czasopismo geograficzne
POLSKIE TOWARZYSTWO GEOGRAFICZNE · POLISH GEOGRAPHICAL SOCIETY
geographical journal

GŁĘBOKIE UCZENIE W PROCESIE TELEDETEKCYJNEJ INTERPRETACJI PRZESTRZENI GEOGRAFICZNEJ – PRZEGLĄD WYBRANYCH ZAGADNIEŃ

DEEP LEARNING IN THE GEOGRAPHICAL SPACE REMOTE SENSING INTERPRETATION – REVIEW OF SELECTED ISSUES

Maciej Adamiak

SoftwareMill, Warszawa, Polska, https://orcid.org/0000-0002-8229-9661,
e-mail: maciej.adamiak@softwaremill.com

Streszczenie

Zastosowanie uczenia maszynowego (ML, ang. *machine learning*) oraz uczenia głębokiego (DL, *deep learning*), zwłaszcza głębokich konwolucyjnych sieci neuronowych (DCNN, ang. *deep convolutional neural network*), w przetwarzaniu oraz interpretacji obrazu jest obecnie szeroko omawianym zagadnieniem wśród przedstawicieli aktywnie rozwijającego się środowiska naukowego skupionego w ramach takich dziedzin, jak teledetekcja oraz geoinformacja. Niniejsza publikacja jest próbą usystematyzowania wiedzy dotyczącej DL oraz jego zastosowania we wspomaganiu interpretacji przestrzeni geograficznej na podstawie zobrazowań lotniczych i satelitarnych. Przeglądem zagadnień DL zainteresowani będą przede wszystkim geografowie, którzy chcieliby wzbogacić prowadzone badania naukowe o metody oparte na sztucznych sieciach neuronowych. W tekście przedstawione zostały główne koncepcje DL oraz metody wraz z przykładowymi kategoriami zadań, które można zrealizować za ich pomocą. Mowa o semantycznej segmentacji, klasyfikacji, augmentacji materiału badawczego i inżynierii cech. Prezentacja każdej z tych kategorii została wzbogacona o opis przypadku użycia i przegląd literatury, co umożliwiło wykonanie pierwszego kroku ku zastosowaniu danej techniki w przyszłych projektach badawczych. Zakończenie artykułu stanowi dyskusja nad nowymi kierunkami rozwoju DL w ramach dyscypliny nauk o Ziemi i środowisku.

Adamiak M. (2021). Głębokie uczenie w procesie teledetekcyjnej interpretacji przestrzeni geograficznej – przegląd wybranych zagadnień. *Czasopismo Geograficzne* 92(1): 49–72. https://doi.org/10.12657/czageo-92-03

CC BY 4.0 OPEN ACCESS

Otrzymano/Received: 24.05.2021
Zaakceptowano/Accepted: 31.08.2021

Słowa kluczowe: konwolucyjna sieć neuronowa, przestrzeń geograficzna, teledetekcja, uczenie głębokie, uczenie maszynowe.

Abstract
The use of machine learning (ML) and deep learning (DL), especially deep convolutional neural networks (DCNN) in image processing and interpretation is currently a widely discussed topic among representatives of actively developing remote sensing and geoinformation scientific community. This article is an attempt to systematize the knowledge of DL in its supportive role in the aerial and satellite imagery interpretation of geographical space. The target audience of this overview are geographers who would like to enrich their research with methods based on artificial neural networks. The text presents main concepts and methods of DL along with example tasks that can be completed with their help i.e.: semantic segmentation, classification, augmentation of research dataset and feature engineering. The description of each task category was enriched with a use case and a literature review, thus making it possible to take the first step towards applying the specified technique in future research. The article conclusion includes a discussion on new directions and opportunities of applying DL in the discipline of Earth and environmental sciences.

Keywords: convolutional neural networks, deep learning, geographical space,machine learning, remote sensing.

WPROWADZENIE

Ostatnią dekadę należy uznać za wybitnie interesujący okres w dziedzinie badań nad uczeniem maszynowym (ML, ang. *machine learning*), a w głównej mierze nad uczeniem głębokim (DL, ang. *deep learning*). W relatywnie krótkim czasie wiele wcześniej opracowanych koncepcji matematycznych zostało zmaterializowanych w postaci ogólnodostępnych narzędzi informatycznych, by wreszcie znaleźć praktyczne zastosowanie w aktywnie rozwijających się dyscyplinach nauki i techniki.

Mówimy, że program komputerowy uczy się podczas nabywania doświadczenia w odniesieniu do pewnej klasy zadań przy jednoczesnym pomiarze skuteczności, jeśli jego skuteczność w trakcie wykonywania tych zadań poprawia się wraz z doświadczeniem (Mitchell, 1997). Innymi słowy, model ML, będący głównym elementem składowym programu, dopasowuje się do danych wejściowych. Uczenie głębokie (znane również jako uczenie głębokie strukturalne) jest częścią rodziny metod ML opartych na złożonych sztucznych sieciach neuronowych i z zastosowaniem zestawu technik, które umożliwiają systemowi automatyczne wykrywanie cech (Bengio i in., 2014; Schmidhuber, 2015). Do najważniejszych osiągnięć należą tu głębokie konwolucyjne sieci neuronowe (DCNN, ang. *deep convolutional neural networks*), które dzięki upowszechnieniu się dostępu do znacznych zasobów ob-

liczeniowych na stałe zasiliły arsenał metodyczny naukowców na całym świecie. Innym znamienitym przykładem, obok DCNN, jest *Visual Transformer* (Vaswani i in., 2017). Jedną z dyscyplin, która w wyraźny sposób zyskała na rozwoju DL, jest widzenie komputerowe (CV, ang. *computer vision*), a wraz z nim teledetekcja (RS, ang. *remote sensing*) i tym samym nauki o Ziemi i środowisku (Chen Y. i in., 2014; Kussul i in., 2017; Kamilaris, Prenafeta-Boldu, 2018).

Rozwój technologii teledetekcji, wsparty znaczną poprawą dostępności i jakości repozytoriów danych przestrzennych, doprowadził do wzrostu zainteresowania wykorzystaniem zebranego materiału w badaniach naukowych i pracach rozwojowych. Doskonałym przykładem takich zbiorów są wysokiej jakości obrazy udostępniane w ramach programów Landsat 8 (*Landsat 8 – Landsat Science*, 2020), Sentinel-2 A/B (*Sentinel-2 – ESA Operational EO Missions – Earth Online – ESA*, 2020) oraz Pléiades (*Pleiades – eoPortal Directory – Satellite Missions*, 2020), a w Polsce uwolnienie zasobów ortofotomap w ramach Tarczy antykryzysowej 4.0 (Dziennik Ustaw, 2020). Systematycznie odbierany i gromadzony jest olbrzymi zasób danych wytworzony w trakcie pracy zaawansowanych sensorów optycznych oraz radarów zainstalowanych na pokładach platform operujących w przestrzeni kosmicznej, atmosferze ziemskiej i na Ziemi. Różnorodność nagromadzonego materiału ze względu na jego tematykę, typ, format, czas przydatności do użycia, rozdzielczość przestrzenną, czasową i spektralną stawia przed badaczami niezliczoną liczbę wyzwań, z których znacząca większość ma na celu automatyzację procesu przetwarzania obrazu i wspomaganie procesu jego interpretacji. Nie budzi zaskoczenia skierowanie uwagi badaczy i inżynierów na możliwości strumieniowego przetwarzania wielkich zbiorów danych (ang. *big data*) oraz interpretacji z użyciem wysokowydajnych algorytmów i modeli opartych na ML i DL (Boyd, 2009).

Zobrazowania lotnicze i satelitarne Ziemi są źródłem cennych informacji o przestrzeni geograficznej. Oferują rozszerzenie analizy poza to, co widoczne (Lillesand i in., 2015), dzięki dostępności do kanałów spoza spektrum światła widzialnego i zwiększenie rozdzielczości spektralnej dzięki obrazowaniu hiperspektralnemu. Ich wykorzystanie pozwala na lepsze zrozumienie zjawisk zachodzących w skali makro bez ograniczeń wynikających z tradycyjnych obserwacji w skali lokalnej (Tang i in., 2009). Z technicznego punktu widzenia na uwagę zasługują następujące cechy danych pochodzących z obrazowania: wysoka jakość związana ze stale zwiększaną rozdzielczością przestrzenną, spektralną, radiometryczną i czasową, standaryzacja materiału w obrębie jednego lub kilku systemów teledetekcyjnych umożliwiająca porównywanie różnych obszarów, systematyczność transmisji materiału sprzyjająca prowadzenia badań w dłuższym horyzoncie czasowym, dostępność danych aktualnych oraz historycznych w ramach usług komercyjnych i repozytoriów publicznych, wsparcie i integracja z ogólnodostępnymi systemami informacji geograficznej (GIS, ang. *geographic information system*), a także ciągły rozwój i ulepszenie metod oraz technologii obrazowania.

Niniejsza publikacja jest próbą zaprezentowania aktualnego stanu wiedzy dotyczącej DL oraz jego zastosowania w interpretacji przestrzeni geograficznej. Autor podejmuje się w niej:

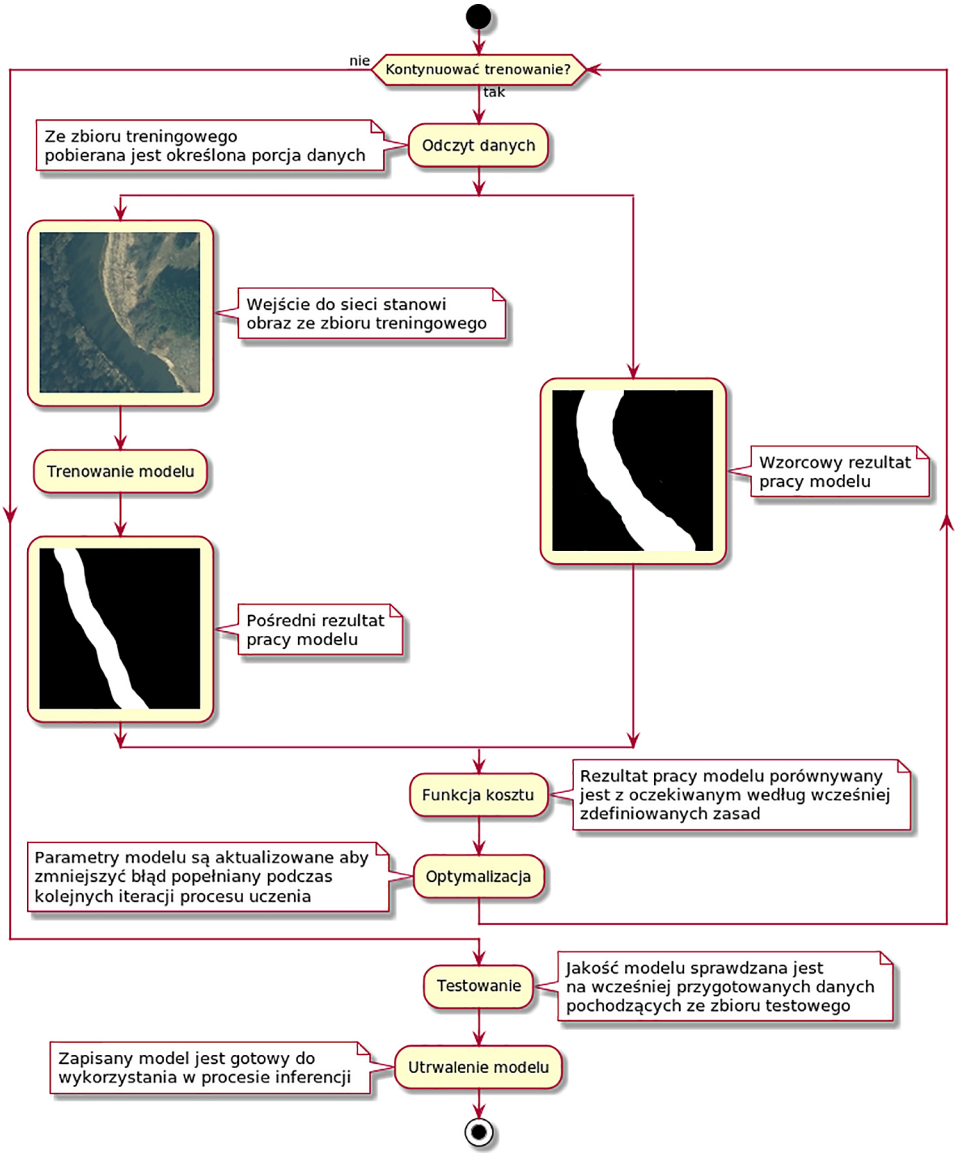
1. Przedstawienia w syntetyczny oraz przystępny sposób koncepcji leżących u podstaw DL i ich powiązania z CV. Szczególną uwagę zwrócono na pojęcia związane z głębokimi konwolucyjnymi sieciami neuronowymi (DCNN).
2. Przybliżenia czytelnikowi metod i technik interpretacji materiału pozyskanego podczas obrazowania lotniczego i satelitarnego z wykorzystaniem metod DL wraz z przykładami ich zastosowania.
3. Podjęcia dyskusji o przyszłych kierunkach w zastosowaniu DL w ramach dyscypliny nauk o Ziemi i środowisku, a przede wszystkim w geografii.

PRZEGLĄD METOD I TECHNIK DL

Uczenie sztucznej sieci neuronowej

Uczenie maszynowe nierozłącznie związane jest z danymi. To dzięki odpowiednio dobranemu i zarządzanemu zbiorowi danych sieć neuronowa może doskonalić się w procesie uczenia. Tradycyjnie zbiór ten dzielony jest na dwa podzbiory. Pierwszy, zwany treningowym, używany jest w procesie uczenia. Drugi testowy, zawiera próbki, które nie wezmą udziału w trenowaniu i posłużą jedynie do sprawdzenia skuteczności modelu.

Proces trenowania sieci neuronowej (ryc. 1) sprowadza się, w znacznym uproszczeniu, do zadania optymalizacyjnego polegającego na minimalizacji zdefiniowanej funkcji kosztu. Zadaniem funkcji kosztu jest pomiar błędu, jaki popełnia sieć w porównaniu do przyjętego wzorca. Przykładem funkcji kosztu są: średni błąd kwadratowy (ang. *mean squared error*), indeks Jaccarda (ang. *Jaccard Index*), dywergencja Kullbacka-Leiblera i inne. Informacja o błędzie propagowana jest z każdą iteracją procesu w głąb sieci. Sieć neuronowa ma budowę warstwową, w której każda z warstw przechowuje skończony zestaw parametrów. Parametry te, inicjalizowane losowymi wartościami, zostają dostosowane zgodnie z informacją o obliczonym błędzie. Konsekwencją jest ewolucja modelu, który w kolejnych etapach coraz lepiej radzi sobie z postawionym przed nim problemem. Uczenie sieci przerywane jest, gdy model osiąga założoną skuteczność albo upłynął czas przeznaczony na jego trenowanie, nie poprawia swoich rezultatów lub przeuczy się *(*ang. *overfitting)*. Skuteczność modelu mierzona jest na zbiorze testowym według wcześniej wybranych miar, np.: dokładności (ang. *accuracy)*, czułości (ang. *recall)*, precyzji (ang. *precision)* i wielu innych metryk dostosowanych do konkretnego zadania. Częstym przypadkiem jest przeuczenie modelu, w którym mimo doskonałej skuteczności mierzonej na zbiorze treningowym model osiąga słabe rezultaty na zbiorze testowym. Jest to oznaka modelu o niskiej jakości, który



Ryc. 1. Proces trenowania modelu uczenia maszynowego (nadzorowany)
Źródło: opracowanie własne.

będzie miał problem z pracą z nowymi obserwacjami. Istnieje również możliwość, że model nie jest w stanie osiągnąć postawionych przed nim wymagań i nie poprawia swoich osiągów lub uczy się na tyle wolno, że uzyskanie zadowalającego rezultat jest kosztowne lub niemożliwe. Wytrenowany model jest gotowy do użycia w procesie inferencji, tj. wnioskowania w oparciu o dane, które badacz zdecy-

dował się poddać transformacji z jego użyciem. Model utrwalany jest w postaci pliku i może zostać wykorzystany wielokrotnie, a także udostępniany w postaci usługi w sieci web.

Zagadnienia leżące u podstaw uczenia maszynowego i konstrukcji sieci neuronowych są złożone. Doskonałym wprowadzeniem do tematu są podręczniki: *Pattern Recognition and Machine Learning* (Bishop, 2006) *Hands-On Machine Learning* (Gron, 2019) oraz kurs *Deep Learning Specialization* (*DeepLearning.AI*, 2021) dostępny w serwisie Coursera.

Reprezentacja obrazu w głębokiej konwolucyjnej sieci neuronowej (DCNN)

Przykładem wykorzystania głębokiej konwolucyjnej sieci neuronowej (DCNN) jest analiza fragmentu terenu zurbanizowanego, znajdującego się w granicach Łodzi. Utworzono w tym celu kwadratowy wycinek przez selektywne scalenie dwóch arkuszy ortofotomapy wykonanej w 2019 r. (*Geoportal*, 2020): M-34-3-D-b-4-2-3 CIR oraz M-34-3-D-b-4-2-3 RGB (ryc. 2). Bok wycinka ma długość 5000 px, co przekłada się na 250 m w terenie przy rozmiarze piksela równym 0,05 m. Jego środek znajduje się w punkcie 51°46'24.3 N 19°28'52.9 E. Obraz zawiera informacje pochodzące z czterech kanałów spektralnych: czerwonego (R), zielonego (G), niebieskiego (B) oraz bliskiej podczerwieni (NIR). Obraz poddano wstępnej interpretacji, ograniczając się jedynie do wyrazistych cech badanego obszaru. Mamy do czynienia ze zróżnicowanym terenem zabudowanym zlo-

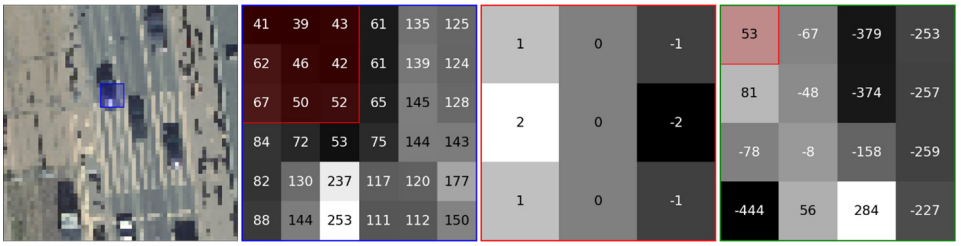
Ryc. 2. Wycinek arkusza M-34-3-D-b-4-2-3 (Łódź) w kompozycji RGB (lewy) oraz CIR (prawy)
Źródło: opracowanie własne na podstawie geoportal.gov.pl.

kalizowanym w pobliżu ruchliwego ciągu komunikacyjnego. Budynki tworzą uporządkowany układ poprowadzony wzdłuż kilku prostopadłych do siebie osi. Zabudowa jest zwarta. Pokrywa roślinna jest szczątkowa. Występują pojedyncze drzewa. Analizę można poprowadzić dalej, wyznaczając powierzchnię danego typu pokrycia terenu, a nawet szacować liczbę kondygnacji budynków oraz określić liczbę samochodów. Zrozumienie treści przedstawionego obrazu jest dla człowieka procesem naturalnym. Z jakich mechanizmów musi skorzystać maszyna w celu przeprowadzenia podobnej analizy obrazu?

Jedną z form reprezentacji cyfrowej obrazu oraz sposobu zaopatrzenia algorytmu ML w dane wejściowe jest wielowymiarowa tablica o zadanej wysokości, szerokości oraz głębokości wynikającej z wymiarów analizowanego obrazu oraz liczby dostępnych kanałów. Dla przykładowego obrazu jest to $5000 \times 5000 \times 4$. Każdemu elementowi tablicy przypisana jest dokładnie jedna wartość liczbowa, pochodząca z pewnego skończonego zakresu, reprezentująca intensywność odbicia światła w danym kanale i położeniu. Innymi słowy, obraz tworzony jest przez zbiór uporządkowanych wartości. Algorytmy ML, takie jak: maszyna wektorów nośnych (SVM, ang. *support vector machine*), lasy losowe (RF, ang. *random forests*), algorytm k-średnich (ang. *k-means*), pracują w oderwaniu od narzuconego porządku przestrzennego, używając jedynie informacji zawartych w wartościach pikseli. Inne algorytmy, takie jak grupowanie przestrzenne (ang. *spatial clustering*), używają informacji o lokalizacji wartości w celu określenia sąsiedztwa. Bez względu na podejście algorytmy te osiągają interesujące rezultaty (Dabija i in., 2021; Kałużna, Będkowski, 2020; Malinowski i in., 2020) i są szeroko stosowane w interpretacji przestrzeni geograficznej. Analizowanie obrazu w formie tablicy ma jednak istotne ograniczenie. Staje się ono widoczne podczas przetwarzania obrazu o dużej rozdzielczości. Nie chodzi jedynie o wydajność, która drastycznie spada wraz ze zwiększeniem rozmiaru tablicy, ale o bogactwo treści, którego opisanie nie jest możliwe jedynie za pomocą analizy bezpośrednio sąsiadujących ze sobą pikseli. Podobnie jak człowiek miałby problem z analizą obrazu piksel po pikselu, bez szerszego kontekstu, jakim są złożone kształty i wzajemna relacja między nimi, tak i maszyna może jedynie ograniczyć się do rozwiązania w obrębie pewnej wybranej klasy problemów. Uczenie z wykorzystaniem DCNN rozwiązuje ten problem przez wprowadzenie dodatkowego kroku przed rozpoczęciem procesu analitycznego. Tak jak w przypadku klasycznych algorytmów na wejściu znajduje się tablica zawierająca informacje o wartości pikseli. Jednak przed podjęciem właściwych działań przekształcana jest ona w reprezentację umożliwiającą zapis cech przestrzennych występujących na obrazie. Sam obraz nie podlega modyfikacji. Zmienia się sposób, w jaki jest on zapisywany oraz przekazywany do algorytmu zajmującego się właściwą interpretacją. W zależności od zaawansowania reprezentacji może ona przechowywać informacje o prostych kształtach, liniach o różnym układzie oraz złożonych cechach obiektów, takich jak dachy budynków, drzewa

czy pojazdy. Reprezentacja obrazu opisana jest w wielu, następujących po sobie, warstwach ukrytych sieci neuronowej (ang. *hidden layers*).

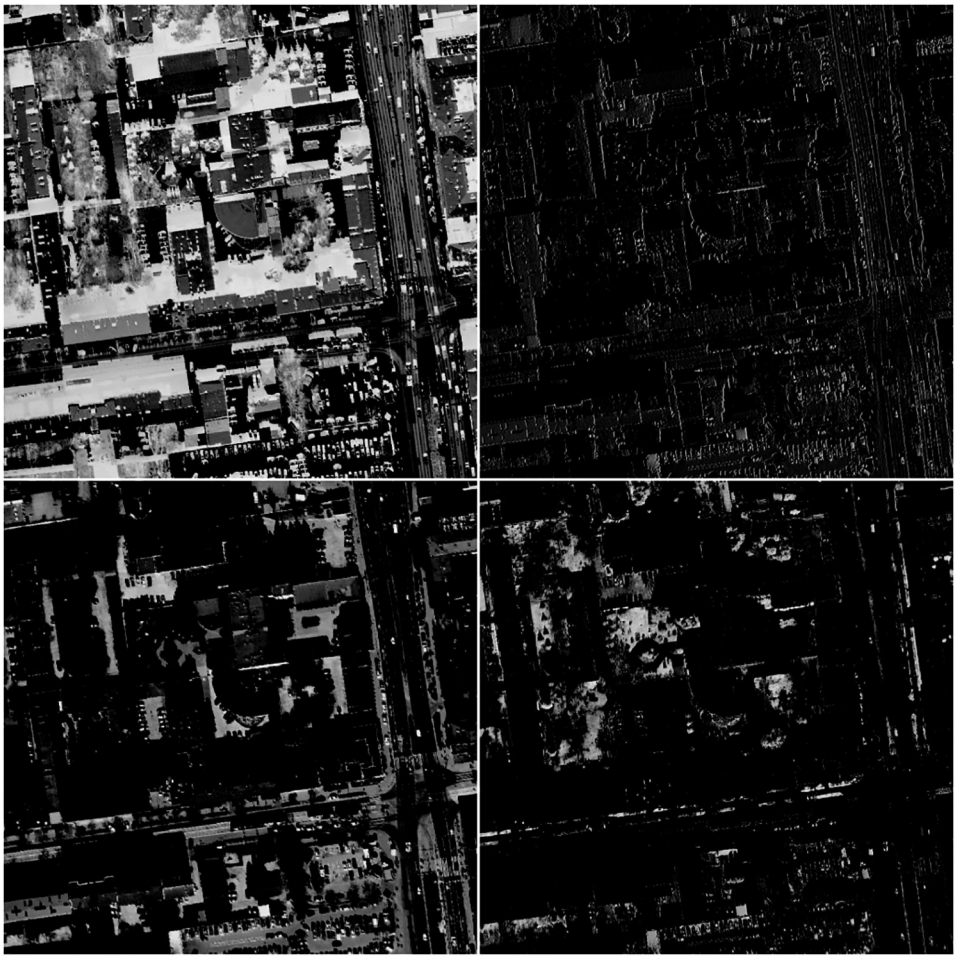
Kodowanie cech obrazu jest możliwe przez zastosowanie operacji konwolucji (splotu), w której zestaw filtrów używany jest, aby poddać obraz wejściowy charakterystycznej, z punktu widzenia sieci neuronowej, transformacji. W zależności od doboru wartości macierzy opisującej dany filtr może on uwydatnić pewne cechy obrazu lub je wygaszać. Przykładem może być operacja, której rezultatem jest wykrycie krawędzi obiektów widocznych na obrazie. Rycina 3 przedstawia zastosowanie wybranego filtra konwolucyjnego o rozmiarze 3 × 3 (Sobel, 1968) na wybranym wycinku pojedynczego kanału przykładowego obrazu o rozmiarze 6 × 6. Rezultatem jest tablica 4 × 4, w której każda wartość została obliczona jako suma iloczynów wartości parametrów filtra oraz odpowiedniej części badanego wycinka.



Ryc. 3. Zastosowanie filtra konwolucyjnego (od lewej: obraz poddawany transformacji, wycinek obrazu, filtr konwolucyjny, rezultat transformacji wycinka)
Źródło: opracowanie własne.

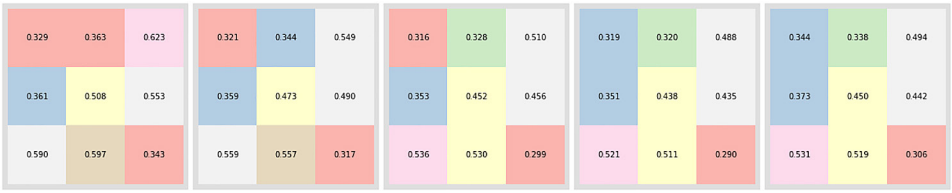
Aby trafnie opisać obraz przy użyciu zbioru jego cech, musi on zostać zakodowany wieloma odmiennymi filtrami. Co więcej, sposób opisu musi być spójny dla całego materiału badawczego trafiającego do algorytmu, co oznacza, że każdy przetworzony obraz przejdzie przez ten sam zestaw transformacji. Intuicyjnie o filtrach można myśleć w kategorii mechanizmów tworzenia nowych kanałów, w których nie są zakodowane intensywności odbicia, lecz różnorodne cechy obrazu (ang. *feature map*). Rycina 4 przedstawia przykładowy obraz poddany działaniu czterech różnych filtrów, uwydatniających specyficzne cechy obrazu: lewy górny – obszary kontrastujące i cienie, prawy górny – krawędzie biegnące wertykalnie, lewy dolny – obszar pokryty specyficznym typem nawierzchni, prawy dolny – tereny zielone. Filtry zostały wytrenowane podczas procesu uczenia sieci neuronowej opartej na architekturze Xception (Chollet, 2016).

Kluczowym zagadnieniem jest odkrycie takich parametrów transformacji, aby jak najlepiej pasowały one do problemu badawczego oraz danych wejściowych. Nie istnieje uniwersalny zestaw filtrów, ich określona liczba czy konfiguracja, która sprawdzi się w każdych warunkach. Odkrycie odpowiednich wartości



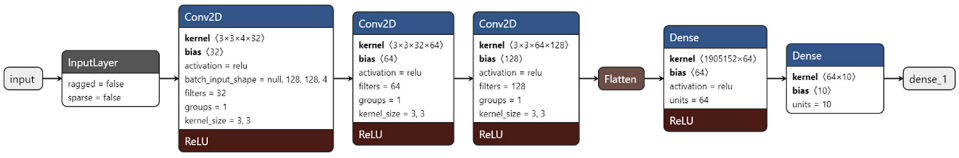
Ryc. 4. Obraz (por. z ryc. 2) poddany transformacji z wykorzystaniem czterech różnych filtrów.
Źródło: opracowanie własne.

i kombinacji filtrów jest niezmiernie trudne. Rozwiązaniem są sieci neuronowe, które dzięki możliwości trenowania są w stanie w iteracyjnym procesie odkryć odpowiednie wartości parametrów, ze względu na wybraną funkcję kosztu, i tym samym stworzyć dogodną dla procesu analizy reprezentację obrazu. Rycina 5 przedstawia postęp w trenowaniu wybranego filtra o rozmiarze 3 × 3 dla pojedynczego kanału w sieci Xception. Pierwsza macierz od lewej przedstawia filtr zainicjalizowany losowymi wartościami. Każda kolejna to udoskonalona w procesie uczenia wersja poprzedniego stanu. Komponentem sieci neuronowej umożliwiającym odnajdywanie optymalnego zestawu filtrów konwolucyjnych jest warstwa konwolucyjna (ang. *convolutional layer*).



Ryc. 5. Trenowanie filtra konwolucyjnego 3 × 3
Źródło: opracowanie własne.

Warstwy konwolucyjne można ze sobą łączyć. W przypadku pierwszego poziomu za wejście posłuży surowy obraz zapisany w postaci tablicy. Na kolejnych poziomach warstwy konsumują rezultaty wcześniejszych transformacji i uzyskując możliwość wydobycia coraz to większej liczby różnorodnych cech. Mówimy, że rośnie głębokość sieci, która tworzy w ten sposób trzon DCNN. W tym stanie sieć jest gotowa, aby poddać ją konfiguracji i przygotować do badania. Dla przykładu może stać się ona klasyfikatorem umożliwiającym przypisanie wejściowemu obrazowi jednej z kilku predefiniowanych klas (ryc. 6), detektorem odnajdującym wybrane obiekty znajdujące się na obrazie, rekonstruktorem poprawiającym jakość materiału wejściowego i wiele więcej. Warstwy konwolucyjne są głównym, ale nie jedynym komponentem, z których powstaje architektura sieci neuronowej. W celu nadania sieci pewnych właściwości należy uzbroić ją w dodatkowe komponenty umożliwiające modelowi skupienie się na pewnych aspektach obrazu, wytworzenie rezultatu w odpowiednim formacie, łączenie ze sobą kilku wcześniejszych warstw, redukcji wymiarowości i regularyzacji. Istnieje cały zestaw operacji oprócz konwolucji, tj.: *max pooling, dropout, concatenate, flattening, convolutional transpose, upsampling* i więcej, będących częściami składowymi modeli DL. Niektóre charakterystyczne złożenia operacji tworzą moduły, takie jak *residual Block* (He i in., 2015), *inception* (Szegedy i in., 2014) czy *attention* (Vaswani i in., 2017), które stosowane są w celu nadania ostatecznemu rozwiązaniu wymaganych właściwości. Dodatkowe informacje odnośnie do powszechnie stosowanych operacji czytelnik znajdzie w dokumentacji bibliotek wspierających DL, takich jak TensorFlow czy PyTorch.



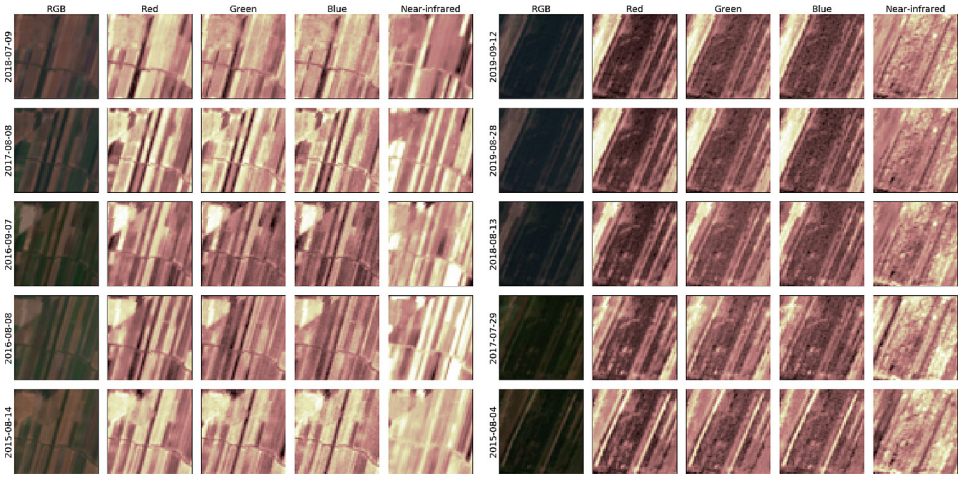
Ryc. 6. DCNN o trzech warstwach konwolucyjnych, o liczbie filtrów 32, 64, 128, umożliwiający klasyfikację obrazu do 10 kategorii
Źródło: opracowanie własne przy użyciu aplikacji netron.app.

Klasyfikacja

Celem klasyfikacji jest pogrupowanie materiału wejściowego przez przyporządkowanie każdego z jego elementów do pewnej kategorii. Zestaw kategorii lub inaczej klas jest ściśle określony. Klasyfikacja może odbywać się na poziomie całego obrazu, jego fragmentu albo każdego piksela z osobna. Praca na pojedynczym pikselu, często zawierającym informacje o wielu kanałach spektralnych, weszła do kanonu przetwarzania zdjęć lotniczych oraz obrazu satelitarnego i jest z powodzeniem stosowana w rozwiązaniach opartych na klasycznych metodach ML. Na uwagę zasługuje klasyfikator oparty na maszynie wektorów nośnych (ang. *support vector classifier*, SVC) oraz regresja logistyczna (ang. *logistic regression*) (Chen Y. i in., 2014). Wspomniane metody używane są również w odniesieniu do całego obrazu. Wymagają one jednak zaaplikowania potężnego warsztatu z zakresu inżynierii cech (ang. *feature engineering*). Poddany analizie zbiór pikseli reprezentujących zdjęcie musi ulec takiemu przetworzeniu, aby jednocześnie spełniał wymagania algorytmu i był adekwatny do postawionego problemu badawczego. Zadanie to sprowadza się do przedstawienia cech obrazu w postaci liczbowej. Dla przykładu już proste policzenie długości obiektów liniowych na obrazie, zliczenie punktów charakterystycznych czy wyrażenie obiektów przez zajmowaną przez nie powierzchnię jest przykładem inżynierii cech. Praktyka pokazuje, że w tym przypadku lepiej sprawdza się podejście z użyciem DCNN. Zakłada ono, że poznanie cech obrazu nastąpi w trakcie procesu trenowania sieci neuronowej. Zakres problemów badawczych możliwych do rozwiązania za pomocą klasyfikacji jest olbrzymi. Metody i techniki DL nie tylko wspomagają prace z zakresu analizy pokrycia terenu (Zhang i in., 2020), ale również umożliwiają rozpoznawanie gatunków flory i fauny (Kellenberger i in., 2019), detekcję zjawisk ekstremalnych (Anantrasirichai i in., 2019), a także identyfikowanie form geomorfologicznych (Ghorbanzadeh i in., 2019).

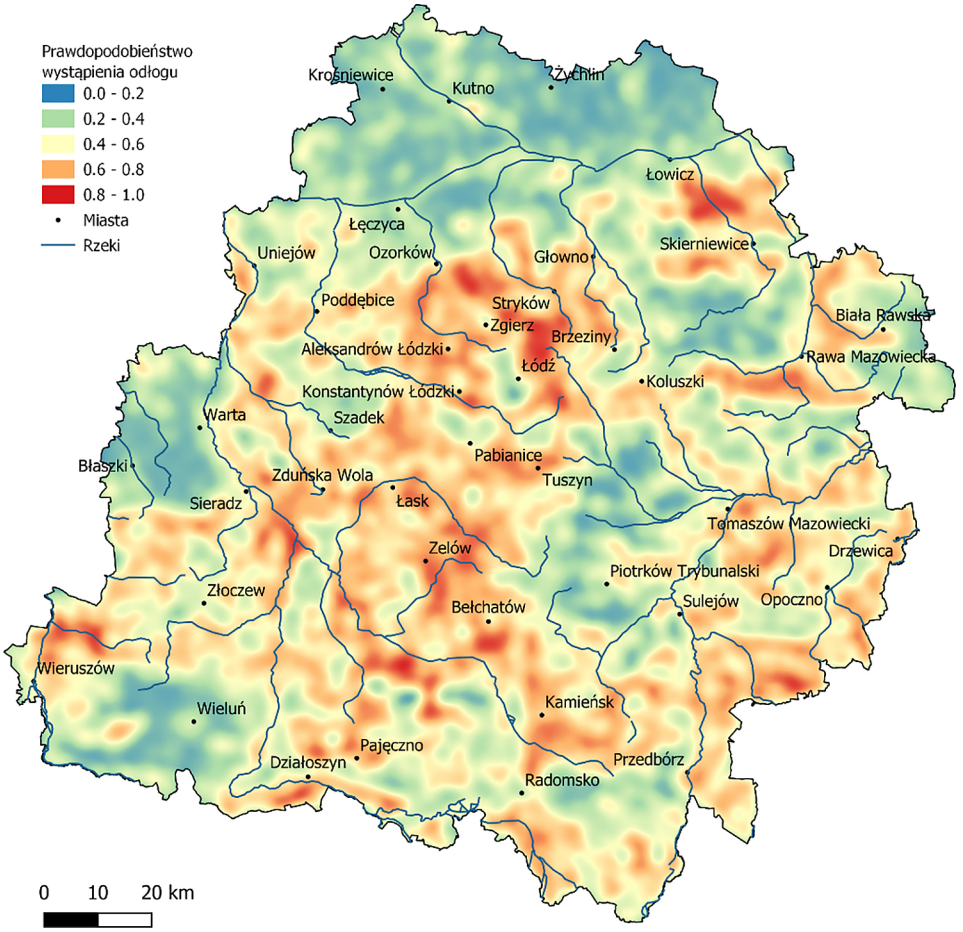
Przykładem zastosowania DCNN do wspomagania interpretacji przestrzeni geograficznej jest badanie występowania odłogów w województwie łódzkim (Krysiak i in., 2020). Zaprojektowano klasyfikator umożliwiający określenie, z jakim prawdopodobieństwem obraz obszaru o powierzchni 15÷25 ha przedstawia odłóg. Wykorzystano ogólnodostępne dane pochodzące z obrazowania satelitarnego (ryc. 7) oraz rezultaty projektu badawczego dotyczącego ekologicznej roli odłogów (Krysiak i in., 2016), na podstawie których opracowano szczegółową mapę ich występowania w strefach otaczających parki krajobrazowe w województwie łódzkim. Cyfrowa wersja mapy posłużyła do przypisania obrazom klas wymaganych podczas procesu uczenia nadzorowanego.

Automatycznej analizie poddano obszar przekraczający 18 000 km^2. Zrealizowanie podobnego projektu w sposób niezautomatyzowany nie byłoby możliwe ze względu na czas oraz koszty. Badanie to jest ciekawym przykładem



Ryc. 7. Dane wejściowe (Sentinel-2) do klasyfikatora odłogów w kompozycji barwnej RGB oraz kanałów spektralnych: R, G, B, NIR (od lewej: pola uprawne bez śladów odłogów i odłogi)
Źródło: Krysiak i in. (2020).

powrotu do wcześniej rozpatrywanego problemu badawczego w celu powtórzenia badań przy użyciu innej metody i jednoczesnym rozszerzeniu obszaru badawczego. Głównym zadaniem modelu DL jest wspomaganie badacza w interpretacji oraz dostarczenie mu dodatkowego materiału umożliwiającego opracowanie problemu badawczego. W badaniu wykorzystano standardową architekturę klasyfikatora opartą na DCNN. Różnica między przykładowym klasyfikatorem (ryc. 6) a użytym w badaniu polegała na głębokość sieci, tj. liczbie wykorzystanych warstw konwolucyjnych. Wejście do sieci stanowiły czterokanałowe wycinki obrazu satelitarnego o rozmiarze 50 px × 50 px. Rezultatem jest przypisanie każdemu wycinkowi wartości z przedziału 0 i 1 reprezentującego prawdopodobieństwo bycia odłogiem. Każdy z wycinków posiada georeferencję, dzięki czemu rezultat można nanieść na mapę (ryc. 8), a także poddać dalszemu procesowi analizy przestrzennej, jak to miało miejsce w cytowanym artykule, w którym obliczono statystykę lokalną Getisa i Orda (Gi*). Jest to przykład zastosowania ML jako kroku pośredniego umożliwiającego zdobycie lub przygotowanie danych do finalnego badania. Należy podkreślić, że rozwiązanie to nie zastępuje eksperta, którego wiedza posłużyła do poprawnego wytrenowania modelu. Wymaga również zbioru treningowego o dużej liczebności.



Ryc. 8. Mapa prawdopodobieństwa, że dany superpiksel jest odłogiem w województwie łódzkim wykonana z użyciem klasyfikatora opartego na DCNN
Źródło: Krysiak i in. (2020).

Semantyczna segmentacja

Celem segmentacji jest podział obrazu na charakterystyczne, rozłączne obszary, zwane regionami. Każdy region przypisany jest do co najmniej jednej ze ściśle określonych kategorii. W ramach pojedynczej kategorii może znaleźć się więcej niż jeden region lub może pozostać ona bez przypisania do niej regionu. W przypadku gdy dodatkowo można rozróżnić regiony wyznaczone w obrębie jednej kategorii, mówimy o segmentacji instancji. Dla przykładu proces wyznaczania obszaru zabudowanego widocznego na zdjęciu satelitarnym można nazwać segmentacją instancji, gdy istnieje możliwość wyznaczenia granicy między dwo-

ma sąsiadującymi lub nachodzącymi na siebie budynkami. W przypadku gdy informacja o indywidualnych jednostkach nie jest dostępna i analiza przebiega tylko w obrębie kategorii, inaczej klasy, mowa jest o semantycznej segmentacji. Segmentacja może zostać przeprowadzona w sposób nadzorowany (gdy liczba kategorii oraz warunki przynależności do każdej z nich są znane) lub nienadzorowany (gdy poznanie warunków przypisania do klasy jest również celem badania). Semantyczna segmentacja jest szczegółowym typem klasyfikacji, w której opracowaniu podlegają pojedyncze piksele. Wytrenowany model DL umożliwia wtedy przypisanie prawdopodobieństwa, z jakim każdy z pikseli należy do danej klasy. W zależności od oczekiwanego rezultatu końcowego dokonuje się wyboru kategorii z najwyższym prawdopodobieństwem, ustala wartości progowe lub zakłada, że dany obszar może należeć do więcej niż jednej kategorii. Poniżej zaprezentowano rezultat segmentacji obszarów leśnych w postaci mapy prawdopodobieństwa, tj. rastra o wartościach piksela z zakresu 0÷1 (ryc. 9). Im wyższa wartość, tym większa szansa, że dany piksel przedstawia las.

Ryc. 9. Mapa segmentacyjna (po lewej) oraz mapa prawdopodobieństwa (po prawej) uzyskana podczas analizy zdjęcia satelitarnego (Sentinel-2) obszaru Polski z użyciem modelu semantycznej segmentacji (U-Net)
Źródło: opracowanie własne.

Segmentacja wspiera rozwiązanie szeregu bardzo ciekawych problemów badawczych. Doskonałymi przykładami jej zastosowania są: rozpoznawanie typów pokrycia terenu, wykrywanie stopnia oraz typu zachmurzenia, delimitacja obszarów zabudowanych lub pomiar i kontrola zasiewów. Intuicyjnym wykorzystaniem segmentacji we wspomaganiu interpretacji przestrzeni geograficznej jest uzyskiwanie informacji o powierzchni zajmowanej przez określony typ pokrycia terenu. Rozwiązanie to należy brać pod uwagę w przypadku, gdy badany obszar

jest znacznych rozmiarów, jak obszar województwa czy kraju. Ze względu na dużą złożoność cech pokrycia terenu, wielkość i różnorodność informacji oferowanych przez obrazy satelitarne, opracowanie modelu zdolnego do określenia przynależności do danej kategorii jest zadaniem nietrywialnym. Tradycyjnie wykorzystywanymi algorytmami są SVM oraz *maximum likelihood classification* (MLC) (Hu i in., 2018). W porównaniu do modeli opartych na DCNN klasyfikatorom tym brakuje jednak dokładności. Olbrzymią zaletą DCNN jest tworzenie bogatego kontekstu powiązań między wydobytymi cechami obrazu. Jednocześnie DCNN wymaga dłuższego czasu trenowania oraz często większego zbioru danych wejściowych niż klasyczne algorytmy.

Spójrzmy, w jaki sposób semantyczna segmentacja oparta na modelu DL została zastosowana w badaniu związanym z wykrywaniem przyczyn rozbieżności między ewidencją lasów a stanem faktycznym (Adamiak i in., 2020). Główną metodą wykorzystaną w badaniu jest regresja ważona geograficznie (ang. *geographically weighted regression*, GWR), dla której zmienną objaśnianą jest różnica w wartościach wskaźnika lesistości obszaru gminy otrzymanego przez zastosowanie semantycznej segmentacji oraz jego wartości dostępnej w rejestrze gruntów. Badanie obejmowało obszar Polski. W procesie uczenia maszynowego użyto danych pochodzących z bazy danych obiektów topograficznych (BDOT) (Główny Urząd Geodezji i Kartografii, 2021) oraz obrazów satelitarnych Sentinel-2 z 2018 r., które to zostały wykorzystane odpowiednio do wykonania map segmentacyjnych oraz

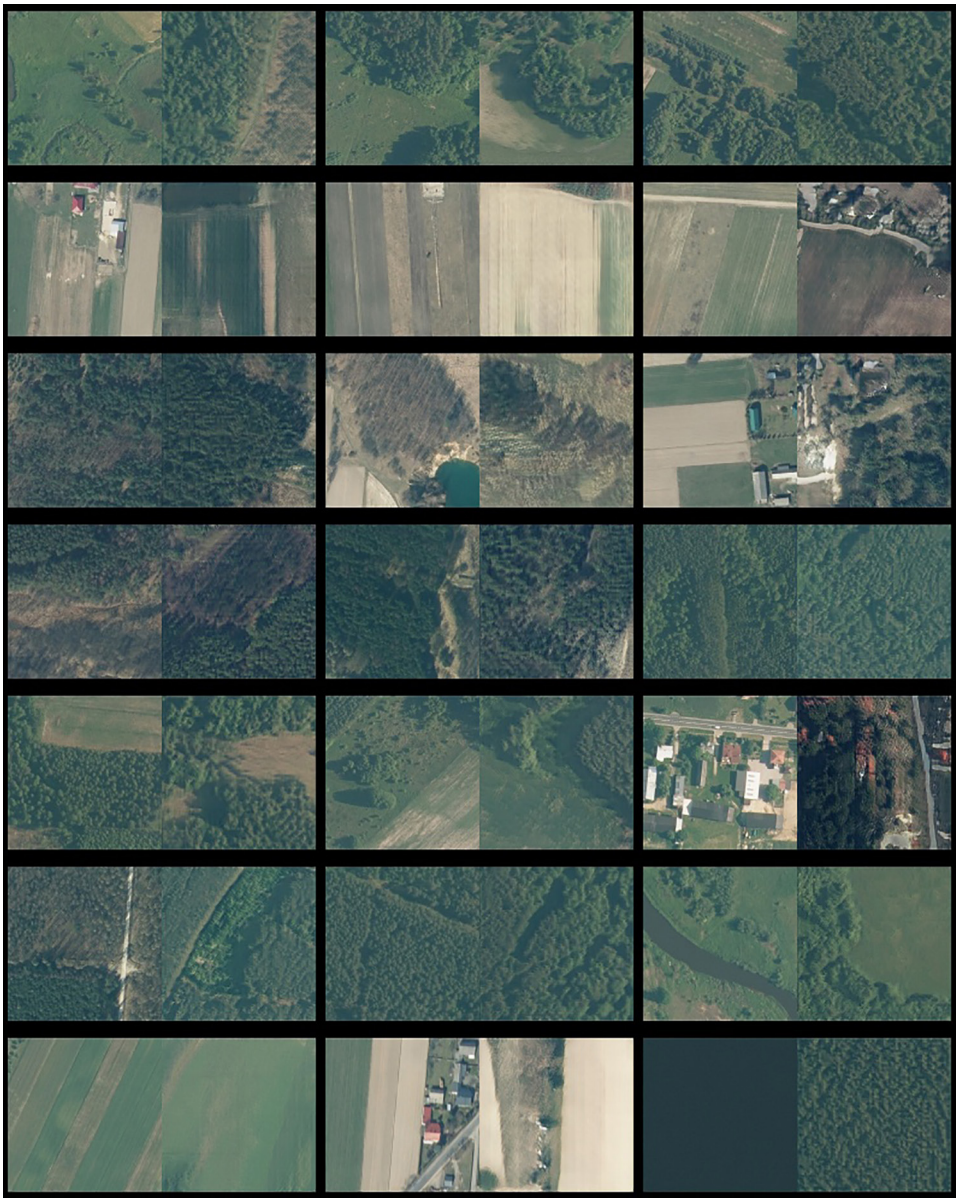
Ryc. 10. Przykład danych wejściowych segmentacyjnej sieci neuronowej będących połączeniem wycinka arkusza obrazu satelitarnego Sentinel-2 (po lewej) obszaru Polski oraz mapy segmentacyjnej powstałej przez przetworzenie odpowiedniej warstwy BDOT (po prawej)
Źródło: Adamiak i in. (2020).

jako materiał wejściowy do sieci neuronowej (ryc. 10). Mapa segmentacyjna (ang. *segmentation map*) zawiera informacje o przypisaniu każdego piksela badanego rastra do wybranej kategorii. Każdy z obrazów stanowiący zbiór uczący model ma odrębną mapę segmentacyjną. Piksel reprezentujący las oznaczono jedynką, a zerem w innym przypadku. Do segmentacji zdecydowano się wykorzystać model sieci neuronowej U-Net (Ronneberger i in., 2015). U-Net dostosowuje koncepcję DCNN do procesu wyznaczania homogenicznych regionów. Składa się z dwóch ścieżek: kontrakcyjnej i ekspansywnej. Ścieżka kontrakcyjna zbudowana jest z wielu następujących po sobie warstw konwolucyjnych, które konsekwentnie zmniejszają rozmiar obrazu i reprezentują jego cechy w coraz to większej liczbie kanałów (ang. *feature map*) (Çiçek i in., 2016). Ścieżka ekspansywna odwraca proces zmniejszania i stara się przy tym doprowadzić obraz do pierwotnego rozmiaru przy jednoczesnym przypisaniu jednolitych obszarów do osobnych kanałów. Każdy wyjściowy kanał tożsamy jest z inną kategorią pokrycia terenu. Bardziej złożoną alternatywą dla U-Net są: FPN (Lin i in., 2017), Linknet (Chaurasia, Culurciello, 2017), PSPNet (Zhao H. i in., 2017) oraz DeepLabV3+ (Chen L.-C. i in., 2017).

Augmentacja

Wydobywanie cech obrazu jest wpisane w mechanizm uczenia modelu DL. Informacje uzyskane w procesie tworzenia reprezentacji można wykorzystać dwojako. W sposób pośredni opis cech obrazu może być użyty jako forma przejściowa między wejściem do modelu, czyli surowym obrazem, a właściwym zadaniem analitycznym, jakim dla przykładu są klasyfikacja i segmentacja. Istnieje również możliwość wykorzystania reprezentacji jako docelowego obiektu badania. Warstwy ukryte sieci neuronowej udostępniają mechanizm kodujący obraz. Kodowanie jest możliwe, ponieważ sieć neuronowa w procesie uczenia opartego na zbiorze treningowym (ang. *training set*) pozyskuje z każdą iteracją coraz to więcej informacji o badanej domenie. Dzięki temu jest w stanie poddać interpretacji również obiekty, z którymi nie miała styczności podczas uczenia się. Sieć nasycona informacjami o przestrzeni geograficznej pochodzącymi z obrazowania lotniczego lub satelitarnego zapewnia nie tylko możliwość klasyfikacji czy segmentacji, ale stanowi również źródło nowych próbek. Innymi słowy, może stać się generatorem. Właściwość ta umożliwia realizację szeregu niezmiernie interesujących zadań, jakimi są: augmentacja, czyli wzbogacenie zbioru przez tworzenie nowych, syntetycznych, ale realistycznie wyglądających obrazów (Adamiak, Nalej, 2020; Zhao W. i in., 2020), usunięcie szumów i artefaktów (Burdziakowski, 2020), redukcja pokrycia obrazu chmurami przez wrysowanie w ich miejsce danych (*inpainting)* i wreszcie nienadzorowana segmentacja oraz inżynieria cech.

Siecią neuronową, której architektura oparta jest na DCNN i która umożliwia generowanie obrazu, jest GAN (ang. *generative adversarial network*). GAN zbudo-



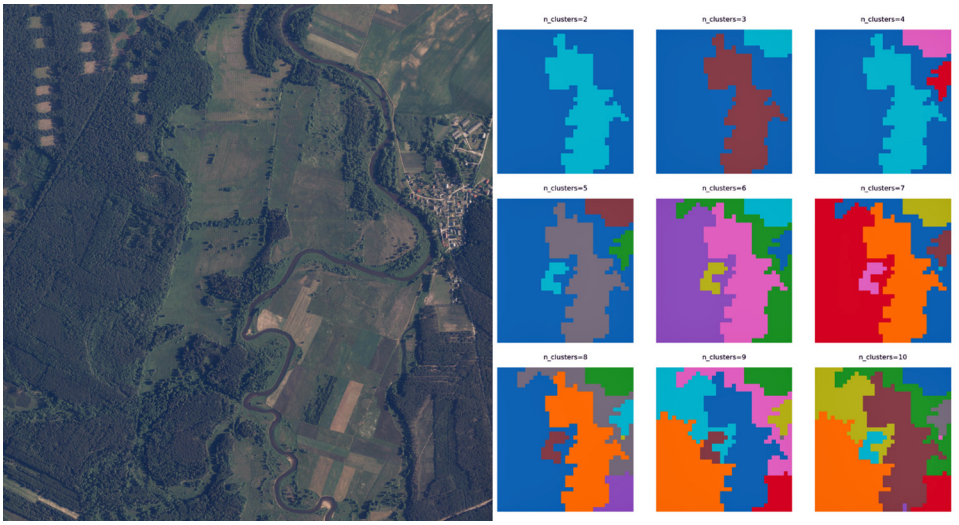
Ryc. 11. Rezultat rekonstrukcji 21 próbek testowych ortofotomapy przedstawiającej okolice rzeki Pilicy. Każda z sekcji zawiera oryginalny obraz (po lewej) oraz jego rekonstrukcję (po prawej)
Źródło: Adamiak i in. (2021), geoportal.gov.pl.

wany jest z co najmniej jednego dyskryminatora i jednego generatora. Zadaniem generatora jest wytworzenie takiego rezultatu, aby oszukać dyskryminator. Jednocześnie dyskryminator stara się być coraz lepszy w odgadywaniu zamiarów generatora. W trakcie procesu uczenia GAN obie sieci konkurują ze sobą, tworząc z czasem coraz doskonalsze rezultaty (Goodfellow i in., 2014). Rycina 11 jest rezultatem pracy modelu BigBiGAN (Donahue, Simonyan, 2019) wytrenowanego w oparciu o wycinki ortofotomapy przedstawiającej okolice rzeki Pilicy. Pierwszy wycinek każdej pary zawiera obraz referencyjny (prawdziwy). Drugi to rezultat procesu rekonstrukcji, w którym zdjęcie referencyjne zostało sprowadzone do uproszczonej reprezentacji w formie kodu liczbowego. Kod ten został następnie użyty, aby stworzyć syntetyczny obraz. Należy zauważyć, że w tym konkretnym przypadku rekonstrukcja nie zachodzi na poziomie pojedynczych piksel, lecz cech badanego obszaru. Innymi słowy, jej rezultatem nie jest identyczny wycinek względem referencyjnego, ale obraz o podobnej treści i tematyce. Możliwość tworzenia syntetycznego obrazu jest nie do przecenienia w przypadku ograniczonej dostępności danych wynikającej z kosztu ich pozyskania lub opracowania. Wygenerowane próbki mogą stać się częścią zbioru treningowego innego modelu ML lub zostać przeznaczone jedynie do jego testowania. Ostatnie rozwiązanie jest popularne ze względu na możliwość wytworzenia próbek przedstawiających rzadko występujące zjawiska, umożliwiając w ten sposób sprawdzenie modelu w skrajnych warunkach.

Inżynieria cech

Inżynieria cech to złożony proces, którego zadaniem jest transformacja zbioru danych i jego utrwalenie w formie adekwatnej do wybranej metody analitycznej. Przykładami tego procesu są: analiza obrazu w celu ekstrakcji cech lub wyodrębnienia informacji o występowaniu pewnego zjawiska, redukcja wymiarowości zbioru z użyciem analizy głównych składowych (ang. *principal component analysis,* PCA), usuwanie zbędnych lub kreowanie nowych atrybutów i wiele innych. Od jakości zabiegów przeprowadzonych na zbiorze w ramach inżynierii cech bezpośrednio zależy skuteczność opracowywanego modelu ML.

Dostęp do zaawansowanej inżynierii cech dostarcza GAN. W projekcie badawczym dotyczącym nienadzorowanego procesu segmentacji ortofotomap obszaru położonego nad środkowym biegiem Pilicy opracowano rozwiązanie (Adamiak i in., 2021), które w oparciu o GAN umożliwia wyrażanie badanego obszaru jako zbioru, wspominanych wcześniej kodów liczbowych. Kody te są ściśle powiązane z reprezentacją obrazu opisaną wewnątrz generatora. Pojedynczy kod zawiera informacje o sile, występowaniu i interakcji cech przestrzeni, której obraz posłużył do wytrenowania sieci. Kod ten wraz z możliwością przypisania georeferencji jest przydatny do wykorzystania w dalszej analizie. W cytowanym arty-



Ryc. 12. Obszar testowy 72961_840430_M-34-40-B-a-2-3 (Bobrowniki, województwo świętokrzyskie) oraz rezultat grupowania oparty na sztucznie wytworzonych kodach
Źródło: Adamiak i in. (2021), geoportal.gov.pl.

kule na zbiorze georeferencjonowanych kodów liczbowych przeprowadzono proces segmentacji z wykorzystaniem aglomeracyjnego grupowania metodą Warda (ryc. 12). Rezultat jest o tyle interesujący, że powstał w pełni automatycznie tylko na podstawie materiału wejściowego, jakim było 39 826 kwadratowych wycinków 512 px × 512 px przy rozmiarze piksela równym 0,25 m. Tego typu rozwiązania, choć mniej dokładne niż ręczna lub nadzorowana segmentacja, umożliwiają pracę z dużymi obszarami badawczymi. Mogą również prowadzić do odkrywania nieuchwytnych dla ludzkiej percepcji zależności, co ma wielkie znaczenie w analizie skomplikowanego materiału z obrazowania hiperspektralnego, a także stanowić specyficzny rodzaj kompresji. Mają też ograniczenia, z których główne związane jest ze stopniem skomplikowania procesu i trudności precyzyjnego wytłumaczenia, w jaki sposób dany model, zastosowany w inżynierii cech, doszedł do otrzymanego rozwiązania.

PODSUMOWANIE

Uczenie maszynowe udostępnia badaczom potężny zestaw metod przetwarzania oraz interpretacji przestrzeni geograficznej. Współcześnie możliwość wieloaspektowej i precyzyjnej analizy materiału pochodzącego z obrazowania lotniczego i satelitarnego jest nie do przecenienia. Możliwości pracy niewspieranej przez zautomatyzowany mechanizm badawczy są znacząco utrudnione w przypadku

częstej konieczności przetwarzania olbrzymiego wolumenu wysokiej jakości danych. Staje się ona jeszcze trudniejsza, gdy opracowanie danych wymagane jest w czasie rzeczywistym. Wykorzystanie ML w teledetekcji oraz naukach o Ziemi nie jest nowym zagadnieniem. Jednak dopiero współcześnie, głównie dzięki możliwościom obliczeniowym komputerów personalnych zaopatrzonych w wydajne karty graficzne (GPU, ang. *graphics processing unit)*, można wykorzystać ich cały potencjał. Kryje się on bezsprzecznie w uczeniu głębokim z użyciem sztucznych sieci neuronowych. Zaletą DL jest możliwość wsparcia pracy badacza w obrębie szerokiego spektrum problemów badawczych. Niniejsza praca traktuje o trzech powszechnie stosowanych technikach, tj.: klasyfikacji, semantycznej segmentacji, augmentacji oraz inżynierii cech w kontekście pracy z materiałem pochodzącym z obrazowania lotniczego i satelitarnego. Wśród innych wartych uwagi zagadnień są również: segmentacja instancji, rozpoznawanie aktywności, wykrywanie obiektów, detekcja nowości i anomalii, redukcja szumów i wiele więcej. Przykłady przybliżone w pracy potwierdzają, że elastyczność rozwiązań opartych na DL jest na tyle duża, że z powodzeniem można wpisać je w najbardziej zaawansowane projekty badawcze. Co więcej, umożliwiają one podejmowanie problemów, które wcześniej były niedostępne ze względu na swoją złożoność lub zakres przestrzenny opracowania.

DL stanowi nierozłączną część warsztatu narzędziowego specjalistów zajmujących się przetwarzaniem obrazu. Jest tylko kwestią czasu, gdy na stałe wpisze się w kanon geoinformacji, szeroko pojętej analizy przestrzennej oraz geografii. Nim to jednak nastąpi, DL przejdzie istotną transformację dostosowującą do wymogów współczesnego procesu naukowego. Mowa o możliwości przejrzystego wytłumaczenia działania sieci neuronowej czy algorytmu (*explainable AI*), tworzeniu rozwiązań deterministycznych (*deep evidential learning*) i zapewnieniu o uczciwości modelu (*bias and fairness*) (Amini, 2021). Jednocześnie intensywnie rozwijać się będą dziedziny wspierające, tj.: inżynieria danych (*data engineering*), inżynieria oprogramowania (*software engineering*), geograficzne systemy informacji (GIS), a także wizualizacja i prezentacja. Jeśli chodzi o bezpośredni wpływ na nauki o Ziemi i środowisku to wydaje się, że najwięcej korzyści zastosowanie DL może przynieść takim obszarom nauki, jak geografia fizyczna kompleksowa, hydrografia, gleboznawstwo czy geografia miast, które będą dzięki niemu mogły zwiększyć obszar badań przy jednoczesnym zachowaniu wysokiej precyzji wyników.

Oprogramowanie wykorzystywane współcześnie w pracach badawczych z zakresu geografii, tj.: ArcGIS, QGIS czy SAGA GIS, od dawna umożliwia rozwiązywanie przytoczonych w tekście problemów. Wspomniane aplikacje oferują olbrzymi, jednak skończony, zestaw funkcjonalności, dostępnych w ramach konkretnych wersji oprogramowania. Często dostępne pakiety okazują się niewystarczające. Rozszerzenie analizy o elementy DL możliwe jest wtedy przez zaimplementowanie odpowiedniej wtyczki lub pośrednio przez utworzenie sa-

modzielnego skryptu lub usługi współpracującej z popularnymi rozwiązaniami bazodanowymi czy formatami plików, takimi jak TIFF, SHP, CSV czy GeoJSON, np.: numpy, pandas, geo-pandas czy rasterio. Popularnymi bibliotekami wspierającymi uczenie maszynowe i głębokie uczenie są: Scikit-Learn, Apache Spark DL, Tensorflow i PyTorch. Oprócz bogatej dokumentacji popartej wieloma przykładami, badacz znajdzie dostęp do szerokiej oferty szkoleń. Warto również zapoznać się z osiągnięciami dziedziny inżynierii oprogramowania zajmującej się integracją i wdrożeniem rozwiązań uczenia maszynowego, którą jest Machine Learning Operations (MLOps). Doskonałym przykładem mogą być usługi Amazon SageMaker czy Google Vertex AI, które w równym stopniu nadają się do zastosowań komercyjnych, jak i w pracach badawczych.

Mimo przydatności metod DL oraz teledetekcji w interpretacji przestrzeni geograficznej nie należy zapominać o roli, jaką w pracy badawczej odgrywają badania terenowe. Zrozumienie badanego obszaru oraz zjawisk panujących w jego obrębie jest pierwszym i najważniejszym krokiem do wypracowania wysokiej jakości zbioru danych, który posłuży jako materiał do wytrenowania złożonego modelu DL. Proces budowania i trenowania sieci neuronowej bez względu na to, czy odbywa się w sposób nadzorowany czy nienadzorowany, wymaga udziału eksperta. Zadaniem eksperta jest poddać rezultat pracy modelu interpretacji oraz weryfikacji, a także dostarczyć wiedzy, która posłuży modelowi za wzorzec. Innymi słowy, metody DL są tak dobre jak specjaliści zajmujący się ich opracowaniem. Wprawne połączenie współcześnie dostępnych metod uczenia maszynowego z wiedzą z zakresu nauk o Ziemi pozwoli na podejmowanie nowych ekscytujących problemów badawczych oraz powracanie do wcześniej postawionych pytań.

Autor ma nadzieję, że niniejszy artykuł umożliwi czytelnikom ogólne zapoznanie się z tematem uczenia głębokiego we wspieraniu interpretacji przestrzeni geograficznej oraz będzie impulsem do wnikliwej eksploracji zagadnienia.

Podziękowania

Serdecznie podziękowania dla prof. Krzysztofa Będkowskiego oraz dr Anny Majchrowskiej za nadzór merytoryczny nad pracą oraz dla dr. Stanisława Krysiaka za wartościowe uwagi odnośnie do pracy.

LITERATURA

Adamiak M., Będkowski K., Majchrowska A. (2021). Aerial Imagery Feature Engineering Using Bidirectional Generative Adversarial Networks: A Case Study of the Pilica River Region, Poland. *Remote Sensing*, 13(2): 306. https://doi.org/10.3390/rs13020306

Adamiak M., Biczkowski M., Leśniewska-Napierała K., Nalej M., Napierała T. (2020). Impairing Land Registry: Social, Demographic, and Economic Determinants of Forest Classification Errors. *Remote Sensing*, 12(16): 2628. https://doi.org/10.3390/rs12162628

Adamiak M., Nalej M. (2020). *Generative adversarial networks in satellite image datasets augmentation*. Medium (https://blog.softwaremill.com/generative-adversarial-networks-in-satellite-image-datasets-augmentation-b7045d2f51ab).

Amini A. (2021). *MIT 6.S191. Deep Learning New Frontiers* (https://youtu.be/-boCMDouF2g).

Anantrasirichai N., Biggs J., Albino F., Bull D. (2019). A deep learning approach to detecting volcano deformation from satellite imagery using synthetic datasets. *Remote Sensing of Environment*, 230: 111179. https://doi.org/10.1016/j.rse.2019.04.032

Bengio Y., Courville A., Vincent P. (2014). Representation Learning: A Review and New Perspectives. *arXiv:1206.5538 [cs]* (http://arxiv.org/abs/1206.5538).

Bishop C.M. (2006). *Pattern Recognition and Machine Learning*. Springer-Verlag (https://www.springer.com/gp/book/9780387310732).

Boyd D.S. (2009). Remote sensing in physical geography: A twenty-first-century perspective. *Progress in Physical Geography: Earth and Environment*, 33(4): 451–456. https://doi.org/10.1177/0309133309346645

Burdziakowski P. (2020). A Novel Method for the Deblurring of Photogrammetric Images Using Conditional Generative Adversarial Networks. *Remote Sensing*, 12(16): 2586. https://doi.org/10.3390/rs12162586

Chaurasia A., Culurciello E. (2017). LinkNet: Exploiting encoder representations for efficient semantic segmentation. *2017 IEEE Visual Communications and Image Processing (VCIP)*, 1–4. https://doi.org/10.1109/VCIP.2017.8305148

Chen L.-C., Papandreou G., Schroff F., Adam H. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv:1706.05587 [cs]* (http://arxiv.org/abs/1706.05587).

Chen Y., Lin Z., Zhao X., Wang G., Gu Y. (2014). Deep Learning-Based Classification of Hyperspectral Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6): 2094–2107. https://doi.org/10.1109/JSTARS.2014.2329330

Chollet F. (2016). Xception: Deep Learning with Depthwise Separable Convolutions. *arXiv:1610.02357 [cs]* (http://arxiv.org/abs/1610.02357).

Çiçek Ö., Abdulkadir A., Lienkamp S.S., Brox T., Ronneberger O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. *arXiv:1606.06650 [cs]* (http://arxiv.org/abs/1606.06650).

Dabija A., Kluczek M., Zagajewski B., Raczko E., Kycko M., Al-Sulttani A.H., Tardà A., Pineda L., Corbera J. (2021). Comparison of Support Vector Machines and Random Forests for Corine Land Cover Mapping. *Remote Sensing*, 13(4): 777. https://doi.org/10.3390/rs13040777

DeepLearning.AI. (2021). Coursera (https://www.coursera.org/specializations/deep-learning?).

Donahue J., Simonyan K. (2019). Large Scale Adversarial Representation Learning. *arXiv:1907.02544 [cs, stat]* (http://arxiv.org/abs/1907.02544).

Dziennik Ustaw (2020). Dziennik Ustaw 2020 r. poz. 1086 (https://www.dziennikustaw.gov.pl/DU/2020/1086).

Geoportal (2020) (http://geoportal.gov.pl).

Ghorbanzadeh O., Blaschke T., Gholamnia K., Meena S.R., Tiede D., Aryal J. (2019). Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional

Neural Networks for Landslide Detection. *Remote Sensing*, 11(2): 196. https://doi.org/10.3390/rs11020196

Główny Urząd Geodezji i Kartografii (2021). *Zintegrowane kopie baz danych obiektów topograficznych BDOT10k*. Główny Urząd Geodezji i Kartografii (http://www.gugik.gov.pl/pzgik/zamow-dane/baza-danych-obiektow-topograficznych-bdot-10k).

Goodfellow I.J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. (2014). Generative Adversarial Networks. *arXiv:1406.2661 [cs, stat]* (http://arxiv.org/abs/1406.2661).

Gron A. (2019). *Hands-On Machine Learning with Scikit-Learn and Tensor Flow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Wyd. 2. O'Reilly Media, Inc.

He K., Zhang X., Ren S., Sun J. (2015). Deep Residual Learning for Image Recognition. *arXiv:1512.03385 [cs]* (http://arxiv.org/abs/1512.03385).

Hu Y., Zhang Q., Zhang Y., Yan H. (2018). A Deep Convolution Neural Network Method for Land Cover Mapping: A Case Study of Qinhuangdao, China. *Remote Sensing*, 10(12): 2053. https://doi.org/10.3390/rs10122053

Kałużna U., Będkowski K. (2020). Klasyfikacja pokrycia terenu z wykorzystaniem obrazów Sentinel-2A przetworzonych za pomocą metody głównych składowych (PCA). *Teledetekcja Środowiska*, 61 (http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-2e165147-5471-45fb-9cbe-69e90347749c).

Kamilaris A., Prenafeta-Boldu F. (2018). *A review of the use of convolutional neural networks in agriculture*. https://doi.org/10/gf4w4k

Kellenberger B., Marcos D., Lobry S., Tuia D. (2019). Half a Percent of Labels is Enough: Efficient Animal Detection in UAV Imagery Using Deep CNNs and Active Learning. *IEEE Transactions on Geoscience and Remote Sensing*, 57(12): 9524–9533. https://doi.org/10.1109/TGRS.2019.2927393

Krysiak, S., Adamczyk, J., Majchrowska, A., Papińska, E., i Sieradzki, J. (2016). *Ekologiczna rola odłogów stref otaczających parki krajobrazowe w województwie łódzkim*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Krysiak S., Papińska E., Majchrowska A., Adamiak M., Koziarkiewicz M. (2020). Detecting Land Abandonment in Łódź Voivodeship Using Convolutional Neural Networks. *Land*, 9(3): 82. https://doi.org/10.3390/land9030082

Kussul N., Lavreniuk M., Skakun S., Shelestov A. (2017). Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*, 14(5): 778–782. https://doi.org/10.1109/LGRS.2017.2681128

Landsat 8 – Landsat Science. (2020) (https://landsat.gsfc.nasa.gov/landsat-8/).

Lillesand T., Kiefer R., Chipman J. (2015). Remote Sensing and Image Interpretation. W *The Geographical Journal*, 146. John Wiley & Sons. https://doi.org/10.2307/634969

Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. (2017). Feature Pyramid Networks for Object Detection. *arXiv:1612.03144 [cs]* (http://arxiv.org/abs/1612.03144).

Malinowski R., Lewiński S., Rybicki M., Gromny E., Jenerowicz M., Krupiński M., Nowakowski A., Wojtkowski C., Krupiński M., Krätzschmar E., Schauer P. (2020). Automated Production of a Land Cover/Use Map of Europe Based on Sentinel-2 Imagery. *Remote Sensing*, 12(21): 3523. https://doi.org/10.3390/rs12213523

Mitchell T.M. (1997). *Machine Learning*. McGraw-Hill.

Pleiades – EoPortal Directory – Satellite Missions (2020) (https://earth.esa.int/web/eoportal/satellite-missions/p/pleiades).

Ronneberger O., Fischer P., Brox T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597 [cs]* (http://arxiv.org/abs/1505.04597).

Schmidhuber J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61: 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Sentinel-2 – ESA Operational EO Missions – Earth Online – ESA (2020) (https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/sentinel-2).

Sobel I. (1968). An Isotropic 3×3 Image Gradient Operator. *Presentation at Stanford A.I. Project 1968*.

Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. (2014). Going Deeper with Convolutions. *arXiv:1409.4842 [cs]* (http://arxiv.org/abs/1409.4842).

Tang Q., Gao H., Lu H., Lettenmaier D.P. (2009). Remote sensing: Hydrology. *Progress in Physical Geography: Earth and Environment*, 33(4): 490–509. https://doi.org/10.1177/0309133309346650

Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. (2017). Attention Is All You Need. *arXiv:1706.03762 [cs]* (http://arxiv.org/abs/1706.03762).

Zhang X., Han L., Han L., Zhu L. (2020). How Well Do Deep Learning-Based Methods for Land Cover Classification and Object Detection Perform on High Resolution Remote Sensing Imagery? *Undefined* (/paper/How-Well-Do-Deep-Learning-Based-Methods-for-Land-on-Zhang-Han/e71c80ed8bb174faed43e7a28a4e07dc2620dbb3).

Zhao H., Shi J., Qi X., Wang X., Jia J. (2017). Pyramid Scene Parsing Network. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6230–6239. https://doi.org/10.1109/CVPR.2017.660

Zhao W., Chen X., Chen J., Qu Y. (2020). Sample Generation with Self-Attention Generative Adversarial Adaptation Network (SaGAAN) for Hyperspectral Image Classification. *Remote Sensing*, 12(5): 843. https://doi.org/10.3390/rs12050843