



Inteligentny system identyfikacji zanieczyszczenia powietrza

Intelligent air pollution identification system

Mirosław Szwed 

Uniwersytet Jana Kochanowskiego w Kielcach
Instytut Geografii i Nauk o Środowisku
ul. Uniwersytecka 7, 25-406 Kielce
miroslaw.szwed@ujk.edu.pl

Zarys treści. Celem opisanego w pracy działania jest opracowanie inteligentnego systemu identyfikacji zanieczyszczeń powietrza. Zastosowanie sztucznej inteligencji, wykorzystującej analizę obrazów powierzchni wybranych indykatorów zanieczyszczenia powietrza do budowy algorytmu uczenia maszynowego umożliwiło opracowanie taniej i skutecznej metody identyfikacji niebezpiecznych substancji. Do budowy modelu zostały wykorzystane zdjęcia skaningowej mikroskopii elektronowej dwuletnich igieł sosny zwyczajnej *Pinus sylvestris* L., z reprezentatywnych zlewni badawczych krajowej sieci Zintegrowanego Monitoringu Środowiska Przyrodniczego. Zdjęcia mikroskopowe zostały przetworzone w programie graficznym, tak aby zaklasyfikowane na podstawie wielkości, kształtu i składu chemicznego cząstki posiadały jednakowy atrybut (barwę). Wykonane warstwy (maski) stanowiły element właściwy do opracowania algorytmu uczenia maszynowego identyfikującego zanieczyszczenia z podziałem na zdefiniowane wcześniej kategorie. Zastosowanie sieci neuronowych do budowy samouczącego się algorytmu pozwoliło zoptymalizować analizę zdeponowanych zanieczyszczeń zobrazowanych na powierzchni igieł sosny. Opracowany system identyfikacji naturalnych i antropogenicznych cząstek w postaci skategoryzowanych warstw daje skuteczność predykcji na wysokim poziomie.

Słowa kluczowe: identyfikacja zanieczyszczeń powietrza, sztuczna inteligencja, uczenie maszynowe, mikroskopia elektronowa, sieci neuronowe, analiza obrazów.

Keywords: *air pollution identification, artificial intelligence, machine learning, electron microscopy, neural networks, image analysis.*

Wstęp

Jakość powietrza atmosferycznego w wielu rejonach świata stanowi poważny problem (Akimoto, 2003; Jiménez et al., 2014; Samet i Chung, 2017). Ma on wymiar zarówno środowiskowy jak i społeczny (Keywood et al., 2023). Przekłada się na obniżenie warunków do życia i stanowi czynnik ograniczający dalszy rozwój (WHO, 2021; Shaddick et al., 2023). Lokalne oraz zdalne emisje zanieczyszczeń mogą być transportowane na dalsze odległości (Konwencja LRTAP, 1979), a wskutek niekorzystnej sytuacji synoptycznej sprzyjającej depozycji w danym regionie mogą powodować przekroczenia dopuszczalnych poziomów stężeń poszczególnych substancji w powietrzu (Kozáková et al., 2018; Huszar et al., 2021; WHO, 2021). Szczególnie niebezpieczne dla życia i zdrowia ludzi są zjawiska nagromadzenia zanieczyszczenia w powietrzu w wyniku spalania paliw kopalnych w okresie grzewczym, występujące w wielu rejonach świata, w tym w Europie Środkowej (Allegrini i Pirrone, 2007; Fowler et al., 2020; Susanto, 2020). Urządzenia pomiarowe umożliwia-

jące automatyczny pomiar zanieczyszczeń gazowych i pyłowych stanowią skomplikowaną i stosunkowo drogą infrastrukturę (Ayaz et al., 2018), dlatego poszukiwane są tanie i proste rozwiązania w ocenie zanieczyszczenia powietrza (Arroyo et al., 2021; Bousiotis et al., 2021), w tym wykorzystujących metody bioindykacyjne (Fan, 2014). Wyniki przeprowadzonych badań wskazują, że błąd pomiaru kompaktowych czujników (niskokosztowych), który może być znaczny, zależy od technologii (neflometr czy optyczny licznik cząstek), konkretnych parametrów danego czujnika oraz różnic pomiędzy aerozolem użytym do kalibracji czujnika a aerozolem mierzonym (Hagan i Kroll, 2020). Z kolei metody bioindykacyjne mają zastosowanie w przypadku braku zorganizowanej infrastruktury pomiarowej oraz badań skutków środowiskowych zdarzeń o charakterze katastrofalnym (Świercz et al., 2024). Szybki rozwój sztucznej inteligencji AI (z ang. *Artificial Intelligence*), zdeterminował próbę jej zastosowania w detekcji zanieczyszczeń powietrza. AI w zaproponowanej metodzie badań, pozwoliło znacznie zoptymalizować proces przetwarzania i analizy uzyskanych danych. Materiał wejściowy stanowiły mikrofotografie przetwarzane przez głęboką sieć neuronową, w taki sposób, aby na wyjściu (wynik) otrzymywać obrazy z zaznaczonymi obiektami (zanieczyszczeniami). W analizie mikroskopowej wykorzystano igliwie sosny zwyczajnej *Pinus sylvestris* L., pobrane w lokalizacjach o zmiennych warunkach antropopresji. Niniejsza praca skupia się na wykorzystaniu klasyfikacji zanieczyszczenia z podziałem na zdefiniowane kategorie (barwy), a następnie odtwarzaniu maski dla zupełnie nowych danych wejściowych, które nie wymagają wcześniejszego przetworzenia zdjęcia. Dzięki zastosowaniu sieci neuronowej do predykcji obiektów, został opracowany inteligentny system identyfikacji zanieczyszczeń powietrza oparty na samouczącym się algorytmie maszynowym.

Koncepcja i metodyka badań

W zaproponowanej metodzie identyfikacji zanieczyszczeń powietrza zobrazowaniu podlegały cząstki na powierzchni popularnego biowskaźnika – igieł sosny zwyczajnej *Pinus sylvestris* L. Zastosowanie skaningowej mikroskopii elektronowej – SEM (ang. *Scanning Electron Microscopy*) pozwoliło uzyskać mikrofotografie, które następnie zostały przetworzone w programie graficznym, tworząc warstwy (maski) z nadanymi odpowiednimi barwami (atrybutami). Charakterystyczne cząstki zanieczyszczeń były następnie przetwarzane i analizowane przez głęboką sieć neuronową. Sieci neuronowe, inspirowane strukturą połączeń w ludzkim mózgu (choć w rzeczywistości różniące się znacząco od jego działania), stanowią zaawansowaną technikę modelowania, umożliwiającą odwzorowanie skomplikowanych, nieliniowych funkcji. Ich zasada działania opiera się na interakcjach pomiędzy danymi wejściowymi a punktami pośrednimi, zwanymi neuronami. W każdym neuronie dane wejściowe są mnożone przez odpowiednie wagi, następnie sumowane, a wynik jest przetwarzany przez funkcję aktywacji. Wyjście z neuronu służy jako wejście dla neuronów w kolejnej warstwie, kontynuując ten proces w całej sieci (Tadeusiewicz, 1993, 2015). Proces uczenia się w tym przypadku miał za zadanie osiągnięcie rezultatów opartych na wiedzy fragmentarycznej, umożliwił doskonalenie się, tworzył nowe pojęcia oraz wnioskował indukcyjnie, zgodnie z założeniami teorii samouczących się algorytmów (Мелехин, 1984; Thacore, 1998; Ailon et al., 2011). Metody te można zasadniczo podzielić na nadzorowane (ang. *supervised learning*) i nienadzorowane (ang. *unsupervised*

learning), a także mniej powszechnie stosowane półnadzorowane (ang. *semi-supervised learning*) oraz uczenie przez wzmacnianie (ang. *reinforcement learning*). Uczenie nadzorowane wymaga posiadania referencyjnych danych wyjściowych. W dziedzinie uczenia maszynowego dominują dwa podstawowe rodzaje problemów: klasyfikacja oraz regresja. Klasyfikacja polega na przyporządkowywaniu danych do dwóch lub więcej kategorii, podczas gdy problem regresji zakłada identyfikację trendów i na ich podstawie uzupełnienie brakujących elementów lub wykonanie predykcji (Chollet, 2021). Do budowy modelu (wejście) zostały wykorzystane zdjęcia skaningowej mikroskopii elektronowej (SEM) dwuletnich igieł sosny zwyczajnej *Pinus sylvestris* L., wykazujących potwierdzone cechy wskaźnikowe (Eriksson, 1989; Alaimo et al., 2000; Romanić i Krauthacker, 2007; Cindrić et al., 2018; Szwed et al., 2021) ze zlewni badawczych krajowej sieci Zintegrowanego Monitoringu Środowiska Przyrodniczego (ZMŚP) (Kostrzewski et al., 2022). Stacje Bazowe ZMŚP zlokalizowane są w reprezentatywnych dla obszaru Polski typach krajobrazu i przedstawiają odrębne geoekosystemy: od wybrzeży i krajobrazów młodogłacjalnych na północy kraju, przez niziny i doliny rzeczne w środkowej części, aż po wyżyny i góry na południu (ryc. 1). Różnorodność krajobrazowa zlewni ZMŚP pozwala obserwować zmiany środowiska przyrodniczego w skali regionalnej i krajowej (Kostrzewski i Majewski, 2021).



Ryc. 1. Lokalizacja punktów poboru próbek środowiskowych
Location of environmental sample collection points

Miejsce poboru próbki	Współrzędne geograficzne
1. Beskid Niski	49°38' 03.01" N, 21°07' 02.23" E
2. Kampinos	52°17' 16.44" N, 20°26' 56.77" E
3. Karkonosze	50°48' 31.69" N, 15°36' 46.58" E
4. Łysogóry	50°52' 15.67" N, 21°03' 18.25" E
5. Parsęta	53°44' 52.67" N, 16°33' 13.27" E
6. Pogórze Karpackie	49°57' 55.26" N, 20°29' 54.60" E
7. Poznań-Morasko	50°27' 55.37" N, 16°55' 57.95" E
8. Puszcza Borecka	54°06' 47.78" N, 22°01' 12.74" E
9. Roztocze	50°35' 23.70" N, 22°59' 47.30" E
10. Wigry	54°03' 34.74" N, 23°00' 47.12" E
11. Wolin	53°56' 48.90" N, 14°28' 34.20" E

Metody badań i zastosowana aparatura

Próbki igieł pobrano z drzew w wybranych lokalizacjach z wysokości ok. 2-3 m n.p.g. w ilości około 200 g, po zakończeniu sezonu wegetacyjnego (X-XI 2023). Dostarczone w papierowych kopertach do Laboratorium Badań Środowiska (LBŚ) Uniwersytetu Jana Kochanowskiego w Kielcach (UJK) zostały wysuszone w temperaturze 40 °C przez 24 godziny, a następnie zmielone młynkiem analitycznym. Do mielenia wykorzystano młynek IKA A11 basic (Staufen, Niemcy). Do zmielonego igliwia dodano ultraczystą wodę (w stosunku 1:10), a następnie przeanalizowano pH oraz przewodnictwo elektrolityczne właściwe EC (z ang. *Electrical Conductivity*). Pomiaru dokonano z wykorzystaniem wieloparametrowego urządzenia HACH HQ2200 (Loveland, USA) wyposażonego w elektrodę pehametrycz-

ną PHC101 oraz konduktometryczną CDC401. Wyniki badań skontrolowano roztworami wzorcowymi Hamilton (Bonaduz, Szwajcaria) o wartościach pH 4,01 (Ref. No: 238917), 7,00 (Ref. No: 238218) i 9,21 (Ref. No: 238919). Do kontroli elektrody konduktometrycznej użyto dwóch wzorców: 15 $\mu\text{S}/\text{cm}$, CPAChem (Ref. No: CS15MOS.L5) 30% n-propanol (Stara Zagora, Bułgaria oraz 1000 $\mu\text{S}/\text{cm}$, YSI 3167 (Yellow Springs, USA) z chlorkiem potasu 0.053% (Ref. No: CAS 7447407).

Zdjęcia skaningowej mikroskopii elektronowej (SEM) zostały wykonane mikroskopem Quanta 250 FEI (FEI Company, Czechy) z analizą składu chemicznego wykonaną techniką EDS (z ang. *Energy Dispersive Spectroscopy*) mikroanalizatorem EDAX Genesis (Mahwah, NJ, USA) w LBŚ, UJK. Do mikroskopowania wybrano kilka egzemplarzy z każdej lokalizacji. Przyklejono je za pomocą krążków węglowych do stolików aluminiowych $\varnothing 12$ mm. Stoliki wraz z próbkami umieszczono w Leica EM SC050, a następnie napyłono 24 – karatowym złotem w atmosferze argonu z kontrolą grubości warstwy (10 nm Au) w układzie próżniowym (0,05 mBar) z wagą kwarcową. Po zakończonym napyłaniu próbki zostały umieszczone w skaningowym mikroskopie elektronowym SEM, FEI Quanta 250, a następnie poddane analizie chemicznej EDS (ang. *Energy Dispersive Spectroscopy*) z wykorzystaniem analizatora EDAX Genesis (spektrometr długości fali promieniowania rentgenowskiego). Obrazy graficzne uzyskane w wyniku analizy EDS powstają równolegle z tworzeniem obrazu SEM powierzchni badanej próbki. Informacje o składzie powierzchni i rozmieszczeniu poszczególnych elementów są wynikiem pomiarów intensywności i energii promieniowania rentgenowskiego, które jest efektem wzbudzania elektronów rdzeni elektronowych pierwiastków znajdujących się na powierzchni lub blisko powierzchni badanego materiału. Elektrony służące do tworzenia obrazu SEM (i jednocześnie uzyskiwania danych EDS) wnikały w materiał na różne głębokości w zależności od cech matrycy, ale nie więcej niż na kilka mikrometrów. Pierwotna wiązka elektronów wnika w warstwę wierzchnią materiału i wzbudza w niej różne sygnały pochodzące z badanej warstwy. Wzbudzony i zanalizowany sygnał elektronów wtórnych SE (z ang. *Secondary Electrons*) daje możliwość obrazowania obserwowanej powierzchni. Dodatkowo, na podstawie surowych danych dotyczących widma X przypisanego do każdego piksela obrazu SEM można uzyskać dane ilościowe, pozwalające na porównanie udziałów masowych pierwiastków w przypowierzchniowej warstwie skanowanego obszaru. W tym celu wybierano do tej analizy cały obszar, a następnie wybrane punkty (tab. 3), wewnątrz którego całkowano otrzymane dane o intensywności promieniowania X dla poszczególnych długości fali. Następnie dane po scałkowaniu przeliczano na ilości pierwiastków i dalej wyniki były normalizowane tak, aby suma udziałów wynosiła 100%. Przy interpretacji tak uzyskanych wyników ilościowych należy wziąć pod uwagę charakter badanego materiału, morfologię i źródła pochodzenia poszczególnych składników próbki jak i charakter samej metody EDS. Wykonane obrazy SEM przetworzono w programie graficznym (GIMP wersja 2.10.34), tak, aby zaklasyfikowane na podstawie wielkości, kształtu i składu chemicznego cząstki posiadały jednakowy atrybut (barwę). Wykonane warstwy stanowiły element właściwy do opracowania algorytmu uczenia maszynowego identyfikującego zanieczyszczenia z podziałem na zdefiniowane wcześniej kategorie (barwy). Zieloną barwą zamalowano powierzchnię aparatów szparkowych, żółtą – ostrokrawędziste cząstki przemysłowe, czerwoną – magnetyczne sferule, a niebieską – drobne cząstki mineralne.

Opis modelu

Pierwsza warstwa budowanego systemu identyfikacji zanieczyszczeń powietrza zawiera 32 filtry o rozmiarze 3x3 i korzysta z funkcji aktywacji ReLU, wprowadzającej nieliniowość do modelu, co pozwala na efektywne modelowanie złożonych zależności między cechami na zdjęciach, a następnie jest aktywowany proces poolingowy, który redukuje wymiarowość danych. Ten schemat powtarza się w kolejnych warstwach, zwiększając liczbę filtrów w każdej z nich. Po spłaszczeniu danych następuje warstwa gęsta z 128 neuronami aktywowanymi funkcją ReLU, z dodatkową warstwą Dropout, która losowo wyłącza części neuronów w trakcie uczenia, co ma na celu zmniejszenie zjawiska przeuczenia poprzez uniemożliwienie sieci polegania na pojedynczych neuronach lub ich kombinacjach (Kurenkov, 2020). Model kończy się warstwą wyjściową z 3 kategoriami zanieczyszczeń oraz funkcją aktywacji sigmoidalną, która przekształca wyniki modelu na wartości z zakresu [0, 1], co interpretowane jest jako prawdopodobieństwo przynależności do poszczególnych klas. Aby przystąpić do trenowania modelu konwolucyjnej sieci neuronowej, skonfigurowano proces optymalizacji oraz zdefiniowano funkcję straty i metryki oceny wydajności modelu. Optymalizator Adam (*Adaptive Moment Estimation*) wykorzystuje stochastyczną, adaptacyjną metodę optymalizacji, która dostosowuje współczynniki uczenia dla każdego parametru na podstawie oszacowanego pierwszego momentu (średnia ruchoma gradientu) oraz drugiego momentu (średnia ruchoma kwadratów gradientu). Wzory aktualizacji parametrów zostały zaprezentowane poniżej (Kingma i Ba, 2014):

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$$

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{\hat{v}_t} + \epsilon} \cdot \hat{m}_t$$

gdzie:

θ_t – parametr modelu w chwili t ,

g_t – gradient funkcji straty względem parametru w chwili t ,

m_t i v_t – pierwszy i drugi moment,

\hat{m}_t i \hat{v}_t – korygowane momenty,

α – współczynnik uczenia (ang. learning rate),

β_1 i β_2 – współczynniki wygładzania dla momentów,

ϵ – bardzo mała liczba zapobiegająca dzieleniu przez zero,

t – numer iteracji.

Funkcja straty to rzadka krzyżowa entropia kategoriowa (ang. *sparse categorical cross-entropy*), co jest odpowiednie w przypadku klasyfikacji wieloklasowej, gdzie etykiety są podane w postaci liczb całkowitych. Wzór na stratę można zapisać jako (Russakovsky et al., 2015):

$$L(y, y^{\wedge}) = -\frac{1}{N} \sum_{i=1}^N y_i \log(y^{\wedge}_i)$$

gdzie:

N – liczbę przykładów,

y – rzeczywiste etykiety,

y^{\wedge} – przewidywane etykiety.

Dokładność jest obliczana jako stosunek liczby poprawnie sklasyfikowanych przykładów do liczby wszystkich przykładów:

$$accuracy = \frac{\text{liczba poprawnie sklasyfikowanych przykładów}}{\text{łączna liczba przykładów}}$$

Odpowiedź sieci oscyluje na poziomach pomiędzy 80% a 90%, co stanowi optymalny wynik opracowanego systemu. W następnym kroku stosowane jest progowanie dla kolorów czerwonego, żółtego i zielonego w przestrzeni kolorów BGR. Znalezione maski są wykorzystywane do identyfikacji konturów, które są następnie przeskalowane i poddane operacji progowania w celu poprawy jakości segmentacji. Następnie funkcja identyfikuje kontury dla każdego koloru i oblicza współrzędne graniczne obszarów na obrazie. Dla każdego obszaru obrazu, w którym wykryto zanieczyszczenie, obliczana jest predykcja za pomocą zadanego modelu. Jeśli wartość ufności predykcji przekracza ustalony próg, współrzędne oraz przypisany poziom zanieczyszczenia są zapisywane. Ostatecznie funkcja zwraca listę zanieczyszczeń w postaci krotek, zawierających współrzędne graniczne które są naniesione na nowy obraz (ryc. 5).

Wyniki badań

Zebrane igliwie po wysuszeniu w szafie termostatycznej w temperaturze 40°C przez 24 godziny zostały poddane analizom. W zmielonych igłach oznaczono pH_{H_2O} i EC. Średnia wartość pH_{H_2O} wyniosła 5,42 z minimum pH_{H_2O} 4,95 (Łysogóry) i maksimum pH_{H_2O} 5,80 (Wolin). Wartości EC mieściły się w przedziale od 578 $\mu S/cm$ (Roztocze) do 1067 $\mu S/cm$ (Pogórze Karpackie) ze średnią 722 $\mu S/cm$ (tab. 1).

Analiza powierzchni igieł, pozwoliła na określenie składu pierwiastkowego (analiza jakościowa) oraz procentowego udziału mikroskładników (analiza ilościowa) zdeponowanych cząstek (ryc. 2-4).

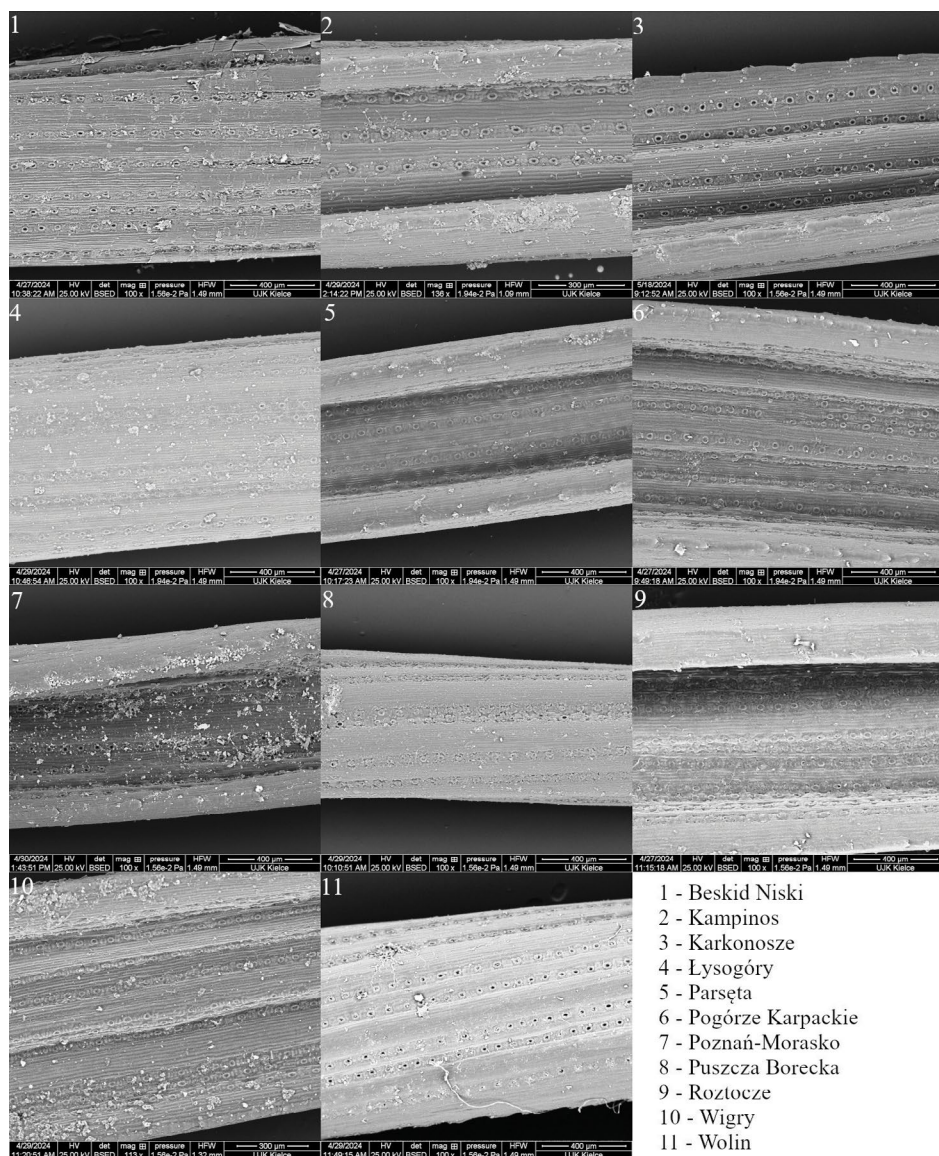
Powiększenie wykonanych mikrofotografii SEM (ok. $\times 100$) dobrano w taki sposób, aby zaprezentować cały przekrój igły z widocznymi rzędami aparatów szparkowych oraz cząstkami stałymi o zróżnicowanym kształcie i rozmiarze. W kolejnym etapie wykonano analizę SEM/EDS w znacznie większym powiększeniu powierzchniową ($\times 600$, ryc. 3) i punktową (powyżej $\times 1200$, ryc. 4) EDS w celu przygotowania materiału wejściowego do treningu samouczącego algorytmu maszynowego.

Tabela 1. Właściwości fizyczno-chemiczne badanych igieł sosny
Physico-chemical properties of the analyzed pine needles

Stacja bazowa ZMŚP	pH	EC ($\mu\text{S}/\text{cm}$)
1. Beskid Niski	5,43	840
2. Kampinos	5,55	581
3. Karkonosze	5,67	628
4. Łysogóry	4,95	625
5. Parsęta	5,51	754
6. Pogórze Karpackie	5,32	1067
7. Poznań-Morasko	5,65	751
8. Puszcza Borecka	5,14	616
9. Rostocze	5,58	578
10. Wigry	5,04	831
11. Wolin	5,80	672

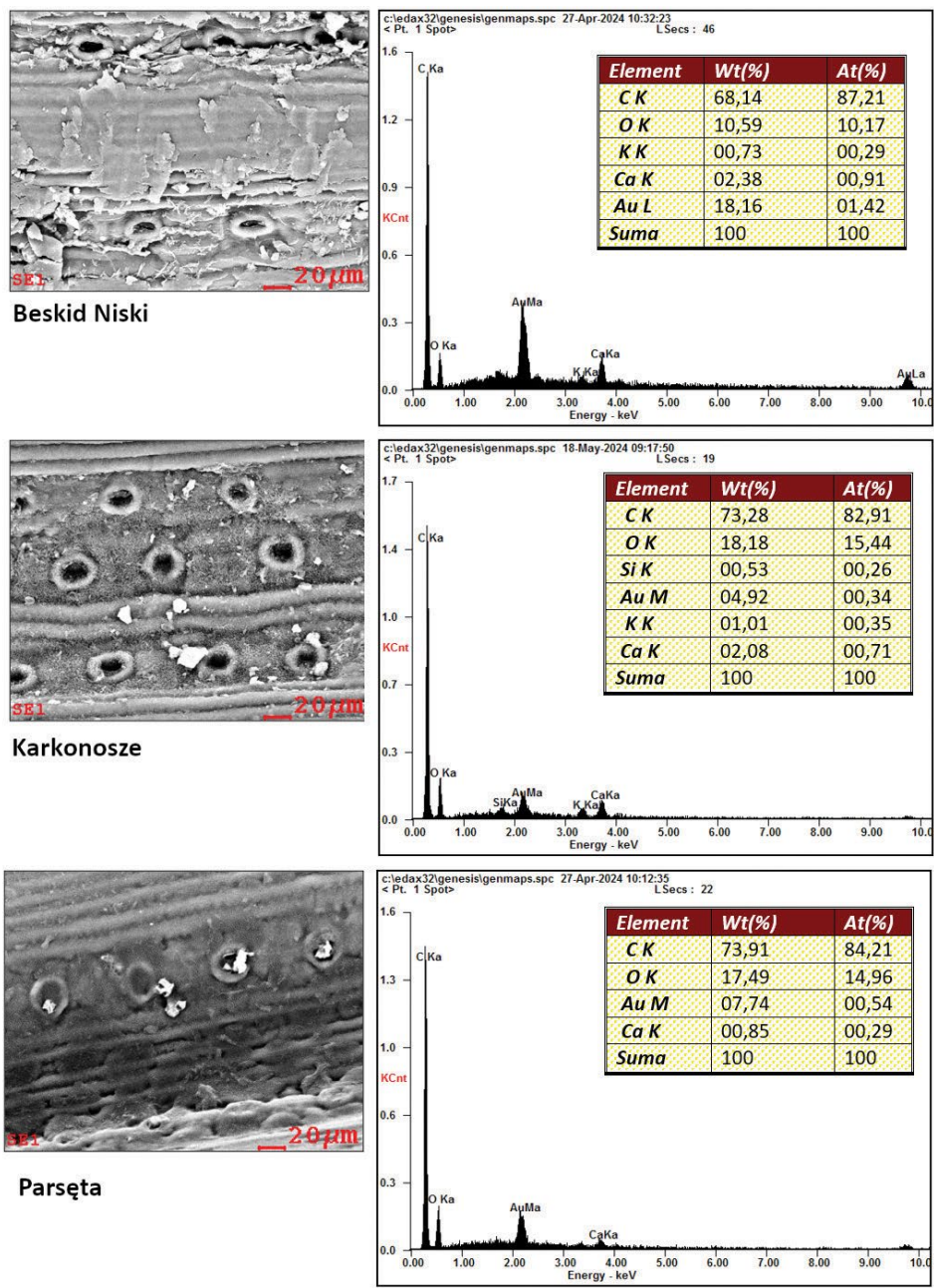
Powierzchnia analizowanych igieł sosny wykazała zróżnicowany skład ilościowy i jakościowy zobrazowanych cząstek w zależności od lokalizacji punktu pobrania próbek. Do budowy algorytmu wykorzystano zdjęcia SEM w powiększeniu $\times 600$ z wykonaną analizą EDS. Zawartość C wahała się od 51,8% wag. (Poznań-Morasko) do 73,9% wag. (Parsęta), przy średniej dla wszystkich próbek wynoszącej 66,1% wag. Koncentracja O zawierała się w przedziale od 10,6% wag do 18,8% wag., ze średnią 14,7%. Najwyższą zawartość Ca (2,9% wag.) stwierdzono w próbce z Łysogór, a Fe (2,7% wag.) i Si (2,6% wag.) – w próbkach z Poznań-Morasko. Udział Na, Mg i Cl w analizowanych próbkach nie przekroczył 1% wag., a pierwiastki te zdiagnozowano tylko w próbkach z Poznania-Morasko. Analizując wybrane powierzchnie igieł w znacznie większych powiększeniach (powyżej $\times 3000$) dokładniej zobrazowano cząstki pochodzenia naturalnego i antropogenicznego o zróżnicowanych rozmiarach i kształtach (ryc. 4). Cząstki mineralne z przewagą Si i Al były obecne na igłach we wszystkich badanych lokalizacjach, głównie jako efekt wietrzenia skał i minerałów (Kabata-Pendias i Pendias, 1999; Hartmann et al., 2013). Z obszarami wydobywania skał węglanowych i ich przetwórstwa były związane cząstki o charakterystycznej ostrokrawędzistej budowie z przewagą Ca (Magiera et al., 2013), zobrazowane na próbkach z Łysogór. Na igłach z Wolina podobny kształt miały cząstki halitu z Na, Cl i K w składzie. Sferule żelazisto-krzemianowe były obecne we wszystkich próbkach. Skład chemiczny (EDS) cząstek pozwolił je pogrupować zgodnie z klasyfikacją Maňkovská et al. (2004):

1. cząstki mineralne (nieregularny kształt i zróżnicowana wielkość od 5 do 20 μm , w składzie Ca, Al, Si, Fe) – kolor niebieski;
2. cząstki wapienno-cementowe (większe konglomeraty wapienne zbudowane z Ca i Mg powyżej 20 μm) – kolor żółty;
3. magnetyczne sferule (kuliste, złożone z Fe, Al, Si, o średnicy poniżej 5 μm) – kolor czerwony;
4. powierzchnia aparatów szparkowych – kolor zielony.

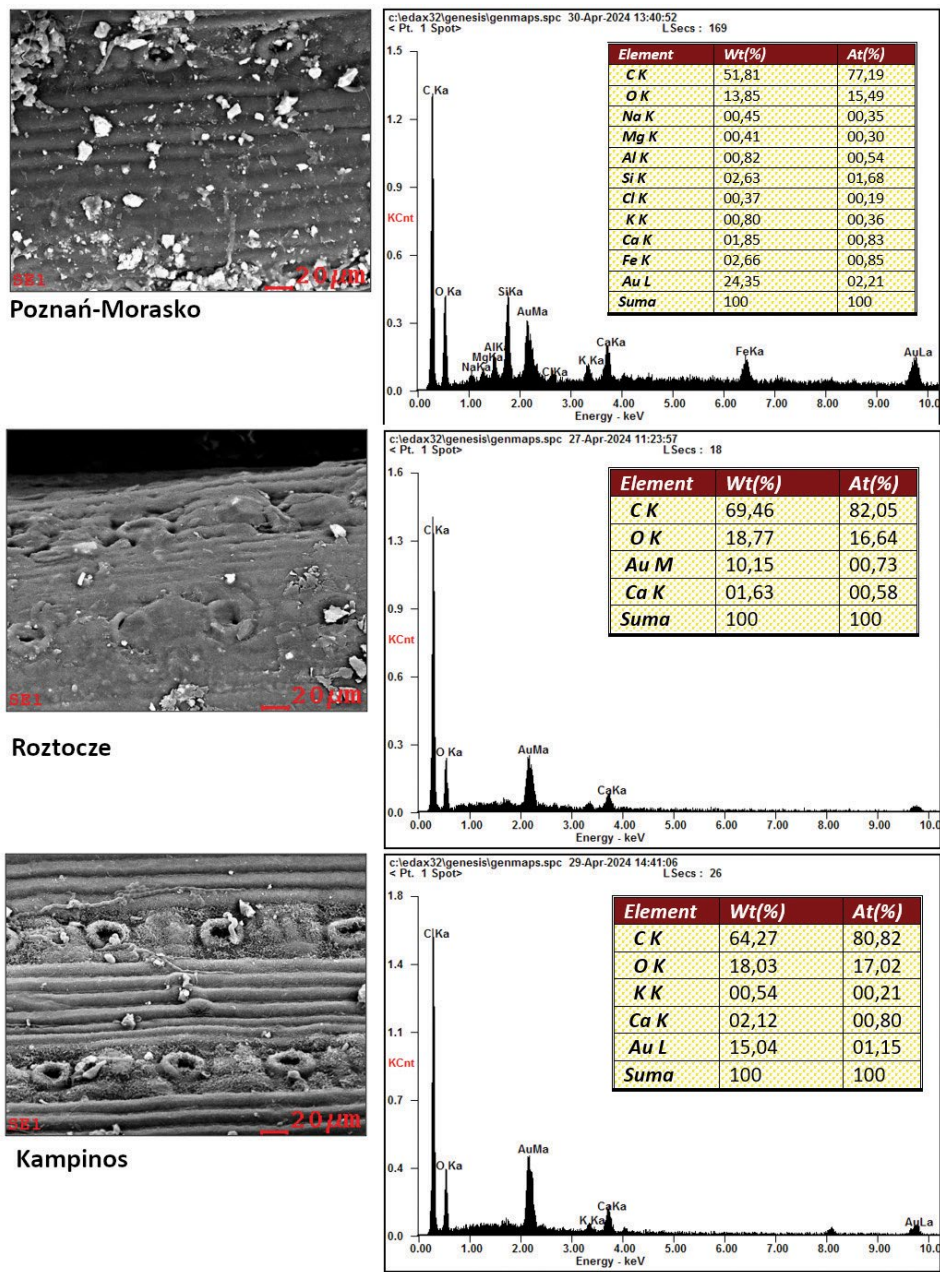


- 1 - Beskid Niski
- 2 - Kampinos
- 3 - Karkonosze
- 4 - Łysogóry
- 5 - Parsęta
- 6 - Pogórze Karpackie
- 7 - Poznań-Morasko
- 8 - Puszcza Borecka
- 9 - Roztocze
- 10 - Wigry
- 11 - Wolin

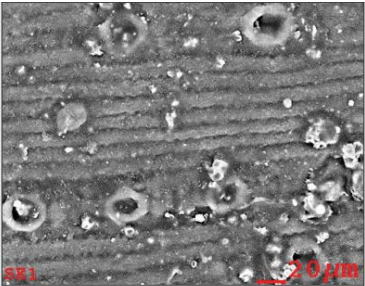
Ryc. 2. Zdjęcia SEM badanych igieł sosny
SEM images of selected pine needles



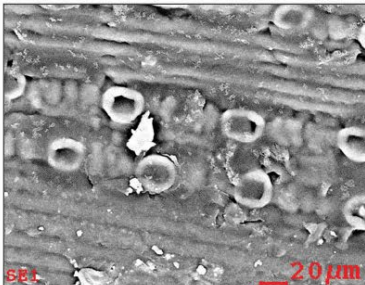
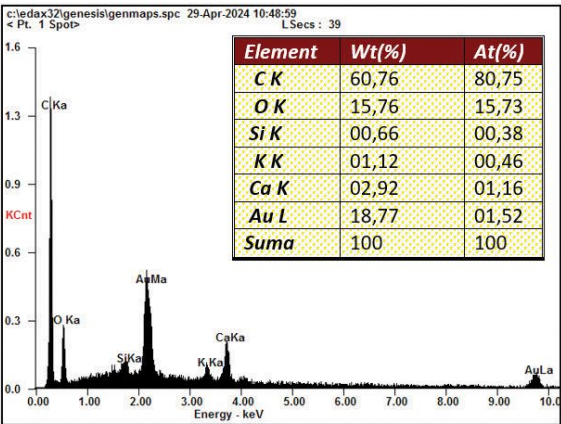
Ryc. 3. →



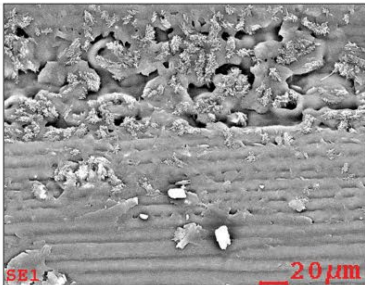
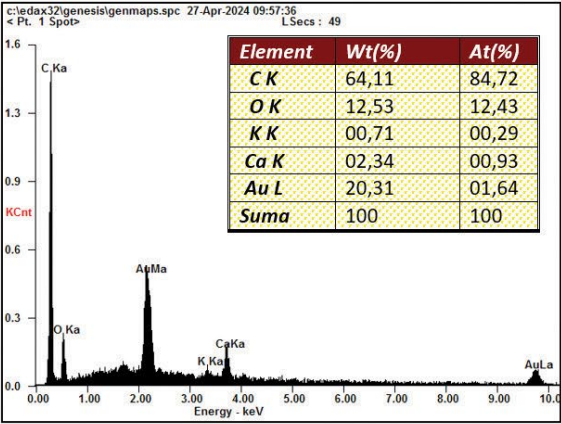
Ryc. 3. →



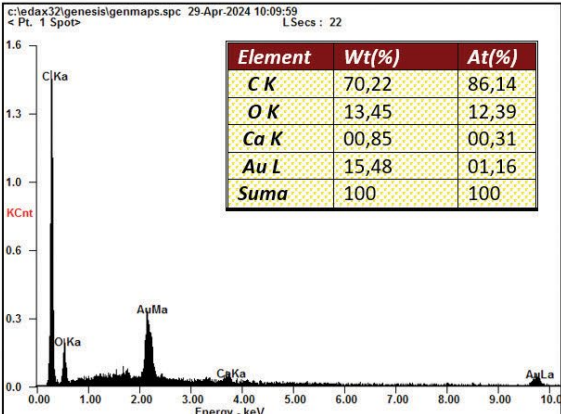
Łysogóry



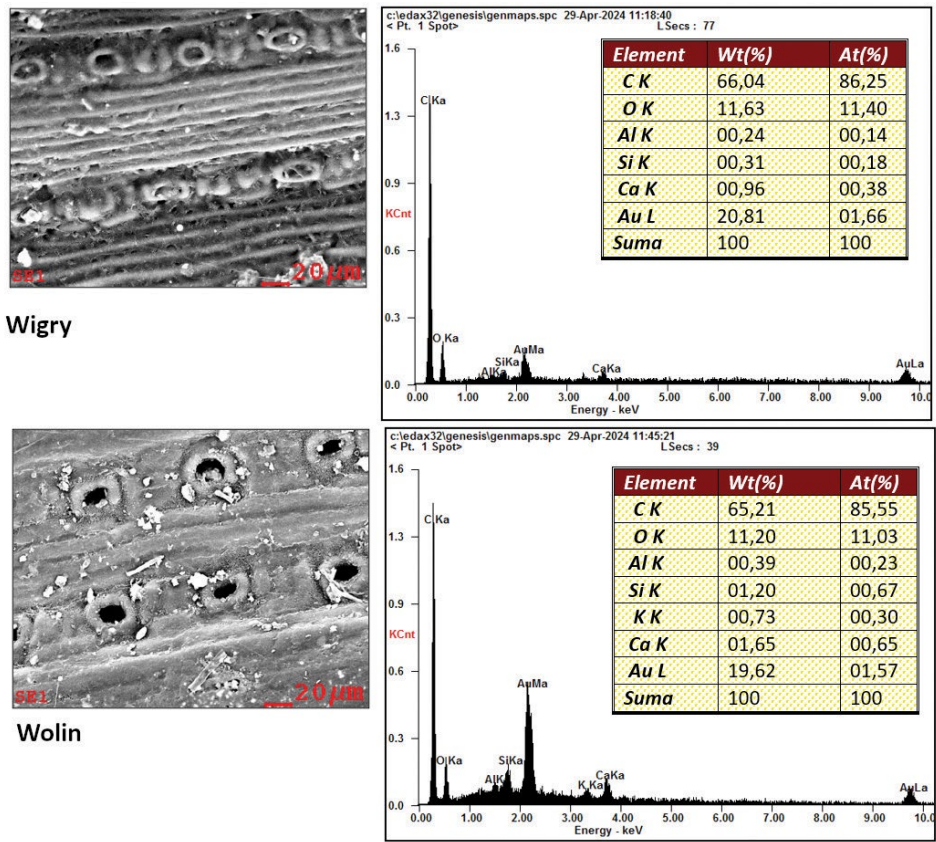
Pogórze Karpackie



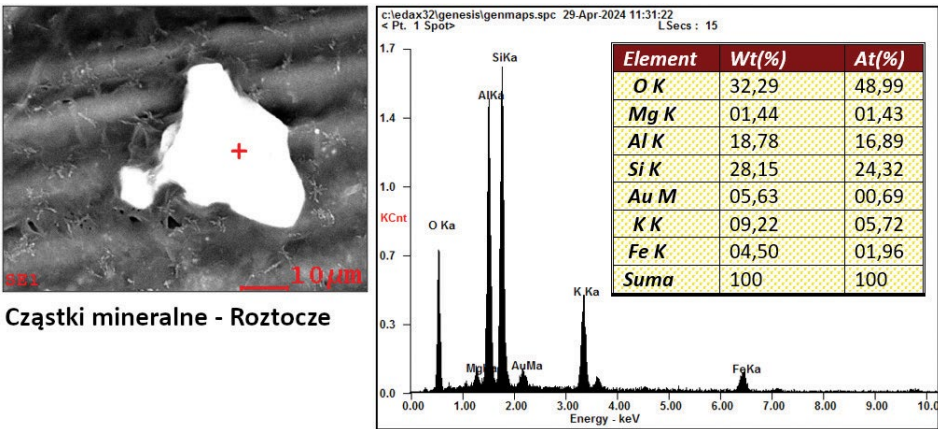
Puszcza Borecka



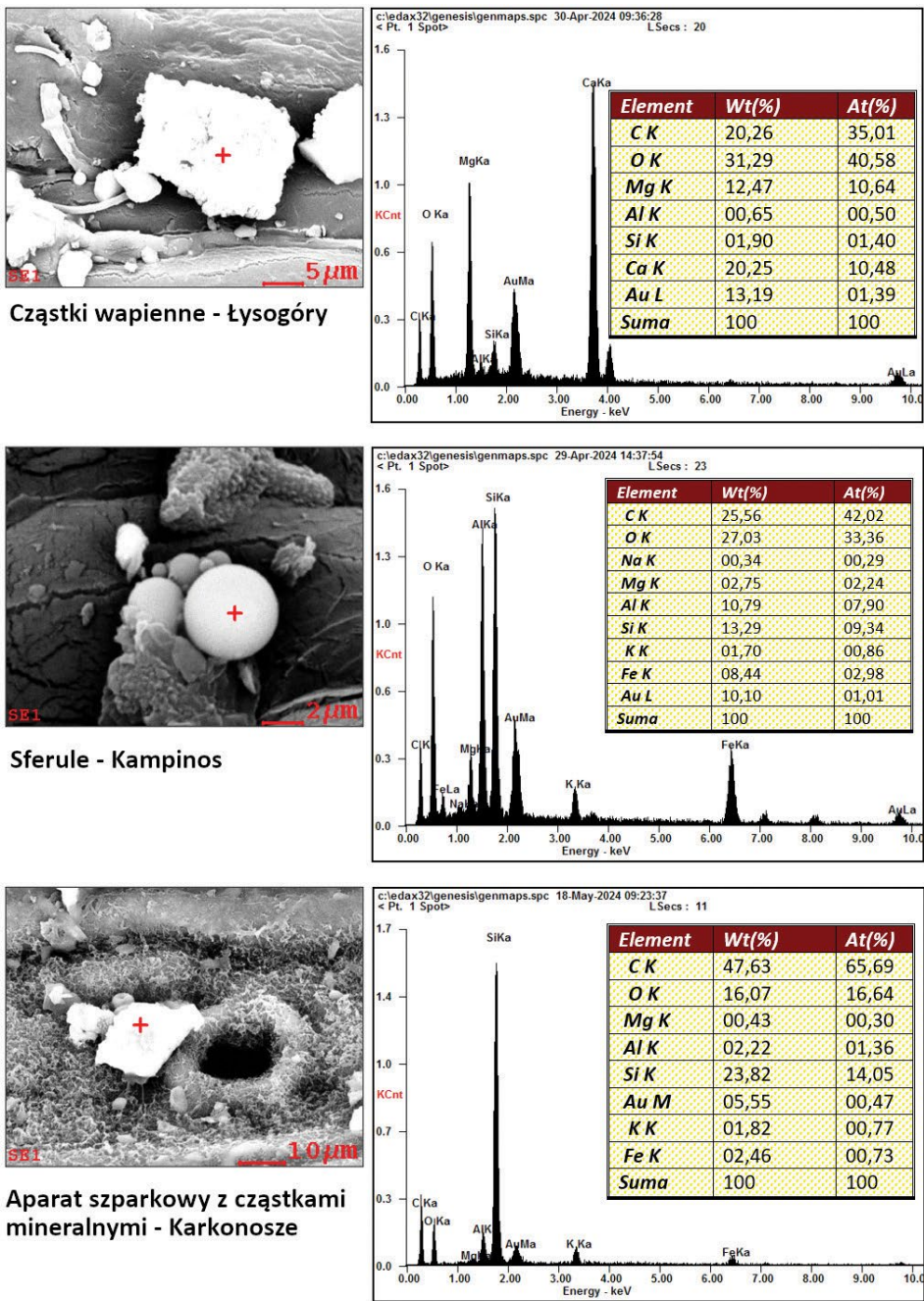
Ryc. 3. →



Ryc. 3. Analiza EDS wybranych igieł sosny (element - pierwiastki wykryte w próbce, Wt(%) - procent wagowy, At(%) - procent atomowy)
EDS analysis of selected pine needles (element - elements detected in the sample, Wt(%) - weight percent, At(%) - atomic percent)



Ryc. 4. →

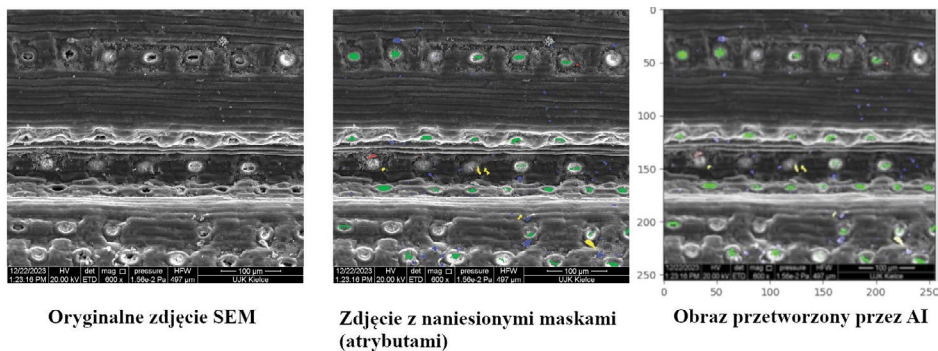


Ryc. 4. Wyniki punktowej analizy EDS (oznaczone czerwonym krzyżykiem). Objasnienia skrótów na rycinie- patrz ryc. 3
Results of EDS point analysis (marked with a red cross). For abbreviations, see Fig. 3

Dyskusja wyników

Określenie pochodzenia głównych składowych czynników obniżających jakość powietrza na danym terenie jest podstawowym zadaniem w skutecznej ochronie ludności i środowiska przed skutkami ewentualnego skażenia (Kozłowski et al., 2024; Świercz et al., 2024). Dostępność aparatury pomiarowej może stanowić w tym zakresie ograniczenie, szczególnie w przypadku zdarzeń o charakterze katastrofalnym (pożary, awarie, wycieki). Dlatego istotne są badania pośrednie oparte na analizach geo- i biowskaźników wrażliwych na zanieczyszczenie, takich jak opady atmosferyczne, mchy, porosty czy organy asymilacyjne (Szwed et al., 2020; Takano in., 2024). Metody te umożliwiają rozróżnienie naturalnych i antropogenicznych źródeł emisji zanieczyszczeń wykorzystując analizę składu chemicznego. Sukcesywny rozwój sztucznej inteligencji w wielu dziedzinach nauki pozwala na efektywny sposób zastosowania sieci neuronowych w identyfikacji zanieczyszczenia powietrza atmosferycznego. Zastosowanie maszynowego algorytmu klasyfikującego cząstki stałe na podstawie wyodrębnionych cech fizycznych (wielkość, kształt) oraz składu chemicznego pozwoliło na optymalizację procesu rozpoznawania potencjalnego zagrożenia. W pracy zaprezentowano zastosowanie sieci neuronowej opartej na maszynowym uczeniu do identyfikacji źródeł zanieczyszczeń powietrza na podstawie analizy mikrofotograficznej powierzchni igieł sosny zwyczajnej pobranych z lokalizacji o różnorodnym nasileniu presji antropogenicznej.

W pobranych igłach ze zlewni reprezentatywnych ZMŚP stwierdzono zróżnicowane właściwości fizyczno-chemiczne: odczyn i przewodnictwo elektrolityczne właściwe oraz skład chemiczny. Na powierzchni badanych organów asymilacyjnych zobrazowano cząstki zawierające: Ca (wszystkie stacje), K (Beskid Niski, Kampinos, Karkonosze, Łysogóry, Pogórze Karpackie, Poznań-Morasko, Wolin), Si (Karkonosze, Łysogóry, Poznań, Wigry, Wolin) i Al (Poznań-Morasko, Wigry, Wolin) oraz Cl, Fe, Mg i Na (zidentyfikowane tylko w próbkach z Poznania-Morasko). W znacznie większych powiększeniach zidentyfikowano sferyczne struktury oraz drobne cząstki mineralne. Ich wielkość umożliwia wnikanie do wnętrza aparatów szparkowych i oddziaływanie na procesy metaboliczne badanych drzew, w tym ich degradację. Opracowywany algorytm jest w stanie identyfikować charakterystyczne elementy z dużą dokładnością na nowych danych (ryc. 5). Podobne działanie ma metoda



Ryc. 5. Graficzny efekt działania algorytmu
Graphical effect of the algorithm

i system detekcji zmian komórkowych oparty na działaniu głębokiej sieci neuronowej. Był on już z powodzeniem wykorzystywany w badaniach biologicznych i medycznych (Chklovskii et al., 2010; Jones et al., 2013; Cetina et al., 2018, Ikeno et al., 2018) oraz analizach przestrzennych (Benbahria et al., 2021).

Podsumowanie

Zróznicowanie lokalizacji zlewni Stacji Bazowych daje podstawę do ocen, porównań i możliwości wszechstronnych badań reakcji środowiska przyrodniczego na różnokierunkową antropopresję. Przeprowadzone badania (pH, EC, skład chemiczny) potwierdziły odmienny charakter tego oddziaływania na organy asymilacyjne jakimi, były igły sosny. Analiza SEM/EDS wykazała znaczne zróznicowanie i nagromadzenie zidentyfikowanych na powierzchni próbek cząstek stałych. Zastosowanie sieci neuronowych do budowy samouczącego się algorytmu pozwoliło zoptymalizować analizę zdeponowanych zanieczyszczeń. Dzięki zastosowaniu wielu warstw konwolucyjnych, sieć neuronowa podczas treningu wychwytiła z obrazu najistotniejsze cechy, aby następnie używać ich do przewidywania masek segmentacyjnych interesujących obiektów. Na podstawie skojarzenia pikseli wejściowych i wyodrębnionych z nich cech oraz pikseli prawdziwej maski segmentacyjnej, algorytm dostosował swoje parametry tak, aby później odtwarzać maski dla zupełnie nowych danych wejściowych. Odpowiedź sieci wyniosła 80%, co stanowi optymalny wynik opracowanego systemu. Opracowany model stanowi innowacyjne rozwiązanie w zakresie diagnostyki igieł sosny. Wytrenowany algorytm ma uniwersalne zastosowanie, co oznacza, że może być wykorzystany do identyfikacji zanieczyszczeń na zdjęciach igieł pobranych z dowolnej lokalizacji. W kolejnym etapie badań zakłada się obliczenia poszczególnych składowych zanieczyszczeń służących identyfikacji źródeł emisji.

Prace sfinansowano ze środków Narodowego Centrum Nauki w ramach konkursu MINIA-TURA 7 (nr grantu 2023/07/X/ST10/00114). Prace dofinansowano ze środków Ministerstwa Nauki i Szkolnictwa Wyższego (nr tematu SUPB.RN.24.107).

Specjalne podziękowania dla Kierownictwa Zintegrowanego Monitoringu Środowiska Przyrodniczego oraz Pracowników Stacji Bazowych ZMŚP za pomoc w realizacji badań.

Ryciny i tabele, pod którymi nie zamieszczono źródeł, są opracowaniami własnymi Autora artykułu.

Piśmiennictwo

- Ailon, N., Chazelle, B., Clarkson, K.L., Liu, D., Mulzer, W., & Seshadhri, C. (2011). Self-Improving Algorithms. *SIAM Journal on Computing*, 40(2), 350-375. <https://doi.org/10.1137/090766437>
- Akimoto, H. (2003). Global air quality and pollution. *Science*, 302(5651), 1716-1719. <https://doi.org/10.1126/science.1092666>

- Alaimo, M.G., Dongarra, G., Melati, M.R., Monna, F., & Varrica, D. (2000). Recognition of environmental trace metal contamination using pine needles as bioindicators. The urban area of Palermo (Italy). *Environmental Geology*, 39(8), 914-924. <https://doi.org/10.1007/s002549900071>
- Allegrini, I., & Pirrone, N. (2007). 19 Global scale atmospheric pollution: a regional problem. W: R. Baudo, G. Tartari, & E. Vuillermoz (red.), *Developments in earth surface processes*, 10 (s. 145-146). [https://doi.org/10.1016/s0928-2025\(06\)10019-x](https://doi.org/10.1016/s0928-2025(06)10019-x)
- Arroyo, P., Gómez-Suárez, J., Suárez, J.I., & Lozano, J. (2021). Low-Cost Air Quality Measurement System Based on Electrochemical and PM Sensors with Cloud Connection. *Sensors*, 21(18), 6228. <https://doi.org/10.3390/s21186228>
- Ayaz, M., Tasdemirci, E., Yuksel, H., & Aygul, V. (2018). Comprehensive study on automation of toxic gas measurement. *Arabian Journal of Geosciences*, 11(22). <https://doi.org/10.1007/s12517-018-4066-5>
- Benbahria, Z., Sebari, I., Hajji, H., & Smiej, M.F. (2021). Intelligent mapping of irrigated areas from Landsat 8 images using transfer learning. *International Journal of Engineering and Geosciences*, 6(1), 40-50. <https://doi.org/10.26833/ijeg.681312>
- Bousiotis, D., Singh, A., Haugen, M., Beddows, D.C.S., Diez, S., Murphy, K.L., Edwards, P.M., Boies, A., Harrison, R.M., & Pope, F.D. (2021). Assessing the sources of particles at an urban background site using both regulatory instruments and low-cost sensors – a comparative study. *Atmospheric Measurement Techniques*, 14(6), 4139-4155. <https://doi.org/10.5194/amt-14-4139-2021>
- Cetina, K., Buenaposada, J.M., & Baumela, L. (2018). Multi-class segmentation of neuronal structures in electron microscopy images. *BMC Bioinformatics*, 19(1). <https://doi.org/10.1186/s12859-018-2305-0>
- Chklovskii, D.B., Vitaladevuni, S., & Scheffer, L.K. (2010). Semi-automated reconstruction of neural circuits using electron microscopy. *Current Opinion in Neurobiology*, 20(5), 667-675. <https://doi.org/10.1016/j.conb.2010.08.002>
- Chollet, F. (2021). *Deep Learning with Python, Second Edition*. Simon and Schuster.
- Cindrić, I.J., Zeiner, M., Starčević, A., & Stinger, G. (2018). Metals in pine needles: characterisation of bio-indicators depending on species. *International Journal of Environmental Science and Technology*, 16(8), 4339-4346. <https://doi.org/10.1007/s13762-018-2096-x>
- Eriksson, G., Jensen, S., Kylin, H., & Strachan, W. (1989). The pine needle as a monitor of atmospheric pollution. *Nature*, 341(6237), 42-44. <https://doi.org/10.1038/341042a0>
- Fan, A.M. (2014). Biomarkers in toxicology, risk assessment, and environmental chemical regulations. W: R.C. Gupta (red.), *Elsevier eBooks* (s. 1057-1080). <https://doi.org/10.1016/b978-0-12-404630-6.00064-6>
- Fowler, D., Brimblecombe, P., Burrows, J., Heal, M.R., Grennfelt, P., Stevenson, D.S., Jowett, A., Nemitz, E., Coyle, M., Liu, X., Chang, Y., Fuller, G.W., Sutton, M.A., Klimont, Z., Unsworth, M.H., & Vieno, M. (2020). A chronology of global air quality. *Philosophical Transactions – Royal Society. Mathematical, Physical and Engineering Sciences/Philosophical Transactions – Royal Society. Mathematical, Physical and Engineering Sciences*, 378(2183), 20190314. <https://doi.org/10.1098/rsta.2019.0314>
- Hagan, D.H., & Kroll, J.H. (2020). Assessing the accuracy of low-cost optical particle sensors using a physics-based approach. *Atmospheric Measurement Techniques*, 13, 6343-6355. <https://doi.org/10.5194/amt-13-6343-2020>

- Hartmann, J., West, A.J., Renforth, P., Köhler, P., De La Rocha, C.L., Wolf-Gladrow, D.A., Dürr, H.H., & Scheffran, J. (2013). Enhanced chemical weathering as a geoengineering strategy to reduce atmospheric carbon dioxide, supply nutrients, and mitigate ocean acidification. *Reviews of Geophysics*, 51(2), 113-149. <https://doi.org/10.1002/rog.20004>
- Huszar, P., Karlický, J., Marková, J., Nováková, T., Liaskoni, M., & Bartík, L. (2021). The regional impact of urban emissions on air quality in Europe: the role of the urban canopy effects. *Atmospheric Chemistry and Physics*, 21(18), 14309-14332. <https://doi.org/10.5194/acp-21-14309-2021>
- Ikeno, H., Kumaraswamy, A., Kai, K., Wachtler, T., & Ai, H. (2018). A segmentation scheme for complex neuronal arbors and application to vibration sensitive neurons in the honeybee brain. *Frontiers in Neuroinformatics*, 12. <https://doi.org/10.3389/fninf.2018.00061>
- Jiménez, E., Tapiador, F.J., & Sáez-Martínez, F.J. (2014). Atmospheric pollutants in a changing environment: key issues in reactivity and monitoring, global warming, and health. *Environmental Science and Pollution Research International*, 22(7), 4789-4792. <https://doi.org/10.1007/s11356-014-3850-3>
- Jones, C., Sayedhosseini, M., Ellisman, M., & Tasdizen, T. (2013). Neuron Segmentation in Electron Microscopy Images Using Partial Differential Equations. *2013 IEEE 10th International Symposium on Biomedical Imaging, San Francisco, CA, USA, 2013*, 1457-1460. <https://doi.org/10.1109/ISBI.2013.6556809>
- Kabata-Pendias, A., & Pendias, H. (1999). *Biogeochemia pierwiastków śladowych*. Warszawa: Wydawnictwo PWN.
- Keywood, M., Paton-Walsh, C., Lawrence, M., George, C., Formenti, P., Schofield, R., Cleugh, H., Borgford-Parnell, N., & Capon, A. (2023). Atmospheric goals for sustainable development. *Science*, 379(6629), 246-247. <https://doi.org/10.1126/science.adg2495>
- Kingma, D.P., & Ba, J.L. (2014). Adam: A method for stochastic optimization. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1412.6980>
- Konwencja LRTAP. 1979. Konwencja w sprawie transgranicznego zanieczyszczania powietrza na dalekie odległości, sporządzona w Genewie dnia 13 listopada 1979 r. (Dz.U. z 1985 r. Nr 60, poz. 311).
- Kostrzewski, A., & Majewski, M. (2021). *Zintegrowany monitoring środowiska przyrodniczego: organizacja, system pomiarowy, metody badań, wytyczne do realizacji*. Warszawa: Biblioteka Monitoringu Środowiska.
- Kostrzewski, A., Majewski, M., & Szpikowski, J. (2022). *Współczesne przemiany naturalne i antropogeniczne środowiska przyrodniczego zlewni rzecznych i jeziornych*. Storkowo: Biblioteka Monitoringu Środowiska.
- Kozáková, J., Pokorná, P., Vodička, P., Ondráčková, L., Ondráček, J., Křůmal, K., Mikuška, P., Hovorka, J., Moravec, P., & Schwarz, J. (2018). The influence of local emissions and regional air pollution transport on a European air pollution hot spot. *Environmental Science and Pollution Research International*, 26(2), 1675-692. <https://doi.org/10.1007/s11356-018-3670-y>
- Kozłowski, R., Szwed, M., Kozłowska, A., Przybylska, J., & Mach, T. (2024). Quality Management System in Air Quality Measurements for Sustainable Development. *Sustainability*, 16(17), 7537. <https://doi.org/10.3390/su16177537>
- Kurenkov, A. (2020). A Brief History of Neural Nets and Deep Learning. Skynet Today. <https://skynet-today.com/overviews/neural-net-history>
- Magiera, T., Gołuchowska, B., Jabłońska, M. (2013). Technogenic magnetic particles in alkaline dusts from power and cement plants. *Water Air & Soil Pollution*, 224 (1389), 1-17. <https://doi.org/10.1007/s11270-012-1389-9>

- Maňkovská, B., Godzik, B., Badea, O., Shparyk, Y., & Moravčík, P. (2004). Chemical and morphological characteristics of key tree species of the Carpathian Mountains. *Environmental Pollution*, 130(1), 41-54. <https://doi.org/10.1016/j.envpol.2003.10.020>
- Мелехин, В.Б. (1984). Self-learning algorithm for an integrated robot with active and passive behavioral logic. *Cybernetics*, 20(4), 600-606. <https://doi.org/10.1007/bf01068936>
- Romanič, S., & Krauthacker, B. (2007). Are pine needles bioindicators of air pollution? Comparison of organochlorine compound levels in pine needles and ambient air. *Arhiv Za Higijenu Rada I Toksikologiju*, 58(2), 195-199. <https://doi.org/10.2478/v10004-007-0012-8>
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211-252. <https://doi.org/10.1007/s11263-015-0816-y>
- Samet, J.M., & Chung, Y.S. (2017). Air Quality, Atmosphere & Health: the 10-year anniversary. *Air Quality, Atmosphere & Health*, 11(1), 1-2. <https://doi.org/10.1007/s11869-017-0541-5>
- Shaddick, G., Salter, J.M., Peuch, V., Ruggeri, G., Thomas, M.L., Mudu, P., Tarasova, O., Baklanov, A., & Gumy, S. (2020). Global Air Quality: An Inter-Disciplinary Approach to exposure assessment for burden of disease analyses. *Atmosphere*, 12(1), 48. <https://doi.org/10.3390/atmos12010048>
- Susanto, A.D. (2020). Air pollution and human health. *Medical Journal of Indonesia*, 29(1), 8-10. <https://doi.org/10.13181/mji.com.204572>
- Szwed, M., Kozłowski, R., Żukowski, W. (2020). Assessment of Air Quality in the South-Western Part of the Świętokrzyskie Mountains Based on Selected Indicators. *Forests*, 11, 499. <https://doi.org/10.3390/f11050499>.
- Szwed, M., Żukowski, W., & Kozłowski, R. (2021). The Presence of Selected Elements in the Microscopic Image of Pine Needles as an Effect of Cement and Lime Pressure within the Region of Białe Zagłębie (Central Europe). *Toxics*, 9(1), 15. <https://doi.org/10.3390/toxics9010015>
- Świercz, A., Szwed, M., Bąk, Ł. (2024). Environmental consequences of a galvanising plant fire. *Journal of Water and Land Development*, 62(7-9), 1-9. <https://doi.org/10.24425/jwld.2024.151552>
- Tadeusiewicz, R. (1993). Sieci neuronowe. Warszawa: Akademicka Oficyna Wydawnicza
- Tadeusiewicz, R. (2015). Neural networks as a tool for modeling of biological systems. *Bio-Algorithms & Med-Systems (Online)/Bio-Algorithms and Med-Systems*, 11(3), 135-144. <https://doi.org/10.1515/bams-2015-0021>
- Takano, A.P.C., Rybak, J., Veras, M.M. (2024). Bioindicators and Human Biomarkers as Alternative Approaches for Cost-Effective Assessment of Air Pollution Exposure. *Frontiers in Environmental Engineering*, 3. <https://doi.org/10.3389/fenve.2024.1346863>.
- Thacore, S. (1998). An evolutionary self-learning methodology: Some preliminary results from a case study. In *Lecture notes in computer science* (pp. 387-396). <https://doi.org/10.1007/bfb0040791>
- WHO. 2021. WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. Geneva: World Health Organization. Pobrane z: <https://www.who.int/publications/i/item/9789240034228> (19.05.2024).

Summary

The primary objective of this study is to develop an intelligent system for air pollution identification using neural networks. The implementation of artificial intelligence (AI), leveraging the analysis of surface images of selected air pollution indicators, has enabled the creation of a low-cost-effective and efficient method for detecting hazardous substances. The study employs scanning electron microscopy (SEM) images of two-year-old Scots pine needles (*Pinus sylvestris* L.) collected from representative research catchments of the national Integrated Environmental Monitoring network. These images serve as input data for a machine learning algorithm designed to classify and segment air pollution particles based on predefined attributes such as size, shape, and chemical composition. To prepare the data for machine learning, SEM images were processed using graphic software, where classified particles were assigned distinct colours, corresponding to their characteristics. The processed layers, referred to as masks, became essential components in training a deep learning model to automatically recognize and categorize pollutants. The key innovation of this study lies in the use of a self-learning algorithm, which optimizes the analysis of contaminants deposited on the surface of pine needles, offering a reliable approach to air pollution assessment without the need for costly traditional measurement devices. The neural network employed in this study was structured with multiple convolutional layers, allowing it to capture intricate details within the images during training. These layers extract essential features from the images and apply them to generate segmentation masks that highlight the presence of pollution particles. By associating input pixels with their corresponding features and comparing them to pre-labeled masks, the model continuously adjusts its parameters, thereby improving its predictive accuracy over time. The system achieved an 80% prediction accuracy, demonstrating its potential as a reliable tool for identifying air pollutants. The analysis of pine needles serves as a bioindicator-based approach to monitoring air pollution, complementing existing measurement techniques. The pine needles, sampled from various locations experiencing different levels of anthropogenic pressure, provide insight into the regional variations in pollution levels. The chemical composition of the particles adhered to the needles was analysed using Energy Dispersive Spectroscopy (EDS), which allowed for a more detailed understanding of the pollutants' origins. The study identified key contaminants such as mineral dust, industrial particulate matter, metallic spherules, and fine atmospheric aerosols. One of the most significant findings of this research is the ability of the AI-powered system to differentiate between natural and anthropogenic pollution sources. The model was trained to recognize four distinct categories of pollution, each represented by a specific colour in the processed masks: green for stomatal surfaces, yellow for sharp-edged industrial particles, red for magnetic spherules, and blue for fine mineral dust. This classification system ensures precise pollutant identification, which is critical for environmental monitoring and policy-making. The integration of deep learning techniques into environmental studies marks a shift toward more advanced and automated pollution assessment methods. Traditional air quality monitoring relies on complex and expensive measurement infrastructure, whereas the proposed AI-based approach provides a low-cost and scalable alternative. The ability of the system to analyse new datasets without requiring manual preprocessing highlights its practical applicability for real-world environmental monitoring. In future research, additional efforts will focus on refining the model to further improve its accuracy.

and expand its application to other vegetation types and environmental indicators. The next phase of development will also explore the quantification of specific pollutants within each identified category, enabling a more precise determination of pollution sources. By incorporating additional datasets and refining neural network architectures, this AI-driven approach could play a vital role in global air quality monitoring initiatives.