) jest szczególnie istotny w dość prostych algorytmach takich jak np. algorytm K -najbliższych sąsiadów, gdzie do poprawnego działania należy policzyć dystans pomiędzy sąsiednimi obserwacjami. Uwzględniając dużą liczbę nieistotnych cech jako argumenty funkcji dystansu uzyskuje się wyniki utrudniające lub nawet uniemożliwiające poprawną klasyfikację zbioru. Konieczna jest wówczas adaptacja wpływu poszczególnych cech np. poprzez wprowadzenie wektora wag.

Problem (2) ma duże implikacje w praktyce stosowania sieci neuronowych, gdyż posiadanie odpowiednio dużego, ustrukturyzowanego zbioru danych stanowi przeważnie wyzwanie. W poprzedniej podsekcji omówiono możliwość sztucznego powiększania zbioru danych. Inną opcją jest zmniejszenie rozmiaru wektora cech, co może być wykonane na dwa sposoby:

- wybór podzbioru istotnych cech o liczności $n' \ll n$,
- przekształcenie oryginalnych n zmiennych na nowy zbiór n' cech, gdzie ponownie $n' \ll n$.

W pierwszym przypadku, wybór podzbioru istotnych cech polega na określeniu minimalnego podzbioru, dla którego rozkład prawdopodobieństwa różnych klas obiektów jest jak najbliższy oryginalnemu rozkładowi uzyskanemu z wykorzystaniem wszystkich cech. Do tych zagadnień wykorzystywane są metody takie jak:

- *miary siły związku* – określające podobieństwo między rozkładami zmiennych losowych. Najczęściej stosowana jest *korelacja Pearsona*, której współczynnik r

dla dwóch zmiennych losowych X i Y zapisywany jest następującym wzorem:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}, \quad (4.13)$$

gdzie X_i, Y_i to wartości kolejnych obserwacji, a \bar{X} i \bar{Y} to ich średnie. Innymi słowy tak zapisany współczynnik r_{XY} jest ilorazem *kowariancji* i iloczynu *odchylen standardowych* zmiennych X i Y .

- *miary entropii względnej* – określające rozbieżności między rozkładami zmiennych losowych. Najczęściej stosowana jest *dywergencja Kullbacka-Leiblera*, której współczynnik d_{KL} dla dwóch rozkładów prawdopodobieństwa p i q zapisany jest wzorem:

$$d_{KL}(p, q) = \sum_i p(i) \log_2 \frac{p(i)}{q(i)} \quad (4.14)$$

- *teoria zbiorów przybliżonych* – wykorzystująca informacje o elementach zbioru i klasyczną teorię zbiorów do porównywania rozkładów (zob. [68]).

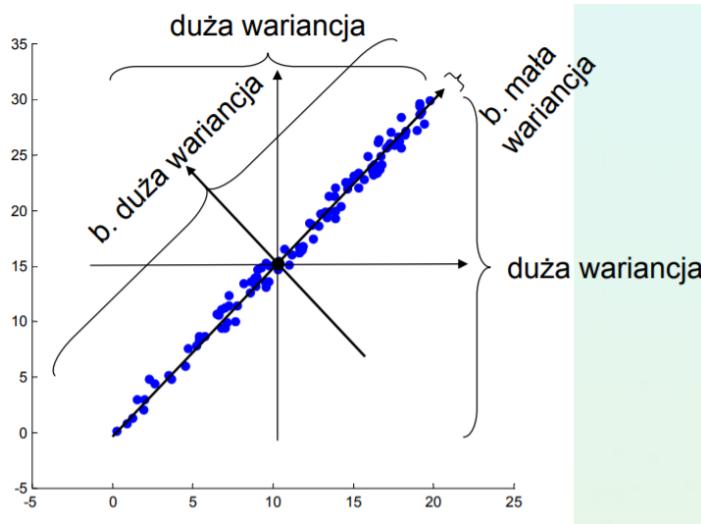
Podobieństwo rozkładów ocenione na podstawie powyższych miar daje możliwość skutecznej redukcji wymiarowości z zachowaniem efektywności algorytmu.

W drugim przypadku, gdy zaistnieje potrzeba przekształcenia przestrzeni cech i wyliczenia nowego zbioru, metodą *state-of-the-art* jest algorytm *analizy składowych głównych* (ang. *Principal Component Analysis*, w skr. *PCA*) opracowany przez Karla Pearsona w 1901 r (zob. [96]).

Istotą PCA jest przekształcenie początkowych, skorelowanych cech w nowy zbiór nieskorelowanych zmiennych. Nowe zmienne, tzw. *składowe główne*, powstają z przekształcenia oryginalnych zmiennych skorelowanych, w taki sposób aby w maksymalnym stopniu wyjaśniać całkowitą wariancję w próbie cech oryginalnych. Wariancje składowych głównych są *wartościami własnymi* macierzy kowariancji oryginalnych zmiennych. Innymi słowy są to *wartości skalarne* (pojedyncze liczby) opisujące poziom wariancji w danych w stosunku do zadanych *ortogonalnych wektorów*, tj. gdzie iloczyn skalarny równy jest 0, co w przestrzeni Euklidesowej oznacza, że kąt pomiędzy wektorami równy jest 90° .

Dla przykładu pierwsza składowa główna redukuje największą część zróżnicowania, druga kolejną, której nie redukowała poprzednia itd. Składowych głównych jest zawsze tyle ile wymiarów danych. Przykład dwuwymiarowy został przedstawiony na Rys. 4.6.

Można zaobserwować, że nowa przestrzeń charakteryzuje się znacznie większym zróżnicowaniem.



Rysunek 4.6: Dwuwymiarowa przestrzeń cech wraz z zaznaczonymi poziomami wariancji i składowymi głównymi.

cowaniem poziomu wariancji niż przestrzeń oryginalna. Wybierając zatem tylko pierwszą składową uzyskana zostanie reprezentacja *skompromitowana stratnie* (tzn. taka która nie daje gwarancji odtworzenia oryginalnych wartości), ale za to zachowująca znaczną część wariancji.

Opis szczegółowej procedury PCA wygląda następująco:

1. Oblicz macierz kowariancji: $S_x = X^T X$, gdzie X to macierz danych zawierająca obserwacje w wierszach. Macierz S_x jest symetryczna i pozwala ocenić: (1) wariancje zmiennych poprzez analizę elementów na głównej przekątnej; (2) zależności pomiędzy zmiennymi poprzez analizę elementów poza główną przekątną.
2. Dokonaj diagonalizacji: $S_x = K L K^{-1}$, gdzie L to macierz diagonalna, a K to macierz odwracalna składająca się z wektorów własnych odpowiadających kolejnym wartościom własnym.
3. Utwórz macierz nowych zmiennych wykonując operację: $Y = XK$

Oprócz PCA, w kontekście redukcji wymiarowości można również wyróżnić następujące algorytmy:

1. *Multidimensional scaling*, MDS – bazuje na wyliczeniu odpowiednio zdefiniowanego dystansu między wartościami cech w oryginalnej przestrzeni i próbie zachowania tego dystansu w zredukowanej wymiarowości (zob. [31]).
2. *T-distributed stochastic neighbor embedding*, t-SNE – stosowany do wizualizacji struktur cech w różnych skalach (zob. [128]).
3. *Isomap* – algorytm nieliniowej redukcji wymiarowości pozwalający na gwarantowane znalezienie globalnego minimum (zob. [127]).
4. *Independent Component Analysis*, ICA – stosowany do znalezienia reprezentacji danych składających się z niezależnych elementów (zob. [73]).
5. *Latent Semantic Analysis*, LSA – stosowany głównie w problemach dotyczących przetwarzania języka naturalnego, zwłaszcza do znalezienia kontekstu użycia danego słowa poprzez analizę statystyczną dużych bloków tekstów (zob. [134]).
6. *Self Organizing Maps*, SOM – bazuje na zachowaniu własności topologicznych przestrzeni cech oryginalnych w zredukowanej wymiarowości (zob. [132]).

Zarówno przekleństwo wymiarowości, jak i wcześniej opisany problem nadmiernego dopasowania są zjawiskami bardzo często wystającymi w praktycznych zastosowaniach algorytmów głębokiego uczenia się. Nie wyczerpują one jednak tematu. Szereg kolejnych wyzwań i rozwiązań został opisany w kolejnej sekcji przy okazji przedstawienia współczesnych topologii.

4.3 Przykłady współczesnych topologii

Współczesne sieci konwolucyjne służące do klasyfikacji obrazów, podobnie jak u protoplasty tj. architektury LeNet, składają się z dwóch głównych części: ekstraktora cech i klasyfikatora. W porównaniu jednak do pierwszych topologii, różnorodność warstw uległa zwiększeniu. Do najczęściej stosowanych komponentów można zaliczyć:

- *warstwy konwolucyjne* – grupują filtry używane do ekstrakcji cech obrazowych różnego poziomu (np. krawędzie, skupiska, obiekty);

- *warstwy max-pool* – stosowane do nielinowej redukcji wymiarowości. Z obszaru $n \times n$ w obrazie wyliczana jest największa wartość ² $I(x, y)$. W praktyce zazwyczaj $n=2$, podobnie jak wartość *kroku* tj. odstępu pomiędzy kolejnymi próbkowaniemiami funkcji I .
- *warstwy aktywacji* – zawierające funkcje aktywacji neuronów. We współczesnych architekturach zamiast funkcji sigmoidalnych, czy progowych stosuje się najczęściej funkcję ReLU:

$$f(x) = \max(0, x). \quad (4.15)$$

Jest ona mniej wymagająca obliczeniowo niż funkcja sigmoid, a dla współczesnych topologii o dużej liczbie parametrów testy empiryczne wykazały, że równie dobrze wpływa na uzyskiwane wyniki dokładności (zob. [70]).

- *warstwy fully connected*, FC – składające się z neuronów, z których każdy jest połączony z każdym neuronem kolejnej warstwy. Występują zazwyczaj w ostatniej części sieci konwolucyjnych tj. w klasyfikatorze.
- *warstwy normalizacji*³ – normalizujące wyjście poprzedzającej warstwy do pewnego określonego przedziału, zazwyczaj $(-1,1)$. Przykładem jest operacja *Local Response Normalization* dana wzorem:

$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} (a_{x,y}^j)^2 \right)^\beta, \quad (4.16)$$

gdzie $a_{x,y}^i$ jest aktywacją danego neuronu, a k , α , n i β to stałe dobierane empirycznie na podstawie zbioru walidacyjnego (za [70]). Warstwy normalizacji wykorzystywane są w celu zrównoważenia poziomu wpływu poszczególnych neuronów na wynik końcowy.

Wprowadzone w ostatnich latach dedykowanych dla głębokiego uczenia się rozwiązania zarówno w warstwie sprzętowej jak i oprogramowania umożliwiło tworzenie architektur o wysokim stopniu komplikacji. W głównej mierze rozwiązania te umożliwiły optymalizację fazy szkolenia omówionej w poprzedniej sekcji jak również *fazy wnioskowania*, gdzie wytrenowana sieć użyta jest do przetwarzania kolejnych obserwacji. W

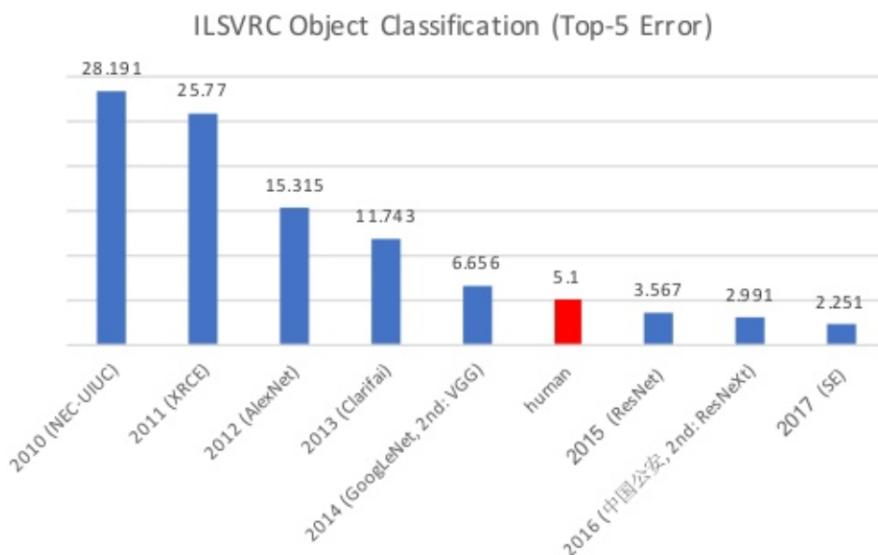
²W uzasadnionych przypadkach stosuje się również obliczanie średniej, normy $L2 = \|x\| = \sqrt{\sum_{k=1}^n |x_k|^2}$ i innych współczynników.

³Coraz rzadziej stosowane w praktyce z uwagi na zniski wpływ na rezultaty.

warstwie sprzętowej wyróżnić można dedykowane akceleratory do operacji macierzowych takie jak [86], stworzone przez firmę NVIDIA *Tensor Processing Unit* (w skr. *TPU*) czy opracowane przez firmę Intel procesory Nervana [71].

W warstwie oprogramowania należy wspomnieć o rozwiązaniach takich jak TensorRT [11], które minimalizują czasy przetwarzania sygnału przez wytrenowaną sieć np. stosując optymalizację reprezentacji liczb zmiennoprzecinkowych lub przetwarzanych macierzy. Dynamicznie rozwijają się również frameworki do zoptymalizowanych obliczeń z udziałem głębokich sieci neuronowych np.: Caffe, Caffe2, TensorFlow, Theano, PyTorch czy MXNet. Ich porównanie osadzone w kontekście aplikacji medycznych można znaleźć w [46].

Postęp w rozwoju współczesnych sieci konwolucyjnych doskonale odzwierciedla progres w rezultatach konkursu ImageNet Large Scale Visual Recognition Competition (w skr. ILSVRC) przedstawiony na Rys. 4.7. Prezentowane wyniki dotyczą wartości błędu



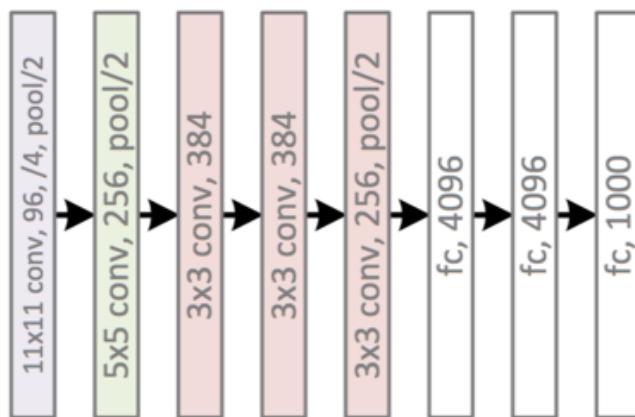
Rysunek 4.7: Błąd top-5 klasyfikacji obiektów w kolejnych latach uzyskiwany przez zwycięzców konkursu ILSVRC

du klasyfikacji na zbiorze danych [40] o nazwie ImageNet, uzyskiwanego w kolejnych latach przez zwycięskie algorytmy biorące udział w konkursie. *Błąd top-n* należy rozumieć jako zdarzenie, w którym dla danego obrazka w n wskazanych przez algorytm najbardziej prawdopodobnych etykietach nie było poprawnej. W zakresie zmniejszenia wartości błędu top- n , znaczący progres dokonał się w 2012 roku, gdzie błąd top-5 zmalał o 10.4 punktów procentowych, co było rezultatem działania nowej architektury sieci

konwolucyjnej nazwanej AlexNet. W kolejnych latach sieci konwolucyjne deklasowały inne podejścia doprowadzając w 2015 roku do spadku błędu top-5 do poziomu 3,5%, co jest uznawane za poziom lepszy niż możliwości ludzkiej klasyfikacji zbioru ImageNet. Lata 2016 i 2017 to intensywne prace nad synergią i złożeniami różnego rodzaju modeli, które w konsekwencji doprowadziły do obniżenia wartości błędu top-5 do poziomu 2,2%. W kolejnych podsekcjach zostanie dokładniej omówiona ewolucja zwycięskich architektur z konkursu ILSVRC.

4.3.1 AlexNet

Sieć AlexNet, której nazwa pochodzi od imienia głównego twórcy tej architektury Alexa Krizhevsky, zawiera blisko 60 milionów parametrów i 650 tysięcy neuronów. Architekturę zaprezentowano na Rys. 4.8



Rysunek 4.8: Schemat architektury AlexNet.

W skład topologii wchodzi pięć warstw konwolucyjnych i trzy typu fully-connected. Po pierwszej, drugiej i piątej warstwie konwolucyjnej występują operacje typu max-pool z maską o wymiarach 2×2 ⁴.

Pierwsza warstwa konwolucyjna przyjmuje na wejściu dane o wymiarze $227 \times 227 \times 3$ ⁵, na których wykonywana jest operacja splotu z 96 filtrami z maską o wymiarach $11 \times 11 \times 3$ i krokiem 4. W rezultacie (uwzględniając również operację max-pool) objętość wynikowa przekazywana do kolejnej warstwy ma wymiar $27 \times 27 \times 96$. W drugiej warstwie

⁴Autorzy pracy podają też przykłady użycia masek o wymiarze 2×3 , które nakładają się w przestrzeni funkcji obrazowej. Nie znalazły one jednak miejsca w finalnej implementacji.

⁵gdzie 3 jest liczbą kanałów kodujących kolor obrazka.

konwolucyjnej wykonywana jest operacja splotu z 256 filtrami z maską o wymiarach $5 \times 5 \times 96$. Wymiar objętości wynikowej zostaje ponownie zredukowany poprzez operacje max-pool do $13 \times 13 \times 256$. Kolejne 3 warstwy konwolucyjne są połączone bezpośrednio ze sobą. Trzecia warstwa zawiera 384 filtry z maską o wymiarze $3 \times 3 \times 256$, w skład czwartej wchodzą 384 filtry z maską o wymiarze $3 \times 3 \times 384$, a w piątej znajduje się 256 filtrów również z maską o wymiarze $3 \times 3 \times 384$. Końcowe dwie warstwy typu FC zawierają po 4096 neuronów, a ostatnia zawiera tyle neuronów ile klas występuje w ostatecznym podziale – w oryginalnej pracy było to 1000 (por. [70]).

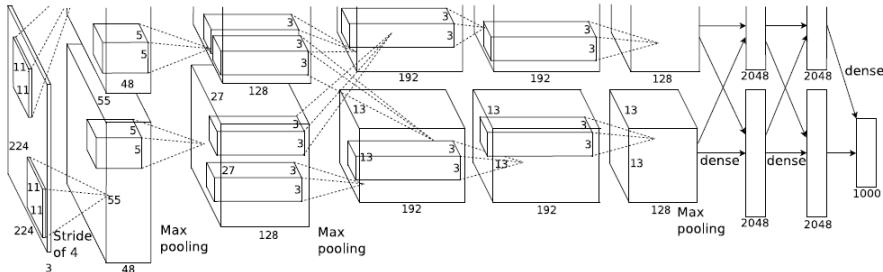
W celu lepszego zrozumienia przetwarzania sygnału wejściowego przez sieć poniżej przedstawiono przykład algorytmu wykorzystywanego dla pierwszej warstwy konwolucyjnej opisywanej topologii:

1. Z danych wejściowych o wymiarze $[227 \times 227 \times 3]$ wybierany jest co czwarty blok o wymiarach $[11 \times 11 \times 3]$ (zarówno wzduż wysokości jak i szerokości). W rezultacie, nie uwzględniając krawędzi obrazu, otrzymywanych jest 217 punktów w każdym rzędzie i w kolumnie, w których mieści się $[55 \times 55]$ tj. 3025 bloków.
2. Zarówno $11 \times 11 \times 3 = 363$ wagi znajdujące się w 96 filtrach jak i wartości 363 punktów obrazowych znajdujących się 3025 blokach są przedstawiane w postaci macierzy A o wymiarach $[96 \times 363]$ i B o wymiarach $[363 \times 3025]$.
3. Następnie liczony jest iloczyn skalarny w postaci $A^T B = C$, gdzie nowa, wyjściowa macierz C ma wymiar $[96 \times 3025]$.
4. Rezultat w postaci macierzy C ponownie przewymiarowany jest na postać $[55 \times 55 \times 96]$.

W architekturze jako funkcję aktywacji neuronów wykorzystano ReLU, co znaczaco przyspieszyło trening sieci. Dla przykładu uzyskano 6-krotne przyspieszenie treningu dla danych [69], CIFAR-10 w stosunku do tej samej topologii wykorzystującej sigmoidalną funkcję aktywacji. Ponieważ funkcja ReLU nie posiada górnego ograniczenia neurony teoretycznie mogą posiadać nieograniczone wartości funkcji aktywacji. W celu polepszenia kontrastu pomiędzy neuronami i wydobycia tych, które na tle innych się wyróżniają, zastosowano normalizację zgodną ze wzorem 4.16. W wyniku czego uzyskano redukcję błędu klasyfikacji top-5 o wartość 1,2 punktu procentowego.

W kontekście zwiększenia efektywności treningu zastosowano również powiększenie rozmiaru danych poprzez rotacje i modyfikacje funkcji obrazowej z wykorzystaniem czynników głównych (zob. [70]) zmniejszając błąd top-1 o 1%. Zastosowano też technikę

dropout opisaną w 4.2.1. Ostatecznie wprowadzono także trening z wykorzystaniem wielu GPU (zob. Rys. 4.9).



Rysunek 4.9: Topologia architektury AlexNet z podziałem na dwa akceleratory GPU.

Topologia z podziałem na 2 karty zwiększyła dwukrotnie sumaryczną pamięć i pozwoliła na kolokację parametrów sieci.

Praca Alexa Krizhevsky, Ilya Sutskever i Geoffrey'a Hinton zapoczątkowała wzrost zainteresowania technikami głębokiego uczenia się, co doprowadziło do publikacji kolejnych podobnych architektur. Do najbardziej znanych należą ZFNet z 2013 roku [139], gdzie m.in. zastosowano zmniejszenie wymiaru maski stosowanego w filtrach pierwszej warstwy konwolucyjnej do 7×7 oraz VGGNet [114] z 2014 roku, gdzie zastosowano większą liczbę warstw konwolucyjnych z mniejszym wymiarem maski. Również w 2014 roku zaprezentowano innowacyjną koncepcję modułów sieci konwolucyjnych, co doprowadziło do zwycięstwa w ILSVRC. Idea ta została dokładniej opisana w kolejnej podsekcji.

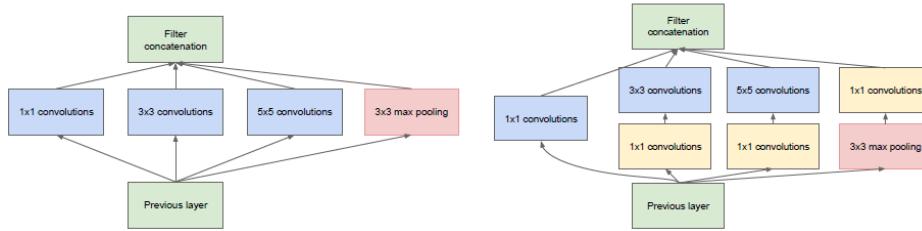
4.3.2 GoogLeNet

Architekturę o nazwie GoogLeNet zaprezentowano w 2014 r. w pracy [122]. Nazwa architektury pochodzi od nazwy zwycięskiego zespołu startującego w ILSVRC 2014, składającego się z pracowników firmy Google. Oryginalnie topologia składała się z 22 warstw i zawierała około 5 mln parametrów (12 razy mniej niż w przypadku sieci AlexNet).

Redukcję liczby parametrów przy jednoczesnym podwyższeniu dokładności klasyfikacji obiektów udało się uzyskać poprzez poszukiwania konstrukcji optymalnych lokalnych topologii i ich połączeń. Mianowicie, wiadomo że duża część funkcji aktywacji neuronów przyjmuje wartość 0 lub jest redundantna z powodu wysokiej korelacji międ-

czy sobą (zob. [20]). Matematyka dotycząca przetwarzania *macierzy rzadkich*, tj. gdzie przeważająca liczba elementów przyjmuje wartość 0, jest dobrze znana (zob. np. [90]). Jednak implementacje bibliotek do obliczeń związanych z algebrą liniową są zoptymalizowane pod kątem *macierzy gęstych*, gdzie przeważająca liczba elementów przyjmuje wartości różne od 0 (zob. [70, 116]).

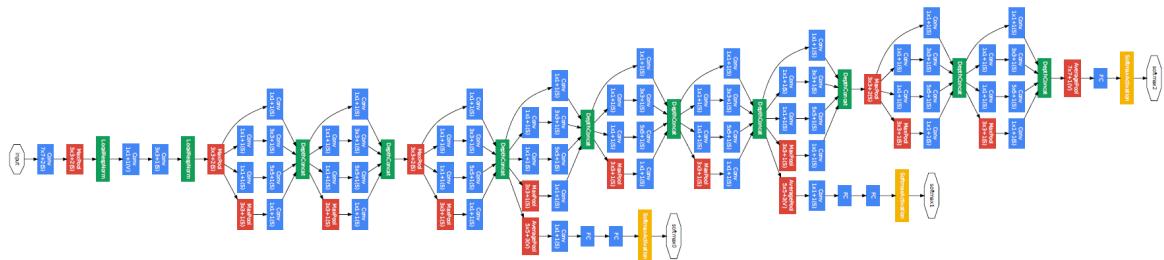
Ideą modułu incepcji zaproponowanego przez twórców GoogLeNet jest aproksymacja rzadkich macierzy z użyciem komponentów o gęstej strukturze. Takie komponenty nazwano *modułami incepcji* (ang. *inception modules*), a ich przykłady pokazano na Rys. 4.10.



Rysunek 4.10: Topologia architektury AlexNet z podziałem na dwa akceleratory GPU.

(a) przedstawia naiwną formę modułu incepcji, gdzie grupowane są operacje filtrów z maską o wymiarach 5×5 , 3×3 , 1×1 oraz operacja max-pool. (b) prezentuje koncepcję zoptymalizowaną obliczeniowo gdzie filtry z maską 1×1 służą do redukcji wymiarowości i używane są bezpośrednio przed splotami z bardziej wymagającymi obliczeniowo splotami 5×5 i 3×3 .

Przy pomocy złożenia różnego rodzaju modułów incepcji otrzymano topologię zaprezentowaną na Rys. 4.11



Rysunek 4.11: Topologia architektury GoogleNet

Dokładne zestawienie parametrów znajduje się w Tabeli 4.1

type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								

Tabela 4.1: Parametry architektury GoogleNet

Ważną cechą sieci GoogleNet jest brak warstw typu FC na zakończeniu, gdzie w przypadku sieci AlexNet znajduje się około 90% parametrów. Końcowe wnioskowanie jest realizowane na podstawie wartości średniej z dwuwymiarowych map cech.

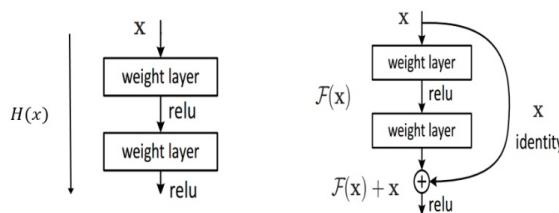
Dla lepszego zrozumienia idei redukcji wymiarowości realizowanej przez moduły incepcji, podobnie jak w przypadku sieci AlexNet, przeanalizowane zostanie działanie pierwszego modułu w topologii z Rys. 4.11. Moduł zawiera 128 filtrów z maskami o wymiarach 3×3 i 32 filtry z maskami o wymiarach 5×5 . Dane na wejściu modułu mają 192 kanały (zob. Tabela 4.1). Dla przykładu, rząd wielkości obliczeń operacji splotów 32 filtrów 5×5 wynosi $25\times 32\times 192=153\ 600$ i dalej wzrastałby z głębokością sieci. W celu zapobiegnięcia nadmiarowi obliczeń stosowana jest redukcja z użyciem 16 filtrów z maską o wymiarach 1×1 . W efekcie rząd wielkości obliczeń spada do $16\times 192 + 25\times 32\times 16=15\ 876$, co pozwala na dalsze budowanie wielowarstwowych struktur.

Topologia GoogLeNet jest wciąż rozwijana. Po pierwszej prezentacji pojawiły się kolejne modernizacje wprowadzające dodatkowe faktoryzacje modułów jak w Inception-v2 lub normalizacje wartości wynikowych poszczególnych warstw jak w Inception-v3. Obie sieci zostały przedstawione w [123]. Kolejny innowacyjny pomysł, bazujący na dodatkowych połączeniach między blokami, został wprowadzony w 2015 roku w sieci ResNet opisanej w kolejnej podsekcji.

4.3.3 ResNet

Jednym z najbardziej oczywistych pomysłów na polepszenie dokładności działania sieci neuronowych jest zwiększenie liczby warstw. Jednak wraz ze wzrostem liczby warstw, trening takich architektur z użyciem tradycyjnych metod gradientowych (takich jak algorytm wstępnej propagacji błędu) staje się mniej wydajny. Problem wynika z faktu, że zmiana wartości sygnału na wyjściu sieci w odpowiedzi na sygnał wejściowy jest mniejsza wraz ze wzrostem liczby warstw. W takiej sytuacji gradient wyliczany na podstawie sygnału będącego różnicą pomiędzy sygnałem wejściowym a wyjściowym może przyjmować wartości bliskie 0 uniemożliwiając dalszy postęp uczenia się. Problem zanikającego gradientu (ang. *vanishing gradient problem*) rozwiązywany jest poprzez zastosowanie normalizacji oraz nieliniowych funkcji aktywacji. Dzięki tym mechanizmom algorytm treningu głębszych sieci neuronowych w większej liczbie przypadków zbiega do użytecznego minimum lokalnego.

W momencie znalezienia takiego minimum dodanie kolejnych warstw i parametrów sieci jest redundantne, a nawet prowadzi do pogorszenia wyników treningu sieci. Zjawisko to nosi nazwę *degradacji treningu* (ang. *degradation problem*). Twórcy architektury ResNet, przedstawionej w [56], zaproponowali rozwiązanie tego problemu poprzez implementację bloków rezydualnych (ang. *Residuum Units*) zawierających dodatkowe, skrótowe połączenia (ang. *skip connections*) pomiędzy wejściem a wyjściem bloków. Porównanie schematów funkcjonalnych nowych bloków i wcześniej istniejącego rozwiązania stosowanego np. w AlexNet zostało przedstawione na Rys. 4.12.



Rysunek 4.12: Schemat funkcjonalny pojedynczego bloku w architekturze ResNet.

Ogólną postać równania bloku rezydualnego można zapisać następująco:

$$\begin{aligned} y_l &= h(x_l) + F(x_l, W_l), \\ x_{l+1} &= f(y_l), \end{aligned} \tag{4.17}$$

gdzie x_l i x_{l+1} stanowią sygnał wejściowy i wyjściowy l -tego bloku. F stanowi funkcję

rezydualną optymalizowaną podczas treningu sieci, $h(x_l)$ stanowi funkcję przekształcenia sygnału x_l przekazywanego skrótowym połączeniem, f jest funkcją ReLU, a W stanowi macierz wag.

Funkcja $h(x_l)$ jest funkcją tożsamościową, a zatem $h(x_l) = x_l$. Źeby uzasadnić ten wybór należy rozważyć propagację gradientu wewnątrz sieci składającej się z bloków rezydualnych. Dla każdego L -tego bloku zachodzi równanie:

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i) \quad (4.18)$$

Korzystając z reguły łańcuchowej można zapisać równanie na gradient funkcji kosztu ε :

$$\frac{\partial \varepsilon}{\partial x} = \frac{\partial \varepsilon}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L} \left(1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i) \right) \quad (4.19)$$

z czego wynika, że gradient może być podzielony na dwie addytywne składowe: (1) $w = \frac{\partial \varepsilon}{\partial x_L}$ propagowaną bez wpływu na warstwy zawierające wagę i (2) $\lambda = \frac{\partial \varepsilon}{\partial x_L} \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i)$ propagowaną przez nie.

Przykład propagacji gradientu w sieci składającej się z trzech bloków wygląda zatem następująco:

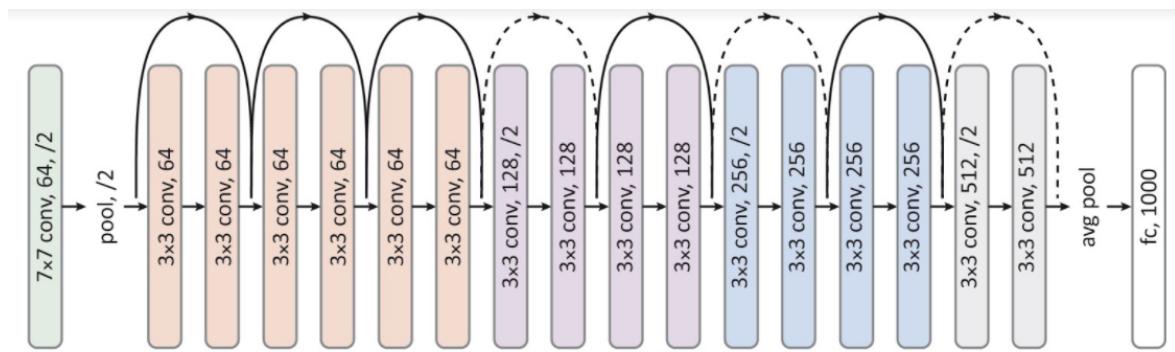
$$\frac{\partial \varepsilon}{\partial x_0} = \frac{\partial \varepsilon}{\partial x_3} * (w_2 + \lambda_2) * (w_1 + \lambda_1) * (w_0 + \lambda_0) \quad (4.20)$$

Wartości w są zazwyczaj znormalizowane do przedziału $(-1;1)$ można więc rozważyć 4 istotne przypadki równania 4.20:

1. $\lambda = 0$ – nie ma skrótowych połączeń, co odpowiada płaskiej strukturze sieci.
Ponieważ wartości w są z przedziału $(-1;1)$ dodawanie kolejnych warstw wzmacnia wcześniejszy omówiony efekt zanikającego gradientu
2. $\lambda > 1$ – z każdą warstwą, sumaryczna wartość gradientu zwiększa się inkrementalnie, co nazywane jest *problemem eksplozji gradientu* (ang. *exploding gradient problem*).
3. $\lambda < 1$ – przy założeniu, że $w + \lambda < 1$, dla sieci składających się z wielu warstw występuje problem zaniku gradientu, jak w przypadku 1. Natomiast, gdy $w + \lambda > 1$ podobnie jak w przypadku 2 może występować problem eksplozji gradientu
4. $\lambda = 1$ – wartości w są inkrementowane dokładnie o 1, co eliminuje problemy

podane w przypadkach 1, 2 i 3 i stanowi uzasadnienie dla wyboru funkcji tożsamościowej $h(x_l)$ w architekturze ResNet.

Dokładny opis matematyczny funkcjonowania bloków rezydualnych wraz z dowodami znajduje się w [57]. Przykład topologii sieci składającej się z 8 bloków i łącznie 18 warstw tzw. ResNet-18, przedstawiono na Rys. 4.13.



Rysunek 4.13: Topologia architektury ResNet-18.

Pierwsza warstwa konwolucyjna zawiera filtry z maską o wymiarach 7×7 . W kolejnych zastosowano wymiar 3×3 . Zastosowanie mniejszych wymiarów maski niż w AlexNet oraz podobnie jak w przypadku sieci GoogLeNet wyliczenie na końcu wartości średniej z dwuwymiarowych map cech zredukowało liczbę parametrów.

Architektura ResNet-18 jest najmniejszą z pojawiających się w literaturze przykładów tego typu. W praktyce, z powodzeniem wykorzystywano topologie składające się nawet z 1202 warstw (zob. [56]). W 2016 roku zaprezentowano w [121] hybrydę sieci GoogleNet i ResNet. Pracowano również nad bardziej złożonymi blokami, co w konsekwencji doprowadziło w 2017 roku do zaprezentowania architektury ResNetX w [137], która w wielu testach klasyfikacji różnych zbiorów okazała się być lepsza niż poprzednicy. Przegląd dotyczący historii tych prac można znaleźć w [12].

Sieć ResNet i jej warianty dla wielu testowych zbiorów danych takich jak ImageNet, CIFAR czy COCO [78] osiągnęły dokładność klasyfikacji porównywalną z możliwościami ludzkiego obserwatora. Dalszy progres był możliwy m.in. dzięki zastosowaniu synergii wielu modeli, co zostało opisane w kolejnej podsekcji.

4.3.4 Złożenia

Uczenie złożień sieci (ang. *ensemble learning*) polega na wykorzystywaniu kilku modeli bazowych i wybranej metody ich synergii. W kontekście głębokiego uczenia się stosowane są różne metody kombinacji modeli bazowych (zob. [64]). Jako często stosowane przykłady można podać: uśrednianie, głosowanie, klasyfikacja Bayesa, generalizację stosów. Zostaną one kolejno omówione:

Uśrednianie

Uśrednianie jest prostą metodą kombinacji wyników predykcji. Najczęściej stosowane jest uśrednienie bez wag, gdzie suma wyników predykcji modeli bazowych podzielona jest przez ich liczbę. Uśredniać można bezpośrednio wyniki ostatecznej klasyfikacji jak również prawdopodobieństwa przynależności do odpowiednich klas, które są np. wynikiem *funkcji softmax*:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \quad (4.21)$$

używanej często bezpośrednio przed ostatnią warstwą sieci neuronowych dla z sygnałów wejściowych i j wyjściowych.

Główna zaleta uśredniania jest redukcja wariancji. Jest ona tym większa im bardziej nieskorelowane są wyniki predykcji modeli bazowych. Pomimo prostoty, tego rodzaju koncepcja odnosiła już sukcesy m.in. w lasach losowych (zob. [28]).

Zastosowanie uśredniania przy silnie odstających od średniej najgorszych predykcjach znaczająco obniża dokładność całego złożenia. Dlatego przy tak nieheterogenicznych modelach bazowych dających bardzo różne wyniki poszukiwane są inne metody.

Głosowanie

W głosowaniu stosuje się mechanizm zliczania przewidzianych przez modele bazowe etykiet. Etykieta, która została wybrana przez największą liczbę modeli bazowych jest obierana jako wynik ostatecznej predykcji. Jest to tzw. *głosowanie większościowe*.

W porównaniu do uśredniania, głosowanie jest mniej czułe na predykcje pojedynczych modeli. Wykorzystuje jednak jedynie informacje o przewidzianych etykietach, co utrudnia konstrukcję bardziej wyszukanych rozwiązań.

Klasyfikacja Bayesa

W przypadku tej metody, każdy model bazowy j postrzegany jest jako hipoteza h_j . Każda z hipotez posiada wagę proporcjonalną do prawdopodobieństwa zdarzenia, w którym dany zbiór trenujący zostałby wybrany z ogółu danych gdyby dana hipoteza była prawdziwa. Jest to tzw. optymalna klasyfikacja Bayesa, którą można zapisać następującym równaniem:

$$y = \operatorname{argmax}_{c_j \in C} \sum_{h_i \in H} P(c_j | h_i)P(T | h_i)P(h_i) \quad (4.22)$$

gdzie y to przewidziana etykieta, C jest zbiorem wszystkich możliwych klas, H to przestrzeń hipotez, a T to zbiór danych trenujących.

W praktyce z uwagi na dużą złożoność obliczeniową nie stosuje się optymalnej klasyfikacji Bayesa, a jedynie aproksymacje tej metody np.: BPA (od ang. *Bayesian parameter averaging*) [37], BMA (od ang. *Bayesian model averaging*) [104], czy też BMC (od ang. *Bayesian model combination*) [87].

Generalizacja stosów

Idea generalizacji stosów oryginalnie została zaproponowana w [135]. Wykorzystana została koncepcja *meta-uczenia*, a zatem konstrukcja nadzawanego klasyfikatora, którego zadaniem jest wybór optymalnego wektora wag a dla stosu s predykcji dla danych x :

$$s(x) = \sum_{i=1}^m a_i s_i(x) \quad (4.23)$$

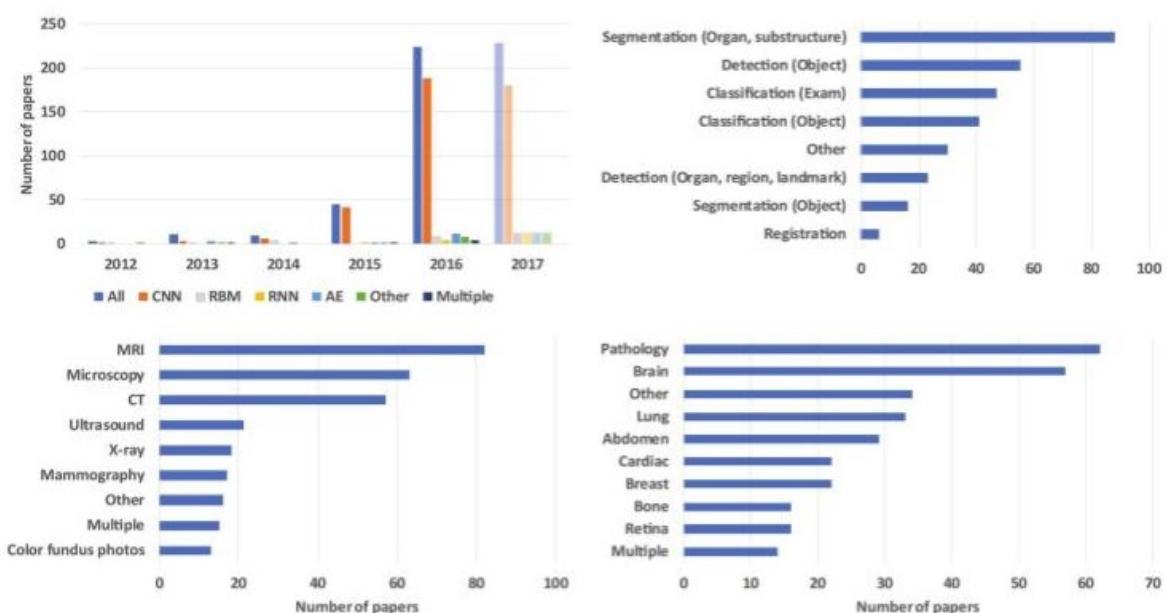
W praktyce predykcje z modeli bazowych składowane są na stosie, a następnie klasyfikator nadzawany wykorzystuje je jako dane do treningu poprawnych wartości a wykorzystując jako odniesienie znane, poprawne etykiety.

4.4 Zastosowania w medycynie

W 1994 roku ukazała się pierwsza praca, która w praktyce wykorzystywała mechanizmy związane z głębokim uczeniem się do przetwarzania obrazów medycznych (zob. [140]). Użyte wówczas sieci nazywano sieciami typu *shift-invariant*. Zastosowanie ich pozwoliło na eliminację 55% FP otrzymywanych przy wcześniejszych metodach sto-

sowanych do detekcji skupisk mikro-zwapnień w mammografach. *Shift-invariant* oznaczało, że przesunięcie obrazu wejściowego nie powodowało zmian w klasyfikacji, co jest istotną wartością dodaną, z uwagi na specyfikę implementacji toru akwizycji danych w praktyce radiologicznej.

Po roku 2012 nastąpił znaczący wzrost zainteresowania metodami głębokiego uczenia się w medycynie. Obrazuje to praca [79] z 2017 roku, w której przytoczono statystyki medycznych publikacji zawierających słowa kluczowe związane z deep learning. Wybrane dane przedstawiono na Rys. 4.14.

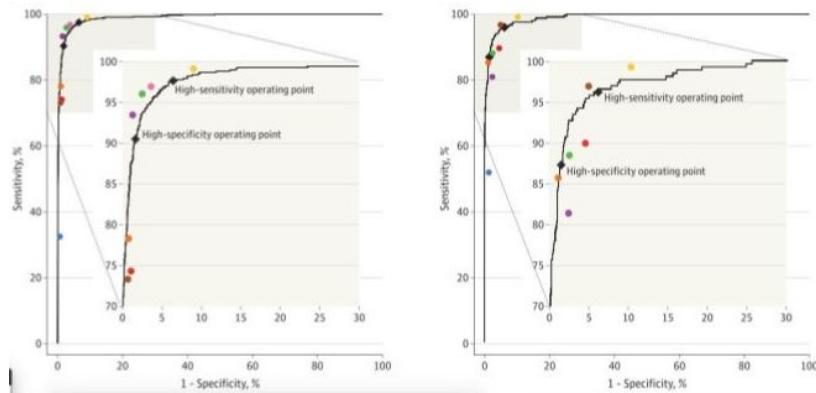


Rysunek 4.14: Statystyki dotyczące publikacji medycznych zawierających słowa kluczowe związane z głębokim uczeniem się.

Widoczny wzrost liczby publikacji nastąpił począwszy od 2015, co związane było z kilkuletnią adaptacją nowych metod w dziedzinie przetwarzania obrazów medycznych i gromadzeniem odpowiednich zbiorów danych. Lata 2016 i 2017 były pod pewnym względem przełomowe gdyż pojawiło się coraz więcej prac naukowych, w których przedstawiano rezultaty dokładności klasyfikacji na poziomie dorównującym ekspertom dziedzinowym.

Dla przykładu, w Listopadzie 2016 ukazała się praca [53] grupy Google Research z Mountain View w Kalifornii, gdzie zastosowano sieć GoogLeNet w wersji inception-v3 do zautomatyzowanej detekcji retinopatii cukrzycowej i cukrzycowego obrzęku plamki w obrazach dna oka. Wyniki porównano z panelem składającym się z 7 ekspertów,

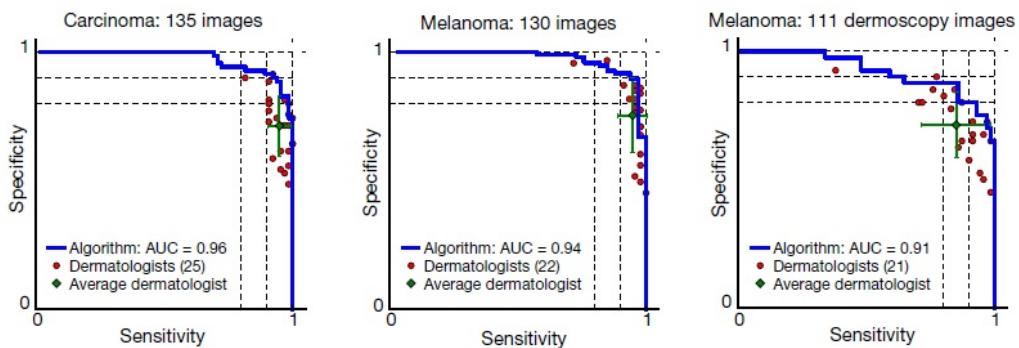
okulistów. Porównanie przedstawiono na Rys. 4.15.



Rysunek 4.15: Porównanie automatycznej klasyfikacji retinopatii cukrzycowej i cukrzycowego obrzęku plamki z oceną panelu ekspertów.

Na wykresach dla dwóch zadań klasyfikacyjnych umieszczono krzywe reprezentujące zależność swoistości od czułości dla algorytmu automatycznego oraz 7 punktów oznaczających wynik oceny każdego z okulistów. Ogółem mniej niż połowa ekspertów uzyskała lepszy wynik niż algorytm sztucznej inteligencji.

Kolejna ciekawa praca tj. [48], pojawiła się w czasopiśmie Nature w styczniu 2017 roku i traktowała o automatycznej detekcji nowotworów skóry na podstawie zdjęć. Autorzy wykorzystali dane składające się z 129.450 obrazów klinicznych, na których zobrazowano 2.032 różne schorzenia skóry. Ponownie do klasyfikacji wykorzystano sieć GoogleNet w wersji inception-v3. Wyniki klasyfikacji automatycznej porównano z oceną przeprowadzoną przez 21 certyfikowanych dermatologów. Przykład porównania zaprezentowano na Rys. 4.16.



Rysunek 4.16: Porównanie automatycznej klasyfikacji 3 chorób skóry z oceną ekspertów dermatologów.

Wykres przedstawia zależność czułości od swoistości. Czerwonymi punktami oznaczono wynik oceny poszczególnych ekspertów, a zielonym krzyżykiem wynik uśredniony. W każdym przypadku średnia ocena była gorsza od automatycznej klasyfikacji.

Obrazy medyczne nie są jedynymi danymi, które z powodzeniem są przetwarzane za pomocą metod głębokiego uczenia się. W lipcu 2017, przez grupę ze Stanford University została opublikowana praca [98] dotycząca klasyfikacji arytmii na podstawie szeregiów czasowych zapisanych na elektrokardiogramach. Autorzy wykorzystali dane z 64.121 elektrokardiogramów, próbkowanych z częstotliwością 200 Hz, pochodzących od 29.163 pacjentów. Zaprojektowano dedykowaną, 34-warstwową sieć konwolucyjną do detekcji 12 różnych dysfunkcji pracy serca, pracy prawidłowej i szumów (łącznie 14 klas). Wyniki klasyfikacji porównano z oceną prowadzoną przez 3 kardiologów. Średnia dokładność z oceny automatycznej wyniosła 80%, natomiast manualnej 72%.

Podobnych przykładów zostało opublikowanych dużo więcej. Architektura AlexNet z sukcesem była użyta do detekcji polipów w kolonoskopii w [125]. Sieć ResNet sprawdziła się w badaniach zrealizowanych w [47] w Mayo Clinic Rotschester. Dotyczyły one radio-genomiki i rozróżnienia zmian w mózgu bez konieczności biopsji. Złożenia natomiast z sukcesem zaaplikowano w pracach dotyczących detekcji nowotworów płuc, gdzie modele bazowe analizowały różne skale problemu (zob. [43]). W wielu pracach dotyczących radiologii raportuje się dokładność klasyfikacji automatycznej znacząco przewyższające możliwości dziedzinowych ekspertów np. [35, 50, 106, 133].

Powyższe przykłady pokazują, że dla szczególnych przypadków pewien element pracy eksperta zajmującego się danymi medycznymi (np. radiologa) może być z sukcesem wspomagany (lub nawet zastąpiony) przez algorytmy głębokiego uczenia się. Należy jednak podkreślić, że jest również szereg problemów wiążących się z wykorzystaniem tego rodzaju sztucznej inteligencji w medycynie. Do najważniejszych należą:

1. Gromadzenie dużych zbiorów danych z odpowiednimi etykietami.
2. Wykorzystanie heterogenicznych danych pochodzących np. z wielu urządzeń lub modalności.
3. Kalibracja i szacowanie niepewności wyników modeli.
4. Unifikacja modeli wykonujących podobne zadania.
5. Minimalizacja liczby parametrów modelu przy zachowaniu satysfakcjonującego poziomu dokładności.

Więcej na temat ograniczeń metod głębokiego uczenia się można przeczytać również w [85].

Dyskusja dotycząca tych problemów wciąż jest tematem wielu paneli dyskusyjnych i debat konferencyjnych (zob np. [6]). Najbardziej zaawansowane prace dotyczą problemu gromadzenia dużych zbiorów danych medycznych, co wymaga bliskiej współpracy ekspertów medycznych z ekspertami od uczenia maszynowego. Często konieczna jest również modyfikacja bądź tworzenie dedykowanych programów do akwizycji danych medycznych. Jako przykłady takich inicjatyw można wymienić programy Stanford Medicine [9], Harvard School of Medicine [33] czy Massachusetts General Hospital, które w swoich repozytoriach zgromadziły już dziesiątki milionów zdjęć radiologicznych (za [4]). Ponadto w roku 2018, na konferencji NVIDIA GTC (GPU Technology Conference) w San Jose (Kalifornia), Amerykańskie Stowarzyszenie Radiologii i stowarzyszenie MICCAI (od ang. *Medical Image Computing and Computer Assisted Intervention*) ogłosili porozumienie, co do wspólnej współpracy mającej na celu eliminację barier legislacyjnych związanych ze współpracą przy pozyskiwaniu danych i wykorzystywaniem algorytmów uczenia maszynowego.

Autor tej rozprawy jest świadom ograniczeń jakie są związane z wykorzystaniem algorytmów głębokiego uczenia się. Jednocześnie jednak nowe metody i ostatnie sukcesy zastosowań współczesnych sztucznych sieci neuronowych w aplikacjach medycznych stanowią silną motywację do przeprowadzenia własnych badań z ich wykorzystaniem.

Rozdział 5

Nowa metoda oceny procesu gojenia ścięgna Achillesa

W tym rozdziale zostanie zaprezentowana autorska propozycja metody oceny procesu gojenia się ścięgna Achillesa bazująca na badaniach obrazowych. W szczególności przedstawiony zostanie ilościowy opis, umożliwiający w zobietywizowany sposób ocenę morfologii tkanek widocznych w obrazach Rezonansu Magnetycznego i Ultrasonografii. Finalnie zostanie również zaproponowane nowatorskie podejście do automatycznego wyliczania wskazanych w opisie procesu parametrów, co jest najważniejszym osiągnięciem całości tej pracy.

Wskazana automatyzacja została przez autora wykonana dla danych z Rezonansu Magnetycznego, które zostały przez specjalistów wskazane jako klinicznie bardziej istotne niż USG. Niemniej jednak dane Ultrasonografii zostały również przeanalizowane pod kątem wartości dodanej do całego proponowanego procesu.

W chwili pisania tej pracy nie istnieje wedle najlepszej wiedzy autora podejście umożliwiające zobietywizowany i co ważniejsze automatyczny sposób oceny badań obrazowych prezentujących gojące się ścięgno Achillesa. Stąd implikacje dotyczące trudności z integracją subiektywnych interpretacji ze skalami testów funkcjonalnych takich jak ATRS i w rezultacie zmniejszona efektywność całego procesu oceny rehabilitacji.

Podczas próby rozwiązania wskazanego problemu autor tej pracy w szczególności chce skupić się na dwóch aspektach. Pierwszym z nich jest jakość generowanej automatycznie oceny odniesiony do jakości wzorca tj. oceny doświadczonego radiologa. Drugim natomiast jest czas akwizycji danych, a zatem wybór praktycznego protokołu,

który zapewni możliwie krótki udział pacjenta w badaniu. W obu przypadkach celem jest poszukiwanie punktu optimum, w którym maksymalizowana będzie jakość oceny, a minimalizowany czas, zarówno radiologa jak i pacjenta, niezbędny do wykonania koniecznych czynności. Szczegóły podejścia zostały opisane w kolejnej sekcji.

5.1 Metodyka

W tej sekcji zostanie szczegółowo opisana proponowana metoda automatycznej oceny procesu gojenia się ścięgna Achillesa widocznego w badaniach obrazowych Rezonansu Magnetycznego. Dodatkowo scharakteryzowany zostanie zbiór danych, który posłużył do opracowania rozwiązania, jak również ankietą walidacyjną stanowiącą wzorzec odniesienia dla przedstawionej metody.

Autor tej pracy, w proponowanym podejściu skorzystał z metod widzenia komputowego, a dokładniej z fuzji algorytmów sztucznej inteligencji i przetwarzania obrazów. W kontekście tej pracy, pierwsze znajdują swoje zastosowanie do ekstrakcji wektora cech dla danej reprezentacji obrazowej. Drugie, pozwalają uwzględnić wiedzę dziedzinową w procesie numerycznej oceny.

W przypadku algorytmów sztucznej inteligencji zastosowano opisane w Rozdziale 4 konwolucyjne sieci neuronowe, a dokładniej AlexNet, GoogLeNet (inceptionV3) i ResNet-18. W pierwszej kolejności wykonane zostało szkolenie podanych sieci dla problemu binarnej klasyfikacji tj. odróżniania obrazów chorego od zdrowego ścięgna. Następnie część klasyfikująca została usunięta z topologii sieci, pozostawiając ekstraktor cech z parametrami zoptymalizowanymi pod kątem wydobycia istotnej informacji opisującej różnice między zdrową i chorą tkanką. Na tak otrzymanym wektorze, przeprowadzono redukcję wymiarowości z wykorzystaniem metody PCA (zob. 4.2.2). W wyniku przeprowadzonych eksperymentów, ostatecznie zdecydowano się uwzględnić 200 pierwszych czynników głównych w końcowym rozwiążaniu.

W przypadku metod przetwarzania obrazów zastosowano obliczenia cech z wydzielonego przez specjalistę radiologa ROI (od ang. *Region of Interest*) reprezentującego tkanki otoczone ościęgnem. Sumarycznie wyliczono 46 klasycznych cech obrazowych, w tym pole powierzchni ROI, 9 cech opisujących statystykę wartości pikseli i 36 cechy Haralick'a opisujące teksturę (zob. [55]). W ramach statystyk scharakteryzowano: min, max, średnią, odchylenie standardowe, skośność, kurtozę, 25-percentyl, medianę oraz 75-percentyl. Natomiast w ramach cech Haralick'a: drugi moment kątowy, kontrast,

korelacja, wariancja, odwrotny moment różnicowy, suma średnich, suma wariancji, suma entropii, entropia, różnica wariancji i maksimum prawdopodobieństwa. Cechy Haralick'a wyliczono dla 3 dystansów separacji $d=1,5,10$. Dokładna metoda selekcji powyższych cech została przedstawiona w pracy [94].

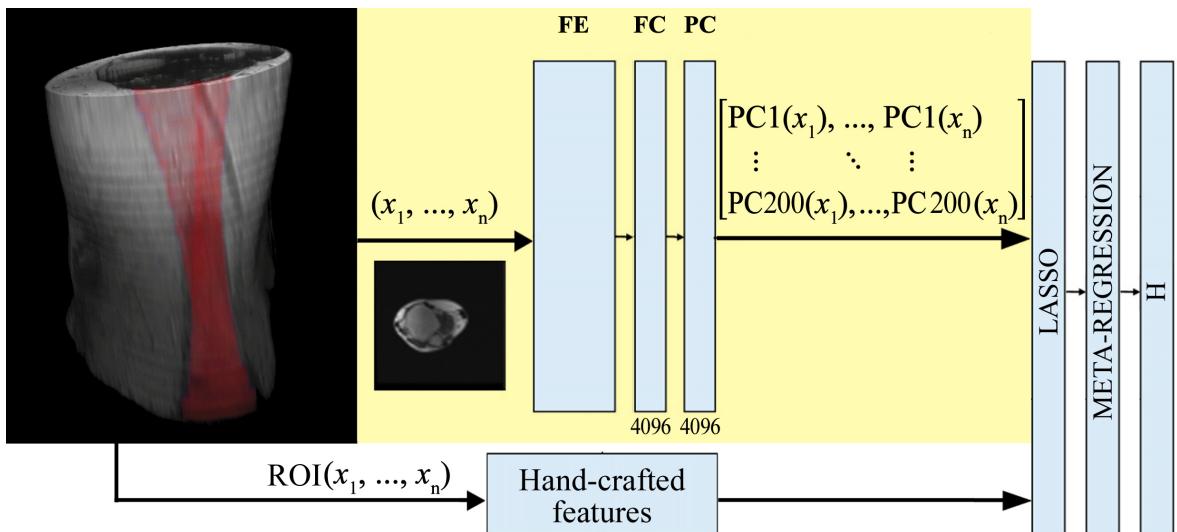
Spośród tak wyliczonych 246 cech dokonano selekcji metodą LASSO (zob. [opis]). W rezultacie dla wszystkich ocenianych parametrów uzyskano optymalną mieszankę poniżej 20 cech ekstrahowanych z wykorzystaniem sieci konwolucyjnych oraz klasycznych metod obrazowych. Wybrane predyktory posłużyły do szkolenia algorytmu meta-regresji, przy użyciu którego wykonana została fuzja.

Końcowa ocena badania została wykonana następującą metryką H zaproponowaną przez autora tej pracy:

$$H = TM(R(x_1), R(x_2), \dots, R(x_n)) \quad (5.1)$$

gdzie TM jest średnią trymowaną z marginesami 2.5%, a $R(x_i)$ jest wynikiem regresji obliczonym dla przekroju osiowego x_i , gdzie i to indeks przekroju w trójwymiarowym badaniu RM.

Schemat powyższej metody został zaprezentowany na Fig. 5.1. W ogólnym ujęciu



Rysunek 5.1: Schemat automatycznej metody oceny procesu gojenia się ściegna Achillesa.

trójwymiarowe badanie RM dzielone jest na pojedyncze przekroje osiowe, które stanowią dane wejściowe do oznaczonej na żółto części opartej o metody głębokiego uczenia

się. Początkowo dane przetwarzane są przez ekstraktor cech sieci AlexNet wybranej w wyniku eksperymentów (FE). Następnie wstępnie grupowane są z wykorzystaniem warstwy fully connected (FC) i poddane redukcji z 4096 pojedynczych wyjść aktywacyjnych do 200 czynników głównych (PC).

Równolegle, ROI z każdego z przekrojów stanowi dane wejściowe do części obliczeń klasycznych cech obrazowych. Końcowa fuzja wyselekcjonowanych cech z wykorzystaniem meta-regresji i metryki H zapewnia numeryczny opis składający się z pojedynczej wartości dla całego badania RM w każdym z ocenianych parametrów.

5.1.1 zbiór danych

Dane wykorzystane w tej pracy zostały zebrane w ramach projektu START "Wykorzystanie autologicznych mezenchymalnych komórek macierzystych w procesie regeneracji rekonstruowanego ścięgna Achillesa" finansowanego z konkursu STRATEGMED1 przez Narodowe Centrum Badań i Rozwoju. W ramach projektu do badanej grupy zakwalifikowano 60 pacjentów po całkowitym zerwaniu ścięgna Achillesa i 29 ochotników stanowiących grupę odniesienia.

Kryteria wykluczające z grupy badanej zostały zdefiniowane tak aby kwalifikowani pacjenci nie posiadali wcześniejszej historii urazów ścięgna Achillesa ani przewlekłych chorób, które mogłyby spowodować degenerację tkanki ścięgnistej. Uzważydlniono również warunki odbiegające od normy takie jak: otyłość ciąża, infekcje. Ostatecznie, w ramach grupy pacjentów znalazło się 49 mężczyzn i 11 kobiet w wieku pomiędzy 24 a 49 lat i ze średnią wieku 36 lat. Wszyscy pacjenci podpisali stosowne zgody wymagane do przeprowadzenia badań.

52 z pośród 60 pacjentów zerwało ścięgna podczas uprawiania sportu. Dokładniej były to: piłka nożna (19), squash/tenis (8), koszykówka (4), bieganie (3), podskoki (2), siatkówka (2) oraz badminton, box, piłka ręczna, fitness, cross-fit, skateboard, rugby i skakanka. Urazy w 31 przypadkach dotyczyły prawego ścięgna, a w 29 lewego. Średnia pozycja zerwania umiejscowiona była 57.11 mm ponad górnym konturem kości piętowej.

W okresie do tygodnia po urazie pacjenci przeszli zabieg rekonstrukcji wykonany metodą szycia na otwartym ścięgnie (zob. 3.1.4). Następnie rozpoczęto rehabilitacje trwającą 12 miesięcy monitorowaną z wykorzystaniem Rezonansu Magnetycznego, Ultrasonografii. Wykonano również badania biomechaniczne na początku i na końcu

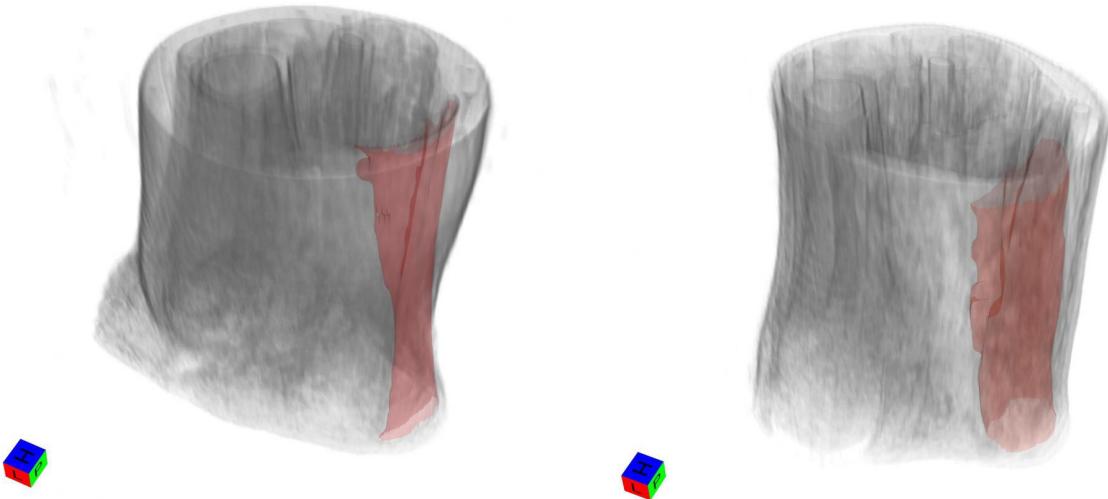
okresu rehabilitacji zgodnie zgodnie z protokołem opisany w sekcji 3.4.

Rezonans magnetyczny:

Przy akwizycji danych z wykorzystaniem RM posłużono się aparatem GE Signa HDxt 1.5T wyposażonym w cewkę Foot & Ankle dedykowaną do pomiarów w rejonie dolnej kończyny. Każde z badań RM było wykonane z użyciem 7 sekwencji i łącznie 10 modalności opisanych szczegółowo w sekcji 3.2 tj.:

1. T1 zależne
2. T2 zależne
3. PD
4. T2 mapping
5. T2 * GRE
6. T2 * GRE TE_MIN
7. 3D FSPGR
 - In Phase Ideal
 - Out Phase Ideal
 - Fat Ideal
 - Water Ideal

Z wykorzystaniem powyższych sekwencji na 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 po, odpowiednio w: 1, 3, 6, 9, 12, 20, 26, 40 i 52 tygodniu. W procedurze zbierania danych wystąpiły problemy wynikające z braku obecności poszczególnych osób na badaniach lub niespójnością danych. Dlatego finalnie skompletowany zbiór składa się z homogenicznych badań 27 zdrowych ochotników (270 trójwymiarowych skanów RM, w skład których wchodzi wspomniane 10 modalności) oraz badań 59 pacjentów (5900 trójwymiarowych skanów RM, w skład których wchodzi 10 modalności i 10 kroków czasowych). Przykładowe wizualizacja danych pacjenta i osoby zdrowej



Rysunek 5.2: Wizualizacja przykładowych danych RM w sekwencji PD dla zdrowego (po lewej) i gojącego się (po prawej) ścięgna Achillesa.

znajduje na Fig. 5.2. Obszar zajmowany przez tkankę ścięgnistą został oznaczony na czerwono. Należy zwrócić uwagę na różnicę w kształcie, która dla zdrowego przykładu jest zgodna z opisem anatomicznym ujętym w sekcji 3.1.1 natomiast dla przykładu chorego zawiera liczne deformacje.

Powyższe zbiory trójwymiarowe posłużyły do opracowania zbiorów dwuwymiarowych zawierających przekroje. W ten sposób udało się zwiększyć liczbę próbek wejściowych potrzebnych do szkolenia sieci neuronowych. Dane RM charakteryzują się wysoką anizotropią i ich rozdzielcość wynosi przykładowo $512 \times 512 \times Z$, gdzie $Z \in (10 : 50)$ (wyjątkiem jest sekwencja 3D FSPGR charakteryzująca się niższą anizotropią, jednak większą ilością artefaktów wynikających z szybkiego czasu akwizycji). Dlatego zbiory oparte o RM zostały stworzone w oparciu o przekroje osiowe 512×512 . W zależności od eksperymentów służących do doboru parametrów i komponentów metody wykorzystano:

- powiększony zbiór treningowy RM – zawiera przekroje osiowe ze wszystkich badań (z wyłączeniem pacjentów testowych) realizowanych wszystkimi sekwencjami. Dodatkowo liczba przekrojów została powiększona poprzez zastosowanie transformacji odbicia lustrzanego i w przypadku chorych przykładów 10 obrotów w zakresie od -10 do 10 stopni. Finalnie zbiór zawiera 234.502 przekrojów oznaczonych jako zdrowych i 277.208 oznaczonych jako chorych.

- ograniczony zbiór treningowy RM – zawiera przekroje osiowe ze zbioru oznaczonych 44 pacjentów jedynie w sekwencji T2 * GRE TE_MIN (18.863 próbek).
- zbiór testowy - zawiera przekroje pochodzące z danych zebranych dla 4 testowych pacjentów.

Ultrasonografia:

W przypadku danych Ultrasonografii stosowano się do identycznych odstępów czasowych, co w przypadku badań RM i akwizycji poddano tych samych pacjentów. Z przyczyn praktycznych zmniejszyła się jedynie grupa odniesienia, która w tym przypadku wyniosła 18-stu zdrowych ochotników. Badania zrealizowano z wykorzystaniem aparatu GE 3D high-resolution Voluson E8 Expert z liniową sondą (5–18 MHz). Jako dane wejściowe wykorzystano informacje z trybu B (zob. 3.3), których ostateczna liczba wyniosła 565 3D skanów.

Dane USG są zbliżone do izotropowych dlatego utworzono zbiory zarówno w oparciu o przekroje w płaszczyźnie osiowej jak i strzałkowej:

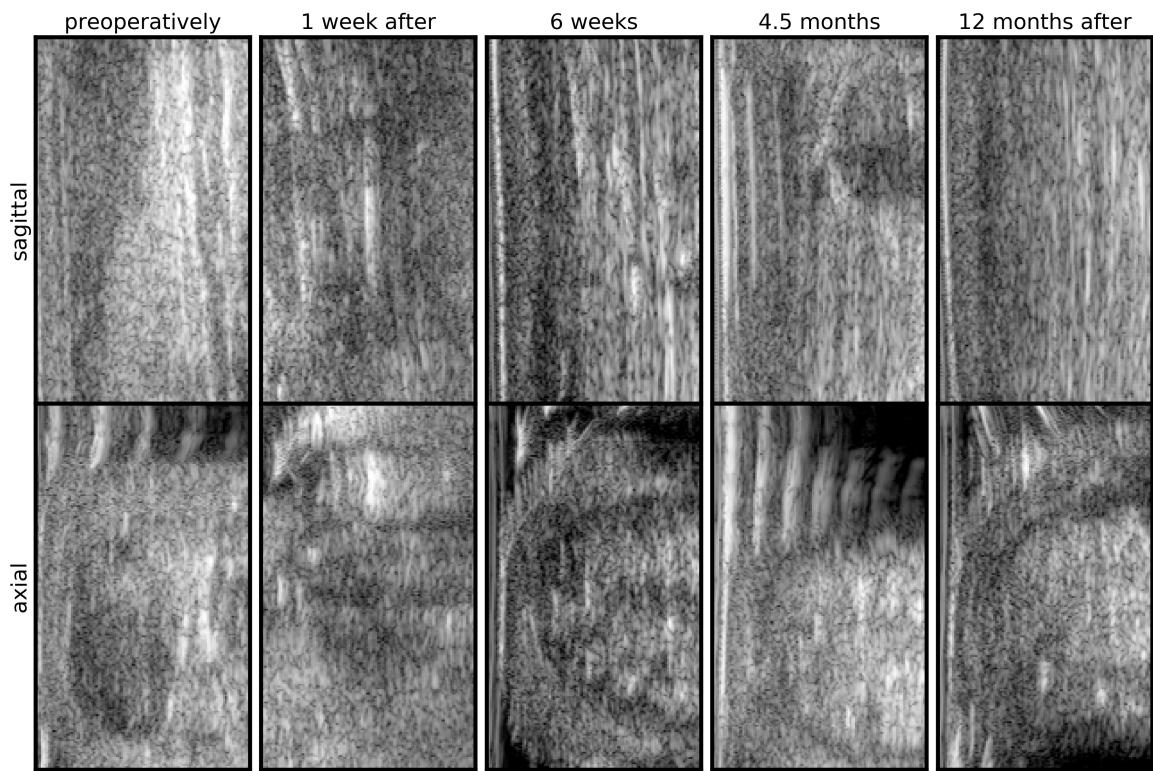
- zbiór treningowy USG (strzałkowy) – zawiera 253.639 2D przekrojów w płaszczyźnie strzałkowej, w tym 245.366 pochodzących od chorych 44 pacjentów oznaczonych przez radiologa i 8.273 pochodzących od zdrowych ochotników.
- zbiór treningowy USG (osiowy) – zawiera 467.548 2D przekrojów w płaszczyźnie osiowej, w tym 450.816 pochodzących od chorych 44 pacjentów oznaczonych przez radiologa i 16.732 pochodzących od zdrowych ochotników.

Wizualizacja przykładowych danych USG znajduje się na Rys. 5.3

Szczegółowa analiza załączonych obrazów może być wykonana jedynie przez specjalistę radiologa, jednak w ogólności można zaobserwować ułożenie włókien ścięgnistych na przekrojach strzałkowych i teksturę oraz tkanki otaczające ścięgno na przekrojach osiowych.

5.1.2 wzorzec odniesienia

Do stworzenia wzorca odniesienia, w ramach projektu START, została powołana specjalna grupa robocza składająca się z eksperta radiologa, specjalistów ortopedów i



Rysunek 5.3: Wizualizacja przykładowych danych USG w kolejnych tygodniach po zszyciu ścięgna w przekrojach osiowych i strzałkowych.

ekspertów od komputerowo wspomaganej diagnostyki. Celem grupy było opracowanie ilościowego opisu, który odzwierciedla elementy, na podstawie których radiolog podejmuje subiektywną opinię. W wyniku prac stworzono zestaw sześciu parametrów oceniany w skali 0–7 opisujący proces gojenia się ścięgna Achillesa widoczny w obrazach RM i Ultrasonografii:

1. Uszkodzenia śródścięgniste (SCT od ang. *Structural changes within the tendon*) – informuje o spójności włókien w obszarze ścięgna. Zawiera się w ROI. Ocena wykonywana zarówno w płaszczyźnie osiowej jak i strzałkowej. 0 – całkowity brak uszkodzeń, 1 – bardzo małe uszkodzenia. 2 – małe uszkodzenia, 3 – uszkodzenia średniej wielkości, 4 – dość duże uszkodzenia. 5 – duże uszkodzenia, 6 - bardzo duże uszkodzenia, 7 - ekstremalnie duże uszkodzenia.
2. Pogróżenie ścięgna (TT od ang. *Tendon thickening* – najgrubsze miejsce w kierunku strzałkowym na widoczne na przekroju osiowym. Zawiera się w ROI. 0 – całkowity brak pogrubienia (3mm-5mm), 1 – bardzo małe pogrubienie - 6mm, 2

- małe pogrubienie - 9mm, 3 – średnie pogrubienie- 12mm, 4 – dość duże pogrubienie -15mm, 5 duże pogrubienie - 18mm, 6 – bardzo duże pogrubienie - 21mm, 7 – ekstremalnie duże pogrubienie – 24mm.
3. Ostrość granic ścięgna/rozgraniczenie (STE od ang. *Sharpness of the tendon edges*) – informuje o jakości granicy między tkankami ścięgnistymi i otoczeniem, w szczególności, czy brzeg jest fraktalny. Oceniane na zewnątrz ścięgna w płaszczyźnie osiowej. 0 – bardzo duża (idelana) ostrość, 1 – duża ostrość, 2 – dość duża ostrość, 3 – średnia ostrość, 4 – mała ostrość, 5 – bardzo mała ostrość, 6 – minimalna ostrość, 7 – całkowicie nieostre.
 4. Obrzęk ścięgna (TE od ang. *Tendon edema*) – informuje o anomaliiach w gromadzeniu się płynów w obszarze ścięgna. Zawiera się w ROI i jest oceniany na przekrojach osiowych. 0 – całkowity brak obrzęku, 1 – bardzo mały obrzęk, 2 – mały obrzęk, 3 – obrzęk średniej wielkości, 4 – dość duży obrzęk, 5 – duży obrzęk. 6 – bardzo duży obrzęk, 7 - ekstremalnie duży obrzęk.
 5. Jednorodność ścięgna (TU od ang. *Tendon uniformity*) – informuje o prawidłowej teksturze poszczególnych przekrojów osiowych jak i podobieństwu sąsiednich, mierzonym w płaszczyźnie strzałkowej. Zawiera się w ROI. 0 – całkowity brak niejednorodności (jednorodne), 1 – bardzo małe niejednorodności, 2 – małe niejednorodności, 3 – niejednorodności średniej wielkości, 4 – dość duże niejednorodności, 5 – duże niejednorodności, 6 – bardzo duże niejednorodności, 7 – ekstremalnie duże niejednorodności.
 6. Obrzęk tkanek (TisE od ang. *Tissue edema*) – informuje o powiększonym przedziale powięziowym ścięgna. 0 – całkowity brak obrzęku, 1 – bardzo mały obrzęk, 2 – mały obrzęk, 3 – obrzęk średniej wielkości, 4 – dość duży obrzęk, 5 – duży obrzęk. 6 – bardzo duży obrzęk, 7 - ekstremalnie duży obrzęk.

Obrazowania 3D zostały opisane przez eksperta radiologa przy użyciu ankiety zawierającej powyższe parametry. W ramach projektu START udało się zgromadzić opisy dla badań RM i USG 48 pacjentów (480 badań). 4 z nich zostało losowo wydzielonych na początku eksperymentów jako pacjenci testowi. Szczegółowy opis eksperymentów znajduje się w kolejnej sekcji.

5.2 Eksperymenty

W tej sekcji zestawiono eksperymenty mające na celu dobór poszczególnych komponentów i parametrów finalnej metody. W pierwszej kolejności zostało opisane badanie dotyczące szkolenia sieci neuronowych celem rozróżnienia ścięgien zdrowych od tych po zerwaniu (chorych). Następnie zaprezentowano badanie dotyczące obliczania wektora cech na wyjściu sieci neuronowej i redukcji jego wymiarowości. W kolejnej części przedstawiono wybór sekwencji RM, które stanowiły dane wejściowe do proponowanego w tej pracy algorytmu. Finalnie, w dwóch ostatnich podsekcjach opisano dobór optymalnego zestawu mieszanki cech ekstrahowanych z użyciem sieci neuronowej i metod przetwarzania obrazów oraz fuzję cech z wykorzystaniem meta-regresji.

5.2.1 Rozróżnienie ścięgna zdrowego i po zerwaniu

W tym badaniu porównane zostały współczesne architektury sieci neuronowych w zadaniu klasyfikacji binarnej, rozróżniającej obrazy zdrowego od chorego ścięgna. W porównaniu uwzględniono topologie AlexNet (zob. 4.3.1), GoogLeNet InceptionV3 (zob. 4.3.2) i ResNet-18 (zob. 4.3.3). W celu doboru hiperparametrów i optymalizacji parametrów sieci zastosowano szkolenie z wykorzystaniem kroswalidacji z podziałem na 5 segmentów (zob. 4.2.1). W każdym z 5 cykli, 3 segmenty były wykorzystane do treningu, 1 do walidacji i 1 do finalnych testów. Do celów tego badania wykorzystano powiększony zbiór treningowy RM. Wyniki z testów zebrane w Tab. 5.1.

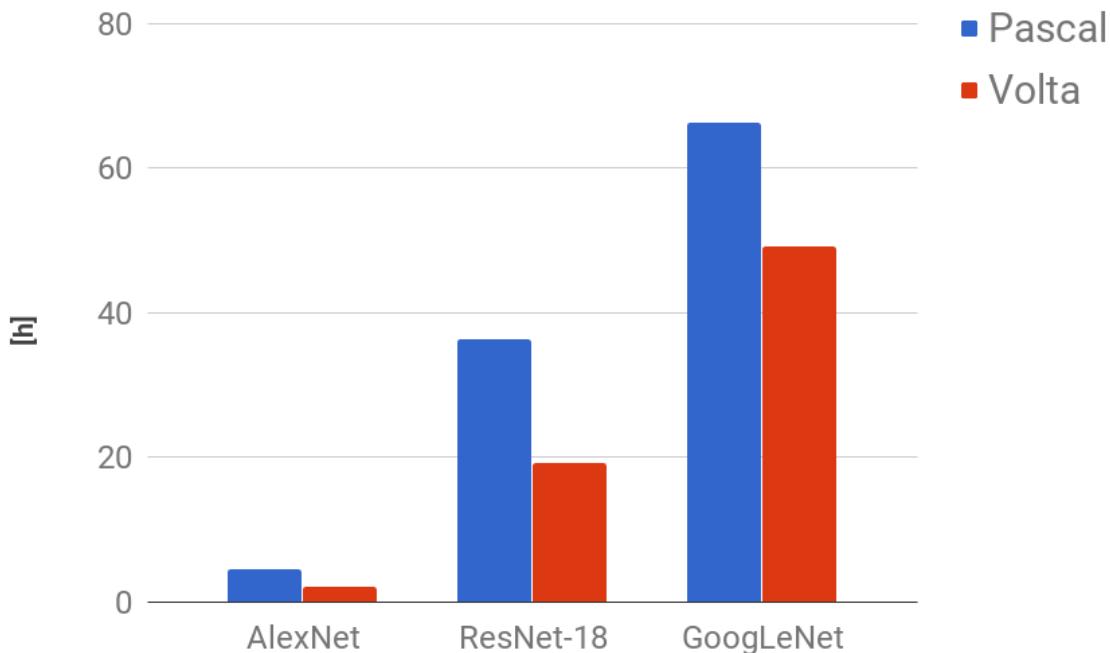
Tabela 5.1: Wyniki szkolenia sieci neuronowych z wykorzystaniem kroswalidacji z podziałem na 5 segmentów dla problemu klasyfikacji obrazów zdrowego i chorego ścięgna.

	Average	Min	Max	SD
AlexNet	99.19	99.15	99.24	0.04
ResNet-18	95.98	92.78	99.04	2.5
GoogleNet	99.83	99.68	99.91	0.1

Dla wszystkich architektur sieci, wyniki dokładności klasyfikacji najlepszych modeli osiągnęły wartość powyżej 99%. Niemniej jednak w przypadku architektury ResNet-18 najgorszy model klasyfikował z dokładnością 92.78%, co przełożyło się również na najgorszy średni wynik 95.98%. Fakt ten może wskazywać na tendencje do nadmiernego dopasowywania się architektury ResNet-18 wynikający z niedoboru liczby danych

wejściowych w odniesieniu do liczby optymalizowanych parametrów w warstwach konwolucyjnych.

Dla wszystkich trzech architektur porównano również czasy szkolenia się dla zadanego problemu. Na Rys. 5.4 zestawiono łączny czas 5 iteracji procesu kroswalidacji.



Rysunek 5.4: Porównanie czasów treningu architektur na dwóch architekturach GPU: Pascal i Volta.

Porównano dwie architektury GPU tj. starszą Pascal P100 i nowszą Volta V100, zaprezentowaną w 2017 roku, charakteryzującą się udoskonalonym osobnym układem optymalizowanym pod kątem operacji tensorowych realizowanych w warstwach splotu. Do szkolenia wykorzystano framework Caffe. Można zaobserwować, że wykorzystanie nowszej architektury GPU dało rezultaty około dwukrotnego przyspieszenia czasu szkolenia w przypadku AlexNet i ResNet oraz około 30% w przypadku GoogLeNet. AlexNet zawiera najmniej warstw konwolucyjnych i optymalizowaną implementację klasyfikatora. Szkolenie tej architektury zajęło mniej niż jedną godzinę na GPU Volta. Dla porównania, sieci z bardziej skomplikowaną topologią tj. ResNet-18 i GoogLeNet charakteryzują się odpowiednio czasami szkolenia w okolicach 20 i 48 godzin.

Biorąc pod uwagę zestawienie wyników dokładności klasyfikacji i czasu szkolenia najlepszą architekturą okazała się sieć AlexNet, dlatego została ona wybrana do reali-

zacji kolejnych eksperymentów. Należy również wspomnieć, że rezultaty wnioskowania mogą wciąż być poprawione. Dla przykładu, uzupełniające badanie dotyczące możliwości fuzji sieci AlexNet i ResNet-18 zostało opublikowane przez autora tej pracy w Acta of Bioengineering and Biomechanics [92]. Wyniki mogą być również poprawione w wyniku analizy lokalizacji przekrojów w przestrzeni trójwymiarowej, a w szczególności poprzez uwzględnienie, że częściej przekroje oddalone od miejsca zerwania bardziej przypominają tkanki zdrowe. W tym zakresie można np. zastosować metody logiki rozmytej lub miękkiej klasyfikacji (zob. []). Niemniej jednak, z uwagi na wysokie wyniki dokładności możliwe prostego podejścia w tej pracy autor nie decyduje się na dalsze usprawnienia kosztem komplikacji algorytmu. Ewentualne anomalie i wyniki odosobnione zostaną wyeliminowane poprzez zastosowanie średniej trymowanej w zaproponowanej metryce H (zob. 5.1).

5.2.2 Obliczanie redukcja wymiarowości wektora cech z sieci neuronowej

5.2.3 Wybór sekwencji rezonansu magnetycznego

5.2.4 Optymalna konkatenacja cech

5.2.5 Fuzja z wykorzystaniem meta-regresji

5.3 Ocena procesu gojenia

$$H = \alpha + \sum_{i=1}^3 \beta_i X_i + \sum_{i=1}^3 \gamma_i X_i^2 + \sum_{\substack{i,j=1 \\ i < j}}^3 \lambda_{i,j} X_i X_j \quad (5.2)$$

gdzie $X_i = TM(PC_n(x_1), PC_n(x_2), \dots, PC_n(x_n))_i$ to kolejne predyktory, TM to średnia trymowana z marginesami 2.5%, $PC_n(x_k)$ to n -ty czynnik główny otrzymany przy wnioskowaniu sieci dla przekroju osiowego x_k , gdzie k jest indeksem przekroju danego protokołu w trójwymiarowym badaniu RM.

Rozdział 6

Wyniki i walidacja

- 6.1 Ocena procesu gojenia z użyciem nowej metody**
- 6.2 Porównanie z wynikami z rezonansu magnetycznego**
- 6.3 Porównanie z wynikami ultrasonografii**
- 6.4 Porównanie z wynikami badań biomechanicznych**

Rozdział 7

Podsumowanie

Bibliografia

- [1] Adams msc software corp. Accessed: 2018-07-23.
- [2] Anybody tech. Accessed: 2018-07-23.
- [3] Bts systems. Accessed: 2018-07-23.
- [4] Deep learning: The next step in applied healthcare data. <https://insights.samsung.com/2016/07/12/deep-learning-the-next-step-in-applied-healthcare-data/>. Accessed: 2018-22-05.
- [5] Fizyka wobec wyzwań xxi wieku – obrazowanie medyczne. Accessed: 2018-07-30.
- [6] From challenges to impact of machine learning in clinical practice. <http://on-demand.gputechconf.com/gtc/2018/video/S8897/>. Accessed: 2018-22-05.
- [7] Ge healthcare. Accessed: 2018-06-08.
- [8] Kistler systems. Accessed: 2018-07-23.
- [9] Medical image net. <http://langlotzlab.stanford.edu/projects/medical-image-net/>. Accessed: 2018-22-05.
- [10] Meet iq. whole body imaging. under 2k. usd. Accessed: 2018-07-30.
- [11] Nvidia tensorrt. <https://developer.nvidia.com/tensorrt>. Accessed: 2018-22-05.
- [12] An overview of resnet and its variants. <https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035>. Accessed: 2018-22-05.

- [13] Project clara: Nvidia supercomputing platform redefines medical imaging. <https://blogs.nvidia.com/blog/2018/03/28/ai-healthcare-gtc/>. Accessed: 2018-07-30.
- [14] Queastions and answears in mri. Accessed: 2018-07-30.
- [15] Typy głowic stosowanych w usg. Accessed: 2018-06-08.
- [16] Us4us – ultrasound r&d in medicine, biology and industry. <http://us4us.eu/>. Accessed: 2018-07-30.
- [17] Vicon systems. Accessed: 2018-07-23.
- [18] Visual 3d - c motion. Accessed: 2018-07-23.
- [19] Salem Saleh Al-amri, Namdeo V. Kalyankar, Khamitkar S. D. Image segmentation by using threshold techniques. *CoRR*, abs/1005.4020, 2010.
- [20] Sanjeev Arora, Aditya Bhaskara, Rong Ge, Tengyu Ma. Provable bounds for learning some deep representations. *CoRR*, abs/1310.6343, 2013.
- [21] D. S. Babcock, H. Patriquin, M. LaFortune, M. Dauzat. Power doppler sonography: basic principles and clinical applications in children. *Pediatric Radiology*, 26(2):109–115, Feb 1996.
- [22] Tito Bassani, Elena Stucovitz, Zhihui Qian, Matteo Briguglio, Fabio Galbusera. Validation of the AnyBody full body musculoskeletal model in computing lumbar spine loads at l4l5 level. *Journal of Biomechanics*, 58:89–96, jun 2017.
- [23] Robert M. Bell, Yehuda Koren. Lessons from the netflix prize challenge. *SIGKDD Explor. Newsl.*, 9(2):75–79, dec 2007.
- [24] Richard Ernest Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- [25] P. K. Beredjiklian. Biologic aspects of flexor tendon laceration and repair. *J Bone Joint Surg Am*, 85-A(3):539–550, Mar 2003.
- [26] F. Bloch. Nuclear induction. *Physical Review*, 70(7-8):460–474, oct 1946.
- [27] A. Bochenek, M. Reicher. *Anatomia człowieka*. Anatomia człowieka. PZWL, 2016.

- [28] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001.
- [29] G. Breit, I. I. Rabi. On the interpretation of present values of nuclear moments. *Physical Review*, 46(3):230–231, aug 1934.
- [30] Carolina Medical Center. Start - wykorzystanie autologicznych mezenchymalnych komórek macierzystych w procesie regeneracji rekonstruowanego ścięgna achillesa, 2014.
- [31] Chun-houh Chen, Wolfgang Hrdle, Antony Unwin, Chun-houh Chen, Wolfgang Hrdle, Antony Unwin. *Handbook of Data Visualization (Springer Handbooks of Computational Statistics)*. Springer-Verlag TELOS, Santa Clara, CA, USA, wydanie 1, 2008.
- [32] Ilkay Zihni Chirali. 14 - sports injuries. Ilkay Zihni Chirali, redaktor, *Traditional Chinese Medicine Cupping Therapy (Third Edition)*, strony 211 – 230. Churchill Livingstone, wydanie third edition, 2014.
- [33] Junghwan Cho, Eunmi Lee, Hyunkwang Lee, Bob Liu, Xinhua Li, Shahein Tajmir, Dushyant Sahani, Synho Do. Machine learning powered automatic organ classification for patient specific organ dose estimation. *Society for Imaging Informatics in Medicine*, wolumen 2017, 2017.
- [34] Anna Choromanska, Mikael Henaff, Michaël Mathieu, Gérard Ben Arous, Yann LeCun. The loss surface of multilayer networks. *CoRR*, abs/1412.0233, 2014.
- [35] Eric M. Christiansen, Samuel J. Yang, D. Michael Ando, Ashkan Javaherian, Gaia Skibinski, Scott Lipnick, Elliot Mount, Alison O’Neil, Kevan Shah, Alicia K. Lee, Piyush Goyal, William Fedus, Ryan Poplin, Andre Esteva, Marc Berndl, Lee L. Rubin, Philip Nelson, Steven Finkbeiner. In silico labeling: Predicting fluorescent labels in unlabeled images. *Cell*, 173(3):792–803.e19, apr 2018.
- [36] Ryan Cunningham, Peter Harding, Ian Loram. Deep residual networks for quantification of muscle fiber orientation and curvature from ultrasound images. María Valdés Hernández, Víctor González-Castro, redaktorzy, *Medical Image Understanding and Analysis*, strony 63–73, Cham, 2017. Springer International Publishing.

- [37] Valentino Dardanoni, Giuseppe De Luca, Salvatore Modica, Franco Peracchi. Bayesian Model Averaging for Generalized Linear Models with Missing Covariates. EIEF Working Papers Series 1311, Einaudi Institute for Economics and Finance (EIEF), dec 2013.
- [38] Scott L. Delp, Frank C. Anderson, Allison S. Arnold, Peter Loan, Ayman Habib, Chand T. John, Eran Guendelman, Darryl G. Thelen. OpenSim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering*, 54(11):1940–1950, nov 2007.
- [39] S.L. Delp, J.P. Loan, M.G. Hoy, F.E. Zajac, E.L. Topp, J.M. Rosen. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Transactions on Biomedical Engineering*, 37(8):757–767, 1990.
- [40] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. *CVPR09*, 2009.
- [41] Denitsa Docheva, Sebastian A. Müller, Martin Majewski, Christopher H. Evans. Biologics for tendon repair. *Advanced Drug Delivery Reviews*, 84:222–239, apr 2015.
- [42] Mahmut Nedim Doral, Mahbub Alam, Murat Bozkurt, Egemen Turhan, Ozgür Ahmet Atay, Gürhan Dönmez, Nicola Maffulli. Functional anatomy of the achilles tendon. *Knee Surgery, Sports Traumatology, Arthroscopy*, 18(5):638–643, feb 2010.
- [43] Qi Dou, Hao Chen, Lequan Yu, Jing Qin, Pheng-Ann Heng. Multilevel contextual 3-d CNNs for false positive reduction in pulmonary nodule detection. *IEEE Transactions on Biomedical Engineering*, 64(7):1558–1567, jul 2017.
- [44] K. J. Dria, J. R. Sachleben, P. G. Hatcher. Solid-state carbon-13 nuclear magnetic resonance of humic acids at high magnetic field strengths. *J. Environ. Qual.*, 31(2):393–401, 2002.
- [45] John Duchi, Elad Hazan, Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.*, 12:2121–2159, jul 2011.

- [46] Bradley J. Erickson, Panagiotis Korfiatis, Zeynettin Akkus, Timothy Kline, Kenneth Philbrick. Toolkits and libraries for deep learning. *Journal of Digital Imaging*, 30(4):400–405, Aug 2017.
- [47] Bradley J. Erickson, Panagiotis Korfiatis, Timothy L. Kline, Zeynettin Akkus, Kenneth Philbrick, Alexander D. Weston. Deep learning in radiology: Does one size fit all? *Journal of the American College of Radiology*, 15(3):521–526, mar 2018.
- [48] Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, jan 2017.
- [49] R. Frisch, O. Stern. Über die magnetische ablenkung von wasserstoffmolekülen und das magnetische moment des protons. i. *Zeitschrift für Physik*, 85(1):4–16, Jan 1933.
- [50] Matthew F. Glasser, Timothy S. Coalson, Emma C. Robinson, Carl D. Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F. Beckmann, Mark Jenkinson, Stephen M. Smith, David C. Van Essen. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, jul 2016.
- [51] G. T. Gomes, S. Van Cauter, M. De Beule, L. Vigneron, C. Pattyn, E. A. Aude-naert. *Patient-Specific Modelling in Orthopedics: From Image to Surgery*, strony 109–129. Springer Netherlands, Dordrecht, 2013.
- [52] Ian Goodfellow, Yoshua Bengio, Aaron Courville. *Deep Learning*. MIT Press, 2016.
- [53] Varun Gulshan, Lily Peng, Marc Coram, Martin C. Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, Ramasamy Kim, Rajiv Raman, Philip C. Nelson, Jessica L. Mega, Dale R. Webster. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 316(22):2402, dec 2016.
- [54] Samuel R. Hamner, Ajay Seth, Scott L. Delp. Muscle contributions to propulsion and support during running. *Journal of Biomechanics*, 43(14):2709–2716, oct 2010.

- [55] R. Haralick, K. Shanmugam, I. Dinstein. Texture features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 1973.
- [56] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.
- [57] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Identity mappings in deep residual networks. *CoRR*, abs/1603.05027, 2016.
- [58] Omar W. Heyward, Lucas M. Rabello, Lucas van der Woude, Inge van den Akker-Scheek, Alli Gokeler, Henk van der Worp, Johannes Zwerver. The effect of load on achilles tendon structure in novice runners. *Journal of Science and Medicine in Sport*, 21(7):661–665, jul 2018.
- [59] A. V. Hill. The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society B: Biological Sciences*, 126(843):136–195, oct 1938.
- [60] F.S. Hill, Stephen M. Kelly. *Computer Graphics using OpenGL*. Prentice Hall, 2006.
- [61] Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *CoRR*, abs/1207.0580, 2012.
- [62] Jacques, Pierre Curie. Development, via compression, of electric polarization in hemihedral crystals with inclined faces. *Bulletin de la Societe de Minerologie de France*, (3):90–93, 1880.
- [63] Chand T. John, Frank C. Anderson, Jill S. Higginson, Scott L. Delp. Stabilisation of walking by intrinsic muscle properties revealed in a three-dimensional muscle-driven simulation. *Computer Methods in Biomechanics and Biomedical Engineering*, 16(4):451–462, apr 2013.
- [64] C. Ju, A. Bibaut, M. J. van der Laan. The Relative Performance of Ensemble Methods with Deep Convolutional Neural Networks for Image Classification. *ArXiv e-prints*, apr 2017.
- [65] Diederik P. Kingma, Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.

- [66] Patrick Koch, Brett Wujek, Oleg Golovidov, Steven Gardner. Automated hyper-parameter tuning for effective machine learning. 2017.
- [67] T Koivunen-Niemelä, K Parkkola. Anatomy of the achilles tendon (tendo calcaneus) with respect to tendon thickness measurements. *Surgical and Radiologic Anatomy*, 17(3):263–268, sep 1995.
- [68] Stanisław Kowalik. Zastosowanie teorii zbiorów przybliżonych do podejmowania decyzji. *Zarządzanie przedsiębiorstwem w erze postindustrialnej – ekonomia, prawo, kultura, etyka*, strony 39–44, Kazimierz Dolny, Polska, 2003.
- [69] Alex Krizhevsky. Learning multiple layers of features from tiny images. Raport instytutowy, 2009. Accessed: 2018-05-21.
- [70] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, strony 1097–1105, USA, 2012. Curran Associates Inc.
- [71] Y. Kwon, M. Rhu. A case for memory-centric hpc system architecture for training deep neural networks. *IEEE Computer Architecture Letters*, 17(2):134–138, July 2018.
- [72] P. C. LAUTERBUR. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*, 242(5394):190–191, mar 1973.
- [73] K. N. Leach. A survey paper on independent component analysis. *Proceedings of the Thirty-Fourth Southeastern Symposium on System Theory (Cat. No.02EX540)*, strony 239–242, 2002.
- [74] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [75] Yann LeCun, Bernhard E. Boser, John S. Denker, Donnie Henderson, R. E. Howard, Wayne E. Hubbard, Lawrence D. Jackel. Handwritten digit recognition with a back-propagation network. D. S. Touretzky, redaktor, *Advances in Neural Information Processing Systems 2*, strony 396–404. Morgan-Kaufmann, 1990.
- [76] D.R. Lide. *CRC Handbook of Chemistry and Physics, 87th Edition*. Taylor & Francis, 2006.

- [77] Tony W. Lin, Luis Cardenas, Louis J. Soslowsky. Biomechanics of tendon injury and repair. *Journal of Biomechanics*, 37(6):865–877, jun 2004.
- [78] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, C. Lawrence Zitnick. Microsoft coco: Common objects in context. David Fleet, Tomas Pajdla, Bernt Schiele, Tinne Tuytelaars, redaktorzy, *Computer Vision – ECCV 2014*, strony 740–755, Cham, 2014. Springer International Publishing.
- [79] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, dec 2017.
- [80] Mengchen Liu, Jiaxin Shi, Zhen Li, Chongxuan Li, Jun Zhu, Shixia Liu. Towards better analysis of deep convolutional neural networks. *CoRR*, abs/1604.07043, 2016.
- [81] Nicola Maffulli, Per Renström, Wayne B. Leadbetter, redaktorzy. *Tendon Injuries*. Springer-Verlag, 2005.
- [82] R. Makarewicz. *Kształtowanie pola akustycznego z uwzględnieniem aspektu akustyki środowiska*. Akustyka - Uniwersytet im. Adama Mickiewicza w Poznaniu. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza, 1978.
- [83] P Mansfield. Multi-planar image formation using nmr spin echoes. *Journal of Physics C: Solid State Physics*, 10(3):L55, 1977.
- [84] J. Ma quirriain. Achilles tendon rupture: avoiding tendon lengthening during surgical repair and rehabilitation. *Yale J Biol Med*, 84(3):289–300, Sep 2011.
- [85] Gary Marcus. Deep learning: A critical appraisal. *CoRR*, abs/1801.00631, 2018.
- [86] Stefano Markidis, Steven Wei Der Chien, Erwin Laure, Ivy Bo Peng, Jeffrey S. Vetter. NVIDIA tensor core programmability, performance & precision. *CoRR*, abs/1803.04014, 2018.
- [87] Miriam Martínez, Luis Enrique Sucar, Hector Gabriel Acosta, Nicandro Cruz. Bayesian model combination and its application to cervical cancer detection.

- Jaime Simão Sichman, Helder Coelho, Solange Oliveira Rezende, redaktorzy, *Advances in Artificial Intelligence - IBERAMIA-SBIA 2006*, strony 622–631, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
- [88] Seymour Papert Marvin Lee Minsky. *Perceptrons: An Introduction to Computational Geometry*. Mit Press, 1969.
- [89] Warren S. McCulloch, Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, Dec 1943.
- [90] Ümit V. Çatalyürek, Cevdet Aykanat, Bora Uçar. On two-dimensional sparse matrix partitioning: Models, methods, and a recipe. *SIAM Journal on Scientific Computing*, 32(2):656–683, 2010.
- [91] Azim Celik Muhammed Elmaoglu. *Rezonans Magnetyczny: Podstawy fizyczne, obrazowanie, ułożenie pacjenta, protokoły*. Springer, 2015.
- [92] Kapinski N., Zielinski J., Borucki B., Trzcinski T., Ciszkowska Łyson B., Zdanowicz U., Smigielski R., Nowinski K. Monitoring of the achilles tendon healing process: can artificial intelligence be helpful? *ACTA OF BIOENGINEERING AND BIOMECHANICS*, 21(1), 2019.
- [93] Katarina Nilsson-Helander, Roland Thomeé, Karin Grävare-Silbernagel, Pia Thomeé, Eva Faxén, Bengt I. Eriksson, Jon Karlsson. The achilles tendon total rupture score (ATRS). *The American Journal of Sports Medicine*, 35(3):421–426, mar 2007.
- [94] Jędrzej Nowosielski, Jakub Zielinski, Bartosz Borucki, Krzysztof Nowinski. Multidimensional haralick's feature space analysis for assessment of the achilles tendon in mr imaging. *International Journal of Computer Assisted Radiology and Surgery*, 12(Supplement 1):218—220, 2017.
- [95] Karl F. Orishimo, Gideon Burstein, Michael J. Mullaney, Ian J. Krementic, Marcus Nesse, Malachy P. McHugh, Steven J. Lee. Effect of knee flexion angle on achilles tendon force and ankle joint plantarflexion moment during passive dorsiflexion. *The Journal of Foot and Ankle Surgery*, 47(1):34–39, jan 2008.
- [96] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.

- [97] M. S. Raikin. Epidemiology of achilles tendon rupture in the us, 2014.
- [98] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, A. Y. Ng. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. *ArXiv e-prints*, jul 2017.
- [99] Piotr Regulski, Jakub Zieliński, Bartosz Borucki, Krzysztof Nowiński. Comparison of noise reducing t2map reconstruction methods in mri imaging of achilles tendon. *International Journal of Computer Assisted Radiology and Surgery*, 12(Supplement 1):15–16, Jun 2017.
- [100] Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.
- [101] Frank Rosenblatt. The perceptron: A perceiving and recognizing automaton. Raport instytutowy, 1957.
- [102] A. Rosińska, M. Syrek. Rehabilitacja ścięgna achillesa - kopletny protokół usprawniania po trójpęczkowej rekonstrukcji ścięgna achillesa., 2017.
- [103] David E. Rumelhart, Geoffrey E. Hinton, Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, oct 1986.
- [104] N. Russell, T. Brendan Murphy, A. E Raftery. Bayesian model averaging in model-based clustering and density estimation. *ArXiv e-prints*, jun 2015.
- [105] Hijazi S., Kumar R., Rowen C. Using convolutional neural networks for image recognition. Raport instytutowy, Cadence, 2015. Accessed: 2018-05-21.
- [106] Saman Sarraf, Ghassem Tofghi. Deep learning-based pipeline to recognize alzheimer disease using fMRI data. *2016 Future Technologies Conference (FTC)*. IEEE, dec 2016.
- [107] W. Sawicki. *Histologia*. Wydawnictwo Lekarskie PZWL, 2008.
- [108] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85 – 117, 2015.
- [109] Ramprasaath R. Selvaraju, Abhishek Das, Ramakrishna Vedantam, Michael Cogswell, Devi Parikh, Dhruv Batra. Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization. *CoRR*, abs/1610.02391, 2016.

- [110] Dipanjan Sengupta, Yida Wang, Narayanan Sundaram, Theodore L. Willke. High-performance incremental svm learning on intel xeon phi processors. Julian M. Kunkel, Rio Yokota, Pavan Balaji, David Keyes, redaktorzy, *High Performance Computing*, strony 120–138, Cham, 2017. Springer International Publishing.
- [111] Qi Shao, Daniel N. Bassett, Kurt Manal, Thomas S. Buchanan. An EMG-driven model to estimate muscle forces and joint moments in stroke patients. *Computers in Biology and Medicine*, 39(12):1083–1088, dec 2009.
- [112] P. Sharma, N. Maffulli. Biology of tendon injury: healing, modeling and remodeling. *J Musculoskeletal Neuronal Interact*, 6(2):181–190, 2006.
- [113] Karen Simonyan, Andrea Vedaldi, Andrew Zisserman. Deep inside convolutional networks: Visualising image classification models and saliency maps. *CoRR*, abs/1312.6034, 2013.
- [114] Karen Simonyan, Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.
- [115] Leon Sixt, Benjamin Wild, Tim Landgraf. Rendergan: Generating realistic labelled data. *CoRR*, abs/1611.01331, 2016.
- [116] Fengguang Song, Jack Dongarra. Scaling up matrix computations on shared-memory manycore systems with 1000 cpu cores. *Proceedings of the 28th ACM International Conference on Supercomputing*, ICS ’14, strony 333–342, New York, NY, USA, 2014. ACM.
- [117] Katherine M. Steele, Ajay Seth, Jennifer L. Hicks, Michael S. Schwartz, Scott L. Delp. Muscle contributions to support and progression during single-limb stance in crouch gait. *Journal of Biomechanics*, 43(11):2099–2105, aug 2010.
- [118] Katarzyna Strojek, Zuzanna Piekorz, Urszula Kaźmierczak, Agnieszka Strączyńska, Walery Zukow. Etiologia i patomechanizm uszkodzenia Ścięgna achillesa = the etiology and patomechanism damage to the achilles tendon. 2016.
- [119] R. M. Summers, P. M. Joseph, H. L. Kundel. Sodium nuclear magnetic resonance imaging of neuroblastoma in the nude mouse. *Invest Radiol*, 26(3):233–241, Mar 1991.

- [120] Yi Sun, Ole Brauckmann, Donald R. Nixdorf, Arno Kentgens, Michael Garwood, Djaudat Idiyatullin, Arend Heerschap. Imaging human teeth by phosphorus magnetic resonance with nuclear overhauser enhancement. *Scientific Reports*, 6(1), aug 2016.
- [121] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke. Inception-v4, inception-resnet and the impact of residual connections on learning. *CoRR*, abs/1602.07261, 2016.
- [122] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott E. Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014.
- [123] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna. Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567, 2015.
- [124] K. Tahara, Z. W. Luo, T. Odashima, M. Onishi, S. Hosoe, A. Kato. Dynamic control and simulation of human musculo-skeletal model. *SICE 2004 Annual Conference*, wolumen 3, strony 2395–2398 vol. 3, Aug 2004.
- [125] Nima Tajbakhsh, Jae Y. Shin, Suryakanth R. Gurudu, R. Todd Hurst, Christopher B. Kendall, Michael B. Gotway, Jianming Liang. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35(5):1299–1312, may 2016.
- [126] Gavin Taylor, Ryan Burmeister, Zheng Xu, Bharat Singh, Ankit Patel, Tom Goldstein. Training neural networks without gradients: A scalable ADMM approach. *CoRR*, abs/1605.02026, 2016.
- [127] Joshua B. Tenenbaum, Vin de Silva, John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319, 2000.
- [128] Laurens van der Maaten, Geoffrey Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.
- [129] H. T. van Schie, E. M. Bakker. Structure-related echoes in ultrasonographic images of equine superficial digital flexor tendons. *Am. J. Vet. Res.*, 61(2):202–209, Feb 2000.

- [130] H. T. van Schie, E. M. Bakker, A. M. Jonker, P. R. van Weeren. Computerized ultrasonographic tissue characterization of equine superficial digital flexor tendons by means of stability quantification of echo patterns in contiguous transverse ultrasonographic images. *Am. J. Vet. Res.*, 64(3):366–375, Mar 2003.
- [131] H. T. M. van Schie, R. J. de Vos, S. de Jonge, E. M. Bakker, M. P. Heijboer, J. A. N. Verhaar, J. L. Tol, H. Weinans. Ultrasonographic tissue characterisation of human achilles tendons: quantification of tendon structure through a novel non-invasive approach. *British Journal of Sports Medicine*, 44(16):1153–1159, aug 2009.
- [132] Juha Vesanto, Johan Himberg, Esa Alhoniemi, Juha Parhankangas. Self-organizing map in matlab: the som toolbox. In *Proceedings of the Matlab DSP Conference*, strony 35–40, 2000.
- [133] D. Wang, A. Khosla, R. Gargya, H. Irshad, A. H. Beck. Deep Learning for Identifying Metastatic Breast Cancer. *ArXiv e-prints*, jun 2016.
- [134] Michael B. W. Wolfe, Susan R. Goldman. Use of latent semantic analysis for predicting psychological phenomena: Two issues and proposed solutions. *Behavior Research Methods, Instruments, & Computers*, 35(1):22–31, Feb 2003.
- [135] David H. Wolpert. Stacked generalization. *Neural Networks*, 5:241–259, 1992.
- [136] Robert H. Wurtz. Recounting the impact of hubel and wiesel. *The Journal of Physiology*, 587(12):2817–2823, jun 2009.
- [137] Saining Xie, Ross B. Girshick, Piotr Dollár, Zhuowen Tu, Kaiming He. Aggregated residual transformations for deep neural networks. *CoRR*, abs/1611.05431, 2016.
- [138] Guang Yang, Benjamin B. Rothrauff, Rocky S. Tuan. Tendon and ligament regeneration and repair: Clinical relevance and developmental paradigm. *Birth Defects Research Part C: Embryo Today: Reviews*, 99(3):203–222, sep 2013.
- [139] Matthew D. Zeiler, Rob Fergus. Visualizing and understanding convolutional networks. *CoRR*, abs/1311.2901, 2013.
- [140] Wei Zhang, Kunio Doi, Maryellen L. Giger, Yuzheng Wu, Robert M. Nishikawa, Robert A. Schmidt. Computerized detection of clustered microcalcifications in

digital mammograms using a shift-invariant artificial neural network. *Medical Physics*, 21(4):517–524, apr 1994.

Dodatek A

**AchillesDL: System komputerowego
wspomagania oceny gojenia ściegien
i więzadeł**