# Automatyczna analiza rastrów metodą Deep Learning w GIS

*Projekt badawczo – rozwojowy GAZ-SYSTEM i Esri Polska nad wykorzystaniem metody Deep Learning w analizie obiektów na ortofotomapach – Etap I.*

Zagadnienia związane z przetwarzaniem obrazów są znane i stosowane od dawna. Zmianom ulegają same zobrazowania – stale zwiększa się ich rozdzielczość terenowa i spektralna, są coraz lepszej jakości, a ich pozyskiwanie staje się łatwiejsze. Przede wszystkim zmienia się jednak sprzęt i oprogramowanie umożliwiające obecnie wykonanie zaawansowanych analiz w krótszym czasie niż jeszcze kilka lat temu. Ponadto obsługa oprogramowania staje się coraz bardziej intuicyjna, co również wpływa na dostępność analiz obrazowych nawet dla osób z podstawową wiedzą w zakresie przetwarzania obrazów.

Wiele technik analizy obrazów wymaga jednak znacznego nakładu pracy ze strony operatora.

W przypadku klasyfikacji konieczne jest przygotowanie danych wejściowych (np. próbek/wzorców poszczególnych klas), które muszą spełniać określone warunki, aby proces klasyfikacji dostarczył zadowalających wyników.

Istnieją jednak narzędzia, które pozwalają znacząco zautomatyzować i przyspieszyć proces identyfikacji obiektów i klasyfikacji obrazów, zapewniając osiągnięcie często lepszych wyników. To narzędzia wykorzystujące Deep Learning, czyli zaawansowane Uczenie Maszynowe (ang. Machine Learning). Wraz z rozwojem i zwiększaniem się możliwości sprzętu komputerowego, narzędzia wykorzystujące Sztuczną Inteligencję (ang. Artificial Intelligence, AI) stają się coraz powszechniej dostępne. Są one również z powodzeniem stosowane w analizach obrazowych łącząc prostotę obsługi i uzyskiwanie bardzo

dobrych, dokładnych wyników.

**Koncepcja projektu**

Inspiracją dla idei projektu badawczo-rozwojowego „Deep Learning w GIS” w Spółce GAZ-SYSTEM były zagadnienia poruszane podczas konferencji programistycznej European DevSummit 2019, organizowanej przez firmę Esri Inc. w Berlinie. W ślad za nimi, podstawą realizacji projektu stało się wykorzystanie narzędzi Deep Learning udostępnionych w oprogramowaniu ArcGIS Pro firmy Esri od wersji 2.5, jako podtypu nauczania maszynowego do automatycznej identyfikacji obiektów na rastrach.

**Problem badawczy**

Spółka GAZ-SYSTEM (dalej: Spółka) posiada bogate, stale poszerzane archiwum danych rastrowych, wśród których główną rolę odgrywają ortofotomapy czyli mapy powstałe wskutek przetworzenia zdjęć lotniczych do postaci zapewniającej ich kartometryczność (proces ortorektyfikacji polega na usunięciu ze zdjęć wszelkich zniekształceń spowodowanych np. wpływem rzeźby terenu, co zapewnia stałą skalę ortofotomapy, a proces mozaikowania pozwala ujednolicić kolorystykę zdjęć i przedstawić cały obszar opracowania na jednym, ciągłym obrazie). Firma pozyskuje te dane m.in. w ramach dedykowanych zleceń przetargowych odbywających się co parę lat. Obecnie wolumen danych wykorzystywany jest jako podkład mapowy w ramach aplikacji GIS, np. Portal GIS lub w systemie paszportyzacji w ramach EAM.

Zauważono, że dzięki analizie ortofotomapy, możliwe byłoby uzyskanie informacji o określonych nieprawidłowościach, które mają bezpośredni wpływ na bezpieczeństwo działania infrastruktury gazowej. W związku z faktem, że obszar działań Spółki ma zasięg ogólnopolski, analizę należałoby przeprowadzać w sposób automatyczny. Analiza manualna jest w tym przypadku zbyt czasochłonna, a wynik zależałby od

doświadczenia i wytrwałości operatora.

Uczenie maszynowe i jej podtypy dostępne są na rynku od wielu dostawców, ale wymagają złożonych działań programistycznych, z różnych powodów niedostępnych obecnie dla GAZ-SYSTEM. Funkcjonująca w Spółce umowa i posiadanie od lat oprogramowania od Esri stanowią najprostszą organizacyjnie, wydajną kosztowo i korzystną pod względem badawczym możliwość zbudowania takiego narzędzia do analiz. Pomysł realizacji wspólnego projektu, ze względu na nowatorstwo metody, dla obu stron jest innowacyjnym na skalę krajową działaniem w tym obszarze. Efektem wspólnych prac badawczych będzie narzędzie dla GAZ-SYSTEM do automatycznej analizy rastrowej.

**Realizacja projektu**

Projekt zakłada wykorzystanie gotowych komponentów informatycznych Esri, a jego realizacja została podzielona na kilka etapów. Obecnie zrealizowano Etap I – teoretyczny będący solidnym fundamentem w realizacji całego projektu.

Realizacja Etapu I:

- Zgromadzenie i scharakteryzowanie danych rastrowych Spółki oraz określenie ich przydatności w kontekście projektu Deep Learning,
- Analiza rynku (firmy, badania naukowe) pod kątem podobnych rozwiązań, tj. stosujących uczenie maszynowe w analizie rastrowej,
- Zbadanie możliwości wykonywania ortofotomapy „siłami własnymi” przy wykorzystaniu bezzałogowych statków powietrznych – dronów,
- Ocena przydatności danych pochodzących z dronów do realizacji projektu,
- Określenie metod i narzędzi do automatycznej analizy rastrowej metodą Deep Learning w środowisku Esri.

***Zgromadzenie i scharakteryzowanie danych rastrowych Spółki oraz określenie ich przydatności w kontekście projektu Deep***

***Learning***

Do testów przekazane zostały 3 zestawy arkuszy ortofotomapy RGB, opracowane na podstawie fotogrametrycznych zdjęć lotniczych różniących się m.in. zakresem przestrzennym oraz rozdzielczością terenową. Ortofotomapy te okazały się być bardzo dobrej jakości, jednorodne radiometrycznie i bez zniekształceń geometrycznych.

Poza opisanymi wyżej ortofotomapami lotniczymi przekazano kilka zestawów ortofotomap wykonanych przez firmę zewnętrzną na podstawie zdjęć z dronów, przedstawiających wybrane obiekty znajdujące się na analizowanym obszarze, które były zmieniane/rozbudowywane w okresie pozyskiwania zdjęć.

Zobrazowania z tego zestawu danych różniły się jakością od zdjęć lotniczych. Zostały one zapisane w postaci jednokanałowego rastra z przypisaną Tabelą Kolorów (ang. Colormap) zawierającą 256 zdefiniowanych kolorów. Są wystarczające jeżeli chodzi o pozyskanie danych geometrycznych czy fotointerpretację obrazów, jednak w przypadku zautomatyzowanych analiz, a zwłaszcza analiz wykorzystujących technologię Deep Learning, będą raczej nieprzydatne, ponieważ znacząco różnią się od zobrazowań RGB zapisanych jako trójkanałowe obrazy 24-bitowe (zawierające ponad 16,7 mln kolorów). Porównanie tych zobrazowań znajduje się na rysunku 1.

Rys.1: Porównanie zobrazowań RGB (po lewej) i zobrazowań wykorzystujących paletę barw Colormap (po prawej)

Poza opisanymi wyżej ortofotomapami przekazano 2 ortofotomapy opracowane na podstawie zdjęć pozyskanych przez pracowników GAZ-SYSTEM przy pomocy bezzałogowych statków lotniczych.

Na podstawie wykonanych zdjęć, używając oprogramowania Esri – ArcGIS Drone2Map – pracownicy GAZ-SYSTEM wygenerowali 2 ortofotomapy „testowe", które zostały przesłane na serwer i udostępnione w ramach Portalu GIS. Produkty były dobrej jakości, zawierając jedynie drobne błędy wynikające z automatycznego przetwarzania – ich przykłady na rysunku 2.

Rys. 2: Przykłady drobnych błędów na ortofotomapach opracowanych przez GAZ-SYSTEM.

a) Błąd mozaikowania (łączenia zdjęć) widoczny na dachu kontenera.

b) Pofalowana krawędź dachu oraz zniekształcenia obiektów wynikające z niedoskonałości modelu powierzchni terenu podczas ortorektyfikacji zdjęć

c) „Duch" – powtórzona część obrazu obiektu (naczepy ciężarówki)

O ile takie błędy nie będą znacząco większe lub będą występowały poza głównym obszarem zainteresowania, (należy rozważyć ewentualną zasadność ich usuwania, aby uniknąć dodatkowego nakładu pracy) zobrazowania tego typu mogą być przydatne zwłaszcza w przypadku wystąpienia konieczności szybkiego lub cyklicznego pozyskiwania danych dla niewielkich obszarów.

***Zbadanie możliwości wykonywania ortofotomapy „siłami własnymi" przy wykorzystaniu bezzałogowych statków powietrznych – dronów***

Jedne z pierwszych wniosków Etapu I określiły potrzebę pozyskiwania dodatkowych zdjęć lotniczych ze względu na znikome zróżnicowanie terenu na dostarczonych ortofotomapach. Gazociągi przesyłowe, jeśli tylko jest to możliwe, budowane są na obszarach słabo zurbanizowanych, a to przekłada się na treść zdjęć, na których przeważają tereny uprawne. W celu wytrenowania modelu wykrywającego np. drogi potrzeba dostarczyć zdjęcia z jak największą liczbą „próbek drogowych". Zdecydowano się zaangażować służby eksploatacyjne Spółki,

które są wyposażone w bezzałogowe statki lotnicze.

Operatorzy „bezzałogowców” wykorzystując dostępne drony pozyskali zdjęcia lotnicze obszarów testowych wyznaczonych do Etapu I, a zdobyty materiał fotograficzny pozwolił na sporządzenie ortofotomap.

Okresowe aktualizowanie ortofotomap pozwala na:

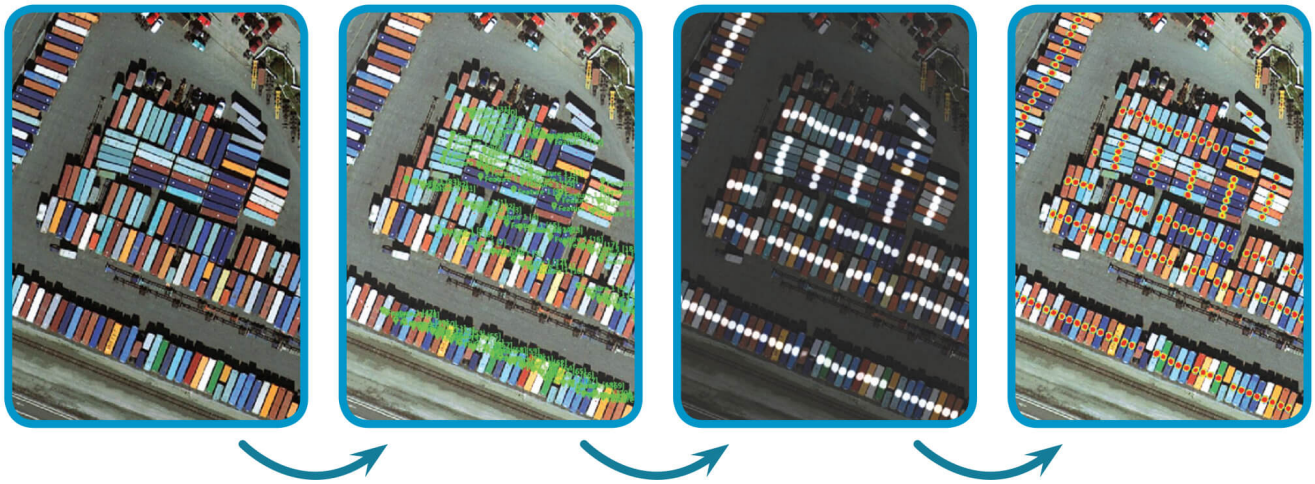
- zobrazowanie zmian na terenach zurbanizowanych w celu zminimalizowania zagrożenia w strefie kontrolowanej gazociągu,
- zarejestrowanie zmiany terenu na gruntach niestabilnych oraz podjęcie działań abezpieczających,
- dokładne udokumentowanie i oszacowanie zajętego terenu w miejscach po zdarzeniach awaryjnych.

Ortofotomapy przekazują informację o podjętych działaniach dużo dokładniej niż pojedyncze zdjęcia. Osobie decyzyjnej ułatwiają podejmowanie odpowiednich działań w stosunku do sytuacji awaryjnej zaistniałej na infrastrukturze gazowej.

**Metody i narzędzia do automatycznej analizy rastrowej techniką Deep Learning w środowisku Esri.**

***Metoda***

Analiza zobrazowania z wykorzystaniem Deep Learning składa się z kilku etapów.



Rys. 3. Analiza zobrazowania z wykorzystaniem Deep Learning. Od lewej: 1. Obraz treningowy; 2. Etykiety obiektów treningowych; 3. Identyfikowane obiekty; 4. Zidentyfikowane obiekty na tle oryginalnego obrazu. Źródło: L3 Harris.

## 1. Etykietowanie przykładowych obiektów (wskazywanie wzorców)

Jest to etap wymagający największego zaangażowania ze strony operatora. Jednakże, w przypadku Deep Learning, praca i czasochłonność nie wynika z konieczności jak najlepszego doboru danych przykładowych, ale z ich masowości i kompletności. W zależności od potrzeb możliwe jest trenowanie modeli mających identyfikować na zobrazowaniach obiekty konkretnego rodzaju (np. białe samochody, drzewa, budynki) lub identyfikować jedną lub więcej klas użytkowania terenu (np. siatka dróg, pola uprawne, lasy). Etykietowanie obiektów wykonuje się na obrazach (lub fragmentach obrazów) przeznaczonych do późniejszego trenowania modeli i zazwyczaj zalecane jest wskazanie wszystkich obiektów danego typu występujących na takim obrazie. Takie oznaczenie może odbywać się poprzez wskazanie lokalizacji obiektu (np. drzewa, dane punktowe), określenie jego przebiegu (np. drogi, dane liniowe) lub dokładne jego obrysowanie (np. budynki, dane poligonowe). Podczas trenowania modelu wszystko, co nie zostało zaetykietowane, traktowane jest jak „wartość tła”, a więc jako element niepożądany. Dlatego tak ważne jest wskazanie na rastrach treningowych wszystkich „pożądanych” obiektów.

## 2. Trenowanie modeli Deep Learning

Etap trenowania modelu jest procesem w pełni automatycznym, poprzedzonym określeniem przez operatora parametrów trenowania takich, jak np.: liczba epok (iteracji procesu), rozmiar próbki, „ostrość” krawędzi obiektów, wartości progowania dla klasyfikowania pikseli. Zestawy parametrów zależą od wybranego środowiska (np. Deep Learning w ArcGIS Pro) oraz rodzaju modelu i wpływa na długość trwania procesu trenowania, dokładność i sposób identyfikacji obiektów, a co za tym idzie późniejszych wyników przetwarzania rastrów za pomocą narzędzi Deep Learning. O ile wspomniany wcześniej proces etykietowania wymaga dużych nakładów pracy i czasu ze strony operatora, o tyle proces trenowania modelu jest najbardziej czasochłonnym procesem z punktu widzenia zasobów sprzętowych. Do jego sprawnego przeprowadzenia niezbędne są komputery wyposażone w wysokowydajne karty i procesory graficzne. Ponieważ analizowane zestawy danych rastrowych mogą przedmiotem zainteresowania może być klasyfikacja obrazów, wykrywanie obiektów lub „zestawów” obiektów, trudno w tym momencie określić konkretną konfigurację sprzętową niezbędną do sprawnego działania wszystkich procesów.

Wydajność systemu zależeć będzie od wielu elementów, wśród których należy wspomnieć szybkość dostępu do zestawów obrazów, wydajność dysków wykorzystywanych bezpośrednio podczas przetwarzania czy ogólną wydajność samych stacji roboczych. Bardzo ważnym elementem konfiguracji sprzętu jest karta graficzna wspierająca technologię CUDA, wykorzystywana głównie podczas trenowania modeli (konfiguracja Deep Learning w ArcGIS Pro) oraz niektórych procesów klasyfikacyjnych wykorzystujących Deep Learning.

**3. Klasyfikacja i identyfikacja obiektów na docelowym rastrze**

Po wytrenowaniu modelu możliwe jest jego wykorzystanie do klasyfikacji zobrazowania lub identyfikacji obiektów na obrazach. W zależności od otrzymanych wyników wytrenowany model może zostać uznany za ostateczny lub możliwe jest jego ponowne trenowanie z wykorzystaniem większej liczby danych

przykładowych (pochodzących np. z wcześniejszego procesu klasyfikacji).

**4. Analiza i ocena wyników**

Analiza otrzymanych wyników polega na weryfikacji ich poprawności oraz ewentualnym „prześledzeniu" produktów procesów cząstkowych i określeniu nowych wartości lub zestawów parametrów wejściowych (parametry przetwarzania, dane etykietowane), które zapewnią otrzymanie lepszych wyników.

**5. Uzupełnienie wzorców i ponowne trenowanie modelu**

Trenowanie modelu jest procesem iteracyjnym nie tylko „wewnętrznie" (te powtórzenia nazywane są epokami, ang. epochs), ale również uwzględniając dodawanie nowych, etykietowanych danych i uruchamiając proces trenowania modelu ponownie. Możliwe jest również zapisanie pewnego „bazowego", wytrenowanego modelu i „dotrenowywanie" go w przypadku konieczności wykorzystania do analizy nieco innych zestawów danych. W takim przypadku takie „dotrenowanie" może być szybsze i łatwiejsze niż przygotowanie modelu od początku.

***Narzędzie***

W badaniach prowadzonych w ramach projektu badawczo□rozwojowego Spółki GAZ-SYSTEM – „Deep Learning w GIS" – wykorzystano narzędzia Deep Learning dostępne w ArcGIS Pro 2.7.

ArcGIS Pro, produkcji Esri Inc., to wszechstronny system GIS, integrujący różnorodne narzędzia do pozyskiwania, przetwarzania oraz analiz danych wektorowych, rastrowych oraz statystycznych pochodzących z różnych źródeł. Wśród nich znajdują się narzędzia wykorzystujące Deep Learning do analizy danych rastrowych.

Narzędzia Deep Learning są dostępne w pakiecie ArcGIS Pro od wersji 2.5, ale możliwość przeprowadzenia całego procesu

(włącznie z trenowaniem modeli) pojawiła się w wersji 2.6 i jest stale rozwijana. W wersji 2.7 lista dostępnych modeli została rozszerzona i oferuje obecnie możliwość trenowania aż 9 różnych typów modeli Deep Learning, opcjonalnie wykorzystując jedną z kilkunastu dostępnych wstępnie skonfigurowanych i wytrenowanych sieci neuronowych (ang. Transfer Learning). Narzędzia Deep Learning w ArcGIS Pro w połączeniu z wydajnymi narzędziami edycyjnymi i bogatą funkcjonalnością analityczną systemu zapewniają szereg możliwości przygotowywania danych wejściowych oraz analizowania wyników.

Obecnie, w zależności od potrzeb, ArcGIS Pro 2.7 oferuje możliwość trenowania modeli Deep Learning przeznaczonych do:

- Wykrywania obiektów na obrazie (ang. Detect Objects Using Deep Learning). W wyniku powstaje warstwa wektorowa zawierająca lokalizację zidentyfikowanych obiektów zapisaną jako obwiednie obiektów (ang. Bounding Box), poligony z obrysami obiektów lub punkty na środku obiektów. Modele przeznaczone do wykrywania obiektów na obrazie to: Single Shot Detector (SSD), RETINANET, MaskRCNN, YOLOv3 oraz FasterRCNN.
- Klasyfikacji obiektów na obrazie (ang. Classify Objects Using Deep Learning). W wyniku, na podstawie danych z obrazu, klasyfikowanym obiektom zostaje przypisana jedna ze zdefiniowanych klas, np. budynek z/bez ogniwami fotowoltaicznymi na dachu, działka z/bez basenem, rodzaj samochodu itp. Informacja taka może być dopisana do tabeli atrybutów warstwy wektorowej, która może być wykorzystywana jako dodatkowa dana wejściowa w procesie klasyfikacji. Model przeznaczony do klasyfikacji obiektów na podstawie danych z obrazu to Feature Classifier.
- Klasyfikacji pikseli obrazu (ang. Classify Pixels Using Deep Learning) W wyniku powstaje nowy, sklasyfikowany obraz zawierający piksele przypisane do poszczególnych

klas. Modele przeznaczone do klasyfikacji pikseli obrazu to: U-Net, Pyramid Scene
Parsing Network (PSPNET) oraz DeepLabV3.

Rys. 4. Wyniki różnego rodzaju przetwarzania obrazów z wykorzystaniem Deep Learning: wykrywanie obiektów (po lewej), klasyfikacja obiektów (w środku) i klasyfikacja pikseli (po prawej). Źródło: www.esri.com

W zależności od wybranego rodzaju modelu, otrzymane wyniki będą różnić się dlatego przed przystąpieniem do trenowania modelu konieczne jest zdefiniowanie oczekiwanych wyników. Taka wiedza jest również niezbędna na etapie przygotowywania i etykietowania danych, ponieważ narzędzia ArcGIS Pro wymagają zapisania etykietowanych danych w konkretnych formatach, w zależności od wybranego modelu Deep Learning.

**Widoki na przyszłość**

Zadania wynikające z założeń Etapu I projektu „Deep Learning w GIS" zostały zrealizowane, a wszelkie materiały i wnioski powstałe w wyniku prowadzonych prac badawczych posłużą do realizacji jego kolejnych kroków. Spodziewanym efektem całego projektu ma być kompletne środowisko do automatycznych analiz rastrowych, działające w oparciu o wypracowane modele na przestrzeni realizacji wspólnego przedsięwzięcia.