







Zastosowania sztucznej inteligencji (AI) w medycynie

Applications of Artificial Intelligence (AI) in medicine

Kacper Niewęglowski^{1,A-F}, Natalia Wilczek^{1,B-D}, Barbara Madoń^{1,B-D}, Jan Palmi^{1,B-D},
Martyna Wasyluk^{1,B-D}

¹ Uniwersytet Medyczny w Lublinie

A – Koncepcja i projekt badania, B – Gromadzenie i/lub zestawianie danych, C – Analiza i interpretacja danych, D – Napisanie artykułu, E – Krytyczne zrecenzowanie artykułu, F – Zatwierdzenie ostatecznej wersji artykułu

Niewęglowski K, Wilczek N, Madoń B, Palmi J, Wasyluk M. Zastosowania sztucznej inteligencji (AI) w medycynie. Med Og Nauk Zdr. doi: 10.26444/monz/142085

■ Streszczenie

Wprowadzenie i cel pracy. Problematyka sztucznej inteligencji jest stosunkowo nowym zagadnieniem w naukach medycznych. Regularnie pojawiają się publikacje dotyczące jej zastosowań w medycynie. Często dotyczą one wykorzystania algorytmów głębokiego uczenia, opartego na sieciach neuronowych, które są w stanie rozpoznać zmiany chorobowe widoczne na obrazie. Celem pracy jest omówienie możliwości wykorzystania sztucznej inteligencji w medycynie, szczególnie w radiologii i patomorfologii, oraz przedstawienie osiągniętych dzięki niej wyników.

Metody przeglądu. W marcu 2021 roku przeszukano bazę danych Medline (PubMed) oraz Google Scholar przy użyciu słów kluczowych: „artificial intelligence”, „deep learning”, „machine learning”, „digital pathology”, „convolutional neural network”. Wybrano prace opublikowane w języku angielskim, w latach 2015–2021.

Opis stanu wiedzy. Istnieje wiele doniesień o zastosowaniach sztucznej inteligencji w medycynie, głównie w dziedzinie radiologii i patomorfologii. Badania pokazują, że samouczące się algorytmy są w stanie z dokładnością zbliżoną do oceny przeprowadzonej przez lekarzy, a niekiedy nawet większą, wykryć zmiany chorobowe na zdjęciu rentgenowskim, tomografii komputerowej czy na zdjęciu preparatu mikroskopowego. W przedstawionych badaniach zauważalne są istotne korzyści wynikające z synergistycznego działania lekarzy i sztucznej inteligencji.

Podsumowanie. Wyniki uzyskiwane przez algorytmy oparte na sztucznej inteligencji świadczą o tym, że może ona usprawniać proces diagnozowania pacjentów, głównie dzięki uzupełnianiu wiedzy i doświadczenia lekarzy. Ważną kwestią jest również to, że korzystanie przez lekarzy z samouczących się algorytmów zmniejsza ryzyko popełnienia błędu ludzkiego, np. niezauważenia zmiany chorobowej widocznej na zdjęciu rentgenowskim.

Słowa kluczowe

sztuczna inteligencja, głębokie uczenie, uczenie maszynowe

■ Abstract

Introduction and objective. Artificial intelligence is a relatively new field of medical sciences. Its application in medicine is described in regularly appearing publications, which often focus on the use of deep learning algorithms based on neural networks that are capable of recognizing pathological changes in images. The aim of this study is to discuss possibilities of application of artificial intelligence in medicine, particularly radiology and pathomorphology, and to present the results achieved.

Review methods. Databases of Medline (PubMed) and Google Scholar were searched using keywords: ‘artificial intelligence’, ‘deep learning’, ‘machine learning’, ‘digital pathology’, and ‘convolutional neural network’. The search was undertaken in March 2021. Studies published in English during 2015–2021 were selected.

Brief description of the state of knowledge. There are many reports concerning the use of artificial intelligence in various fields of medicine, such as radiology and pathomorphology. Multiple research shows that self-learning algorithms are capable of finding pathologies in radiograms, computed tomography scans, or microscopic slides with accuracy equal to or even better than physicians. The study indicates significant advantages resulting from synergic cooperation of artificial intelligence and physicians.

Summary. Results achieved by artificial intelligence based algorithms provide evidence for improvement of patient diagnosis, predominantly by supplementation of physicians’ knowledge and experience. It is also an important fact that the use of AI decreases the risk of medical error, e.g. failure to recognize a pathological change visible on RTG.

Key words

machine learning, deep learning, artificial intelligence

WPROWADZENIE I CEL PRACY

Sztuczna inteligencja (z ang. *artificial intelligence*, AI) nie ma jednej, powszechnie przyjętej definicji, ale bywa definiowana dość ogólnie jako dziedzina wiedzy skupiająca się na badaniu komputerów rozwiązujących zadania, przy

Adres do korespondencji: Kacper Niewęglowski, Uniwersytet Medyczny w Lublinie
E-mail: kniew99@gmail.com

Nadesłano: 12.04.2021; zaakceptowano do publikacji: 9.09.2021; publikacja on-line: 24.09.2021

których wykonywaniu przez człowieka konieczne jest użycie jego inteligencji. Można zatem powiedzieć, że AI jest w zasadzie zespołem nauk, a nie dyscypliną samą w sobie. Sztuczna inteligencja wzorowana jest na naturalnych procesach poznawczych i działaniach, które zachodzą w ludzkim mózgu. Reasumując, AI to dziedzina badań, której celem jest zbudowanie systemu komputerowego działającego na wzór inteligencji człowieka. Stworzenie takiego systemu możliwe jest m.in. dzięki metodom uczenia maszynowego (z ang. *machine learning*). W systemie tym wiedza zapisywana jest w postaci modeli lub reguł obowiązujących dla danego zbioru danych, dzięki czemu można ją wykorzystać później, podczas analizowania nowych obiektów [1, 2]. Wszystkie samouczące się algorytmy są zaprogramowane podobnie do siebie, mianowicie na części bazy danych tworzony jest model klasyfikacyjny, na kolejnych, mniejszych częściach jest on testowany i weryfikowany. Dzięki temu algorytmy mają zdolność samodoskonalenia po ekspozycji na nowe dane. Tego typu algorytmy mają zastosowanie m.in. w rozpoznawaniu obrazów, co już teraz może być wykorzystywane w medycynie, np. do klasyfikowania zmian widocznych na tomografii komputerowej klatki piersiowej w przebiegu COVID-19 [3, 4]. Uczenie głębokie (z ang. *deep learning*) jest podkategorią uczenia maszynowego, która wymaga większych zbiorów danych, i w przypadku której samo uczenie trwa dłużej [5]. Wzorowane jest ono na pracy mózgu – w jego trakcie przetwarzanych jest równocześnie wiele danych. Matematycznym modelem biologicznego układu nerwowego są sieci neuronowe (z ang. *neural networks*). Przetwarzając dane, działają one na wzór podejmowania decyzji w neuronach, chociaż są ich znacznie uproszczonym modelem. Dane wejściowe przetwarzane są przez algorytm, a po zakończeniu tego procesu trafiają do kolejnych warstw algorytmów do momentu uzyskania wyniku [6].

Z AI wiąże się duże nadzieje na przyszłość, które, jeśli szybkość rozwoju tej dziedziny zostanie utrzymana, już niedługo mogą być codziennością w ochronie zdrowia. Sztuczna inteligencja może być używana m.in. przy obliczaniu ryzyka chorób sercowo-naczyniowych, ciągłym kontrolowaniu poziomu glukozy we krwi u pacjentów z cukrzycą czy analizowaniu obrazów USG [7, 8]. Celem pracy jest omówienie możliwości wykorzystania sztucznej inteligencji w medycynie, szczególnie w radiologii i patomorfologii, oraz przedstawienie osiągniętych dzięki niej wyników.

ZASTOSOWANIA SZTUCZNEJ INTELIGENCJI W RADIOLOGII

Diagnostyka zapalenia płuc

U każdego przyjętego do szpitala pacjenta z podejrzeniem zapalenia płuc zaleca się wykonanie badań obrazowych w celu poprawnego rozpoznania. Zazwyczaj jest to zdjęcie rentgenowskie (RTG), zaś w przypadkach wątpliwych wykonuje się tomografię komputerową (CT) klatki piersiowej [9]. Powoduje to jednak konieczność opisanego przez radiologa licznych badań obrazowych, co może opóźnić proces diagnostyczny. Pożądanym wydaje się zatem poszukiwanie rozwiązań, które umożliwiłyby odczytanie radiologów w rozpoznawaniu, często typowych i bardzo do siebie podobnych, zmian obecnych w zapaleniu płuc. Metody uczenia maszynowego wydają się odpowiednie do tego celu. Zaletą korzystania z oceny przeprowadzonej komputerowo jest powtarzalność

otrzymywanych wyników, co ułatwia również porównywanie wyników uzyskanych w różnych ośrodkach [10]. Sztuczna inteligencja może być stworzona w ten sposób, aby stawiła wyłącznie katagoryczne rozpoznania, takie jak obecność lub brak zapalenia płuc [11].

Tworzone obecnie systemy komputerowej analizy otrzymanych badań obrazowych są bardzo zbliżone pod względem czułości i specyficzności do ocen przeprowadzanych przez doświadczonych radiologów. W zależności od metody oraz rodzaju ocenianego badania obrazowego przy zapaleniach płuc niezwiązanych z COVID-19 współczynnik AUC (ang. *area under the curve*) wynosił od 0,81 do 0,923, a nawet 0,97 [10, 12, 13].

W badaniu przeprowadzonym przez lekarzy z Department of Radiology, School of Medicine, Stanford University 33 badanych zostało zakwalifikowanych przez radiologów z 15-letnim doświadczeniem jako pacjenci z niespecyficznym lub zwykłym śródmiąższowym zapaleniem płuc. Następnie ci sami pacjenci zostali zakwalifikowani do tych grup przez algorytm oraz dwóch lekarzy, którzy odbyli jednoroczne szkolenie z tego zakresu. AUC uzyskany przez algorytm wyniósł 0,81, przy czym błędy w diagnozie przeprowadzonej przez szkolących się lekarzy i algorytm dotyczyły różnych pacjentów. Może to wskazywać na dodatkowe korzyści w postaci zmniejszenia ryzyka popełnienia błędu ludzkiego, wynikające z synergii działania lekarzy i algorytmów [12]. Jeden z opracowanych systemów, oparty na uczeniu maszynowym, korzystał z bazy zdjęć RTG klatki piersiowej dzieci w wieku od 1 do 5 lat. Baza danych zawierała 3883 zdjęć, które zostały opisane jako obecne zapalenie płuc, oraz 1349 opisanych jako zdjęcia bez zmian. Zastosowany system uzyskał AUC równy 0,97, a ponadto metoda była od 5,4 do 6,7 razy szybsza niż porównane w badaniu systemy [13]. Biorąc pod uwagę fakt, że zapalenie płuc jest przyczyną nawet 16% zgonów dzieci na całym świecie, poprawne rozpoznanie, ułatwione dzięki zastosowaniu algorytmów sztucznej inteligencji, może pomóc ocalić wiele ludzkich istnień [14]. Może to być szczególnie przydatne w krajach afrykańskich, gdzie pracuje bardzo niewielu lekarzy (ich liczba wynosi nawet zaledwie 0,14 na 10 tys. Mieszkańców), a zatem możliwość konsultacji z radiologami jest utrudniona [15].

Dodatkową zaletą wykorzystywania metod komputerowych i sztucznej inteligencji jest możliwość przygotowywania oprogramowania w taki sposób, by podczas naświetlania pacjentów z niedoborami odporności stosowane były niższe dawki. Przy dwudziestokrotnym zmniejszeniu dawki naświetlania podczas tomografii komputerowej klatki piersiowej (z $1,43 \pm 0,54$ mSv przy dawce standardowej do $0,07 \pm 0,03$ mSv przy dawce zredukowanej) i wykorzystaniu algorytmu klasyfikującego otrzymany obraz jako obecną lub nieobecną konsolidację płuc, wartość AUC spadła z 0,923 przy dawce standardowej do 0,881 przy dawce zredukowanej, a czułość i specyficzność kolejno z 82,9% i 93,8% do odpowiednio 71,0% i 93,3% [16]. Mimo widocznego spadku wartości parametrów ukazuje to możliwość stosowania niższych dawek promieniowania u pacjentów z niewydolnym układem immunologicznym po wcześniejszym odpowiednim dostosowaniu do tego algorytmów.

Diagnostyka zapalenia płuc w przebiegu COVID-19

Metody wykorzystujące sztuczna inteligencję są obecnie stosowane w badaniach związanych z pandemią COVID-19. Jedno z badań wskazuje na możliwość przewidywania przebiegu

klinicznego zapalenia płuc podczas infekcji SARS-CoV-2. Otrzymane obrazy tomografii komputerowej klatki piersiowej pacjentów podczas ich przyjęcia do szpitala zostały przeanalizowane przez algorytm uczenia maszynowego. Spośród wszystkich pacjentów wyodrębnił on tych, którzy biorąc pod uwagę obecne przy przyjęciu zmiany, zapalenie płuc będą przechodzili w stopniu umiarkowanym lub ciężkim, osiągając przy tym AUC wynoszący 0,98 i 0,97 [17].

Sztuczna inteligencja znalazła zastosowanie w wydzieleniu zapaleń płuc spowodowanych SARS-CoV-2 z tych wywołanych innymi patogenami. W badaniu wykorzystano 73 obrazy tomografii komputerowej pacjentów z potwierdzonym COVID-19 oraz 27 pacjentów z zapaleniem płuc spowodowanym innymi czynnikami. Algorytm przy podziale pacjentów na te dwie grupy osiągnął czułość na poziomie 88,62% oraz swoistość na poziomie 100% przy AUC wynoszącym 0,99. Wykorzystanie tego algorytmu może pomóc lekarzom w odróżnianiu zapalenia płuc w przebiegu COVID-19 od zapalenia płuc spowodowanego innymi czynnikami [18]. Badania pokazały, że gdy korzystano z pomocy sztucznej inteligencji, czułość i swoistość diagnozowania COVID-19 przez lekarzy wzrosły odpowiednio o 16,5% i 11,6% [19].

Wykorzystanie algorytmów sztucznej inteligencji w diagnostyce COVID-19 ma jednak pewne ograniczenia. Obrazy z tomografii komputerowej płuc u pacjentów z COVID-19 często są gorszej jakości, a zatem mają mniejszą wartość diagnostyczną, ze względu na czynniki środowiskowe czy ruchy płuc. Powoduje to spadek skuteczności wspomaganego komputerowo diagnozowania tych pacjentów. W celu przeciwdziałania tym czynnikom opracowano metodę rekonstrukcji obrazów TK za pomocą sieci neuronowej SRGAN, a następnie poddano je analizie dzięki użyciu sieci neuronowej VGG16. Umożliwiło to zwiększenie precyzji do 97,87%, co pozytywnie wpłynęło na diagnozowanie COVID-19 przez algorytm [20].

Diagnostyka rozedmy płuc

Rozedma płuc to nieodwracalne poszerzenie przestrzeni powietrznych położonych dystalnie do oskrzelików końcowych, w wyniku destrukcji ścian pęcherzyków. Dwie najważniejsze przyczyny rozedmy płuc stanowią palenie tytoniu oraz niedobór α -1-antytrypsyny. Obrazowanie odgrywa istotną rolę w diagnostyce tej choroby, ze względu na różnice w rozmieszczeniu i tym samym leczeniu zmian rozedmowych w zależności od źródła ich występowania [21]. Oparty na sztucznej inteligencji, w pełni zautomatyzowany kwantyfikator rozedmy płuc wykazuje dobrą korelację ze wskaźnikiem Tiffeneau (TI), bazującym na spirometrii, co świadczy o tym, że może być wykorzystywany do diagnozy i oceny ciężkości przebiegu rozedmy oraz długoterminowej obserwacji pacjentów. Badanie przeprowadzono na grupie 141 pacjentów, u których zostały wykonane tomografia komputerowa klatki piersiowej oraz spirometria. Średnia wartość TI, zastosowanego do oceny stopnia zaawansowania rozedmy, wśród badanych wynosiła $0,57 \pm 0,13$. Do porównania kwantyfikacji rozedmy oraz TI został użyty współczynnik korelacji Spearmana. Wykazano bardzo silną korelację pomiędzy rozpoznaniem rozedmy przez AI a wskaźnikiem TI. Diagnostyka rozedmy oparta na AI, dzięki zastosowaniu metod głębokiego uczenia już na poziomie pacjenta, pozwala na stworzenie indywidualnej formy leczenia pacjentów, a tym samym wcześniejsze rozpoznawanie zmian i podejmowanie odpowiednich działań [22, 23].

Diagnostyka odmy opłucnowej

Techniki wykorzystujące AI znalazły też zastosowanie w diagnostyce, określaniu lokalizacji i oceny zaawansowania odmy opłucnowej. 13 292 prześwietleń rentgenowskich klatki piersiowej zostało użytych do wytrenowania sieci neuronowej, w której odma o wysokim i średnim natężeniu była uznawana za wynik pozytywny, a śladowa odma lub jej brak za wynik negatywny. Zaprojektowano dwa modele: model o wysokiej czułości oraz model o wysokiej swoistości. Pierwszy z nich uzyskał czułość na poziomie 84%, swoistość 90% i AUC 0,94, a drugi – czułość 80%, swoistość 97% i AUC 0,96. Autorzy badania wskazują jednak na ograniczenia metody: musi być ona nadzorowana, a odma o małym natężeniu i czasem również o dużym może być niepoprawnie klasyfikowana [24]. Inne retrospektywne badanie zostało przeprowadzone poprzez wytrenowanie sieci neuronowej, do czego użyto 159 prześwietleń klatki piersiowej pacjentów z odmą i 682 bez odmy. Osiągnęła ona AUC równy 0,90. Jednak przy odniesieniu wytrenowanego wcześniej modelu do własnej bazy danych wartość AUC spadła do 0,58. Autorzy podkreślili, że użycie sieci do codziennego użytku może wymagać ponownego trenowania modelu [25].

Diagnostyka POChP

Wiedząc, że szacunkowo 70% przypadków osób cierpiących na przewlekłą obturacyjną chorobę płuc (POChP) może pozostawać niezdiagnozowane [26], grupa naukowców postanowiła wykorzystać metodę *transfer learning*, aby sprawdzić skuteczność niskodawkowej tomografii komputerowej (z ang. *low-dose CT*, LDCT) w diagnostyce POChP [27]. LDCT jest metodą stosowaną głównie w badaniach skriningowych raka płuca. Wiedząc, że zarówno patogenezą POChP, jak i raka płuca mogą mieć związek z przewlekłym narażeniem na czynniki prozapalne [28], postanowiono wypróbować LDCT w rozpoznaniu POChP.

Badania z zastosowaniem głębokiego uczenia, wykorzystujące tomografię komputerową, wykazały skuteczność POChP w diagnostyce osób palących i określeniu stopnia zaawansowania choroby. Pozwoliły również wskazać jednostki, u których może wystąpić ostra niewydolność oddechowa oraz największa śmiertelność. Sieć neuronowa została wytrenowana przy użyciu wyników tomografii komputerowej 7983 uczestników z COPDGene (badania skupiającego ponad 10 tys. palaczy, mającego na celu znalezienie powiązania między warunkami genetycznymi a POChP) oraz 1672 uczestników z ECLIPSE (*Evaluations of COPD Longitudinally to Identify Predictive Surrogate End-points*). W grupie COPDGene C-statistic diagnozy stopnia ciężkości POChP wynosiło 0,856, czyli 51,1% badanych zostało poprawnie zdiagnozowanych. W grupie ECLIPSE prawidłową diagnozę otrzymało 29,4%. W obu tych grupach prawidłową diagnozę z zakresem błędów jednego stopnia ciężkości otrzymało 74–75% badanych [29].

Mammografia

Rak piersi jest najczęściej spotykanym nowotworem u kobiet, stanowiącym wśród nich główną przyczynę śmierci z powodu raka i odpowiadającym za ok. 500 tys. zgonów w ciągu roku na całym świecie [30]. Obecnie najpowszechniej stosowanym badaniem przesiewowym w diagnostyce raka piersi jest mammografia, której zastosowanie zmniejszyło umieralność z powodu tej choroby nawet o 20–22% [31]. Nie wszystkie jednak zmiany mogą zostać wykryte za pomocą

tej metody. Jednym z powodów takiego stanu rzeczy może być trudność w wykryciu lokalizacji nowotworu, np. wskutek jego przysłonięcia przez gęsto utkaną tkankę gruczołu piersiowego. W części wykonywanych mammografii zmiana może pozostać niezauważona lub źle zinterpretowana przez lekarzy. Problem stanowi również wysoki odsetek wyników fałszywie dodatnich, co naraża pacjentów na wykonywanie niepotrzebnej biopsji piersi. W związku z tym poszukiwanie technik i metod, które pomogą zwiększyć czułość i swoistość wykrywania raka piersi w badaniu mammograficznym wydaje się kluczowe [32].

Jedną z bardziej znanych i stosowanych obecnie metod sztucznej inteligencji stanowi technologia głębokiego uczenia, której wpływ na wykrywanie nowotworu zbadali A.T. Watanabe i wsp., wykorzystując cmAssist™, stanowiące oprogramowanie AI-CAD (diagnostyka wspomagana komputerowo w oparciu o AI) dla mammografii. Radiolodzy, niezależnie od ich doświadczenia, wykazali znaczną poprawę wskaźnika wykrywalności nowotworów (CDR) dzięki zastosowaniu algorytmu, który został wytrenowany przy użyciu bazy danych zawierającej liczne mammogramy z nowotworami oraz bez nieprawidłowości, jednak należy podkreślić, że przypadki oceniane w tym badaniu nie zostały wykorzystane do szkolenia. Procentowy wzrost CDR z pomocą cmAssist™ wyniósł średnio 27%, przy czym większe korzyści zostały zaobserwowane u radiologów z małym doświadczeniem niż u radiologów specjalizujących się w badaniach mammograficznych. Wartość AUC dla specjalistów jako grupy wzrosła z 0,760 do 0,815, co stanowi wzrost AUC o 7,2%. Nastąpił także nieistotny statystycznie wzrost liczby wyników fałszywie dodatnich przy użyciu algorytmu, głównie w grupie początkujących radiologów, i wyniósł on mniej niż 1%. Zastosowanie AI-CAD niesie także liczne korzyści, w przypadku gdy istnieją trudności diagnostyczne wynikające z umiejscowienia zmiany nowotworowej, zwłaszcza u pacjenta z gęstą budową piersi. Za pomocą zastosowanego algorytmu dodatkowo czterech radiologów zmieniło swoją decyzję na temat rozpoznania, natomiast podczas samodzielnej interpretacji zaledwie jeden z siedmiu specjalistów zauważył masę nowotworową. Przy ocenie kolejnego przypadku zastosowanie cmAssist™ wpłynęło na zmianę decyzji dwóch radiologów, z braku rozpoznania na rozpoznanie raka, a pozostałych pięciu specjalistów zauważyło ten nowotwór niezależnie od udziału algorytmu. Istotna poprawa czułości specjalistów w wykrywaniu nowotworów, które zostały pierwotnie przeoczone podczas interpretacji badań mammograficznych, może potencjalnie przyczynić się do usprawnienia pracy radiologów, wcześniejszego wykrywania raka piersi i zmniejszenia liczby wyników fałszywie ujemnych [32].

W celu sprawdzenia skuteczności zastosowania metod sztucznej inteligencji oraz ich wpływu na decyzje diagnostyczne radiologów H.E. Kim i wsp. opracowali algorytm AI służący do wykrywania zmian nowotworowych w badaniach mammograficznych. Ogólna wartość AUC w zbiorze danych pochodzących z trzech krajów dla wydajności systemu AI wynosiła 0,959. Z kolei w badaniu, w którym oceniano skuteczność diagnostyczną radiologów, AUC wyniósł 0,810 i było ono niższe w porównaniu ze skutecznością zastosowanego niezależnie systemu AI, gdzie AUC wyniósł 0,940. W przypadku wykorzystania AI jako narzędzia wspomagającego decyzje lekarzy wyniki specjalistów uległy znaczącej poprawie, AUC wzrósł do 0,881. Różnica w interpretacji

badania przeprowadzonego bez wspomaganie AI i z zastosowaniem metod sztucznej inteligencji była najbardziej widoczna w piersiach o gęstej budowie i wynosiła 0,083 [33].

Ostatnie rezultaty zastosowania sztucznej inteligencji w wykrywaniu raka przy użyciu mammografii cyfrowej i tomosyntezy piersi w porównaniu do wyników uzyskanych przez grupę 101 radiologów zostały ocenione przez A. Rodriguez-Ruiz i wsp. Zastosowane systemy oparte na algorytmach głębokiego uczenia, które zostały użyte podczas badania, cechowały się zbliżoną dokładnością w wykrywaniu nowotworów piersi w porównaniu do dokładności uzyskanej przy interpretacji obrazów przez doświadczonych specjalistów. AUC systemu AI wyniósł 0,840 i był wyższy od AUC uzyskanego przez 61,4% radiologów. Obiecujące wyniki osiągnięte dzięki użyciu systemu AI dopracowanego przy użyciu licznych i heterogenicznych badań mammograficznych pokazują, iż metody sztucznej inteligencji mogą mieć znaczny udział w rozwoju systemu opieki zdrowotnej w zakresie badań przesiewowych nowotworów piersi, zwłaszcza w krajach, w których brakuje doświadczonych radiologów, a gdzie AI mogłaby być wykorzystywana w celu niezależnej interpretacji badań [30].

Gęsta tkanka gruczołu piersiowego stanowi czynnik ryzyka rozwoju raka piersi. Obiecujące są wyniki badań na temat udziału systemów CAD w ocenie gęstości piersi w mammografii. Sugerują one, że zautomatyzowana ilościowa ocena gęstości piersi przewyższa niedoskonałość ludzkiej percepcji wzrokowej, co zmniejsza szansę przeoczenia zmiany. Sztuczna inteligencja, dzięki istotnemu rozwojowi metod automatycznej analizy obrazu, które stają się coraz lepsze, wydaje się być narzędziem, które może odegrać kluczową rolę w interpretacji badań mammograficznych nie tylko przy wykrywaniu raka piersi, ale także w ocenie jego ryzyka, wyborze odpowiedniej terapii a wreszcie przewidywaniu wyników zastosowanego leczenia [31].

Diagnostyka złamań

Obecnie najpowszechniejszym narzędziem służącym do oceny pacjentów ze schorzeniami układu mięśniowo-szkieletowego, z uwagi na dostępność, cenę, szybkość wykonania oraz niski poziom dawki promieniowania przyjmowanej przez pacjenta, pozostają zdjęcia rentgenowskie. W celu diagnostyki złamań często wykonuje się także tomografię komputerową lub rezonans magnetyczny [34]. Metody AI mogą zrewolucjonizować pracę radiologów, ortopedów czy chirurgów w zakresie wykrywania złamań, które są niekiedy trudne do zauważenia czy obecne w nietypowych lokalizacjach. Ponadto ważne jest, by zmniejszyć ryzyko przeoczenia złamania u pacjentów z licznymi urazami lub podczas wtórnej oceny chorych po skomplikowanym urazie. Bezstronna ocena prawdopodobieństwa złamania przeprowadzona przez algorytm, jako uzupełnienie ludzkiej wiedzy i doświadczenia, może pomóc lekarzom w podejmowaniu decyzji i doborze odpowiedniego dalszego postępowania z chorym [35]. Stosowane współcześnie metody sztucznej inteligencji cechują się zbliżoną dokładnością w rozpoznawaniu złamań na podstawie badań obrazowych w porównaniu z oceną złamań dokonywaną przez doświadczonych specjalistów [34, 36–38]. W zależności od zastosowanej metody AUC wyniósł od 0,91 do 1,0 [34, 37, 38].

Ostatnie postępy w dziedzinie sztucznej inteligencji w zakresie wykrywania złamań dały obiecujące rezultaty. Zostały one ocenione przez D.W.G. Langerhuizen i wsp.,

którzy uwzględnili dziesięć badań spełniających odpowiednie kryteria włączenia, spośród których dziewięć dotyczyło wykrywania złamań, a jedno ich klasyfikacji. Najczęściej stosowaną metodą AI były konwolucyjne sieci neuronowe (CNN). Dokładność identyfikacji złamania w siedmiu badaniach wahała się od 83 do 98%, a wartość AUC w pięciu badaniach mieściła się w zakresie 0,95–1,0. W trzech badaniach diagnostyka złamań za pomocą AI została porównana z wynikami uzyskanymi przez pięciu ortopedów, gdzie dokładność w rozpoznawaniu złamań wynosiła 96% dla sieci neuronowej, a 92% w odniesieniu do specjalistów [34].

Badanie przeprowadzone przez K. Gan i wsp. oceniało z kolei umiejętności sieci R-CNN, wyposażonej w algorytm szybkiej identyfikacji obiektów, zastosowanej w celu wykrywania złamań dalszej nasady kości promieniowej (DRF). Sieć neuronowa poradziła sobie lepiej z postawieniem rozpoznania dotyczącego obecności DRF lub braku złamania w porównaniu z grupą trzech radiologów i wykazywała skuteczność podobną do grupy trzech ortopedów z ponad 5-letnim doświadczeniem, zaś AUC wynosił 0,96 [38].

W celu eksperymentalnego zastosowania standardowych sieci opartych na głębokim uczeniu (z ang. *deep convolutional neural network*, DCNN), jednej z gałęzi AI, do identyfikacji złamań J. Olczak i wsp. wyodrębnili 256 458 zdjęć rentgenowskich kończyny górnej, nadgarstka oraz kostki, z czego 56% obrazów ukazywało złamania. Do przeprowadzenia badania wybrano 5 głębokich sieci neuronowych i przeszkolono je pod kątem wykrywania złamań, z których najlepszą dokładność, wynoszącą 83%, prezentowała 16-warstwowa sieć Visual Geometry Group (VGG 16). Sieć wykazała działanie zbliżone do działania doświadczonych ortopedów, którym przedstawiono zdjęcia rentgenowskie o tej samej rozdzielczości obrazu, przy czym specjaliści uzyskali dokładność na poziomie 82% [36].

Niebywałą zaletą udziału sztucznej inteligencji w medycynie jest to, iż może ona pomóc w szybszym postawieniu diagnozy i odpowiednio wczesnym wdrożeniu leczenia. K. Murata i wsp. pochylił się w swoich badaniach nad poważnym problemem, jaki stanowią złamania kręgow (VF), gdyż w tym przypadku opóźnienie procesu diagnostycznego wiąże się często z pogorszeniem rokowania, a samo złamanie często skutkuje obniżeniem jakości życia chorego. Kluczowe wydaje się zatem zminimalizowanie liczby błędów w wykrywaniu VF w badaniach obrazowych. W tym celu przeszkolono DCNN w zakresie identyfikacji złamań. Sieć charakteryzowała się wysoką czułością (84,7%) oraz swoistością (87,3%) w stawianiu diagnozy dotyczącej VF, a AUC z zastosowaniem modelu DCNN wynosił 0,91. Dokładność, z jaką sieć rozpoznawała złamania, wynosiła 86% i była wyższa niż ta uzyskana przez rezydentów ortopedii (77,5%), natomiast znacznie niższa od tej, którą uzyskali chirurdzy kręgosłupa (98,4%). Lekarze dokonali błędnej diagnozy u 27 pacjentów, gdzie DCNN, poza jednym z tych przypadków, trafnie rozpoznało obecność VF lub jego brak. Może to wskazywać na istotne korzyści wynikające z wykorzystywania przez lekarzy algorytmów opartych na AI, ponieważ szybsze wykrycie złamania jest kluczowe dla rokowania i funkcjonowania pacjenta [37].

Badania obrazowe zlecane w sposób niewłaściwy mogą stanowić problem diagnostyczny i powodować trudności związane z leczeniem pacjenta, ale też narazić go na kilkukrotne powtarzanie tych badań, co wiąże się z kolejną ekspozycją na promieniowanie, dodatkowym obciążeniem finansowym systemu opieki zdrowotnej, a także może stanowić dużą

niedogodność dla pacjenta. W celu zapobiegania takim sytuacjom oraz pomocy w wyborze najlepszego w danej sytuacji klinicznej badania obrazowego warto rozważyć zastosowanie algorytmów uczenia maszynowego, które mogą uwzględnić wcześniejszą historię choroby pacjenta, tj. uzyskane wcześniej wyniki badań laboratoryjnych, obrazowych, zgłaszane objawy oraz wyniki badania fizykalnego, umożliwiając dobór optymalnego dla chorego w danej sytuacji badania obrazowego [39].

Zastosowania AI w patomorfologii

Patomorfologia cyfrowa (z ang. *digital pathology*) jest stosunkowo nową gałęzią patomorfologii. Działanie w jej zakresie polega na wykonywaniu serii zdjęć preparatów mikroskopowych za pomocą skanera, dzięki czemu generowane są wysokiej rozdzielczości zdjęcia całego preparatu (z ang. *whole slide imaging* – WSI lub *virtual microscopy*) [40–42]. Obrazy te mogą zostać przesłane za pomocą Internetu do patomorfologów na całym świecie, np. w celu konsultacji lub prowadzenia badań naukowych. Również sam algorytm zdolny jest do analizy WSI [40, 41].

Mimo wielu zalet patomorfologia cyfrowa nie jest jeszcze rutynowo wykorzystywana w procesie diagnostycznym, głównie ze względu na konieczność zakupu kosztownego sprzętu, oprogramowania oraz przeszkolenia personelu. Prowadzonych jest coraz więcej badań, których celem jest porównanie trafności rozpoznań dokonywanych przez sztuczną inteligencję z rozpoznaniem stawianymi przez doświadczonych patomorfologów. Wyniki dowodzą, że w dużej części przypadków sztuczna inteligencja radzi sobie podobnie do specjalistów, a dodatkowo może im pomagać w stawianiu rozpoznania, zmniejszając tym samym ryzyko błędu ludzkiego [43–48].

Badanie przeprowadzone z udziałem 12 patomorfologów, których zadaniem była ocena biopsji kroczułu krokowego według skali Gleasona, wykazało, że patomorfolodzy korzystający ze sztucznej inteligencji osiągnęli lepsze rezultaty niż patomorfolodzy pracujący samodzielnie – ważony współczynnik Kappa Cohena z wykorzystaniem wag kwadratowych wyniósł 0,799 dla patomorfologów oraz 0,872 dla patomorfologów korzystających z AI, a samo AI uzyskało współczynnik Kappa równy 0,854 [43]. W innym badaniu, którego celem również była ocena biopsji gruczołu krokowego według skali Gleasona, sztuczna inteligencja uzyskała współczynnik Kappa równy 0,62, przy czym współczynnik Kappa dla patomorfologów, którzy wzięli udział w badaniu, wyniósł 0,60–0,73 [44].

Innym przykładem wykorzystania AI w patomorfologii jest konwolucyjna sieć neuronowa, która uzyskała ponad 90% dokładności w stosunku do danych walidacyjnych w rozpoznawaniu glejaka mózgu, dokładność wynoszącą 96% przy klasyfikacji GBM lub LGG oraz 71% w identyfikacji stopnia LGG II lub III [45]. Z kolei w innym badaniu zaprojektowano sieć neuronową, do której walidacji i wyszkolenia użyto 378 preparatów mikroskopowych z biopsji węzłów chłonnych w przebiegu chłoniaka grudkowego (197) lub hiperplazji pęcherzykowej (181), a które uzyskała AUC mieszczący się w granicach 0,92–0,99 [47].

Oparty na głębokim uczeniu klasyfikator różniący pod względem histopatologicznym gruczolakoraka płuc, raka płaskonabłonkowego płuc, raka drobnokomórkowego płuc, gruczlicę płuc, organizujące się zapalenie płuc i prawidłowe płuca, przetestowany na 1067 preparatach mikroskopowych

pochodzących z 4 kohort z różnych ośrodków medycznych, uzyskał AUC wynoszące kolejno: 0,970, 0,918, 0,963, 0,978. Współczynnik zgodności z klasyfikacją przeprowadzoną przez doświadczonych patomorfologów wyniósł 0,873 [46].

Przygotowany w ramach *International Symposium on Biomedical Imaging* algorytm oparty na głębokim uczeniu, służący do wykrywania przerzutów raka piersi na WSI z biopsji węzłów wartowniczych, uzyskał AUC 0,925, przy czym AUC uzyskany przez specjalistę był równy 0,966. Z kolei skorzystanie z pomocy sztucznej inteligencji przez patomorfologa zwiększało uzyskany AUC do 0,995, dzięki czemu uzyskano redukcję błędu ludzkiego o ok. 85% [48].

PODSUMOWANIE

Wspomniane metody nie są jeszcze powszechnie stosowane w radiologii i patomorfologii. Podejmowane są liczne próby wykorzystywania systemów opartych na głębokim uczeniu w rozpoznawaniu zmian chorobowych widocznych na zdjęciach RTG, TK czy zdjęciach preparatów mikroskopowych. Samouczące się algorytmy zazwyczaj osiągają wyniki porównywalne lub niekiedy nawet lepsze od wyników uzyskiwanych przez radiologów i patomorfologów. Synergistyczne działanie lekarzy i sztucznej inteligencji może przyczynić się również do zmniejszenia ryzyka popełnienia błędu ludzkiego w diagnozowaniu, np. będącego efektem przeoczenia zmiany przez lekarza, dzięki czemu może usprawniać i ułatwiać proces diagnostyczny.

PIŚMIENICTWO

- Suzuki K. Overview of deep learning in medical imaging. *Radiol Phys Technol.* 2017; 10: 257–273. <https://doi.org/10.1007/s12194-017-0406-5>
- Rajkumar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med.* 2019; 380(14): 1347–1358. <https://doi.org/10.1056/NEJMr1814259>
- Hussain AA, Bouachir O, Al-Turjman F, et al. AI Techniques for COVID-19. *IEEE Access.* 2020; 8: 128776–128795. <https://doi.org/10.1109/ACCESS.2020.3007939>
- Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism.* 2017; 69S: S36–S40. <https://doi.org/10.1016/j.metabol.2017.01.011>
- He J, Baxter SL, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine. *Nat Med.* 2019; 25(1): 30–36. <https://doi.org/10.1038/s41591-018-0307-0>
- Holzinger A, Langs G, Denk H, et al. Causability and explainability of artificial intelligence in medicine. *WIREs Data Mining Knowl Discov.* 2019; 9: e1312. <https://doi.org/10.1002/widm.1312>
- Briganti G, Le Moine O. Artificial Intelligence in Medicine: Today and Tomorrow. *Front Med (Lausanne).* 2020; 7: 27. <https://doi.org/10.3389/fmed.2020.00027>
- Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol.* 2019; 28(2): 73–81. <https://doi.org/10.1080/13645706.2019.1575882>
- Liang G, Zheng L. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Comput Methods Programs Biomed.* 2020; 187: 104964. <https://doi.org/10.1016/j.cmpb.2019.06.023>
- Schwyzler M, Martini K, Skawran S, et al. Pneumonia Detection in Chest X-Ray Dose-Equivalent CT: Impact of Dose Reduction on Detectability by Artificial Intelligence. *Acad Radiol.* 2020; S1076–6332(20)30338-X. <https://doi.org/10.1016/j.acra.2020.05.031>
- Shih G, Wu CC, Halabi SS. Augmenting the National Institutes of Health Chest Radiograph Dataset with Expert Annotations of Possible Pneumonia. *Radiology: Artificial Intelligence.* 2019; 1(1): e180041. <https://doi.org/10.1148/ryai.2019180041>
- Depeursinge A, Chin AS, Leung AN, et al. Automated classification of usual interstitial pneumonia using regional volumetric texture analysis in high-resolution computed tomography. *Invest Radiol.* 2015; 50(4): 261–267. <https://doi.org/10.1097/RLI.000000000000127>
- Luján-García JE, Yáñez-Márquez C, Villuendas-Rey Y, et al. A Transfer Learning Method for Pneumonia Classification and Visualization. *Applied Sciences.* 2020; 10(8): 2908. <https://doi.org/10.3390/app10082908>
- Gabruseva T, Poplavskiy D, Kalinin A. Deep Learning for Automatic Pneumonia Detection. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 1436–1443. <https://doi.org/10.1109/CVPRW50498.2020.00183>
- Warjri L, Shah A. India and Africa: Charting a Post-COVID-19 Future. *ORF Special Report No. 111, June 2020, Observer Research Foundation.* https://www.orfonline.org/wp-content/uploads/2020/06/ORF_Special-Report_111_India-Africa-Health.pdf (access: 2021.03.30).
- Schwyzler M, Martini K, Skawran S, et al. Pneumonia Detection in Chest X-Ray Dose-Equivalent CT: Impact of Dose Reduction on Detectability by Artificial Intelligence. *Academic Radiology.* 2020. <https://doi.org/10.1016/j.acra.2020.05.031>
- Xiong F, Wang Y, You T, et al. The clinical classification of patients with COVID-19 pneumonia was predicted by Radiomics using chest CT. *Medicine (Baltimore).* 2021; 100(12): e25307. <https://doi.org/10.1097/md.00000000000025307>
- Liu C, Wang X, Liu C, et al. Differentiating novel coronavirus pneumonia from general pneumonia based on machine learning. *BioMed Eng OnLine.* 2020; 19: 66. <https://doi.org/10.1186/s12938-020-00809-9>
- Wang H, Wang L, Lee EH, et al. Decoding COVID-19 pneumonia: comparison of deep learning and radiomics CT image signatures. *Eur J Nucl Med Mol Imaging.* 2020; 1–9. <https://doi.org/10.1007/s00259-020-05075-4>
- Tan W, Liu P, Li X, et al. Classification of COVID-19 pneumonia from chest CT images based on reconstructed super resolution images and VGG neural network. *Health Inf Sci Syst.* 2021; 9(1): 10. <https://doi.org/10.1007/s13755-021-00140-0>
- Martini K, Frauenfelder T. Advances in imaging for lung emphysema. *Ann Transl Med.* 2020; 8(21): 1467. <https://doi.org/10.21037/atm.2020.04.44>
- Fischer AM, Varga-Szemes A, van Assen M, et al. Comparison of Artificial Intelligence-Based Fully Automatic Chest CT Emphysema Quantification to Pulmonary Function Testing. *AJR Am J Roentgenol.* 2020; 214(5): 1065–1071. <https://doi.org/10.2214/AJR.19.21572>
- Labaki WW, Han MK. Artificial Intelligence and Chest Imaging. Will Deep Learning Make Us Smarter?. *Am J Respir Crit Care Med.* 2018; 197(2): 148–150. <https://doi.org/10.1164/rccm.201709-1879ED>
- Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. *PLoS Med.* 2018; 15(11): e1002697. <https://doi.org/10.1371/journal.pmed.1002697>
- Kitamura G, Deible C. Retraining an open-source pneumothorax detecting machine learning algorithm for improved performance to medical images. *Clin Imaging.* 2020; 61: 15–19. <https://doi.org/10.1016/j.clinimag.2020.01.008>
- Diab N, Gershon AS, Sin DD, et al. Underdiagnosis and Overdiagnosis of Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med.* 2018; 198(9): 1130–1139. <https://doi.org/10.1164/rccm.201804-0621CI>
- Tang LYW, Coxson HO, Lam S, Leipsic J, Tam RC, Sin DD. Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. *Lancet Digit Health.* 2020; 2(5): e259–e267. [https://doi.org/10.1016/S2589-7500\(20\)30064-9](https://doi.org/10.1016/S2589-7500(20)30064-9)
- Durham AL, Adcock IM. The relationship between COPD and lung cancer. *Lung Cancer.* 2015; 90(2): 121–127. <https://doi.org/10.1016/j.lungcan.2015.08.017>
- González G, Ash SY, Vegas-Sánchez-Ferrero G, et al. Disease Staging and Prognosis in Smokers Using Deep Learning in Chest Computed Tomography. *Am J Respir Crit Care Med.* 2018; 197(2): 193–203. <https://doi.org/10.1164/rccm.201705-0860OC>
- Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. *J Natl Cancer Inst.* 2019; 111(9): 916–922. <https://doi.org/10.1093/jnci/djy222>
- Geras KJ, Mann RM, Moy L. Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives. *Radiology.* 2019; 293(2): 246–259. <https://doi.org/10.1148/radiol.2019182627>
- Watanabe AT, Lim V, Vu HX, et al. Improved Cancer Detection Using Artificial Intelligence: a Retrospective Evaluation of Missed Cancers on Mammography. *J Digit Imaging.* 2019; 32(4): 625–637. <https://doi.org/10.1007/s10278-019-00192-5>
- Kim HE, Kim HH, Han BK, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence:

- a retrospective, multireader study. *Lancet Digit Health*. 2020; 2(3): e138–e148. [https://doi.org/10.1016/S2589-7500\(20\)30003-0](https://doi.org/10.1016/S2589-7500(20)30003-0)
34. Langerhuizen DWG, Janssen SJ, Mallee WH, et al. What Are the Applications and Limitations of Artificial Intelligence for Fracture Detection and Classification in Orthopaedic Trauma Imaging? A Systematic Review. *Clin Orthop Relat Res*. 2019; 477(11): 2482–2491. <https://doi.org/10.1097/CORR.0000000000000848>
35. Michelson JD. CORR Insights®: What Are the Applications and Limitations of Artificial Intelligence for Fracture Detection and Classification in Orthopaedic Trauma Imaging? A Systematic Review. *Clin Orthop Relat Res*. 2019; 477(11): 2492–2494. <https://doi.org/10.1097/CORR.0000000000000912>
36. Olczak J, Fahlberg N, Maki A, et al. Artificial intelligence for analyzing orthopedic trauma radiographs. *Acta Orthop*. 2017; 88(6): 581–586. <https://doi.org/10.1080/17453674.2017.1344459>
37. Murata K, Endo K, Aihara T, et al. Artificial intelligence for the detection of vertebral fractures on plain spinal radiography. *Sci Rep*. 2020; 10(1): 20031. <https://doi.org/10.1038/s41598-020-76866-w>
38. Gan K, Xu D, Lin Y, et al. Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments. *Acta Orthop*. 2019; 90(4): 394–400. <https://doi.org/10.1080/17453674.2019.1600125>
39. Gyftopoulos S, Lin D, Knoll F, et al. Artificial Intelligence in Musculoskeletal Imaging: Current Status and Future Directions. *AJR Am J Roentgenol*. 2019; 213(3): 506–513. <https://doi.org/10.2214/AJR.19.21117>
40. Prochorec-Sobieszek M. Future perspectives of digital pathology. *NOWOTWORY J Oncol* 2016; 66: 277–284. <https://doi.org/10.5603/njo.2016.0054>
41. Prochorec-Sobieszek M, Szumera-Ciećkiewicz A. Aktualny stan wiedzy na temat patomorfologii cyfrowej. *Hematologia*. 2017; 8(1): 1–11. <https://doi.org/10.5603/hem.2017.0001>
42. Farahani N, Parwani A, Pantanowitz L. Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International*. 2015; 7: 23–33. <https://doi.org/10.2147/PLMI.S59826>
43. Bulten W, Balkenhol M, Belinga JA, et al. Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Mod Pathol*. 2021; 34(3): 660–671. <https://doi.org/10.1038/s41379-020-0640-y>
44. Ström P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol*. 2020; 21(2): 222–232. [https://doi.org/10.1016/S1470-2045\(19\)30738-7](https://doi.org/10.1016/S1470-2045(19)30738-7)
45. Ertosun MG, Rubin DL. Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. *AMIA Annu Symp Proc*. 2015; 1: 1899–1908.
46. Yang H, Chen L, Cheng Z, et al. Deep learning-based six-type classifier for lung cancer and mimics from histopathological whole slide images: a retrospective study. *BMC Med*. 2021; 19(1): 80. <https://doi.org/10.1186/s12916-021-01953-2>
47. Syrykh C, Abreu A, Amara N, et al. Accurate diagnosis of lymphoma on whole-slide histopathology images using deep learning. *npj Digit Med*. 2020; 3(1): 63. <https://doi.org/10.1038/s41746-020-0272-0>
48. Wang D, Khosla A, Gargeya R, et al. Deep Learning for Identifying Metastatic Breast Cancer. *arXiv* <http://arxiv.org/abs/1606.05718> (access: 2021.03.29).